UNCERTAINTY ANALYSIS FOR EVENT SEQUENCE DIAGRAMS IN AVIATION SAFETY

Azin Zare Noghabi
George Mason University
Tel: 703-395-9532 Email: azarenog@masonlive.gmu.edu

John F. Shortle
George Mason University
4400 University Drive, Fairfax, VA 22030
Tel: 703-993-3571 Fax: 703-993-1521; Email: jshortle@gmu.edu

Word count: 4,142 + 13 tables/figures x 250 words (each) = 7392 words

Submission Date: 11/15/16

ABSTRACT

The Integrated Safety Assessment Model (ISAM) is being developed by the FAA for analysis and assessment of risk in the National Airspace System (NAS). ISAM includes a collection of Event Sequence Diagrams (ESDs) and their supporting fault trees and hazards. Historical incident and accident data provide point estimates to quantify the probabilities in the event trees. However, because accident occurrences are rare, there is some uncertainty in the point estimates. In particular, many accident event sequences have never been observed and thus are quantified as having zero probability of occurrence, but this does not mean that such events could never occur. Because a large number of the quantified parameters in ISAM are rare-events, it is important to characterize the uncertainty in these estimates in order to estimate the uncertainty in the output produced by model. The objective of this paper is to quantify the uncertainty of the point estimates in the model and to infer the resulting uncertainty in the intermediate pivoting event probabilities. Results indicate that the uncertainty in the pivoting events is driven by the number of accident end states with no historical observations.

Keywords: aviation safety, event trees, event sequence diagrams, uncertainty analysis

INTRODUCTION

The Integrated Safety Assessment Model (ISAM) is being developed by the FAA for analysis and assessment of risk in the National Airspace System (NAS) (*1*). The ISAM safety model includes a collection of Event Sequence Diagrams (ESDs) and supporting fault trees. Each ESD has a unique initiating event (e.g., an engine failure on take-off) that branches into multiple paths terminating at one of the end events. An end event might be an accident (e.g., runway excursion, etc.) or a safe state (e.g., aircraft stops on runway). The intermediate branching points are called pivoting events.

To quantify the probabilities of the pivoting events, historical incident and accident data are used. These data provide point estimates for the initiating-event probabilities and the end-event probabilities, but not typically the intermediate probabilities. These probabilities instead are quantified by solving a system of equations to make the end-state probabilities consistent with the initiating-event probabilities via the structure of the tree.

Currently, the end-state probabilities are quantified as *point estimates*. The point estimate of an accident probability is the number of historical accidents divided by the total number of operations over some time period. This is typically a very small number. In fact, because there are a considerable number of accident event sequences in ISAM, a large number of the end-state probabilities have a point estimate of *zero*. Many event sequences have simply never occurred. However, this does not mean that such events would *never occur*, it just means that they have *not yet occurred*. That is, the point estimate (which is zero) is only an approximation for the "true" underlying accident probability. Because a large number of the quantified parameters in ISAM are rare-events, it is important to characterize the uncertainty in these estimates in order to estimate the uncertainty in the output produced by ISAM.

The objective of this paper is to quantify the uncertainty of the end events in ISAM and to quantify the resulting uncertainty in the intermediate event probabilities. The rest of the paper is organized as follows: The next section gives a brief survey of the ESDs in ISAM and the database used for this study. The methodology is described in two parts including the point estimation process and the uncertainty analysis method. Finally, results on selected ESDs in ISAM are given.

A SURVEY OF EVENT SEQUENCE DIAGRAMS AND COLLECTED DATA

The ISAM (version 3.3) causal model contains 35 ESDs capturing all possible initiating events and their respective accident/incident scenarios. These ESDs are numbered as US-01, …, US-43, where numbers are not necessarily consecutive. For reference, Table 1 lists all initiating events in ISAM. Looking at all 35 ESDs, it is observed that similar ESD structures are repeated for different initiating events. In total, there are 10 different ESD structures which are depicted in Figure 1.

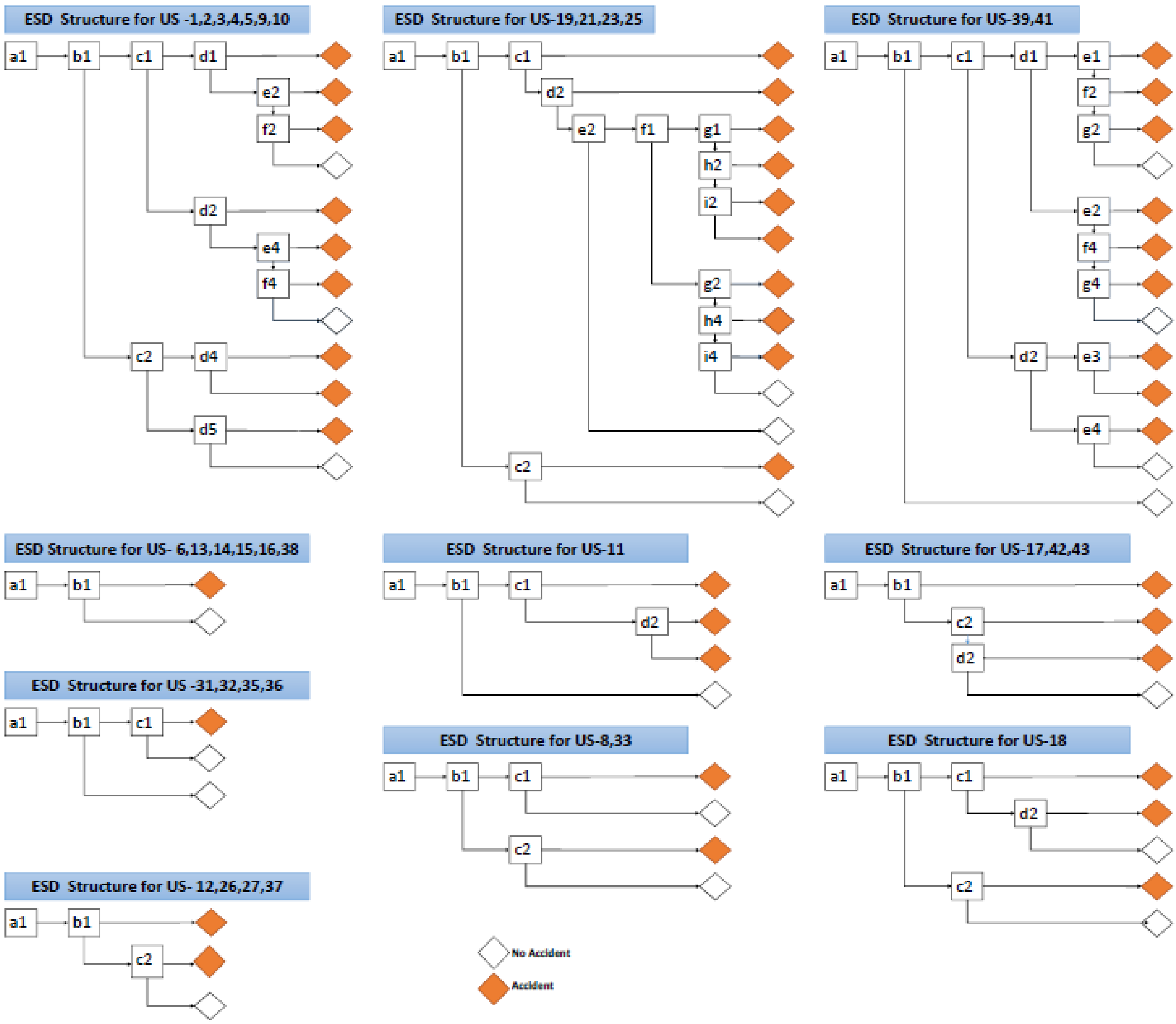

FIGURE 1 Common ESD structures across ISAM

The collected data that is used for quantifying these ESDs is a combination of accident information from the National Transportation Safety Board, the Accident/Incident Data System, service difficulty reports, and post-hoc interpretation of radar surveillance collected by the FAA. The dataset includes accident information on 208,582,368 flights operated in the United States excluding Alaska between 1995 and 2013 consisting of part 121 flights (air carrier operations) and scheduled part 135 flights (charter and air taxi operations); 579 accidents are reported in the dataset for 23 different ESDs, which means that 12 of the 35 initiating events have no reported accidents. Table 1 shows a breakdown of the 35 ESDs. The table shows the number of end states in each ESD and the number of safe end states (i.e., the number of possible end states in an ESD that do not represent an accident). The number of end states corresponding to accidents is the number of end states minus the number of safe state. The table also shows the total number of accident observations and the number of accident end states with no observations. The shading is simply to organize the ESDs into groups of similar structure.

TABLE 1 Summary of ESDs and Data

ESD	Initiating Event	# end States	# safe end states	# total accident observations	# end states with zero observations
US01	Aircraft system failure during take-off	12	3	1	8
US02	ATC event during take-off	12	3	0	9
US03	Aircraft directional control by flight crew inappropriate during take-off	12	3	5	7
US04	Aircraft directional control related system failure during take-off	12	3	0	9
US05	Incorrect configuration during take-off	12	3	2	8
US09	Single engine failure during take off	12	3	2	7
US10	Pitch control problem during take-off	12	3	2	8
US06	Acft takes off w/ contaminated flight surface	2	1	0	1
US13	Flight control system failure	2	1	3	0
US14	Flight crew member incapacitation	2	1	0	1
US15	Ice accretion on aircraft in flight	2	1	1	0
US16	Airspeed, altitude or attitude display failure	2	1	0	1
US38	Loss of control due to poor airmanship	2	1	2	0
US31	Aircraft are positioned on collision course in flight	3	2	1	0
US32	Runway incursion involving a conflict	3	2	1	0
US35	Conflict with terrain or obstacle imminent	3	2	4	0
US36	Conflict on taxiway or apron	3	2	190	0
US12	Flight crew member spatially disoriented	3	1	1	1
US26	Aircraft handling by flight crew inappropriate during landing roll	3	1	17	0
US27	Aircraft directional control related systems failure during landing roll	3	1	22	0
US37	Wake vortex encounter	3	1	5	0
US19	Unstable approach	13	3	32	4
US21	Aircraft weight and balance outside limits during approach	13	3	0	10
US23	Aircraft encounters wind shear during approach or landing	13	3	2	8
US25	Aircraft handling by flight crew inappropriate during flare	13	3	32	6
US40	ATC event during landing	13	3	0	10
US39	Runway incursion involving incorrect presence of single aircraft for takeoff	13	4	1	9
US41	Taking off from a taxiway	13	4	0	9
US17	Aircraft encounters adverse weather	4	1	203	1
US42	Landing on a taxiway	4	1	0	3
US43	Landing on the wrong runway	4	1	0	3
US11	Fire onboard aircraft	4	1	43	0
US08	Aircraft encounters wind shear after rotation	4	2	0	2
US33	Cracks in aircraft pressure boundary	4	0	0	0
US18	Single engine failure in flight	5	2	7	0

METHODOLOGY

We use ESD 17 (US17) as an example to explain the methodology for quantification of the probabilities and the resulting uncertainty analysis. US17 shows the accident and incident scenarios for the initiating event “Aircraft encounters adverse weather” (Figure 2). An adverse weather encounter is defined as an encounter with severe turbulence that could result in occupant injuries, an aircraft upset, or structural damage to the aircraft as a result of overstress of the aircraft’s structure (*2*).

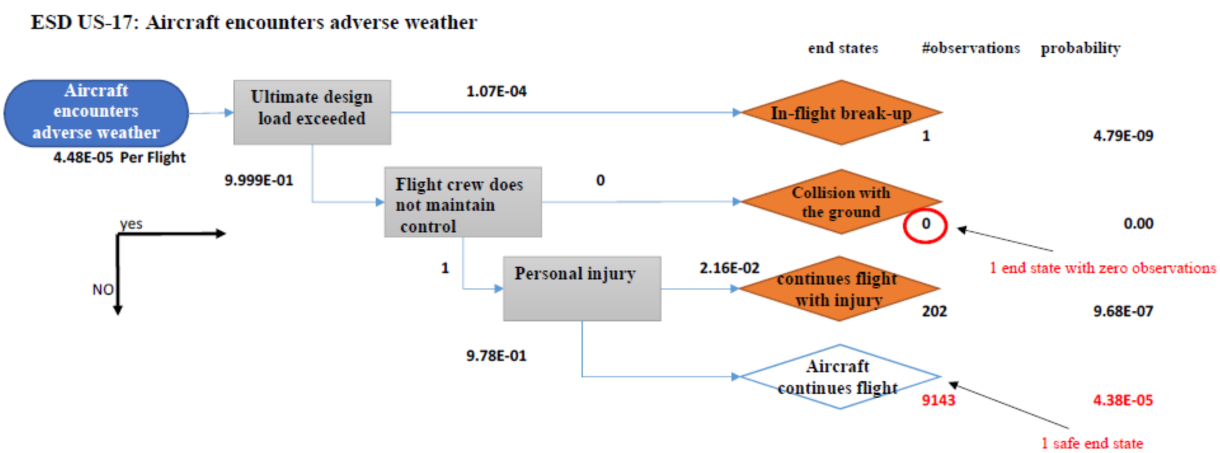

FIGURE 2 ESD US-17 – Aircraft Encounters Adverse Weather

Quantification of ESDs

ESD 17 consists of three pivoting events and four end states. Three of the end states are considered accidents (in-flight break-up, collision with ground, and flight continues with injury), while one end state is considered safe (aircraft continues flight). The pivoting events provide the branching points to the various end states. For example, when an aircraft encounters adverse weather, if the ultimate design load of the aircraft is exceeded, the result is an in-flight break-up of the aircraft. If the ultimate design load is not exceeded, the next pivotal event evaluates if the flight crew is able to control the aircraft, where failure to maintain control of the flight leads to collision with ground. The last pivoting event indicates the occurrence of passenger or crew injury as a result of encountering adverse weather.

According to the dataset, the probability of the initiating event is estimated as 4.48E-05 per flight. The data sample contains one accident of an in-flight break up and 202 encounters with adverse weather resulting in personal injury; the dataset contains no occurrence of a collision with the ground. The accident probabilities can then be estimated by dividing the number of accidents in each case by the number of operations. The point-estimate probability of an in-flight break up is 4.79E-09 per flight, the point-estimate probability of collision with the ground is zero and the point-estimate probability of personal injury is 9.68E-07. With these results, we can then back-calculate the pivoting event probabilities. For example, the conditional probability that the ultimate design load is exceeded when aircraft encounters adverse weather is 4.79E-09 / 4.48E-05 = 1.07E-04. The conditional probability of failure to maintain control in adverse weather is zero, and the conditional probability of occurrence of personal injury is 2.16E-02.

One important issue is that the data set does not contain the number of occurrences of the safe end states. In this particular ESD, that is not a problem, because the number of safe outcomes can be obtained by subtracting the total number of accidents from the total number of initiating events. However, in other ESDs, there may be two or more safe end states, so there is no way to tell how the safe outcomes are divided among these multiple states. When we have more than one safe state (i.e., more than one free parameter), additional constraints are required to uniquely determine the probabilities of the pivoting events. In these situations, we add a constraint that the free pivoting events have equal probabilities. That is, the probability is "evenly spread out" among the pivoting events leading to a particular quantified end state, subject to constraints maintaining consistency with the end-event data.

Another important point is how we treat zeros. In point estimation, when there are zero observations for an end state, the probability of that end state and the conditional probability of the immediate event before the end state are zero. However, such an event might occur in the future, so it is not clear that these values are truly equal to zero. There is uncertainty that comes from a finite number of observations. In turn, uncertainty in the probabilities of the end states leads to uncertainty in the pivoting events.

Uncertainty Analysis

To construct a confidence interval for probabilities of the end states, we choose the Poisson distribution as the appropriate statistical distribution for the number of accidents observed for each scenario. The argument for this selection is that incidents are rare events and events are (approximately) independent between flights. Historical data only gives a point estimate for the end state probabilities. To construct $(1-\alpha)\%$ confidence intervals for the Poisson mean parameter, as many as 19 different methods have been suggested in the literature (*4*) due to the conservative nature of the exact intervals. In this paper, we use an exact method that is based on the relationship between the Poisson and Chi-squared distributions (*5*). In this method, the $(1-\alpha)\%$ confidence interval (CI) for a Poisson distribution with mean λ and k observations can be expressed as:

$$[\frac{1}{2}H^{-1}_{2k}(\frac{\alpha}{2}),\quad \frac{1}{2}H^{-1}_{2k+2}(1-\frac{\alpha}{2})],$$

or equivalently:

$$[G^{-1}(\frac{\alpha}{2},k,1),\quad G^{-1}(1-\frac{\alpha}{2},k+1,1)],$$

where ${H^{-1}}_n(\alpha)$ and ${G^{-1}}_{(\alpha,n,1)}$ are respectively the inverse of the chi-squared distribution with n degrees of freedom and the inverse of the gamma distribution with shape parameter n and scale parameter 1 at the corresponding probability α. Table 2 shows the bounds of the 95% confidence intervals calculated using exact method for different numbers of Poisson observations. To explain with a simple example, suppose that 2 events are observed in 100 flights. The point estimate for λ, the true probability of the event, is 0.02. The 95% confidence interval for 100λ is [0.242,7.225] (from the table); hence, the 95% CI for λ is [0.0024,0.0722], which is a range including 0.02.

TABLE 2 Poisson Confidence Limits for alpha=0.05

Number of Observations	Lower Bound	Upper Bound
0	0.000	3.689
1	0.025	5.572
2	0.242	7.225
3	0.619	8.767
4	1.090	10.242
5	1.623	11.668
10	4.795	18.390
20	12.217	30.888
50	37.111	65.919
100	81.364	121.627
200	173.241	229.722

Now that we have a confidence interval for the end state probabilities, the next step is to infer the resulting uncertainty in the pivoting events. This is accomplished via the following steps using Monte-Carlo simulation:

1. Generate a random point estimate (via Monte-Carlo simulation) for each end-event probability, based on that end event's confidence interval (this is described in more detail below).
2. For a given random set of point estimates for the end events, derive the resulting pivoting event probabilities. This is done in a manner similar to how the pivoting event probabilities were obtained in Figure 2. This can also be viewed as solving a system of *n* equations and *n* unknowns, where *n* is the number of pivoting events.
3. Repeat steps 1 and 2. This gives a sample distribution of point estimates for each pivoting event.

To complete the first step, we need to generate a random rate parameter of the Poisson distribution for each end state in each simulation run. In Bayesian inference, the conjugate prior for the rate parameter λ of the Poisson distribution is the gamma distribution. Let $\lambda \sim \text{Gamma}(\alpha,\beta)$ denote that λ is distributed according to the gamma distribution with shape parameter α and rate parameter β. That is, the probability density function for λ is:

$$g(\lambda\,|\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\,\lambda^{\alpha-1}\,e^{-\beta\lambda} \quad for\ \lambda > 0.$$

Then, given the sample of n measured values k_i (number of observations), and a prior distribution of $Gamma(\alpha,\beta)$, the posterior distribution is $\lambda \sim Gamma(\alpha + \sum_{i=1}^{n} k_i\,, \beta + n)$. The posterior mean E[$\lambda$] approaches the maximum likelihood estimate in the limit as $\alpha \rightarrow 0, \beta \rightarrow 0$. We assume that the uninformed prior probability distribution is $Gamma\ (0.1{,}0.1)$ for all the end states. The posterior mean is:

$$\hat{\lambda} = \frac{\alpha + \sum_{i=1}^{n} k_i}{(\beta+n)} = \frac{\alpha}{(\beta+n)} + \frac{\sum_{i=1}^{n} k_i}{(\beta+n)} = \left[\frac{\beta}{n+\beta}\right]\left(\frac{\alpha}{\beta}\right) + \left[\frac{n}{n+\beta}\right]\left(\frac{\sum_{i=1}^{n} k_i}{n}\right)$$

For the accident events, n is typically very large, while the number of observations (sum of k_i) is relatively small or zero. Because n is large, the posterior distribution is not sensitive to β. However, when the number of observations (sum of k_i) is 0, there is a sensitivity on the prior parameter α, but this is greatly reduced provided there is at least one observation.

In order to get uncertainty intervals on the pivoting events, we use simulation. In each simulation run, we generate random numbers for the end-state probabilities based on their posterior gamma distributions. To get the posterior distribution we need the number of observations for each scenario and the total number of flights, both of which come from historical data. Figure 3 shows the simulated distribution of the end-event probabilities (for ESD US-17), from one million samples. The end event "in-flight break up" has a relatively large uncertainty band (relative to its mean), since it was quantified with only one observation. In contrast, the end event "aircraft continues flight with injury" has a relatively tighter uncertainty band, since it was quantified with 202 events. The safe end state "aircraft continues flight" has a very narrow uncertainty band since most of the observed flights end in this state (it is not a rare event). The large amount of data results in a more precise estimate.

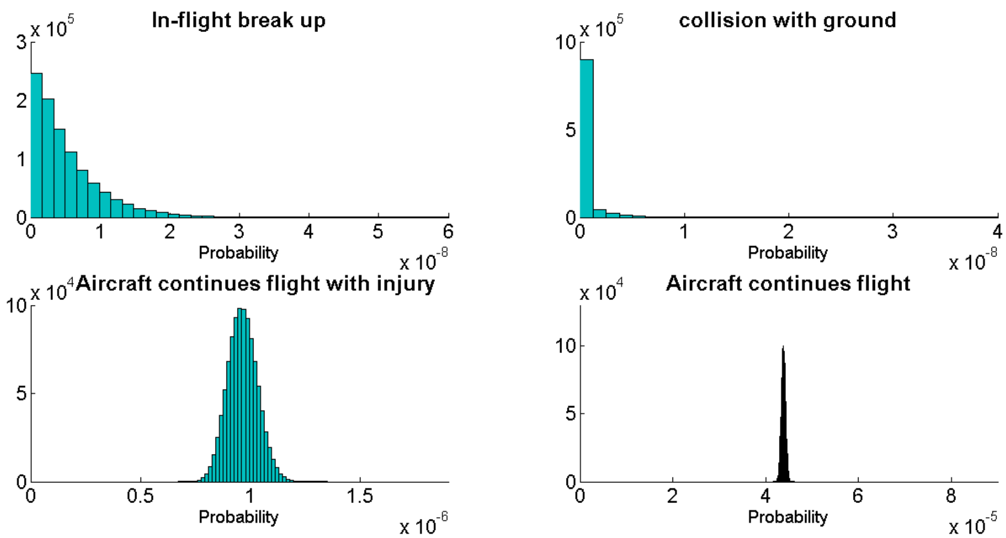

FIGURE 3 ESD US-17 – Histograms on probabilities of end states

After generating random numbers from the posterior distributions, we solve a system of equations to get the pivoting event probabilities. The simulation is repeated one million times. Figure 4 shows the resulting simulated distribution of the pivoting event probabilities.

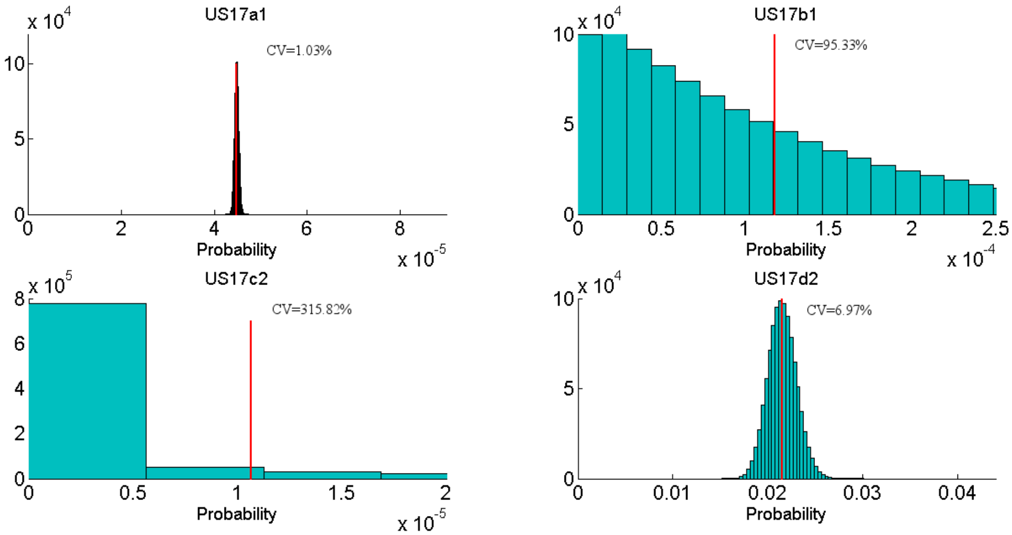

FIGURE 4 ESD US-17 – Histograms on probabilities of pivoting events

One way to measure the uncertainty of each estimate is with its *coefficient of variation* (CV), which is the standard deviation of the distribution divided by the mean. In Figure 4, the vertical line (red) gives the sample mean of the distribution. Since the mean values are approximately in the center of each graph in Figure 4 (i.e., normalizing the *x*-axis scales to the mean), the spread of the distribution is a visual proxy for the CV. Table 3 gives the exact numerical values for the CV. The pivoting event US17c2, "flight crew does not maintain control," (see Figure 2) has the largest CV. This event leads to an end state with zero observations and therefore has a point estimate of zero, resulting in a relatively large CV. The pivoting event US17b1, "ultimate design load exceeded," also has a fairly large CV, since it leads to an end event with only one observation. Event US17a1 is the initiating event in Figure 2 ("aircraft encounters adverse weather") and has a low CV since it is derived from the full set of incident data. In the next section, we present the results of our simulations on different ESDs and we try to establish relationships between the coefficient of variation as a measure of uncertainty and the ESD structure and amount of data. Figure 5 shows an overall graphical depiction of the uncertainty in each node of the event tree.

TABLE 3 Estimated Mean, Standard deviation and CV for US-17

ID	Description	Standard Deviation	Mean	CV
US17a1	Aircraft encounters adverse weather	4.63E-07	4.48E-05	0.01
US17b1	Ultimate design load exceeded	1.12E-04	1.18E-04	0.95
US17c2	Flight crew does not maintain control	3.37E-05	1.07E-05	3.16
US17d2	Personal injury	1.50E-03	2.15E-02	0.07

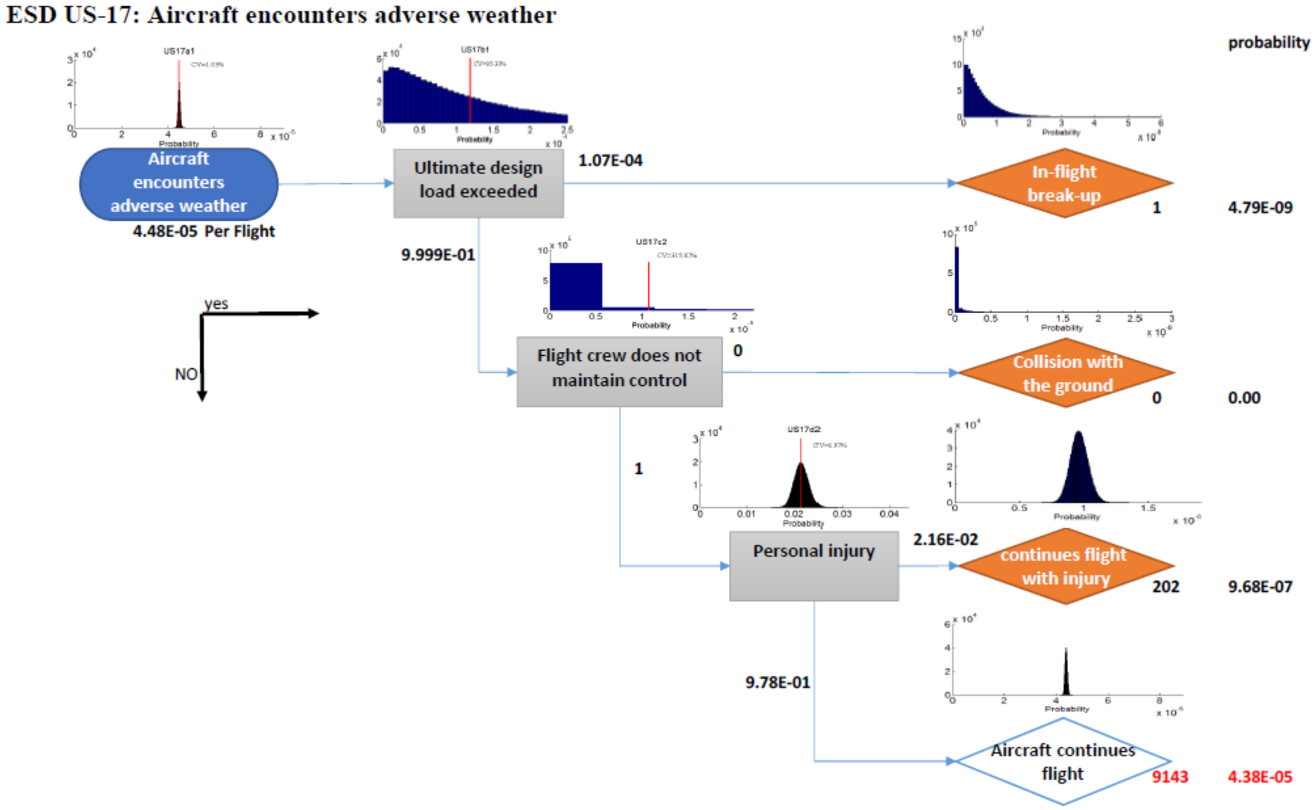

FIGURE 5 ESD US-17 with Histograms on end states and pivoting events

RESULTS

From 35 ESDs in ISAM, a subset of 18 ESDs are selected for simulation. The selected ESDs represent a variety in structure, number of observations, and number of end states with zero observations. Other ESDs are similar to the ones chosen here in terms of structure and the kind of data that we have for them, except for 3 ESDs for which we have no accident observations and no data on the probability of initiating events in our dataset, so we cannot investigate them any further.

Figure 6 shows the coefficient of variation (CV) for the initiating event and all pivoting events of the selected ESDs. A higher CV denotes greater uncertainty in the estimate. The ESDs are grouped (color-coded) based on structural similarity as depicted in Figure 1. We cannot see any meaningful relationship between the structure of the ESDs and the uncertainty in the pivoting events.

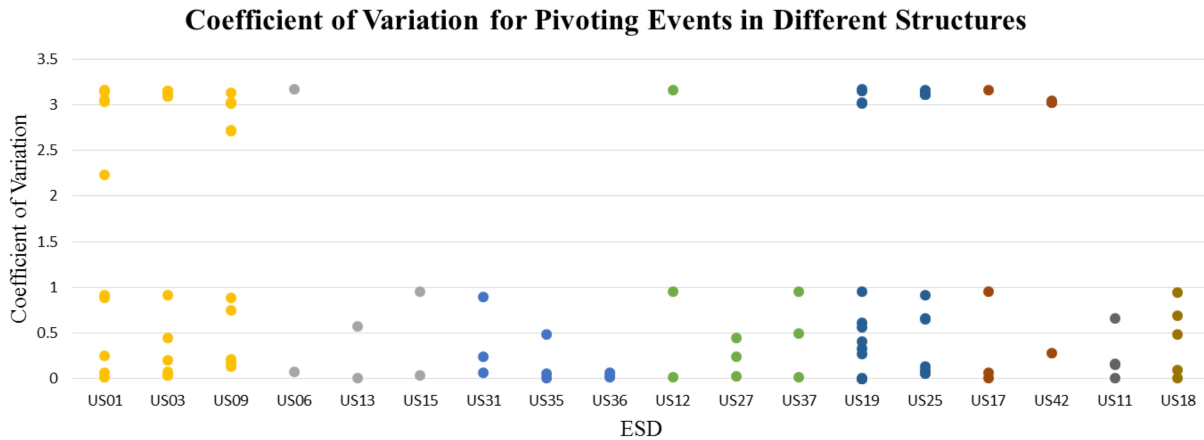

FIGURE 6 Comparison of CVs for pivoting events in ESDs with same structure

The relationship between the maximum CV in each ESD and the number of total observed accidents for ESDs is shown in Figure 7. There is no discernable trend between these variables. This is because, for some ESDs, the accidents may be coming from a single end event, so the pivoting events leading to the other end events (with no observations) have a high CV, even though the ESD as a whole has a relatively large number of observations. For example, ESD US17 has more than 200 accidents, but still has a large maximum CV – more than 3, which is similar to the values observed for many ESDs with zero or one observed accidents.

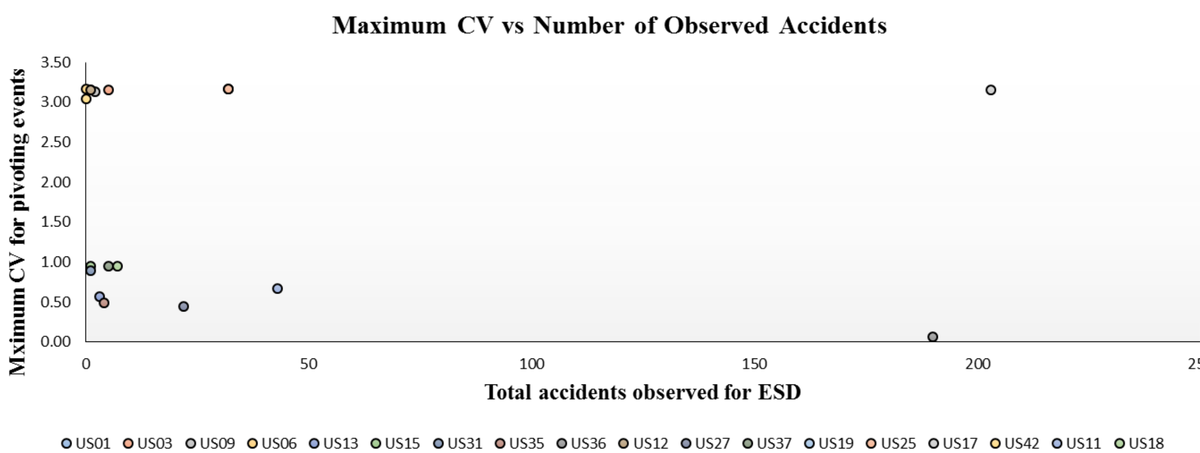

FIGURE 7 Maximum CV vs number of total accident/incident observations

The most consistent results are given when we graph the maximum CV for ESDs against the number of end states with zero observations for that ESD, as shown in Figure 8. When we have end states with zero observations in an ESD, independent of the number of such states, we get a maximum coefficient of variation of about 3. Among the ESDs for which we have non-zero observations for all the end states, the ones with the greater number of accident/incident observations have the least coefficient of variation.

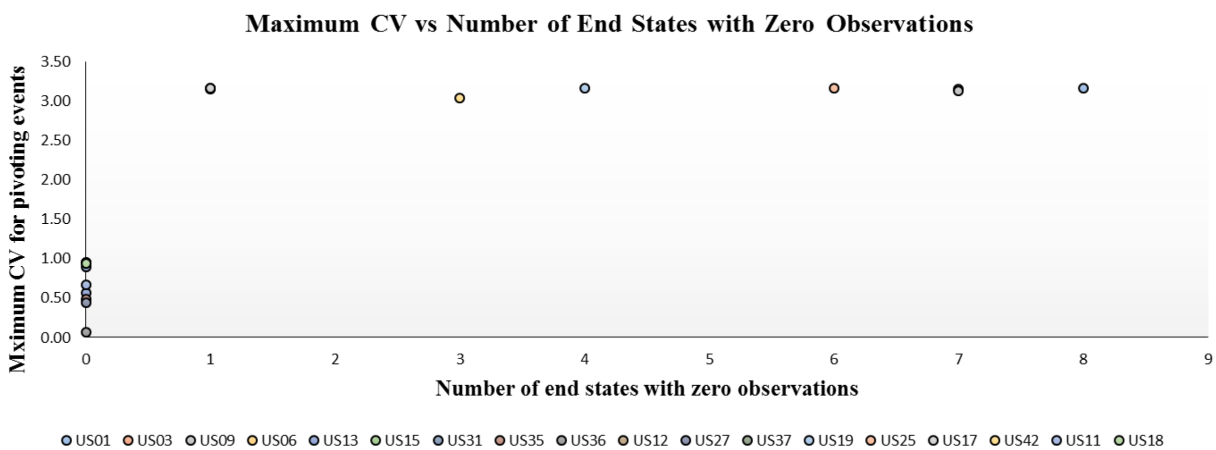

FIGURE 8 – Maximum CV vs Number of end states with zero observations

Figure 9 shows the relationship between the percentage of the end states with zero observations and average CV for each ESD. The increasing trend is clear. When most of the end states have zero observations, there are greater variances for the probabilities of the end states and this results in higher average coefficient of variation of pivoting events.

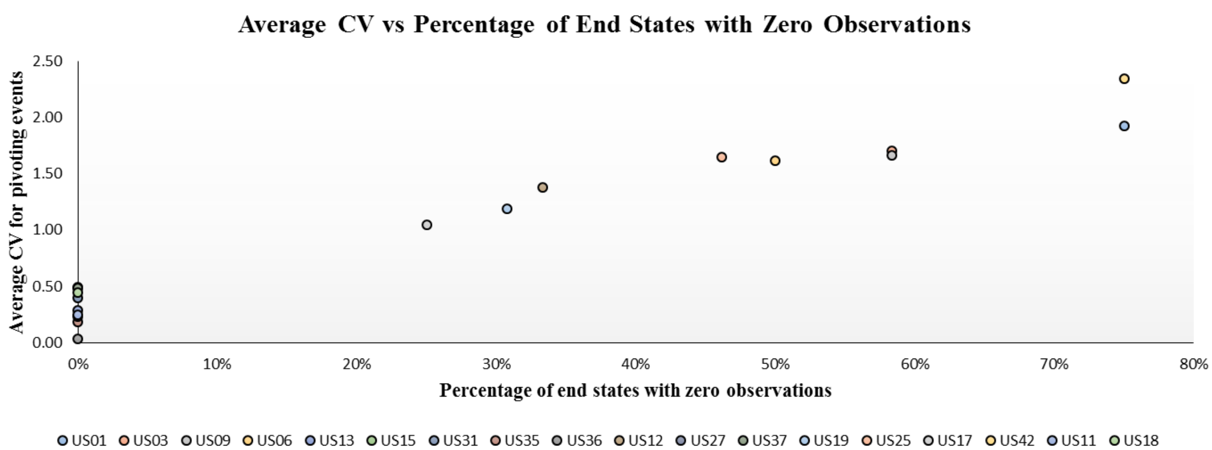

FIGURE 9 – Average CV vs percentage of end states with zero observations

CONCLUSIONS

In this paper, we presented an uncertainty analysis of the Integrated Safety Assessment Model. Because the model is quantified by a large number of rare-event probabilities, there is potentially a large relative uncertainty in the quantified values of the model, which might translate to uncertainty in the overall model performance. This paper attempted to quantify the uncertainty using a simulation-based approach which generated random point estimates of the end-event probabilities based on confidence intervals of the original data and then inferred the resulting probabilities of intermediate probabilities in the model. We then attempted to relate the relative size of the uncertainty in each tree to structural properties of the tree. The event trees with the largest uncertainties generally were trees with the largest number of end events with zero observed events. Having a large number of total observed accidents was not the critical factor yielding a low level of uncertainty in an event tree. Rather, the critical factor was having the observed data be spread out among all possible end events so that there were no “missing” end events in the tree (i.e., with zero data).

This research specifically addressed the Integrated Safety Assessment Model. However, the methodology is somewhat generic and could be applied to other event tree models. The key requirement is a data source to quantify all of the end events (and the initiating event) in the event tree. In this paper, a key limitation is that not all end events are quantified in all of the ESDs, so the pivoting event probabilities for some ESDs are not uniquely determined. For such cases, the uncertainty due to unquantified end events may dominate uncertainty due to the rare-event nature of the data.

Future work can include sensitivity analysis related to the assumed prior distributions of the end event probabilities. Future work may also extend the methodology to cases where some of the intermediate pivoting events are quantified (instead of just the end events). The methodology would be similar, except that a modified algorithm would be used for mapping the input data (known quantities) to the unknown model parameters. This research could also be used as a foundation for establishing a capability within ISAM for providing uncertainty bounds on the parameter quantification, rather than just a point estimate. This would help to communicate which parameters are known fairly accurately and which parameters have a large uncertainty. Results could be used to prioritize data collection efforts that would most effectively improve accuracy of the model.

ACKNOWLEDGMENT

This work was sponsored by Dr. Sherry Borener, Program Manager, System Safety Management Transformation, FAA. The authors also thank Mr. Alan Durston and Mr. Brian Hjelle for providing data, code, and support to run ISAM in the context of this research. The results and opinions expressed in this paper are solely those of the authors.

REFERENCES

1. Borener, S., S. Trajkov, P. Balakrishna. Design and Development of an Integrated Safety Assessment Model for NextGen, International Annual Conference of the American Society for Engineering Management,2012.
2. Noh,S., Shortle, J.F., Sensitivity Analysis of Event Sequence Diagrams for Aircraft Accident Scenarios, Proceedings of Digital Avionic Systems Conference, 2015, 3E2-1 – 3E2-12.
3. Roelen, A.L.C., B.A. van Doorn, J.W. Smeltink, M.J. Verbeek, R. Wever. Quantification of Event Sequence Diagrams for a Causal Risk Model of Commercial Air Transport. *National Aerospace Laboratory NLR, the Netherlands, NLR-CR-2008-646.*
4. Patil ,V.V., Kulkarni,H.V. Comparison of Confidence Intervals for the Poisson Mean: Some New Aspects, *REVSTAT*, Vol. 10, No. 2, 2012, pp. 211–227.
5. Garwood, F. Fiducial Limits for Poisson Distribution. *Biometrika*, Vol. 28, No. 3/4 (Dec., 1936), pp. 437-442.