This method should prove particularly helpful in analyzing not only fare changes, but also such LOS characteristics as minimum-stay requirements, advance-purchase requirements, and preferences for certain types of aircraft. Once these variables have entered into the utility functions and coefficients have been estimated, it is possible to measure the impact of changes in these characteristics. Even more helpful will be the ability of the analyst to identify the reactions of various market segments, which will permit the type of price discrimination necessary to induce new business without diverting revenues from full-fare passengers.

In order to project aggregate demand for a given alternative, it is necessary to have information about the potential travelers and their potential trips. The total number of trips in a market and the characteristics of trip makers must be forecast externally from these ticket-type models by using the carrier's standard methods. This process may involve sophisticated models of trip generation or may be based on something as simple as a projected market-growth trend and an assumption that the socioeconomic distribution of passengers remains unchanged. In either case, aggregation methods require that the carrier's passengers be grouped into some number of relatively homogeneous cells. The model must then be applied to each cell separately; this will forecast the ticket-type choice of its members and accumulate the aggregate shares for each alternative.

CONCLUSIONS

The above discussion has set forth a proposed method for analyzing the many new factors that affect the airline-passenger flight and ticket-selection process. The model relies on a statistical technique that is well tested in other behavioral modeling disciplines and particularly in modeling transportation mode-choice decisions.

A pilot application of the model was performed on a set of survey observations by using a two-alternative choice set--full fare or standby--but is easily extended to any number of alternatives and is adaptable to many types of distinguishing characteristics, such as booking requirements, length-of-stay requirements, and time-of-day restrictions. A complete data set for estimation could easily be obtained by using the on-board surveys made by the carriers. With the richer data base, many of the simplifying assumptions in the pilot application could be relaxed, which would provide a sound model to aid carriers in their complex marketing decisions.

REFERENCES

1. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (P. Zarembka, ed.), Academic Press, New York, 1973.
2. M. Richards and M. Ben-Akiva. A Disaggregate Travel Demand Model. Heath, Lexington, MA, 1975.

Publication of this paper sponsored by Committee on Aviation Demand Forecasting.

Assessing the Safety and Risk of Air Traffic Control Systems: Risk Estimation from Rare Events

ALLEN C. BUSCH, BRIAN COLAMOSCA, J. STUART HUNTER, AND NEIL W. POLHEMUS

To assess the safety and risk of current and proposed air traffic control route-separation standards, it is necessary to estimate the frequency of occurrence of extremely rare events. Since direct estimates of collision risk from historical data require sample periods that are unacceptably long, alternative methods are necessary. This article describes a probabilistic model for collision risk and its use in the North Atlantic airspace; it includes a discussion of sequential testing designed to determine whether current navigational performance meets a specified target level of safety.

A problem frequently encountered by analysts who wish to determine the level of risk associated with a particular transportation system is that of estimating the frequency of rare events. Most catastrophic transportation accidents, such as the midair collision of two commercial airliners, occur so infrequently that estimates of the accident rate are difficult to obtain directly. Consequently, probabilistic models are often constructed to describe the various factors that must occur to cause an accident. Estimates of the rate of occurrence of these factors are then obtained separately and combined later in an overall-risk computation.

An example of such an indirect approach is that of collision-risk methodology, first proposed by Reich (1) to estimate the risk of midair collisions between aircraft strategically separated in the lateral, longitudinal, and vertical dimensions and subsequently applied to determine route spacing in the North Atlantic and Central East Pacific regions (2-4). Essentially, the model factors the occurrence of a collision into three events (lateral overlap, longitudinal overlap, and vertical overlap), all of which must occur simultaneously to create a collision. Since the frequency of each event is several orders of magnitude higher than the frequency of a collision, it can be estimated in a sufficiently short period of time. If we assume that the three events are independent, their probabilities can then be multiplied to estimate the probability of a collision.

This paper examines some of the important estimation problems raised in applying collision-risk methodology to oceanic environments. We concentrate first on the Poisson distribution, long used to model the occurrence of rare events, and discuss the problem of obtaining precise estimates of the rate parameter. The collision-risk model is then introduced and its use in risk estimation described. Emphasis is placed on the questions of validation and monitoring, both of which are necessary ingredients in establishment of a minimum navigational performance standard. The discussion includes both sequential and fixed-sample-size testing.

DIRECT ESTIMATION OF ACCIDENT RATES

The Poisson model for rare events states that the probability of x accidents in a period Δt follows the Poisson distribution

$$p(x) = (\lambda\Delta t)^x \exp(-\lambda\Delta t)x! \quad x = 0, 1, 2, \ldots \tag{1}$$

where λ, the rate parameter, is expected number of accidents per unit of time. The expected value and variance of x are equal and given by

$$E(x) = \text{var}(x) = \lambda\Delta t \tag{2}$$

Figure 1 shows a plot of the Poisson distribution for $\lambda = 0.1 \times 10^{-7}$ and $\Delta t = 10^7$ track-system flying hours. The rate of 0.1 collision/10^7 flying hours has been selected by the International Civil Aviation Organization (ICAO) as the target level of safety (TLS) for use in setting oceanic navigational performance standards. Since the yearly number of track-system flying hours in the North Atlantic organized track system is currently less than 400 000, the chance of a midair collision in any given year at the TLS is extremely small.

Now suppose that one wishes to verify from historical data that the accident rate in the North Atlantic organized track system as currently structured does not exceed the TLS. Further, suppose we assume the rate to remain constant and begin monitoring the system. After observing T track-system flying hours with no midair collisions, the maximum-likelihood estimate for λ is $\hat{\lambda} = 0$ accident/10^7 flying hours. While $\hat{\lambda} = 0$ is the best point estimate for the accident rate, it is clearly unacceptable by itself, since some risk certainly exists. Of considerably more use is an interval estimate for λ that has upper and lower limits ($\hat{\lambda}_L$, $\hat{\lambda}_U$). Classical results for the Poisson distribution show that the limits of a 100 $(1 - \alpha)$ percent confidence interval for λ can be obtained from (5, p. 96)

$$\hat{\lambda}_L = (1/2T)\chi^2_{0,\alpha/2} \tag{3}$$

$$\hat{\lambda}_U = (1/2T)\chi^2_{2,1-\alpha/2} \tag{4}$$

where $\chi^2_{\nu,\alpha/2}$ is that value of a χ^2-distribution with ν degrees of freedom for which the probability of a larger value equals $\alpha/2$.

Figure 2 plots the upper 95 percent confidence limit ($\hat{\lambda}_U$) as a function of the length of monitoring period T. Note that, with no observed accidents, a period of approximately T = 400 million flying hours is required to bring $\hat{\lambda}_U$ below 0.1×10^{-7}. At current traffic levels, this corresponds to about 1000 years, which makes direct validation clearly impractical.

A slightly different perspective is obtained if one takes a Bayesian approach. If we begin with a noninformative prior distribution for λ, the posterior distribution for λ after T hours with no midair collisions is a χ^2-distribution with 1 df. Figure 3 plots the area of the posterior distribution that lies below 0.1×10^{-7} as a function of T, which shows how the degree of belief that the TLS is being achieved increases as the period of midair collisions increases. Again, the time required for that belief to reach an acceptable level is too long for practical purposes.

COLLISION-RISK METHODOLOGY

Because of the extremely long periods required to obtain precise estimates of the rate of occurrence of rare events, alternative formulations are necessary. In a series of articles, Reich (1) describes a probabilistic model for estimating collision risk in a system of parallel routes on which aircraft are strategically separated in the along-track, cross-track, and vertical dimensions. As subsequently developed and applied by the Federal Aviation Administration (FAA) and ICAO, the collision-risk model takes the following form:

$$N_{ay} = 10^7 [P_y(S_y)] P_z(0) \frac{\lambda_x}{S_x} \left\{ E_y(\text{same}) \left[\left|\frac{\overline{\Delta V}}{2\lambda_x}\right| + \left|\frac{\overline{\dot{y}(S_y)}}{2\lambda_y}\right| + \left|\frac{\overline{\dot{z}(0)}}{2\lambda_z}\right| \right] + E_y(\text{opp}) \left[\left|\frac{\overline{V}}{\lambda_x}\right| + \left|\frac{\overline{\dot{y}(S_y)}}{2\lambda_y}\right| + \left|\frac{\overline{\dot{z}(0)}}{2\lambda_z}\right| \right] \right\} \tag{5}$$

where

N_{ay} = expected number of collisions in 10^7 track-system flying hours;
S_y = lateral separation between parallel tracks;
$P_y(S_y)$ = probability of lateral overlap, given lateral separation S_y;
$P_z(0)$ = probability of vertical overlap for co-altitude aircraft;
$\lambda_x, \lambda_y, \lambda_z$ = longitudinal, lateral, and vertical dimensions of a typical aircraft;
S_x = longitudinal dimension of proximity shell for measuring occupancy;
$E_y(\text{same}), E_y(\text{opp})$ = same- and opposite-direction occupancies;
$|\overline{\Delta V}|, |\overline{\dot{y}(S_y)}|, |\overline{\dot{z}(0)}|$ = average relative along-track, cross-track, and vertical closing velocities for same-direction traffic; and
$|\overline{V}|, |\overline{\dot{y}(S_y)}|, |\overline{\dot{z}(0)}|$ = average relative along-track, cross-track, and vertical closing velocities for opposite-direction traffic.

An excellent discussion of the model, which includes its mathematical development, is given by Busch, Colamosca, and Vander Veer (2).

The above model factors the occurrence of a collision into essentially three events--longitudinal overlap, lateral overlap, and vertical overlap--all of which must occur simultaneously to create a collision. Overlap in a given dimension is defined as a situation in which two aircraft deviate from their planned positions in such a manner that their centroids are within some critical distance (such as a wingspan) of each other in that dimension. One then proceeds to estimate the probability of the three types of overlap, which are assumed to be independent (an assumption that has been tested from empirical data and appears reasonable).

In a general context, the most important aspect of the above model is its description of a collision in terms of events whose probability can be estimated with acceptable precision in a reasonable time. For example, consider the probability of

Figure 1. Poisson distribution rate 0.1 x 10^{-7}.

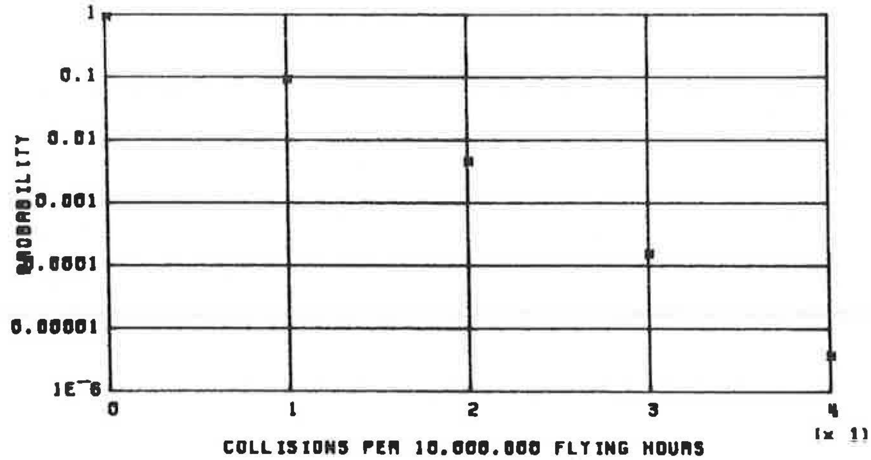


Figure 2. Upper 95 percent confidence limit for collison rate.

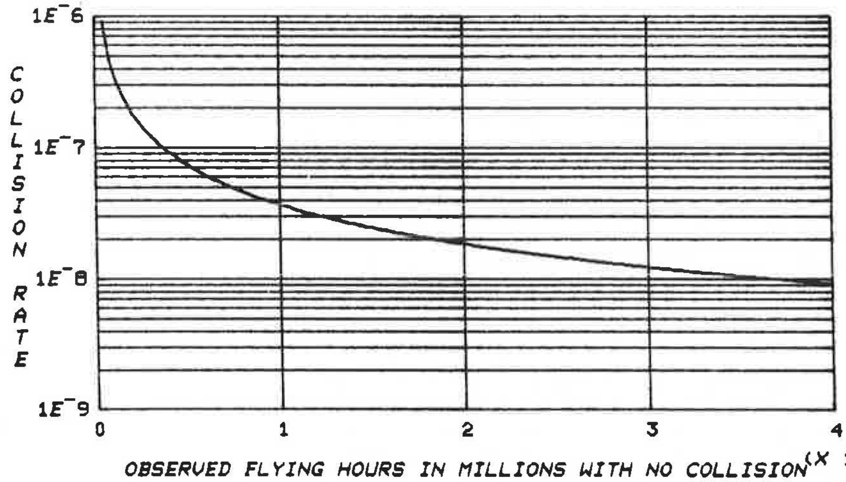
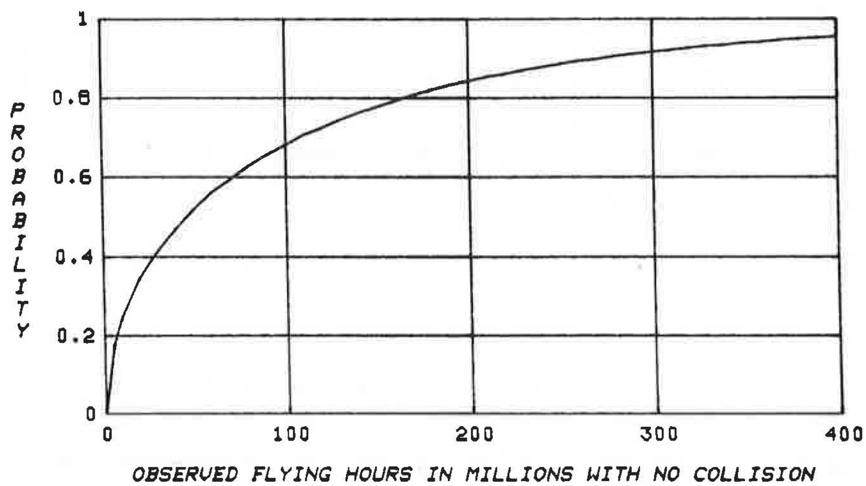


Figure 3. Degree of belief below the TLS.



lateral overlap $[P_y(S_y)]$, which is a function of the lateral distance S_y between two routes. If f(y) is the density function of lateral deviations from the track and if navigation of aircraft on adjacent tracks is independent, this probability is approximately

$$P_y(S_y) \doteq 2\lambda_y \int_{-\infty}^{\infty} f(y)f(S_y + y)dy \qquad (6)$$

which is a convolution of the distributions on the two routes.

Given estimated or concurred values (or both) for the other parameters in the collision-risk model (3), it was determined from Equation 1 that to meet the TLS in the North Atlantic requires a value of

$$C(S_y) = \int_{-\infty}^{\infty} f(y)f(S_y + y)dy \leq 6.45 \times 10^{-6} \qquad (7)$$

Kerstein (6) showed that if $f(y)$ was symmetric with zero mean and unimodal with a small, slowly varying tail, then

$$C(S_y) \doteq 2f(S_y) \quad (8)$$

Use of this result for S_y = 60 nautical miles (a track separation currently under consideration for the North Atlantic) requires that the proportion of absolute lateral deviations between 50 and 70 nautical miles off track be below 1.3×10^{-4}; i.e.,

$$\text{prob}(S_y - 10 \leq |y| \leq S_y + 10) \leq 1.3 \times 10^{-4} \quad (9)$$

In the terminology of earlier reports, this requirement is called the zeta (ζ) criterion.

Since radar coverage is available only at the ends of the routes, only one measurement of the lateral deviation from path is available at the egress point of each flight. By making the conservative assumption that the observed distribution of lateral deviations at the end of a route is applicable across the entire route length, the observed lateral deviations could be used to determine whether the proportion in the band of 50-70 nautical miles meets the requirement given in expression 9. As discussed in the following section and given for some 120 000 flights per year in the organized North Atlantic route system, the time required to decide that such a standard is being met with reasonably high confidence is practical.

The requirement in expression 9 forms a part of the minimum navigational performance standard (MNPS) that is now being used in the North Atlantic and Central East Pacific organized track systems. (Additional specifications have been formulated to ensure that the assumptions employed to derive ζ are being met.) Full details of the MNPS requirement are contained in documentation from ICAO (4).

The next section discusses in detail sampling plans designed to determine compliance with the ζ-requirement.

COMPLIANCE EVALUATION

In testing compliance with a specified navigational performance standard, it is important to consider carefully the characteristics of any statistical test employed. In particular, suppose that it is decided to consider a separation between two adjacent routes of S_y = 60 nautical miles. At this separation, we formulate two hypotheses:

$$H_0: \zeta = 1.3 \times 10^{-4}$$
$$H_1: \zeta = 2.6 \times 10^{-4} \quad (10)$$

The initiating hypothesis (H_0) corresponds to collision risk at the TLS, or 1.3×10^{-4}, while the alternative hypothesis (H_1) corresponds to a level deemed clearly unacceptable.

If S_y = 60 nautical miles, the problem faced by the decision maker is given in the following decision-analysis table:

	True State of Nature	
Decision	H_0 Is True	H_1 Is True
Accept H_0 and reduce separation	Correct decision	Type II error (unsafe)
Reject H_0 and do not reduce separation	Type I error (costly)	Correct decision

A type I error, whose probability is denoted by α, would mean that the decision to establish the separation of 60 nautical miles was rejected even though the risk at that separation met the TLS. Such an error can create serious economic penalties, particularly in use of fuel. A type II error, whose probability is denoted by β, creates an unsafe condition in that the risk, after a separation of S_y = 60 nautical miles has been decided on, exceeds the TLS.

In monitoring compliance with the standard, both α and β must be considered. The fixed-sample-size likelihood ratio test for the above hypotheses has the following decision rule: Reject H_0 if the proportion of deviations in the band of 50-70 nautical miles exceeds some value k. For fixed α and β, k and sample size N can be determined from

$$1 - \alpha = \sum_{x=0}^{k} (N\lambda_0)^x \exp(-N\lambda_0)/x! \quad (11)$$

$$\beta = \sum_{x=0}^{k} (N\lambda_1)^x \exp(-N\lambda_1)/x! \quad (12)$$

where $\lambda_0 = 1.3 \times 10^{-4}$ and $\lambda_1 = 2.6 \times 10^{-4}$. For the hypotheses in Equations 10, if α and β are both set at 5 percent, the solutions to Equations 11 and 12 are k = 22 and N = 120 900. Given current monitoring of 35 percent of the traffic in the North Atlantic, a sample size equal to 120 900 would require a sample period of approximately three years. Such a scheme is practical and could be implemented, although the time to decision is admittedly long.

If the actual proportion of deviations in the band of 50-70 nautical miles is either much less than λ_0 or much greater than λ_1, a decision could be made much sooner by using a sequential-probability ratio test. Such a procedure works as follows: Suppose N flights have been observed, of which x are between 50 and 70 nautical miles off track. Then (a) accept H_0 if $x \leq k_1$, (b) reject H_0 if $x \geq k_2$, and (c) continue sampling if $k_1 < x < k_2$, where k_1 and k_2 are functions of N. The testing starts with N = 1 and continues until a decision is made by means of either step a or step b.

For two simple hypotheses that involve Poisson parameters,

$$k_1 \doteq \{\log[\beta/(1-\alpha)]/\log(\lambda_1/\lambda_0)\} + [N(\lambda_1 - \lambda_0)/\log(\lambda_1/\lambda_0)] \quad (13)$$

$$k_2 \doteq \{\log[(1-\beta)/\alpha]/\log(\lambda_1/\lambda_0)\} + [N(\lambda_1 - \lambda_0)/\log(\lambda_1/\lambda_0)] \quad (14)$$

which are parallel lines with N plotted along the abscissa and x along the ordinate. For the hypotheses in Equations 10 and $\alpha = \beta = 0.05$, sampling continues as long as

$$-4.25 + (1.876 \times 10^{-4})N < x < 4.25 + (1.876 \times 10^{-4})N \quad (15)$$

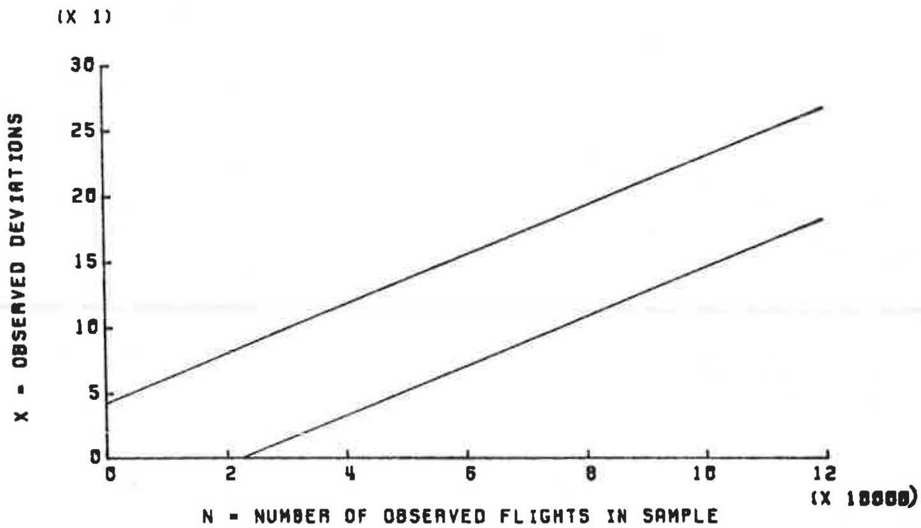
as illustrated in Figure 4. It will be observed that if the risk in the system is very large (that is, if $\zeta >> 1.3 \times 10^{-4}$), H_0 could be rejected very quickly. On the other hand, if the risk is actually much less than the TLS, a decision could be made to declare the 60-nautical-mile route separation acceptable in as few as 23 000 observed flights, given that zero aircraft were seen in the band of 50-70 nautical miles.

The expected sample size to reach a decision between the two hypotheses is given by

$$E(N|H_0) = \{\alpha \log[(1-\beta)/\alpha] + (1-\alpha)\log[\beta/(1-\alpha)]\} \div [\lambda_0 \log(\lambda_1/\lambda_0) - (\lambda_1 - \lambda_0)] \quad (16)$$

$$E(N|H_1) = \{(1-\beta)\log[(1-\beta)/\alpha] + \beta\log[\beta/(1-\alpha)]\} \div [\lambda_1 \log(\lambda_1/\lambda_0) - (\lambda_1 - \lambda_0)] \quad (17)$$

Figure 4. Sequential sampling plan for zeta.



which, in our example, yield $E(N|H_0)$ = 66 431 and $E(N|H_1)$ = 52 770, both of which are quite a bit smaller than the fixed-sample-size solution.

The above sequential-sampling plan was in fact adopted to assess whether route separation in the North Atlantic could safely be reduced to 60 nautical miles. As of the writing of this article, the data collection is in the continue-sampling region.

CONCLUSION

In estimating the risk of catastrophic transportation accidents (which occur very infrequently), direct estimation based on the frequency of occurrence of such events is not practical due to the extremely long sample periods needed to get reasonable estimates. In such cases, alternative methods, such as those described in this article, are necessary to reduce the required sampling period. In doing so, it is usually necessary to employ mathematical models based on assorted assumptions that, it is hoped, are reasonable. One of the basic tenets applied in developing and implementing collision-risk models to be used in practice was always to err when necessary on the conservative side, i.e., to make assumptions that, if not correct, would overestimate rather than underestimate the risk.

Careful design of statistical decision-making procedures is also important to control both types of wrong decisions. Too often, one or the other type of error is neglected, and the emphasis is placed solely on either safety or cost. The testing procedures described in this article and used in the North Atlantic systematically control the probability of both types of errors, which gives the decision maker assurance that the correct decision will be made with high probability.

REFERENCES

1. P.G. Reich. Analysis of Long-Range Air Traffic Systems: Separation Standards I, II, III. Journal of the Institute of Navigation, Vol. 19, 1966.
2. A.C. Busch, B. Colamosca, and J.R. Vander Veer. Collision Risk and Economic Benefit Analysis of Composite Separation for the Central East Pacific Track System. U.S. Department of Transportation, Rept. FAA-EM-77-5, 1977.
3. Methodology for the Derivation of Separation Minima Applied to the Spacing Between Parallel Tracks on ATS Route Structures. International Civil Aviation Organization, Montreal, Circular 120-AN/89/2, 1976.
4. Report of the Limited North Atlantic Regional Air Navigation Meeting, 1976. International Civil Aviation Organization, Montreal, Document 9182, 1976.
5. N.L. Johnson and S. Kotz. Distributions in Statistics: Discrete Distributions. Houghton Mifflin, New York, 1969.
6. A.R. Kerstein. Statistical Methods for the Estimation of Aircraft Collision Risk. Department of Civil Engineering, Princeton Univ., Ph.D. dissertation, 1976.

Publication of this paper sponsored by Committee on Airfield and Airspace Capacity and Delay.