A Bayesian Network Model of Pilot Response to TCAS Resolution Advisories

Edward H. Londner and Robert J. Moss
Massachusetts Institute of Technology Lincoln Laboratory
Lexington, Massachusetts, USA
elondner@ll.mit.edu, robert.moss@ll.mit.edu

***Abstract*—The effectiveness of an airborne collision avoidance system (CAS) is influenced by the manner in which pilots respond to the system's advisories. Current pilot response models used in CAS modeling and simulation are agnostic to parameters affecting pilot response in individual encounters and therefore treat all encounters equally. Simulations using these models can potentially underestimate collision risk in encounters where pilot response probability is low. This paper proposes a parametric pilot response model built from operational data using Bayesian networks. A network was constructed from radar recordings of TCAS encounters and the encounter parameters with the strongest influence on pilot response were identified. These parameters can be used to predict the probability of pilot response for individual encounters. The model was employed in simulation of safety-critical encounters. Results showed that standard pilot response models may underestimate collision risk. These results have implications for the design and performance evaluation of separation advisory systems, including collision avoidance and detect and avoid systems.**

***Key Words*—Traffic Alert and Collision Avoidance System (TCAS), Airborne Collision Avoidance System (ACAS), detect and avoid (DAA), aviation safety, Bayesian networks, pilot response, aircraft separation**

I. Introduction

TCAS is an airborne collision avoidance system (CAS) mandated worldwide on board all large passenger and cargo aircraft [1]. TCAS mitigates collision risk by surveilling and tracking nearby air traffic and issuing avoidance instructions to pilots when a threat is determined.[1] The effectiveness of these instructions—termed *resolution advisories* (RAs)—depends in large part on how pilots respond to them. TCAS' threat logic assumes an initial response delay of 5 seconds and a vertical acceleration of $0.25g$ ($g \approx 32\,\mathrm{ft/s^2}$) and times its advisories accordingly [2]. Any deviation from these assumptions can compromise system effectiveness, increasing collision risk [3,4]. As a result, it is important to understand how pilots respond to RAs and the circumstances that influence pilot response.

The performance of TCAS and other collision avoidance logics is evaluated primarily through fast-time simulation of aircraft encounters [4–6]. A model of pilot response to RAs is a critical component of these simulations. Historically, pilot response models have incorporated parameters such as response probability, delay, and acceleration. For example, ICAO[2] defines a *Standard Pilot Model* whose parameters conform to TCAS logic assumptions [3]. Some models incorporate stochasticity in pilot response delay [7], while others incorporate a response probability based on aggregated operational data [4]. However, in all cases, these models are applied identically across all encounters, even though studies have demonstrated that pilot response is strongly influenced by the properties of individual encounters, such as altitude, vertical rate, and the RA issued [8,9]. This suggests the need for a pilot response model that is sensitive to encounter-specific variables.

This paper introduces a model of pilot response to TCAS RAs where response is a function of the properties of each encounter. Using radar data recorded in US airspace, pilot response to TCAS RAs was characterized across tens of thousands of observed encounters. This response data was then analyzed alongside the geometry and RA profiles of the encounters in a Bayesian network [10]. From the network, the encounter parameters that most strongly influence pilot response were determined. The result is a model that can estimate pilot response probability for any arbitrary encounter for which the influential parameters are known. This model was employed in simulations of safety-critical encounters involving TCAS-equipped aircraft to observe its effect on probability of near mid-air collision (NMAC). The results were then compared to those of other pilot response models.

Although this is a study of pilot response to TCAS advisories, the methodologies introduced here can support an analysis of any separation advisory system. Currently, there is substantial ongoing work to integrate unmanned air vehicles (UAVs) into civil airspace. To facilitate this, large UAVs will be required to carry detect and avoid (DAA) systems to maintain safe separation from other air traffic [11]. The way in which UAVs will respond to DAA advisories is a matter of current study and may incorporate both automated and manual response. As DAA systems are developed and deployed, an understanding of actual unmanned vehicle response to DAA advisories will be critical.

This paper is organized as follows: Section II contains background on TCAS, the data source used in this study, and Bayesian networks; Section III describes the methodology

This work is sponsored by the Federal Aviation Administration under Air Force Contract #FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

[1]In this document, TCAS refers to TCAS II, known internationally as ACAS II.

[2]International Civil Aviation Organization

used to create the pilot response model; Section IV presents the model, including a description of the variables that most strongly influence pilot response; Section V analyzes the impact of the pilot response model on TCAS safety benefit and compares it to other models; and Section VI summarizes this work and describes follow-on activities.

II. Background

A. Traffic Alert and Collision Avoidance System

TCAS issues advisories based on estimated time to closest point of approach (CPA). Its advisories come in two forms:

1) *Traffic Alerts* (TAs), which prepare pilots for subsequent alerting and aid them in visually acquiring intruders
2) *Resolution Advisories* (RAs), which are recommended vertical maneuvers intended to maintain or achieve a desired separation. Advisories that require a change in vertical rate are known as *corrective* and are accompanied by a target rate. For example, TCAS issues *climb* and *descend* RAs that direct pilots to maintain a 1500 feet per minute (fpm) climb or descent, respectively.

In the United States, flight crews are nominally directed to comply with all TCAS RAs. However, they may choose to not respond in cases where they believe doing so would jeopardize the safety of flight or when they can ensure safe separation with definitive visual acquisition of the intruder causing the RA [12]. Studies of radar data have shown that pilot response varies widely. One such study estimated compliance with *climb* and *descend* RAs in the United States at 41% and 59%, respectively [8], while a study of European data estimated overall compliance with *climbs* and *descends* at 59% [13].

Operational studies have shown that when TCAS alerts, it is often during normal and safe procedures. For example, one analysis of United States radar data observed that 51% of TCAS RAs are issued when aircraft are safely separated in altitude by 500 feet and that 12% are issued during approaches to parallel runways [6]. Therefore, it is important to keep in mind that non-compliance with TCAS RAs does not necessarily suggest that safety has been compromised.

B. TRAMS

The recorded radar data analyzed in this study comes from the TCAS RA Monitoring System (TRAMS). TRAMS is a network of 21 secondary-surveillance radars distributed across the contiguous United States (Figure 1 and Table I outline the locations and coverage areas of the TRAMS radars) [2]. When a TCAS RA is issued within the TRAMS coverage area, RA information and other encounter data are downlinked by the transponders of the encountering aircraft and recorded along with the geometry of the encounter as measured by the radar.

The format and content of the data recorded by TRAMS is a function of the transponder type and the version of TCAS on the aircraft receiving the RA. The format associated with version 6.04a of TCAS, the oldest version of the logic deployed in US airspace, contains less information than the format associated with subsequent versions TCAS 7 and 7.1. This legacy format comprises approximately 37% of TRAMS recordings.

TRAMS makes separate recordings for each TCAS-equipped aircraft involved in an encounter. This means, for example, that a single encounter in the airspace between two TCAS-equipped aircraft will be recorded as two unique encounters by TRAMS, assuming an RA is issued by both TCAS units. In this instance, the first recorded encounter will contain RA information for the first aircraft and the second recorded encounter will contain RA information for the second aircraft. TRAMS recordings also include a small number of encounters involving three or more aircraft (approximately 0.3%), although this analysis considers encounters between two aircraft only.

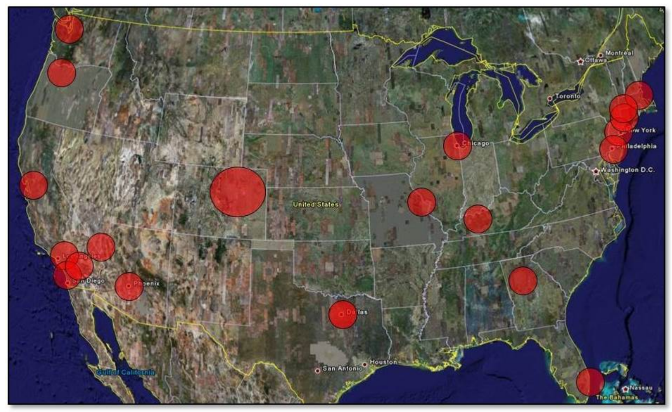

Fig. 1. Coverage areas of the TRAMS radars

TABLE I
LOCATIONS OF THE TRAMS RADARS

Radar	Location	Radar	Location
PHL	Philadelphia, PA	ATL	Atlanta, GA
LAXN	Los Angeles, CA	HPN	White Plains, NY
JFK	New York City, NY	SEA	Seattle, WA
DFW	Dallas-Ft. Worth, TX	ORDA	Chicago, IL
ONT	Ontario, CA	FLL	Ft. Lauderdale, FL
LGB	Long Beach, CA	LAS	Las Vegas, NV
OAK	Oakland, CA	PDX	Portland, OR
SDF	Louisville, KY	EWR	Newark, NJ
STL	St. Louis, MO	DEN	Denver, CO
PHX	Phoenix, AZ	QPK	Parker, CO
ACY	Atlantic City, NJ		

Over 550,000 RA encounters have been recorded by TRAMS since encounter monitoring began in 2008. With the exception of the Parker, Colorado sensor, all TRAMS sensors are terminal radars with a coverage radius of 60 nautical miles (nmi) and a rotational period of approximately 4.6 seconds, which is also the sampling period of data recorded by these sensors. The Parker, Colorado sensor is an en-route radar with a 200 nmi coverage radius and rotational period of approximately 10 seconds.

C. Bayesian Networks

A Bayesian network is a compact graphical representation of a joint probability distribution [10]. Bayesian networks consist of *nodes* connected by arrows. Each node represents a random variable that can be discrete or continuous. Arrows point from *parent* to *child* nodes and indicate statistical correlations between the nodes. Associated with each node is a conditional probability distribution that is a function of the values of the node's parents.

An example Bayesian network is depicted in Figure 2. The nodes of this sample network pertain to encounters involving a TCAS-equipped aircraft and are a subset of the pilot response network created in this study, shown in Figure 4. Abbreviated definitions follow (complete definitions are in Section III-B):

- *AC* represents the category of the TCAS aircraft.
- *RC* represents the relative course between the TCAS aircraft and the intruder.
- *AS* represents the airspace type in which the encounter took place (Class A, Class B, etc.).
- *PL* represents whether the encounter took place during an approach to parallel runways.
- *GR* represents the ground range between the aircraft when the RA was issued on the TCAS aircraft.

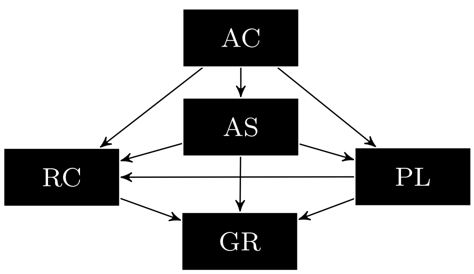

Fig. 2. Example Bayesian network

In this example network, the *GR* node has three parents: *RC*, *AS*, and *PL*, each of which is a child of *AC*. Arrows capture the correlations among the nodes. Because of these correlations, we can make inferences about the state of a node given knowledge of the states of the nodes that it correlates to. In this example network, these correlations include one between *AC* and *GR*, which exists through the connections of the nodes between them and despite the fact that they are not directly connected by arrows themselves. However, if we have knowledge of the states of the *RC*, *AS*, and *PL* nodes, then by virtue of the network structure, any inferences for the state of the *GR* node become independent of the state of the *AC* node. We say that *GR* is *conditionally independent* of *AC* given knowledge of the parents of *GR*, and therefore knowledge of these parents fully defines the probability distribution of *GR*. This notion of conditional independence is an important element of Bayesian networks and this analysis.

The objective of this work is to build a Bayesian network that characterizes the probability of pilot response to TCAS RAs based on encounter parameters. Using the principles of Bayesian network structure learning and conditional independence, this analysis will determine the encounter parameters that pilot response probability conditionally depends on (i.e., its parent nodes) and then define pilot response probability as a function of these parameters.

Bayesian networks are a powerful statistical tool with precedence in aviation research. For example, Bayesian networks were used to construct the *Lincoln Laboratory Correlated Encounter Model* (LLCEM), which used United States radar data to model aircraft trajectories in encounters [14].

III. Methodology

The first task in the construction of the Bayesian network was to define *pilot response* in the context of this analysis. Next, the network nodes were selected. Afterwards, data was collected for each node from the recorded TRAMS data. Finally, the arrows between the nodes were drawn (i.e., the network structure was *learned*) based on the gathered data. This section describes these steps.

A. Pilot Response

In this analysis, the definition of pilot response was constrained by the data source. TRAMS data is sampled at approximately 4.6 second intervals (excluding data recorded by the Parker, Colorado sensor) and TRAMS altitude data, which is acquired from aircraft transponder replies, is quantized to either 25 or 100 foot bins. Additionally, TRAMS data downlinked in the legacy format (see Section II-B) does not distinguish between certain RAs, including the various types of *adjust vertical speed* or *level off* advisories issued by TCAS versions 7 and 7.1.

Because of these limitations, this analysis studies pilot response to *climb* and *descend* advisories issued as the initial RA in a sequence only. Additionally, the definition of pilot response employed in this study considers RA compliance only, where a pilot is said to have responded to (i.e., complied with) the *climb* or *descend* RA if the aircraft achieved a vertical rate of at least 400 fpm in the appropriate direction within 15 seconds after the RA was issued (note that initial *climb* and *descend* RAs advise a rate of 1500 fpm). Response delay and vertical acceleration were not considered, as they would require a data source with finer resolution in time and altitude than TRAMS. This definition of pilot response has precedence in previous studies of TCAS operational data [8].

Note that approximately 31% of TRAMS encounters contain an initial *climb* or *descend* RA and are therefore eligible for inclusion in the Bayesian network. Any conclusions drawn from this analysis must bear this in mind.

B. Node Selection

Nodes were selected based on subject matter expert perception of the factors influencing pilot response. For example, experience suggests that pilots may be more likely to comply with RAs that do not conflict with their current vertical rate—an intuition supported by research into compliance with *climb* RAs [9]. In addition, nodes were constrained to data that could be ascertained from TRAMS recordings. This excluded any potential effects of TAs, for example, as they are not recorded by TRAMS.

A summary of the selected nodes follows. All selected nodes represent discrete quantities, and the discretization cutoffs of the nodes are summarized in Table II.

- **Aircraft Category** *AC*: Category of the TCAS-equipped aircraft, including major air carrier, regional air carrier, business jet, helicopter, other (typically piston engine general aviation), and unknown.

- **Airspace** *AS*: Airspace of the encounter. Potential values include Classes A, B, C, D, and E/G, or Special Use.
- **TCAS Sensitivity Level** *SL*: Sensitivity level of the TCAS unit issuing the RA. Potential values range between 3 and 7, with higher levels corresponding to more sensitive alerting and higher altitudes [2]. Sensitivity level served as a surrogate for aircraft altitude in this analysis.
- **Intruder Beacon Category** *VFR*: Boolean variable that is true when the intruder is squawking 1200 (VFR) and false otherwise. This node is based on the assumption that intruders squawking 1200 are less likely to be receiving separation services from air traffic control, with potential effects on RA compliance by the TCAS aircraft.
- **Parallel** *PL*: Boolean variable that is true if the encountering aircraft are on approach to parallel runways, which was determined based on aircraft course, horizontal range to intruder, and proximity to an appropriate airport.
- **Relative Course** *RC*: Difference in course between the ownship and intruder. A value of $0°$ corresponds to parallel courses, while a value of $90°$ corresponds to an intersection from the right.
- **Relative Altitude** *RH*: Unsigned altitude difference between the ownship and intruder at alerting time.
- **Vertical Rate** *VR*: Unsigned vertical rate of the ownship at alerting time.
- **Rate Reversal** *RR*: Boolean variable set to true if the RA commands a vertical rate in the opposite direction of the aircraft's current vertical rate, which must be in excess of 400 fpm. Note that rate reversals are distinct from RA reversals, which occur, for example, when a *climb* RA transitions to a *descend* RA.
- **Ground Range** *GR*: Horizontal range between the ownship and intruder at alerting time.
- **Climb/Descend** *CD*: Boolean variable set to true if the RA is a *climb* and false if it is a *descend*.
- **Pilot Response** ρ: Boolean variable set to true if the aircraft complied with the *climb* or *descend* RA according to the definition outlined previously.
- **Vertical Miss Distance** *VMD*: Unsigned vertical distance at time of minimum horizontal separation.
- **Horizontal Miss Distance** *HMD*: Horizontal distance at time of minimum horizontal separation.

TABLE II
DISCRETIZED ENCOUNTER VARIABLES

Variable	Discretization	Units
RC	$0, 45, 90, \ldots, 315$	degrees
RH	$0, 400, 800, \ldots, 1600, \geq 2000$	ft
VR	$0, 500, 1000, \ldots, 2000, \geq 2500$	fpm
GR	$0, 1, 2, \ldots, 5, \geq 6$	nmi
VMD	$0, 250, 500, \ldots, 1000, \geq 1250$	ft
HMD	$0, \frac{1}{4}, \frac{1}{2}, \ldots, 2, \geq 2\frac{1}{4}$	nmi

C. Data Collection

Data was collected from a subset of TRAMS encounters recorded between 2008 and 2016. Recorded position and altitude data were smoothed and interpolated to one-second intervals using a collision avoidance simulation tool developed at Lincoln Laboratory that incorporates a dynamic model of aircraft motion. Geometric values such as relative course and ground range were calculated based on this smoothed data. Figure 3 shows an example encounter comparing aircraft trajectories before and after smoothing. In addition, geometric parameters computed at alerting time (see previous subsection) were calculated at the time that the RA was first indicated by the TRAMS sensor—the *time of first downlink*—minus 5 seconds. This is because for any given encounter, the time an RA was issued was actually between the time of the corresponding downlink and the previous radar sweep approximately 4.6 seconds earlier. If parameters such as vertical rate were calculated at the time of first downlink without the 5 second offset, then the results would be affected by any pilot response to the RA occurring between radar sweeps. Consider that at a vertical acceleration of $0.25g$, the standard acceleration assumed by TCAS logic, vertical rate will change by approximately 2400 fpm in only 5 seconds. Calculating these parameters 5 seconds before the time of first downlink eliminates this potential biasing.

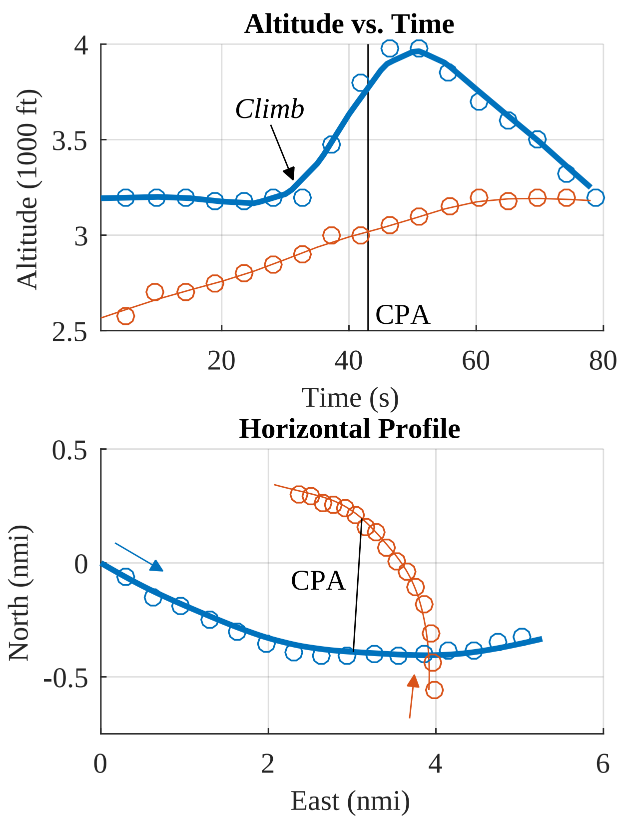

Fig. 3. Example TRAMS encounter between a TCAS aircraft (thick blue lines) and an intruder not equipped with TCAS. Open circles represent the original radar recording; solid lines represent smoothed trajectories. Note the *climb* RA downlinked by the TCAS aircraft at $t \approx 28$ and the subsequent response.

TRAMS encounters were filtered for validity and appropriateness to this analysis. An encounter was included only if it met the following criteria:

- First RA was *climb* or *descend*
- Not a formation or military flight
- Not recorded by the Parker, CO radar
- Longer than two downlinks (approximately 10 seconds)
- Contained no RA reversals (e.g., *climb* transitioning to *descend*)

Steps were also taken to eliminate duplicate encounters caused by overlapping radar coverage. The resulting dataset after applying these criteria consisted of 80,955 encounters.

D. Structure Learning

The final step in the construction of the Bayesian network was to determine the connections between nodes: the network structure. Known as *structure learning*, this step was supported by the GeNIe software environment created by the Decision Systems Laboratory at the University of Pittsburgh [15].

Several candidate networks were created using a variety of commonly used structure learning algorithms appropriate for this application.[3] For each candidate, the nodes were organized into four *temporal layers*. Temporal layers enforce causality between nodes by imposing the constraint that the children of any particular node must be in the same or a lower temporal layer. The temporal layers of this analysis were chosen to capture the causal relationships among encounter parameters and are outlined in Table III.

TABLE III
TEMPORAL LAYERS

Layer	Description	Nodes
1	Aircraft parameters and encounter geometry	AC, AS, SL, VFR, RC, RH, VR, GR, PL
2	RA-related parameters	CD, RR
3	Pilot response	ρ
4	Encounter outcome	VMD, HMD

Network candidates were judged based on several criteria, including a metric known as the *Bayesian score*. A network's Bayesian score measures how well its structure probabilistically represents the data used to build it [10]. It is represented logarithmically, with higher scores corresponding to more representative structures. Other judging criteria included network structure simplicity, with simpler networks preferred, and the ease by which the network could be implemented in simulations of aircraft encounters.

The candidate networks are summarized in Table IV along with their Bayesian scores and the algorithm used to create them. The selected network is marked in bold. This table includes a *Naive Bayes* network, which has the *pilot response* node as the direct parent to all of the other nodes.

[3]Multiple configurations of the *Greedy Thick Thinning* (GTT) [16] and *Bayesian Search* (BS) [17] algorithms were employed. Certain configurations require a maximum number of parents k for each node. Among the GTT algorithms, *K2* and *BDeu* refer to specific search strategies [18,19].

TABLE IV
BAYESIAN SCORES OF CANDIDATE NETWORKS

Algorithm[3]	Bayesian Score
$\text{GTT}_{k=8}^{K2}$	-1.09010×10^6
$\mathbf{GTT_{k=5}^{K2}}$	$\mathbf{-1.09154 \times 10^6}$
GTT^{BDeu}	-1.09708×10^6
$\text{BS}_{k=5}$	-1.19234×10^6
$\text{BS}_{k=8}$	-1.19329×10^6
Naive Bayes	-1.20687×10^6

IV. Selected Bayesian Network

A. Influence of Encounter Parameters on Pilot Response

The selected network optimally balanced the judging criteria outlined above and is depicted in Figure 4. It was generated using the *Greedy Thick Thinning* structure learning algorithm with the constraint that an individual node could have no more than five parents, included to limit network complexity.

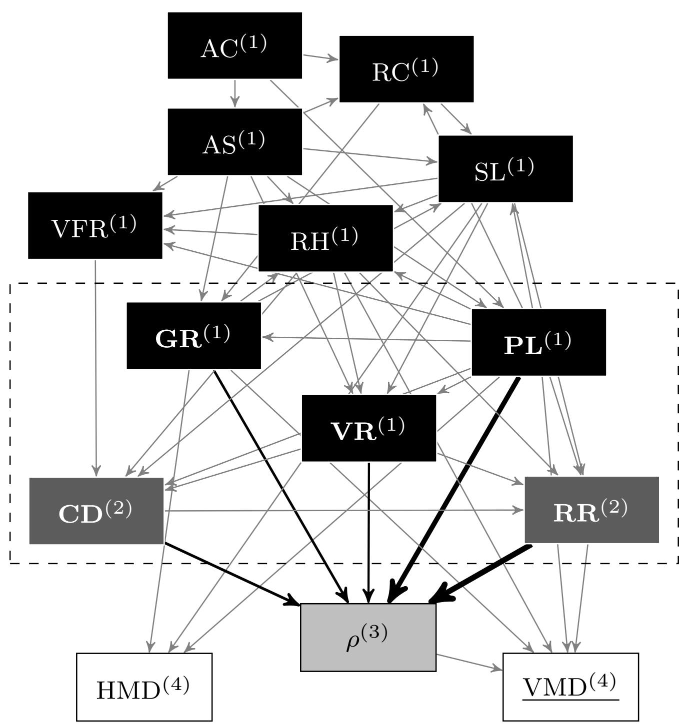

Fig. 4. The selected Bayesian network. The parents of the pilot response node ρ are enclosed in the dashed box and the child of ρ is underlined. Shading and superscripts indicate temporal layer. Black arrows indicate the links between ρ and its parents; arrow thickness correlates to strength of influence.

In this network, pilot response probability is fully defined by the values of its five parents: *rate reversal*, *parallel approach*, *climb/descend*, *ground range*, and *vertical rate*. A *strength of influence* analysis, conducted using techniques described in the literature [20], showed that among the five parents, the existence of a rate reversal or parallel approach encounter have the strongest influence on pilot response. The relative strengths of influence for each parent node are summarized in Table V, which also contains results for the other candidate networks (the selected network is marked in bold).

TABLE V
NORMALIZED STRENGTH OF INFLUENCE ON PILOT RESPONSE

Algorithm[3]	RR	PL	CD	GR	VR	VFR
$\mathrm{GTT}^{K2}_{k=8}$	0.24	0.27	0.12	0.13	0.12	0.13
$\mathbf{GTT}^{\boldsymbol{K2}}_{\boldsymbol{k=5}}$	**0.31**	**0.27**	**0.15**	**0.14**	**0.12**	·
GTT^{BDeu}	0.32	0.26	0.16	0.14	0.13	·
$\mathrm{BS}_{k=5}$	0.67	·	·	0.18	0.16	·
$\mathrm{BS}_{k=8}$	0.67	·	·	0.18	0.16	·

The overall pilot response probability for this dataset is 56%. Considering non-parallel approach encounters only, response probability becomes 62% overall, 58% for *climb* RAs, and 69% for *descend* RAs. These results are close to the response probabilities reported by the previously-referenced studies of radar data.

Figure 5 shows the probability distributions of *pilot response* and its five parent nodes in the dataset. In this figure, the y-axis represents the proportion of the dataset corresponding to each node value. Discussions of each of the five parents follows. Some of these discussions reference sections of Table VI, which outlines pilot response probability for various subsets of the dataset. The values in this table correspond to the probability that each of the included nodes is *true* within the corresponding subset (when RR is true, it indicates a rate reversal; when PL is true, it indicates a parallel approach encounter; when CD is true, it indicates a *climb* RA; and when ρ is true, it indicates a response to the RA—the GR and VR nodes are not included). Values in bold indicate which subset of the encounter set is being examined. For example, if RR is set to 1, then the values of the other nodes correspond to the subset of encounters that contain a rate reversal. Section 1 of the table corresponds to the complete dataset and represents a baseline. The rightmost column indicates the size of each subset represented as a percentage of the complete dataset. A tabular breakdown of the dataset containing values for all nodes is included in the Appendix.

- **Rate Reversal** *RR*: The data supports the notion that pilots are less likely to respond to RAs that oppose their current flight path. As section 2 of Table VI shows, rate reversals are associated with *climb* RAs and a lower probability of pilot response. These associations remain true for both the parallel approach and non-parallel approach subsets of the dataset.
- **Parallel** *PL*: Section 3 of Table VI shows that parallel approaches are associated with a lower probability of pilot response. And the third line of Section 3 shows that in 92% of the parallel approach encounters where the pilot did not respond, the RA was a *climb*, which would notionally necessitate a go-around. Considering the potential disruption caused by go-arounds, TCAS' propensity to alert unnecessarily against parallel approach intruders, and the fact that pilots oftentimes have these intruders in sight, it is reasonable that pilot response rate would be relatively low for these operations.
- **Climb/Descend** *CD*: In keeping with the discussion so far, section 4 of Table VI shows that pilots are less likely to respond to *climb* RAs than *descend* RAs. This is true even when considering only the non-parallel approach subset of the dataset. One potential reason for this trend is the association between *climb* RAs and rate reversals described earlier.
- **Ground Range** *GR*: Although not outlined in Table VI, the data shows that probability of response is lower for RAs issued at smaller ground ranges. This is due in part to the strong correlation between ground range and parallel approaches, though it is also true when considering only non-parallel approach encounters. One plausible explanation for this observation is that lower ground ranges correlate to slower airspeeds (TCAS alerts based on time to CPA) and lower altitudes where visual acquisition of intruders is more likely.
- **Vertical Rate** *VR*: Compared to the other parent nodes, the correlation between vertical rate and pilot response is relatively weak. An examination of the data suggests that the relationship between the two nodes is potentially a consequence of the definition of pilot response used in this study.

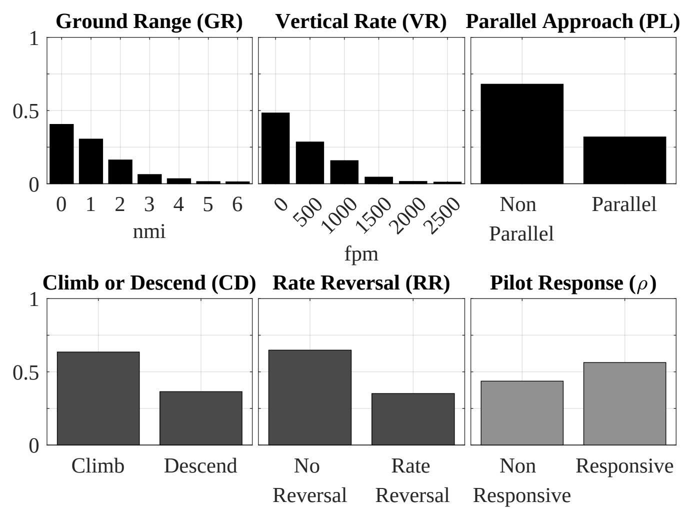

Fig. 5. Distribution of parent nodes and pilot response in the complete dataset. The y-axis represents the proportion of the dataset corresponding to each discretized node value

Given these parent nodes, it is possible to calculate the combinations of node values that result in the highest and lowest probabilities of pilot response. These combinations are outlined in Table VII for both the complete dataset and non-parallel approach encounters.

B. Influence of Pilot Response on Encounter Outcomes

While the previous subsection discussed the influence of encounter parameters on pilot response, this subsection discusses the influence of pilot response on encounter outcomes.

TABLE VI
PROBABILITY OF RESPONSE FOR VARIOUS PARENT NODE VALUES

	RR	PL	CD	ρ	Subset
1	0.35	0.32	0.63	0.56	100%
2	**1**	0.44	0.88	0.29	35.2%
	1	**0**	0.78	0.47	19.5%
	1	**1**	0.99	0.07	15.7%
3	0.29	**0**	0.68	0.62	68.0%
	0.49	**1**	0.54	0.45	32.0%
	0.84	**1**	0.92	**0**	17.6%
4	0.49	0.27	**1**	0.44	63.5%
	0.12	0.40	**0**	0.77	36.5%
	0.34	**0**	**1**	0.58	46.2%
	0.19	**0**	**0**	0.69	21.8%

TABLE VII
PARENT NODE VALUES FOR MAXIMUM AND MINIMUM PROBABILITY OF RESPONSE

	RR	PL	CD	GR	VR	ρ
Complete Dataset	False	True	Descend	0 nmi	−500 fpm	0.9998
	True	True	Climb	0 nmi	−500 fpm	0.0440
Non-Parallels	False	False	Descend	1 nmi	−500 fpm	0.9991
	True	False	Climb	0 nmi	−2000 fpm	0.1875

Note that the following results pertain to non-parallel approach encounters only.

As Figure 4 shows, VMD is a direct descendant of the pilot response node, meaning there is a direct statistical correlation between pilot response probability and VMD. And as Figure 6 shows, pilot response correlates with higher values of VMD. There is no such correlation between pilot response and HMD, and in accordance with this, HMD is not a descendant of pilot response in the selected Bayesian network. This is expected, as TCAS RAs mitigate collision risk by increasing vertical separation and notionally have no effect on horizontal separation.

As mentioned previously, this study assigned a threshold of 400 fpm to pilot compliance with *climb* or *descend* RAs. The actual maximum vertical rates achieved for both compliance and non-compliance are shown in Figure 7. The mean values of these results are close to 1500 fpm: the rate advised by TCAS for *climb* and *descend* RAs.

V. SAFETY IMPACT

The final step in this analysis was to assess the impact of the pilot response model on the calculation of the safety benefit provided by TCAS. To accomplish this, the model was employed in simulations of safety-critical encounters where one or both aircraft were equipped with TCAS. Safety benefit was gauged using the *risk ratio* metric, which measures the effect of collision avoidance advisories on probability of

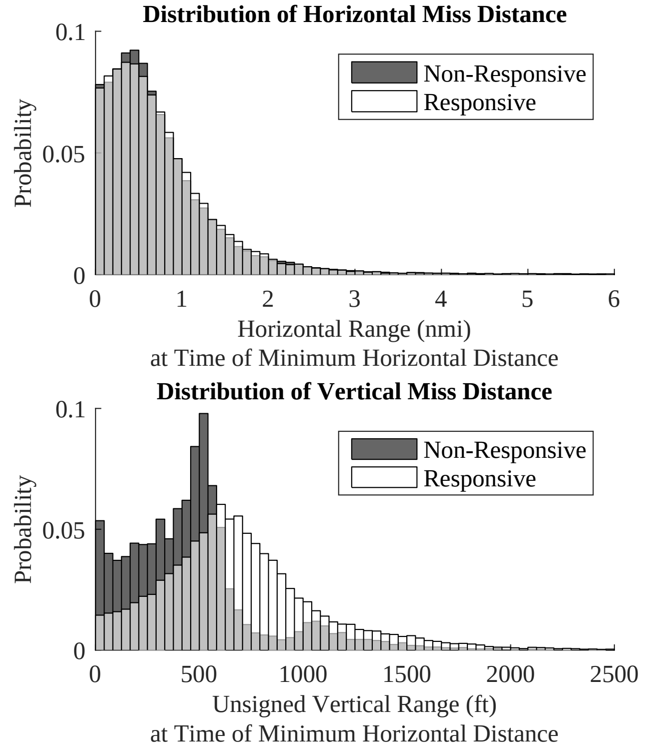

Fig. 6. Pilot response is associated with greater vertical separation. As expected, there is no correlation between pilot response and horizontal separation. Note the spikes in VMD at the procedural vertical separations of 500 and 1000 feet for the non-responsive distribution.

NMAC.[4] Risk ratio is defined as:

$$\textit{Risk Ratio} = \frac{\textit{P(NMAC) with CAS}}{\textit{P(NMAC) without CAS}}.$$

The lower the risk ratio, the greater the safety benefit of the CAS. A risk ratio less than 1 indicates a net safety benefit, a risk ratio of 1 indicates no net effect on safety, and a risk ratio greater than 1 indicates a net safety detriment.

The simulated encounter set consisted of 3,976,080 two-aircraft encounters drawn from the LLCEM. As mentioned, the LLCEM models encountering aircraft trajectories based on US radar data. Whereas TRAMS recordings represent relatively safe encounters where alerting was typically not necessary, LLCEM encounters are by design safety-critical, meaning collision avoidance intervention is oftentimes necessary to avert an NMAC. This makes the LLCEM an ideal encounter model for assessing the safety benefit of CAS advisories. Despite their differences, however, both TRAMS recordings and LLCEM encounters represent operations in US airspace. To simultaneously represent US operations and model safety-critical encounters, which are rare, the LLCEM assigns a likelihood-based weight to each encounter, with relatively high weights assigned to those encounters possessing a relatively high likelihood of occurring [14]. These weights were incorporated into the following analysis.

[4] An NMAC occurs when encountering aircraft come within 500 feet horizontally and 100 feet vertically.

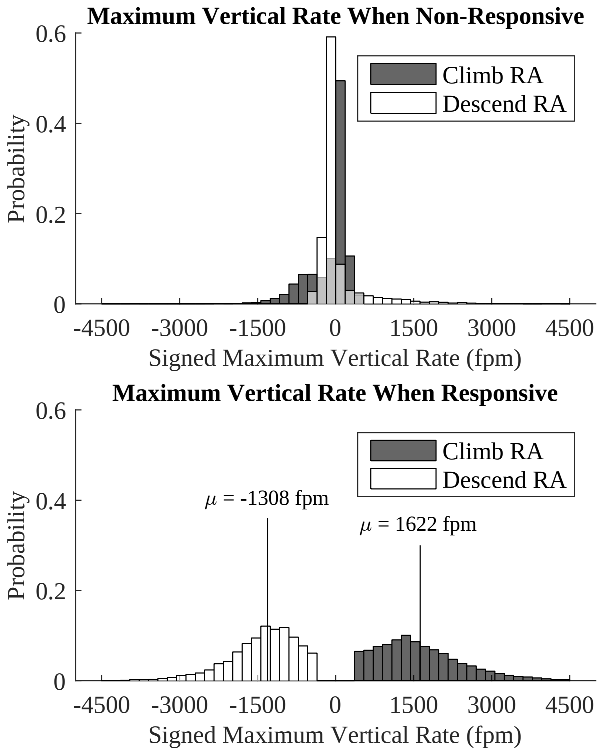

Fig. 7. Maximum vertical rate achieved during the RA, split by RA type and pilot responsiveness

The encounter set was simulated two ways: as encounters between two TCAS-equipped aircraft and as encounters between one TCAS-equipped aircraft and an intruder equipped with a Mode S transponder only. Version 7.1 of the logic was used for the TCAS aircraft. Surveillance noise conforming to standard error models was included, and both aircraft reported altitude with 25 foot quantization. In addition, when responsive to RAs, TCAS-equipped aircraft responded according to the standard model: for initial RAs, with 5 seconds of delay and $0.25g$ vertical acceleration; for subsequent RAs, with 2.5 seconds of delay and $0.35g$ vertical acceleration.

Encounters were divided into two groups: those for which the Bayesian network pilot response model was considered valid and those for which it was not. Model-valid encounters were those where the first corrective advisory issued by TCAS was a *climb* or a *descend* RA. Encounters beginning with *level off* RAs, for example, were not considered model-valid. For each model-valid encounter, pilot response probabilities for each TCAS-equipped aircraft were calculated based on their respective values for the *vertical rate*, *ground range*, *rate reversal*, and *climb/descend* nodes. These probabilities were used to determine whether or not each aircraft would respond to RAs in the encounter, and this responsiveness (or lack thereof) was applied identically to all RAs issued for the aircraft. Note that because the LLCEM does not model parallel approaches, the *parallel approach* node was fixed at *non-parallel* for these calculations.

In the simulated encounter set, the pilot response model was valid in approximately 20% of the encounters, when weighted. For those encounters where the pilot response model was not considered valid ("model-invalid"), pilot response probability was assumed to be 100%.

The primary goal of this analysis is to gauge the importance of including encounter parameters in a pilot response model. To this end, a variety of other pilot response models were also simulated and their resulting risk ratios compared to the model described above. The main variable among these models was the degree to which each model incorporated encounter parameters in the calculation of pilot response probability. In total, four pilot response models were evaluated. They are described below and summarized in Table VIII.

- **Naive 100%**: Response probability was 100% for all encounters.
- **Naive Aggregated**: Response probability was 86% for all encounters. This is a weighted combination of 66%—the average response probability for model-valid encounters in the simulated encounter set, as derived from the Bayesian network—and the 100% response assumption for model-invalid encounters.
- **Climb/Descend Averaged**: Response probability was 66% (the average described above) for all model-valid encounters and 100% for model-invalid encounters.
- **Climb/Descend Lookup**: Response probability was calculated ("looked up") from the Bayesian network encounter parameters for each model-valid encounter and was 100% for all model-invalid encounters. *Climb/Descend Lookup* is the full implementation of the pilot response model developed in this study.

TABLE VIII
PILOT RESPONSE PROBABILITIES OF THE SIMULATED MODELS

Pilot Response Model	Model-Valid Encounters	Model-Invalid Encounters
Naive 100%	100%	100%
Naive Aggregated	86%	86%
Climb/Descend Averaged	66%	100%
Climb/Descend Lookup	Encounter-specific calculation from Bayesian network	100%

The progression of these pilot response models is from lower to higher fidelity and sensitivity to encounter parameters. The first model is completely naive to pilot non-compliance with RAs, assuming 100% response probability. The second model applies a constant, non-perfect pilot response probability identically among all encounters. The third model is sensitive to a single encounter parameter: the RA type issued by TCAS (*climbs* and *descends* are treated differently than other RAs). And the fourth and most specific model incorporates all of the relevant encounter parameters outlined in the Bayesian network of this study and applies the result to *climb* and *descend* encounters.

The risk ratio results for these encounters are shown in Figure 8. From a comparison of *Naive 100%* and *Naive*

Aggregated, we can see that lower pilot response probabilities correspond to higher collision risks, as expected. However, the critical comparisons are among the risk ratios calculated using the *Naive Aggregated*, *Climb/Descend Averaged* and *Climb/Descend Lookup* models. In each of these models, pilot response probability was ultimately derived from the Bayesian network, but with differing levels of averaging and aggregation. Moving from left to right, pilot response probability incorporates more encounter-specific parameters, while at the same time, collision risk increases. This suggests a critical result: applying lower-fidelity, encounter-agnostic pilot response models can result in an underestimation of collision risk.

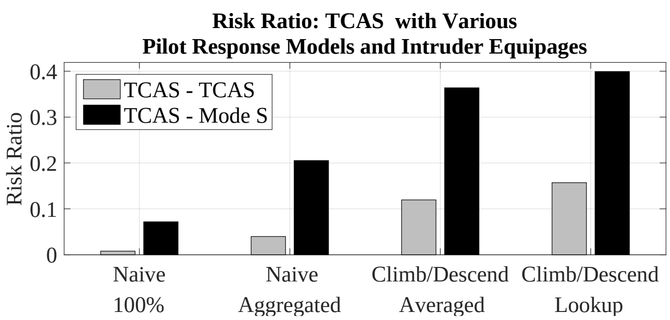

Fig. 8. Risk ratio evaluated for the simulated pilot response models

Figure 9 explores this point through a comparison of the encounter-agnostic *Climb/Descend Averaged* model and the encounter-specific *Climb/Descend Lookup* model for TCAS-TCAS encounters. In this figure, the y-axis represents the normalized, averaged safety benefit of responding to TCAS RAs, calculated based on reduction of probability of NMAC. The left bar corresponds to those encounters where the *Climb/Descend Averaged* model's pilot response probability, which is fixed at 66%, was lower than the probability calculated from the *Climb/Descend Lookup* model, while the right bar corresponds to those encounters where the opposite was true (note that only model-valid encounters were included in these calculations). Put another way: the left bar corresponds to those encounters where the *Climb/Descend Averaged* model overestimated pilot response probability, while the right bar corresponds to those encounters where it underestimated response probability. The greater height of the left bar explains why estimated collision risk was lower for the *Climb/Descend Averaged* model: that model assigned higher pilot response probabilities to those encounters where pilot response has the largest safety benefit, decreasing estimated collision risk.

Underestimating collision risk has many consequences, one of which is the masking of undesired system behavior. For example, given the choice between a *descend* and a *climb* RA in some encounter, a CAS logic may choose *descend* because by some standard response model it results in a safer outcome. However, a higher-fidelity model may reveal that pilots are more likely to respond to the *climb*, making it ultimately safer. If millions of encounters such as this one are incorporated into the development and evaluation of a collision avoidance system, then incorporating a higher-fidelity pilot response model that captures real-world pilot behavior could ultimately result in a safer system.

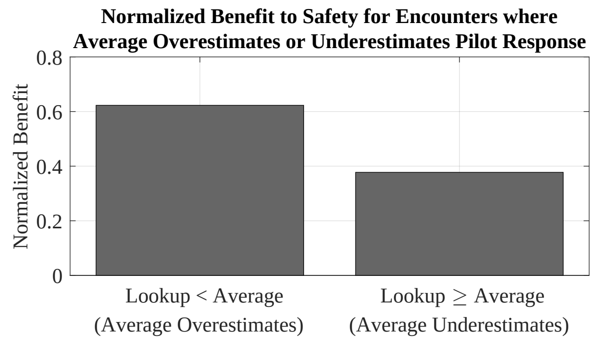

Fig. 9. Normalized, averaged benefit to safety of RA response for TCAS-TCAS encounters where *Climb/Descend Averaged* model overestimates or underestimates pilot response relative to *Climb/Descend Lookup* model

VI. Conclusion

A. Summary

The purpose of this study was to construct and demonstrate the safety impact of a pilot response model that is sensitive to the parameters of individual encounters. A model was built from operational TCAS data incorporated into a Bayesian network. Within this model, pilot response to TCAS *climb* and *descend* RAs was shown to be sensitive to five encounter parameters: parallel approach, rate reversal, vertical rate, RA type (*climb* or *descend*), and ground range. The model was then employed in simulations of safety-critical encounters and compared to other pilot response models. The results demonstrated that encounter-agnostic pilot response models can underestimate collision risk, potentially impacting the design and safety benefit of separation advisory systems.

Any conclusions drawn from this study must recognize its limitations. These limitations include the TRAMS data source, whose coverage area is limited to the terminal areas of large airports and whose contents do not include potentially relevant encounter parameters such as traffic alert timing and instructions from air traffic control. These limitations also include the definition of pilot response used in this study and the application to *climb* and *descend* RAs only.

B. Follow-on Work

The methodology demonstrated in this study can be applied using other data sources, with potential gains in the scope of the resulting pilot response model. For example, a response model incorporating pilot delay and vertical acceleration could be obtained from data possessing finer resolution in time and altitude than TRAMS. Similarly, a response model incorporating corrective RAs other than *climb* and *descend* (e.g., *level off* RAs) could be constructed from a data source with more complete RA information than TRAMS. The selection of model nodes could also be broadened to include encounter parameters not present in this study by using a data source possessing additional relevant information.

This methodology is also transferable to separation advisory systems other than TCAS. As UAVs begin their deployment in civil airspace, observed advisory response data can support the construction of UAV response models.

Acknowledgments

This work was sponsored by the FAA TCAS Program Office AJM-233, and we gratefully acknowledge Neal Suchy for his leadership and support. We also thank Mykel Kochenderfer, Wes Olson, Shirley Phillips, Tom Teller, Robert Klaus, Ian Jessen, and Matthew Edwards for their comments and insight. GeNIe, QGeNIe and SMILE Software are licensed from the University of Pittsburgh.

References

[1] J. Kuchar and A. C. Drumm, "The Traffic Alert and Collision Avoidance System," *Lincoln Laboratory Journal*, vol. 16, no. 2, pp. 277–296, 2007.

[2] FAA, "Introduction to TCAS II Version 7.1," February 2011.

[3] ICAO, "ICAO Annex 10 - Aeronautical Telecommunications, Volume 4 - Surveillance and Collision Avoidance Systems," ICAO, Fifth Edition, 2014.

[4] E. H. Londner, "Collision Avoidance System Effectiveness on Low Performance Unmanned Aircraft," *AIAA Infotech @ Aerospace, San Diego, CA*, 2016.

[5] L. Espindle, J. Griffith, and J. Kuchar, "Safety Analysis of Upgrading to TCAS Version 7.1 Using the 2008 U.S. Correlated Encounter Model," *Massachusetts Institute of Technology, Lincoln Laboratory, Project Report ATC-349*, 2009.

[6] J. E. Holland, M. J. Kochenderfer, and W. A. Olson, "Optimizing the Next Generation Collision Avoidance System for Safe, Suitable, and Acceptable Operational Performance," *Air Traffic Control Quarterly*, vol. 36, 2014.

[7] J. P. Chryssanthacopoulos and M. J. Kochenderfer, "Collision Avoidance System Optimization with Probabilistic Pilot Response Models," in *Proceedings of the American Control Conference*. IEEE, 2011, pp. 2765–2770.

[8] W. A. Olson and J. E. Olszta, "TCAS Operational Performance Assessment in the US National Airspace," in *Proceedings of the IEEE/AIAA Digital Avionics Systems Conference*, 2010, pp. 4.A.2.1–11.

[9] A. Pritchett and E. S. Fleming, "Pilot Compliance to TCAS Resolution Advisories," in *Proceedings of the IEEE/AIAA Digital Avionics Systems Conference*, 2013, pp. 6B6–1.

[10] M. Kochenderfer, *Decision Making Under Uncertainty*. MIT Press, Cambridge, MA, 2015, ch. 2, pp. 11–55.

[11] ICAO, "Manual on Remotely Piloted Aircraft Systems (RPAS)," ICAO Document 10019, Fifth Edition, 2015.

[12] FAA, "Air Carrier Operational Approval and Use of TCAS II," Advisory Circular 120-55C Change 1, March 2013.

[13] Eurocontrol, "TCAS II Performance in European TMAs; Part 1: Analysis," February 2009.

[14] M. Kochenderfer, L. Espindle, J. Kuchar, and J. D. Griffith, "Correlated Encounter Model for Cooperative Aircraft in the National Airspace System Version 1.0," *Massachusetts Institute of Technology, Lincoln Laboratory, Project Report ATC-344*, 2008.

[15] M. J. Druzdzel, "SMILE: Structural Modeling, Inference, and Learning Engine and GeNIe: A Development Environment for Graphical Decision-Theoretic Models," in *AAAI/IAAI*, 1999, pp. 902–903.

[16] J. Cheng, D. A. Bell, and W. Liu, "An Algorithm for Bayesian Belief Network Construction from Data," in *Proceedings of the International Conference on Artificial Intelligence and Statistics*, 1997, pp. 83–90.

[17] D. Heckerman, D. Geiger, and D. M. Chickering, "Learning Bayesian Networks: The Combination of Knowledge and Statistical Data," *Machine Learning*, vol. 20, no. 3, pp. 197–243, 1995.

[18] G. F. Cooper and E. Herskovits, "A Bayesian Method for the Induction of Probabilistic Networks from Data," *Machine Learning*, vol. 9, no. 4, pp. 309–347, 1992.

[19] W. Buntine, "Theory Refinement on Bayesian Networks," in *Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann, 1991, pp. 52–60.

[20] J. R. Koiter, "Visualizing Inference in Bayesian Networks," Ph.D. dissertation, Delft University of Technology, 2006.

Author Biographies

Edward H. Londner (BS'06–MS'09) earned a bachelor's degree in aerospace engineering from the University of Florida and a master's degree in aeronautics and astronautics from Purdue University. He is an associate staff member at MIT Lincoln Laboratory where his research concentrates on aviation safety, including extensive work with the operational tuning and evaluation of ACAS X. He is the flight test director for the ACAS X program and is also a private pilot.

Robert J. Moss (BS'14) earned a bachelor's degree in computer science from the Wentworth Institute of Technology. He is an assistant staff member at MIT Lincoln Laboratory where his research focuses on the development and analysis of ACAS X for manned and unmanned aircraft and specification development of prototype systems using the scientific programming language Julia.

Appendix

TABLE IX
Probability distributions of all nodes in Bayesian network

AC Category		Airspace	
0.08	Major Air Carrier	0.44	Class B
0.29	Regional Air Carrier	0.02	Class C
0.26	Business Jet	0.06	Class D
0.36	Helicopter	0.01	Special Use
>0.00	Other	0.46	Class E, G
0.01	Unknown	0.02	Class A
Rel. Course (degrees)		**HMD (nmi)**	
0.28	0	0.15	0
0.08	45	0.29	0.25
0.08	90	0.20	0.50
0.12	135	0.14	0.75
0.17	180	0.09	1.00
0.12	225	0.05	1.25
0.08	270	0.03	1.50
0.08	315	0.02	1.75
		0.01	2.00
		0.03	2.25+
Vertical Rate (fpm)		**Rel. Altitude (ft)**	
0.48	0	0.54	0
0.28	500	0.37	400
0.16	1000	0.03	800
0.05	1500	0.02	1200
0.02	2000	0.02	1600
0.01	2500+	0.02	2000+
VMD (ft)		**Ground Range (nmi)**	
0.23	0	0.40	0
0.22	250	0.30	1
0.26	500	0.16	2
0.15	750	0.06	3
0.07	1000	0.04	4
0.07	1250+	0.02	5
		0.01	6+
Beacon Code		**Climb/Descend**	
0.71	Discrete	0.63	Climb
0.29	1200	0.37	Descend
Rate Reversal		**Parallel**	
0.65	No	0.68	Non-Parallel
0.35	Yes	0.32	Parallel
Pilot Response		**TCAS SL**	
0.44	No Response	0.31	SL3
0.56	Response	0.29	SL4
		0.28	SL5
		0.10	SL6
		0.02	SL7