technical note technical note

8

Methodology for Establishing A Target Level of Safety

Allen C. Busch

OCT 30

August 1985

DOT/FAA/CT-TN85/36

Document is on file at the Technical Center
Library, Atlantic City Airport, N.J. 08405

U.S. Department of Transportation
Federal Aviation Administration

Technical Center
Atlantic City Airport, N.J. 08405

FAA TECHNICAL CENTER LIBRARY
00013172

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents or use thereof.

The United States Government does not endorse products or manufacturers. Trade or manufacturer's names appear herein solely because they are considered essential to the object of this report.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/CT-TN85/36		

4. Title and Subtitle	5. Report Date
METHODOLOGY FOR ESTABLISHING A TARGET LEVEL OF SAFETY	August 1985
	6. Performing Organization Code
7. Author(s) Allen C. Busch	8. Performing Organization Report No. DOT/FAA/CT-TN85/36
9. Performing Organization Name and Address Federal Aviation Administration Technical Center Atlantic City Airport, New Jersey 08405	10. Work Unit No. (TRAIS) 11. Contract or Grant No. F19-04B
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Aviation Administration Technical Center Atlantic City Airport, New Jersey 08405	13. Type of Report and Period Covered Technical Note 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This paper provides a proposed methodology for the development of an acceptable level of risk for the United States en route aviation system. It provides a description of the historical development of a Target Level of Safety (TLS) used by the International Civil Aviation Organization (ICAO) and tries to show how that approach might be enhanced to be used in the United States.

17. Key Words	18. Distribution Statement
System Safety Risk Analysis	Document is on file at the Technical Center Library, Atlantic City Airport, New Jersey 08405

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	15	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

PREFACE

The author is indebted to the many persons who reviewed this document and provided critical reviews. Particular thanks go to Professor A. Odoni, Massachusetts Institute of Technology; Department of Aeronautics and Astronautics, Cambridge, Massachusetts.

TABLE OF CONTENTS

	Page
PREAMBLE	1
JUDGING THE SAFETY OF CHANGES	1
ACCEPTABLE LEVEL OF RISK	2
PROPOSED METHODOLOGY	5
COMPLICATIONS	8
PRELIMINARY CONCLUSIONS	11
REFERENCES	13

PREAMBLE

Risk and safety as an analytical decision criterion for transportation is something of an intellectual orphan. Scientists/engineers are uncomfortable with it because the methodology and whole construct for quantitative analysis must use relevant information in a way outside the "normal" constraints of the physical sciences. Yet it involves the physical sciences in a way that is outside the "normal" purview of the soft and social sciences (e.g., economics, psychology and political science). In fact, William Ruckelshaus has proposed the term "trans-scientific" for these questions.

Once use of collision risk estimates is accepted as a tool in evaluating changes to ATC system separation standards, it is necessary to decide how to use those estimates. Key to this use is the manner in which a change is judged to be "safe." The purposes of this report are to: (1) describe the various modes in which risk estimates can be employed to judge the safety of a change, (2)review present literature relevant to determination of a "safe" level of risk, (3) propose a methodology for determining a threshold of risk to use as a safety criterion, and (4) identify some important open issues that must be addressed in the process. One of the problems in employing an operational decision tool based on risk analyses and a concept of anacceptable level of safety is that of accommodating the ideal of no-risk to the practical impossibility of accepting no-risk as the criterion of satisfactory performance of an air transportation system. Adopting implicitly or explicitly a no-risk criterion or, put another way, failing to acknowledge a minimally acceptable non-zero risk level as socially/politically acceptable, almost certainly guarantees failure to meet a safety criterion and denies the decision maker the ability to consider tradeoffs among risk, capacity, and costs. There is a definite analytical relationship between system capacity and expected risk. We frequently explicitly constrain capacity because of its implicit impact on risk of collision, yet we today have no quantitative methodology for deciding what is acceptable risk. This report assumes that the methodology for analyzing and estimating risk is valid and has considerable utility in judging the acceptability of separation changes. The emphasis in the report will be to propose a "strawman" methodology for developing an acceptable level of risk, or a safety criterion, for studies of air transportation system capacity, safety/risk, and effectiveness (principally cost and operational efficiency).

Thus, the goal here is to propose a methodology for system safety criteria aware of the fact that it will have regulatory implications. However, to make an informed judgment, the regulator requires a framework for assessing risks systematically. The uncertainties involved in risk assessment, risk management, and system monitoring must be developed in consonance with each other and explained. People (principally the users of the "system") should be informed that "safety" is a social construct, an acceptable criteria which ought to be a part of each participating citizen's duty in our society.

JUDGING THE SAFETY OF CHANGES

We now address the first purpose of the report, description of the various modes in which risk estimates can be employed to judge the safety of a change. There are several approaches to this decision process. However, for now, no cost/risk benefit ratio analyses or utility functions will be discussed.

1. The current and proposed systems can be analyzed and compared to see if the proposed system has a lower estimate of risk. Included in this analysis has to be an estimate of risk for both the current and the proposed system with future demand characteristics because one might have "safer" growth potential than the other. After these analyses are complete, then one selects the more favorable system configuration from the standpoint of risk. This argument avoids the issues of whether either system is safe enough or both are overly conservative and thus potentially unduly limit capacity or whether the risk is distributed equitably.

2. Analyze the proposed system with future demand characteristics and compare it to a predetermined acceptable target level of safety and, if the risk is less than the target level, accept the proposed system.

3. Do analysis (1), but if the current system is so very conservative (possibly due to system improvements--new technology, software, procedures and/or closed loop control dynamics--relative to an acceptable target level) and the proposed system has significant gains in terms of cost and/or efficiency (but with a higher estimated collision risk which is still below a target level) opt for the new proposed system as being acceptably safe even though risk would be increased.

Implicit in the above is the assertion that the state-of-the-art in air traffic systems analysis can reasonably analyze the probability of expected accidents/collisions and that the regulators/administrators of policy will promulgate the use of quantitative risk modeling. A regulatory approach to risk and safety not based on quantified risk must eventually confront economic realities and thus become risk management in name only. It is thus impossible for decision makers to avoid dealing with quantified risk and acceptable levels of risk and the public apprehension this engenders. The only way to mitigate this fear is to make, as part of the decision process, a full disclosure of the methodology used in establishing regulatory decisions.

The use of quantitative risk assessment inevitably provokes arguments about the validity of the estimates. Despite this apprehension, there is no substitute for quantitative risk assessment in that an estimate of risk is what tells us that there is any basis for regulatory action. The alternative to no-risk assessment is to adopt a policy of either reducing all "potential" collisions to the greatest degree technology allows (i.e., making the national aerospace system overly designed) or so restricting the operational flow that the system capacity would only be one aircraft in any local airspace at any one time, thus, a policy that no technological society could survive. Therefore, risk assessment is an irreplaceable tool for setting priorities and making decisions about system changes. In the author's view, we must use and make improvements in analytical/probabilistic risk assessment with full recognition of its current shortcomings.

ACCEPTABLE LEVEL OF RISK

The latter two approaches above have imbedded in them the use of an acceptable level of risk which conceptually has been studied, endorsed, and used by the International Civil Aviation Organization (ICAO) to evaluate the risk of reduced separation standards. The more universal concept of "safe system operation" has

been studied by interested researchers in a number of fields. In particular, endeavors having the potential of low-probability, high-consequence events--such as nuclear reactor design, transportation of hazardous materials and the like--have stimulated a considerable body of scientific literature relevant to the methodology to determine an acceptable level of risk. Thus, we pursue the second purpose of this report, review of research and principles relevant to determining acceptable levels of risk and examination of those principles where they have been applied. The development of a consistent methodological approach for determining an acceptable level of risk is a natural progression from a basic understanding of risk factors and risk assessment. A methodology is an open-ended procedure. Inasmuch as the methodology is a socio/political/technologically driven process and thus is a time dependent derivation; therefore, given a change in time, say 5 to 10 years or some catastrophically related event happens, the assumptions or data for the derivation of the methodology change. Validation of any given methodology comes about only by the pragmatic acceptance of its utility in solving actual problems and to the degree to which the results of its application are adopted. Methodological development is aimed at "proving an existence theorem," not at gaining acceptance of specific methods (reference 1). However, the putting forward of a specific method should provide a "strawman" for others to break down, enhance, or use as a point of departure.

A definition to be used in this report is "A thing (system) is safe if its risk(s) are judged to be acceptable" (reference 2). This statement fits our arguments because it implies two very different and important activities: (1) measuring risk is an objective and analytical pursuit, and (2) judging the acceptability of that risk is a social/political value judgment. This emphasizes that decisions on safety are relativistic and judgmental, and thus subject to operational interpretations and application from a jurisprudence standpoint. Having said that risk/safety assessment is a societal and political process, it is necessary to focus attention on the social/political group under concern; i.e., the aviation community. A semantic distinction needs to be made here, that is, it is necessary to draw a distinction between safety and risk. Safety is herein used to define the qualitative/subjective decision process whereby a determination is arrived at concerning an acceptability criterion for an event (in this case aircraft accidents or aircraft proximity). Risk is a quantitative examination of the same event(s) to determine, by means of the best scientific and technological methods available, the expectation of that event happening. Thus, as Keeney (reference 3) and Hooden (reference 4) state, the assessment of acceptable risk means developing paradigms that transcend political and technical perspectives and begin to integrate technical and behavioral concepts. That is, direct comparisons to perceived individual risk or to risk for other groups (the population at large, for example) do not fit into this decision framework. Rather, risk assessment must "fit" the technology, the social group(s), the political processes--both from a historical and projected time reference--to which the methodological approach of an acceptable risk paradigm is to apply.

By far the most commonly cited and most unimpeachable principle in safety judgments is "reasonable." Panels of experts and law judges frequently invoke a "rule of reason" in rendering a decision (reference 2). "This definition is useful when the consumer understands that risks exist, can appraise their probability and severity, know how to cope with them, and voluntarily accept them to get benefits

that could not be obtained in less risky ways." (Meeting of The National Research Council.) These premises seem to be applicable to the question of assessing safety in the United States (U.S.) air traffic control (ATC) environment.

In 1966, the North Atlantic System Planning Group (NAT/SPG) of ICAO met, derived, and accepted a target level of safety (TLS) for aircraft collision, to help in the assessment of separation standards over the NAT ocean. Those values were slightly modified in 1967 by NAT/SPG and, in 1974, the Review of the General Concepts of Separation Panel (RGCSP) suggested another TLS value incorporating additional data and operational decisions into the process. Then, in 1977, the NAT/SPG, in developing the Minimum Navigation Performance Specification (MNPS), decided that the previously used acceptable range of safety needed to become a single number so that the numerical values of the MNPS criteria could be derived (reference 5).

The metric chosen was expected accidents in 10 million flying hours. The TLS was derived from an accumulation of all accident statistics for all countries (except the Union of Soviet Socialist Republics (U.S.S.R.) and People's Republic of China) and the accumulated flight hours for a 10-year period. From this, the ICAO panel made judgmental approximations as to what proportion of this data could be applicable to jet aircraft in en route flying as compared to arrival/departure and accidents due to something other than a collision. A judgmental factor of two for a desired improvement over the historical average was then applied to the results. They then agreed that, given no other prior information, it might be equally likely to have a collision due to loss of separation in any one of the three dimensions (x, y, or z) corresponding to the longitudinal, lateral, or vertical dimension. This exercise resulted in a consensus for the TLS of an acceptable level of vertical risk in the NAT airspace of 0.2 fatal accidents in 10 million flying hours. The NAT TLS methodology (and its specific value) is not a perfect method of approaching the problem of setting acceptable performance standards, but it has the very considerable advantage of general ICAO acceptance over a reasonable time period. It should be pointed out that the TLS of 0.2 accidents in 10^7 flying hours, established for the NAT Organized Track System (OTS) (MNPS derivation), resulted in the expectation of not more than one collision every 150 years of NAT OTS traffic (based on approximately 300 aircraft per day with 3 1/4 hours of average flight duration).

Risk has been estimated, up to now, in terms of an expected number of collisions and the associated safety criteria have been defined in terms of such rare events. Probably no one reading this report has experienced such an event and very few have even been closely associated with such an event. Thus, it is an inferred experience for which we cannot well express feelings or perceptions about our level of acceptance or avoidance. The risk-assessment process clearly supports the analytical and heuristic argument that there is a significant relationship between the expectation of collisions, near collisions, proximity, and "normal" operations. A collision cannot occur unless an aircraft is at least one-half or more of a separation standard away from its intended path. When constraining the acceptable level of collisions to some desired "rare event" level, the frequency of occurrence of deviations from planned flight path of one-half a separation standard occur frequently enough that those who are users of the ATC system are "likely" to experience it--controllers, pilots, passengers, etc. Thus, a criterion that has facets which can be related to system operation experience, which can be more easily measured and monitored (thus constrained) and whose overall system effects can be validated under reasonable assumptions, would significantly enhance its

utility and public acceptance. Currently, a 500-foot vertical deviation (or 500 feet vertical distance between two aircraft) is defined (under certain conditions) to be a "near miss" and a "system incident." Therefore, including such an event in a safety criterion provides a direct link to the operational system and helps to tie the risk-assessment process, management process, and safety criteria together for validation purposes. This is not to say that the primary function of safety criteria should not still be accidents or collisions but acknowledges a functional relationship between near collision and collisions.

PROPOSED METHODOLOGY

We now pursue the third purpose of this paper, proposition of a possible methodology for determining a safety criterion for U.S. high-altitude airspace. The general methodology used by ICAO in arriving at a TLS seems to be a reasonably pragmatic way of approaching an acceptable level of safety. It has been in use and seriously discussed by prudent people for almost 20 years and has demonstrated utility as an aid in decision making for establishing ATC separation practices. However, it should be examined for its applicability to the decision making process of a 1,000-foot vertical separation standard above Flight Level 290 (FL290) in U.S. airspace.

The ICAO TLS was derived from worldwide statistics for use in international airspace. The accumulated number of flying hours that was used was dominated by U.S. registered aircraft, but the number of accidents used in that derivation occurred with predominantly non-United States aircraft. Thus, the frequency of accidents per flight hour would be considerably different for the U.S. subset of the ICAO TLS, that is, much smaller. If ICAO-type procedures were used with just U.S. accident and exposure data for operations above FL290, then the actuarial data would contain no observed events of midair collision. However, if the general ICAO approach was believed to be practical and prudent, we could scale the relevant U.S. data to arrive at a numerical estimate applicable to U.S. airspace and practices (i.e., upper airspace en route flying and vertical separation only). A cursory examination of 1980 and 1981 data suggests that the total fatal accident rate for en route flying for U.S. domestic airspace might be a factor of 20 to 50 times better than that of international flying. Thus, it seems clear that an acceptance level of safety can and should be individually developed for the environment to which it is to be applied. Oceanic en route environments are distinct from large domestic environments, and procedures for assessment of acceptable levels of approach and departure phases of flight should be uniquely addressed. In fact, for the latter case, maybe a metric of events per flying hour would not even be the most appropriate.

The fact that, to date, no collisions due to loss of planned vertical separation have occurred in the flight regime above FL290 is a very satisfying piece of information that should be used in our deliberations about overall system safety. Thus, we can state that we have had no accidents in 10^x flying hours (e.g., during the past 25 years) and can analyze how many 10^x epochs would be necessary with no accidents to be confident at the 95 percent (or any other) level that no accidents will happen in the next 10^x flying hours. Clearly, having only one epoch with no collisions provides only weak evidence of having satisfied a criterion based on such data. Thus, multiple epochs must be examined to increase the confidence of decision making.

Another approach to this same issue would be to hypothesize how many 10^x epochs it would be necessary to examine if the criterion had been specified as a max tolerance limit of one collision in 10^x flying hours and none were observed. Again, multiple epochs of such experience would need to be examined to assure the meeting of this criterion. What this discussion intends to convey is that current system performance provides data that should be examined in support of the derivation and endorsement of an acceptable performance level. While we should be happy that no collisions in the environment under consideration have occurred, we should be enjoined from any high level of confidence concerning this achieved safety level. However, a "policy" of waiting for no collisions for firm evidence of system safety amounts to using the nation's people as guinea pigs and that is not morally acceptable. Also, it would take a long time. Therefore, continued effort for analytical modeling to support decision criteria of acceptable levels of safety need to be pursued, along with analysis of empirical evidence of "no collisions."

Since risk is itself a stochastic variable changing over time and space, it seems appropriate that the development of a given numerical value(s) should have a limited time duration associated with it. That is, over a decade or so, technology improvements in hardware, control system technology, software, computer processing, etc., should result in reduced risk. Improved reliability, ergonomics, and greater measurement accuracy would also be such that the whole risk assessment and safety picture (along with social/political changes) would need new analysis and assumptions resulting in different acceptable levels of safety.

Quite appropriate for consideration are the numerical values of a TLS and their application in a specific environment. As previously shown, the 0.2 accidents in 10 million hours of flying time indicated for NAT OTS traffic flows leads to an expected accident rate of approximately 1 per 150 years, whereas that same value (TLS), because of the greater traffic density and system flying hours, would indicate an accident rate of approximately 1 per 20 years for the U.S. en route airspace at FL290 and above. A point to consider is the effect of transforming the use of accidents per flight hour into accidents per calendar time (both concepts probably need to be psychologically acceptable and reconciled).

In this respect it should be noted that, since there is a strong perceptual preference for risk and value associated with risk elimination rather than risk reduction within the frame of reference of one's life expectancy, it would be a sound target-setting guideline to have the expected event occur outside of one's personal frame of reference; e.g., one's lifetime. Or, in terms of Tversky and Kahneman's argument (reference 6), "people evaluate acts in terms of a minimal account, which includes only the direct consequence of the act"; i.e., death by aircraft accident outside of one's lifetime. Or, to put it in terms being used by Environmental Protection Agency (EPA), Department of Energy (DOE), and others, the additional risk due to flying should not significantly impact or add to one's "normal" life expectancy of death by natural causes. "Significantly" is interpreted here as at least an "order of magnitude" less.

Viewed in this light, the TLS adopted by NAT/SPG may, in turn, have to be interpreted as aiming at a risk level equivalent to "one collision every 150 years" rather than "one collision for every 0.5×10^8 hours of flight." In other words, a TLS specified in terms of "accidents per calendar time" might be more restrictive than one specified in terms of "accidents per flight hour," in the case of high-altitude flights in the domestic U.S. airspace.

We turn next to a brief discussion of a few peculiar statistical problems that arise in connection with setting a TLS.

First, we wish to point out a common fallacy when using historical/actuarial data for estimating the probability of rare events. The following argument (reference 1) is typical:

> "Historic data always underestimates risk, since low probability events that could occur may not have had time to occur. However, estimates of behavior toward historic and experienced risks do not necessarily err on the low side. This is an important distinction that technologists making analysis of risk for technological systems fail to recognize. Actuarial data will always be necessarily incomplete because very low probability-high consequence events will not be included. Thus, the data collected needs to be complemented by (analytical) risk analysis using predictive techniques."

Unfortunately, this argument is only partially correct. While indeed low probability events "that could occur may not have had time to occur," by the same token, a low probability event may also occur early on--in which case, use of historical data may lead to an overestimate of risk. The latter, in fact, is more likely to be the case in studies of aviation safety which--sad to say--are most often triggered by aviation accidents and immediately follow such accidents in the temporal sense. To take a very simple example, suppose we start counting airline accidents (the "low probability events") on the first day of a year. If the first accident occurs, say, on the tenth day of the year, the inference that there may be 36 accidents per year may seriously overestimate risk; actually, there may not be another accident for the rest of the year. We would support the latter part of the quote concerning the need for data to be complemented by risk analysis using predictive techniques because of the probability of being incorrect using only actuarial data.

Barnett (reference 7) has recently published a most interesting note showing that using the time at which a first low probability event occurs to estimate the frequency (or likelihood) of such events will lead to significant overestimation of risk, if the event of interest has probabilities of the order of 10^{-5} or smaller, as in our case. For events with true probability equal to 10^{-7}, the overestimation of risk could be a factor of almost 20!

A second statistical concern--and a very important one from the practical point of view--is well summarized by the following statement (also from reference 1):

> "If society has perceived an institutional or technological change, effective since older events have taken place, the older data are not valid for the purpose of revealing preferences."

In our case, the ICAO and worldwide actuarial data are dominated by first- and second-generation jets under more basic or rudimentary ATC system structures, navigation capability, and procedures; ergo, caution should be observed in using this data for the U.S. ATC system, which is indeed different from the data base from which the ICAO TLS was derived.

In this same area, there is a suggestion from decision theorists that an agreed-upon level of acceptable risk should be inversely proportional to the uncertainty in the assessment of that risk. In the process of trying to estimate the risk in the system, the analyst typically produces a point estimate of the expected probability of a collision occurring, given all the best data and assumptions available. If things work out nicely for the analyst, it is also possible to produce a confidence interval around that point estimate which is indicative of the uncertainty of the whole modeling process. That is, if the whole process is very precise, the confidence interval will be small. If the process is not precise, the width of this interval grows proportionally to the noise or inversely proportionally to the information. As this interval gets larger, the probability of making a wrong decision, which can adversely impact safety, increases. One reason this point is so important is because it impacts the logic of arriving at a TLS and impacts the data collection program and analysis/modeling of collision of risk. In fact, for awhile, the NAT/SPG used the 95th percentile value of estimated risk to compare against the TLS for system acceptability.

Having said the above should not impede the effort in establishing or agreeing upon a methodology for arriving at an acceptable level of safety, but should only impact the resultant numerical value and both the operational and technological issues for modeling and analysis. Thus, it is the proposal of this report that a TLS for the U.S. airspace being considered be developed along the guidelines put forward above.

COMPLICATIONS

Pate (reference 8) has provided a cogent description of a normative process for setting a target level of safety. She argues that:

> "consistency of standards (TLS) relies on consistency of government principles of law, of information used, and finally of the safety level that is achieved. The characteristics of each specific risk may justify different levels of acceptance of risks and costs. An acceptable decision process is one in which complete information and value judgements of affected individuals are explicitly considered and made public."

In this section we identify certain particular characteristics of the high-altitude collision risk problem that, as suggested by the above quotation, call for special considerations, introduce a number of complications, and eventually necessitate some "value judgments" on the part of decision makers.

One such consideration is the fact that a high-altitude collision would be likely to involve a large number of victims, due to the nature of high-altitude traffic. The complication that this introduces is described by the following (reference 9):

> "We can compare risks as long as no more than one person is involved (in the risk event) at a time. This would include diseases, general mortality, common accidents, etc. However, when an accident involves more than one person--say N persons--then such a comparison may no longer be valid and, I assume, as a guess, that a risk involving N people simultaneously is N^2 (not N) times as important as an accident involving one person. Thus, a bus or airplane (commercial) accident

involving 100 people is as serious as 10,000 not merely 100 automobile accidents killing one person."

While one might argue with the specific "N^2" weight suggested by the author of the above, there is no question that society as a whole is extremely risk-averse with regard to accidents that could potentially involve many victims and are likely to have a small expected frequency of occurrence. In fact, this type of risk-aversion is known as "catastrophe avoidance" in decision theory (see reference 10) and can be succinctly described as follows:

> "It has been observed that most people feel that a small probability of a catastrophic loss of life is worse than a larger probability of a smaller loss of life, if the expected number of fatalities is the same in both cases. For example, given the choice between a 10^{-6} probability that a small private airplane with 4 people on board will crash and a 10^{-8} probability that a Boeing 747 with 400 people on board will crash, most people would apparently select the former."
>
> "More formally, x fatalities with probability p is preferred by y fatalities with probability q for any $x < y$, given that $px = qy$."

Many observations on the phenomenon of catastrophe avoidance are contained in reference 11. In the setting of air traffic above FL290, one can anticipate that catastrophe avoidance would impose (on judgmental grounds) a more "conservative" TLS than might result from an analysis that would examine only expected values (averages).

A second relevant characteristic of the ATC system, in general, and of high-altitude traffic, in particular, is that many of the immediate causes of system overload are transient or ephemeral phenomena and the behavior of the system cannot be understood or analyzed in terms of long-term averages in the time or the spatial domains. Thus, the characterization of safety performance merely through "steady-state" analysis would be inadequate and not very useful for TLS-setting purposes. This has no negative impact on the methodology advocated and espoused here. However, it does suggest a level of complication to the problem because there is the distinct inference that not just average risk should be addressed, but also variability of risk. Also, analysis of collision risk in the ATC environment has indicated that risk itself is a stochastic random variable that is neither normal nor symmetric in its distributional characteristics. Thus, a simple one-dimensional characterization of the variation or distribution of risk (such as average risk or variance) is insufficient even from an analytical standpoint, let alone its use in an operational decision making framework.

Directly related to the above is a third decision-theoretical issue, that of risk-equity. This is a complication which has quite recently been discussed by Keeney (reference 10) and can be succinctly described in terms of the following example:

> "Suppose two technical projects, A and B, expose two individuals, X and Y, to great physical risk. Under project A, the probabilities of being killed are (0.5, 0.5) for X and Y, respectively. Under B, the corresponding probabilities are (0.51, 0.49), meaning that X has probability 0.51 of being killed and Y probability 0.49. In both cases, the event "individual X is killed." (Note that the expected number of fatalities under either project is the same (and equal to 1))."

"There is a lot of evidence that, if forced to vote for either A or B, most people would vote for A on the premise that A is more "equitable" with respect to risk distribution than B. (This also assumes that the "voters" do not know anything about the individuals X and Y who are at risk, i.e., this is a case of "statistical" risk.)"

An interesting behavioral experiment relating to both catastrophe avoidance and risk equity is described in reference 12.

Translated into ATC terms, the risk-equity argument seems to be saying that, given two choices: (1) Route A made up of 6 segments, each of which has an average risk of 0.03 accidents in 10 million flying hours, or (2) Route B, also of 6 segments, where the average risk per segment varies between 0.01 and 0.09 per 10 million hours with the overall average risk, when properly weighted, equal to 0.03 per 10 million hours, most people would opt for, all other things being equal, Route A to get to their destination.

More generally, one would certainly expect that, in the case of high-altitude safety, an informed public (or decision makers) would have a strong preference for risk equity. This has some very important implications; e.g., if a TLS were to be set for the entire CONUS system, one would then expect that this TLS would not be used as an average target for the entire system (meaning that in some parts of the U.S. or during certain hours of the day, the actual level of safety would be above and below the TLS), but rather as an upper level on how much risk can be tolerated at any location and any time. This would mean that the TLS might need to be met at the busiest locations in the country (e.g., the Northeast) during the peak traffic hours of the day.

The principal common thread running through all the issues noted in this section is the need to consider criteria other than, and in addition to, a mere central tendency estimate in setting a TLS. There is ample precedent in practice for doing exactly this: insurance companies that underwrite risk establish the rates for low-probability risk events, not so much on the basis of averages but for some expected level above the average, and apply a multiplier depending on the distance above the average. It is that elevated level that is used when setting rates for policies. Thus, insurance rates will probably be influenced by the methodology employed in the analysis of risk and in setting appropriate standards.

To go one step further, from a market and economics standpoint, it can be argued that the overall system risk (average) should not be regulated, but set by the marketplace. That is, if the public perceives that air traffic is not risky, relative to other transportation modes, and is cost-effective, they will opt to fly and vice versa, provided the marketplace has allowed all modes of transportation to freely compete from a cost/resource/benefit standpoint.

Thus, regulation/legislative action should be used to ameliorate peak risk, inequities or to balance risk. This is probably not viable for today's marketplace, but it does tend to reinforce the argument that the methodology for establishing an acceptable level of risk addresses issues more complex than just the average or central tendency.

Finally, it is clear that an overview of the literature and conferences held on the generic problem of assessment of acceptable safety levels and risk in our environment, that once a tentative or interim decision or consensus has been arrived at by those responsible for promulgating or using the information, the presentation of the results can be critical. Since the individual level of one's own intuitive interpretation and personal acceptance will be conditioned by how it relates to his/her own experiences, a TLS must be conveyed by presenting diverse arguments as broadly derived as possible to satisfy the largest audience possible.

PRELIMINARY CONCLUSIONS

The conclusion or recommendation to be made is that work should proceed in deriving acceptable level-of-safety criteria to be used in the decision making process in assessing the acceptability of a change in the operating procedures to a 1000-foot minimum vertical spacing between aircraft above Flight Level 290 (FL290). These criteria should be derived from the following methodology:

1. Examine the historical accident data and accumulated flight hours associated with that time period, similar to that studied by International Civil Aviation Organization (ICAO) for the North Atlantic (NAT) target level of safety (TLS), and derive an estimate.

2. Examine historical life experiences for the U.S. population and determine what number would result if the acceptable level for air transportation were to be such that flying would not significantly contribute to a reduction in one's life expectancy, with a reasonable level of confidence.

3. There would appear to be something like five orders of magnitude difference between what the current risk for automobile travel and the risk due to residing in the proximity of nuclear power plants. Accidents for en route aircraft carrying hundreds of people would seem to fall somewhere between these two brackets of risk. Thus we should examine this interval to see what might be a prudent target level for air transportation.

4. Examine the relative level of uncertainty of this whole estimation process and the effect of making errors of decision (Type II errors), so as to propose a better process as an alternative to the estimation process of using a classical mean-point estimate.

5. Examine the question of using peak and average levels, or some other multivariable measure, for a decision criterion.

6. Examine the need to relate events per hour of exposure with events per calendar time, as one other way to evaluate how acceptable is the acceptable level of risk.

7. Analyze the functional relationship between the expectation of a deviation of one-half a separation standard and a full-separation deviation and include in the overall "acceptable safety criteria."

8. Examine the currently achieved level of safety and the amount of time necessary to sample so that confidence in the achievement of this level is high.

9. Attempt to synthesize the above attributes into an acceptable level of performance for expected aircraft collisions as an aid to decision making for the acceptability of a 1000-foot separation above FL290.

Some of the points enunciated here are not necessarily to be employed directly in deriving a safety criterion for vertical separation but to provide a broad perspective within which to evaluate the relative acceptability. Also, this methodology should provide the necessary insight as to the relative comparison of an aviation-related criterion with other aspectsof our overall social/political environment.

Further elaboration in this area should await reactions and/or consensus for this facet of the methodology.

Consensus is achieved more by default of criticism than by active testing and espousal.

REFERENCES

1. Rowe, William D., An Anatomy of Risk, John Wiley and Sons, New York, N.Y., 1977.

2. Keeney, R., and Raiffa, H., Decision With Multiple Objectives - Preferences and Value Trade-Offs, John Wiley and Sons, New York, N.Y., 1976.

3. Hooder, J., Accidents in Norway Oslo, Risk Research Committee Royal Norwegian Council for Scientific and Industrial Research, Oslo, Norway, 1980.

4. Lawrence, William W., Of Acceptable Risk - Science and the Determination of Safety, William Kaufamann, Inc., Los Altos, California, 1976.

5. North Atlantic System Planning Group (NAT/SPG of ICAO):

from Working Group Papers, Meeting #2 (1966)

from Working Group Papers, Meeting #3 (1967)

from Working Group Papers, Meeting #10 (1974)

from Working Group Papers, Meeting #13 (1977)

from Working Group Papers, Meeting #15 (1979)

6. Tversky, G., and Kahneman, D., The Framing of Decisions and the Psychology of Choice, Science, Vol. 211, 1981, pp. 453-458.

7. Barnett, A., "Misapplications Review: Mea Culpa," Interfaces, 14, No.2, March-April 1984.

8. Pate, M.E., Acceptable Decision Processes - Acceptable Risks in Public Sector Regulation, IEEE Transaction on System, Man, and Cybernetics, Vol. SMC-13 #2, 1983.

9. Wilson, R., Examples of Risks Benefit Analysis, CHEMTECH, October 1975, pp. 604-607.

10. Keeney, R.L., "Equity and Public Risk," Operations Research, 28, No.3 May-June 1980.

11. Ferreira, J., and Slesin, L., Observations on the Social Impact of Large Accidents, Technical Report No. 122, MIT Operations Research Center, 1976.

12. Hammerton, M., Jones-Lee, M.W., and Abbott, V., " Equity and Public Risk: Some Experimental Results," Operations Research, 30, No.1, January-February 1982.