Title
The Development of An Accident Database to Structure Land Use Regulations in Airport Runway Approach Zones

Permalink
https://escholarship.org/uc/item/37g9f0kg

Authors
Cooper, Douglas L.
Gillen, David

Publication Date
1993-11-01

Institute of Transportation Studies
University of California at Berkeley

The Development of An Accident Database to Structure Land Use Regulations in Airport Runway Approach Zones

Douglas L. Cooper
David Gillen

RESEARCH REPORT
UCB-ITS-RR-93-13

Prepared for the
California Department of Transportation
Division of Aeronautics

November 1993
ISSN 0192 4095

Table of Contents

The preparation of this document was financed in part through a planning grant from the Federal Aviation Administration as provided under Section 505 of the Airport and Airway Improvement Act of 1982 as amended. The document was prepared by the Institute of Transportation Studies at U.C. Berkeley under contract to the California Department of Transportation, Division of Aeronautics. The content of this report reflect the views of the Institute of Transportation Studies, who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policy of the FAA or the Division of Aeronautics.

1.0 Introduction

One of the key issues addressed in preparing the 1993 revision of the Airport Land Use Planning Handbook was the potential effect on safety brought about by changing land use in the vicinity of airports . Traditionally, regulation of land adjacent to airports has been concerned with achieving compatibility between land use and noise levels generated by airport operations. As new generations of aircraft with quieter engines have been introduced, however, noise impact has been appreciably reduced, moving the constant dB noise contours closer to the airport boundary. As a consequence, local communities may receive development pressure to alter land uses around their airport, particularly in the approach and departure corridors. This, in turn, has raised concern regarding the safety implications of permitting development in areas which may have a higher exposure to aircraft accidents. The safety concerns are both for people on the ground as well as those onboard the accident aircraft who may be placed at greater risk with certain land uses.

One solution to these concerns is to keep the areas completely clear of obstructions and development. This, however, may create a sizable cost in terms of forgone land use opportunities. Sound policy requires an assessment of the costs and benefits of alternative land uses and a balancing of the risks with the potential benefits. The costs will depend upon the probability of an accident and the land use being considered and will most likely vary from airport to airport. The benefits will be derived from the development of the land and its associated use and productivity.

In order to provide guidance for local agencies responsible for establishing land use regulations, there needs to be a better understanding of the risks involved and the consequences of particular use restrictions. A first step in this process is to develop a comprehensive database of aircraft accidents on or in the vicinity of airports including such information as location relative to the runway, type of aircraft, phase of flight, and relevant airport characteristics.

This report presents and describes 400 aviation accidents which occurred within five miles of an airport. Section 2 contains a description of the development of the database and a discussion of the criteria used in selecting accidents for the database. Section 3 provides a description of the database itself as well as a set of statistics that provide a comprehensive overview of the accidents. A set of aircraft accident contours developed from the accident data points is presented in section 4. The purpose of these contours is to provide a picture of the distribution of accidents over space. Section 5 contains a brief discussion of modeling aircraft accident location probability. A model is not developed in this report. The section also provides a brief discussion of the use of the database in examining airport operations and management strategy. The final section presents an approach to be used in a subsequent empirical investigation using the database supplemented with aircraft movement data.

2.0 Development of the Accident Database

In order to investigate the safety implications of structuring or altering current land use around airports and to form a basis for weighing risks against potential benefits of development, it was necessary to develop a model useful in assessing the probabilities associated with aircraft accident locations. This model required a substantial and detailed accident database which includes the following:

(1) Precise distance of accident site with respect to runway used
(2) Type of operation (takeoff, landing, etc.)
(3) Aircraft type
(4) Date of occurrence
(5) Runway length
(6) Weather conditions (VFR/IFR; day/night; wind)
(7) Injuries on board aircraft
(8) Injuries on ground
(9) Damage to structures on ground

2.1 Initial Research

The first task in collecting the required data was to review potential sources of information in order to establish the quality, nature, and applicability of the available material. With the exception of precise location of accident site, the desired information listed above is readily from the FAA.[1]

For accident site location, however, the solution is not as simple. As noted in the Reid-Hillview land use study by Hodges & Shutt (March 1991) and "Location of Commercial Aircraft Accidents/Incidents Relative to Runways" (DOT, July 1990), precise information regarding accident sites relative to the associated airport runway is difficult to obtain.

The NTSB is responsible for the investigation of all aircraft accidents (defined as "an occurrence associated with the operation of an aircraft...in which any person suffers death or serious injury...or in which the aircraft receives substantial damage or is

---

[1] An alternative is a commercial product available on CD-ROM from Avantext, Honey Brook, PA, which was created using FAA tapes and is fully formatted and searchable on 161 datafields.

destroyed"), although in many cases the field investigation is delegated to the FAA. Information from the investigation goes into an accident docket, most of which later becomes part of the Factual Report - Aviation (NTSB Form 6120.4). The Factual Report asks for accident location in three separate places: Sections 28 and 29 which call for "Distance From Airport Center" (to the nearest statute mile) and "Direction From Airport"; Supplement I—Crash Kinematics, which requests latitude and longitude to the nearest minute; and Supplement Q—Airport/Airstrip, which requires "Bearing and Distance" in nautical miles (note different unit of measurement than Section 28) from the applicable runway end. Even on those occasions when some or all of this information is supplied, which in the case of Supplement Q is seldom, it is often suspect due to rounding, imprecise origin of measurement (where is an airport's center?), or conflicting information. While frustrating to researchers, it must be remembered that the primary goal of the NTSB investigation is to determine the cause of the accident, not its exact location.

When latitude & longitude are included in the Factual Report, their usefulness is limited. At 40 degrees north latitude (roughly the middle of the U.S.), one degree of longitude is approximately 53 1/2 miles, while one degree of latitude (which remains constant) is approximately 69 1/2 miles. An accident location rounded to the nearest second would, therefore, put the site within a rectangle that is roughly 100x80 feet. While this may be accurate enough for our purposes, it is based on distances estimated from the nearest known fix by people with no specific training in this skill. In addition, Supplement I only requires latitude and longitude to the nearest minute.

Bearing and distance can be equally troublesome, even when the origin point of the measurement is given. Bearings are often given using eight points of the compass (N, NE, E, etc.) which, at a distance of two miles from the airport, for example, could put the reported site as much as 3/4 of a mile from the actual site. Even rounding to the nearest 10°, a common practice, can shift an accident location by as much as 460 feet at a distance two miles. Distance is generally given in rounded form also, usually to the nearest 1/4 mile. While latitude and longitude and bearing and distance are inadequate by themselves, they are occasionally useful in corroborating data contained in other sections of the report.

The search for additional sources of information yielded the following:

State Aeronautics Offices in Other States - Of the fifty states, only seven (Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, Ohio, and Rhode Island) do any of their own investigating. Even among these seven, the investigations are usually quite limited and often done for a specific purpose (e.g., Massachusetts checks to make sure that the aircraft owner is insured or has sufficient net worth to cover damages).

Aircraft Owners and Pilots Association - This organization relies almost entirely on the FAA and the NTSB for the information they collect and publish regarding accidents. There is no new or enhanced information available from this source.

Airline Pilots Association - The union for airline pilots is involved in investigations of commercial air carrier accidents only and would be of no help in general aviation accident locations. Their reports on commercial accidents would shed no new light on location.

Aircraft Insurance Industry - A check with two of the major general aviation insurance companies (USAIG and Associated Aviation Insurance) yielded little in the way of useful results. Except in cases where location is useful in assessing fault (such as a defective navaid or cockpit instrument) specific accident location is not of interest and, therefore, not included. Even in cases where accurate site data is given, two problems arise. First, finding the cases that would be of use would require a hand search through individual accident reports. Second, it would probably be difficult to obtain permission to go through the files as the information is considered proprietary and also could expose the company to lawsuits.

Newspapers - Information from this source is essentially limited to published reports and pictures. As protection against possible erosion of first amendment rights, unpublished notes and photographs are not released to the public, even under court order. The probability of published stories or photographs adding to the information available from other sources is small.

Local Police and Fire Departments - A check with several California emergency agencies regarding specific accidents within their jurisdiction yielded little in the way of official (written) information other than that which is already included in the Factual Report. The only way this source could be useful would be to contact the individuals who went out on the call and ask them to try to pinpoint the accident site.

2.2 Final Search Procedure

The final search procedure and database structure evolved over the course of a review of the NTSB computerized database (which consists of short accident summaries called 'minibriefs') and a week spent in Washington, D.C. looking through complete NTSB accident files where, it was hoped, the required location information could be found buried in parts of the factual report that do not make it to the computer database.

The conclusion drawn from the review was that, while the required information can be found, it is a very labor intensive task to produce it. This is because approximately five 'minibriefs ' must be read for every one that meets our location criteria (accident is at least 100 feet, but no more than five miles, from the runway) and out of those that are acceptable only one out of six has sufficient information in the full report (including investigator, pilot, police, and witness statements) to allow us to establish precise location. None of this can be done through a computer search. This low yield (one out of thirty) forced us to go outside our initial four state search area (California, Texas, Florida, and Arizona) in order to get the desired number of datapoints The final search included all 50 states.

Numerous decisions also had to be made regarding definitions and standardization of data. These decisions included:

- Only accident records will be searched due to lack of information on incidents
- Accidents will be plotted as an X-Y scatterplot with the runway approach threshold as the 0-0 point of the X-Y axes
- Where no other information is available, T/O roll will be assumed to have begun at the beginning of the runway pavement
- Y distances for departures will be computed using distance from T/O roll beginning point to departure end of runway as runway length
- In the case of touch-and-goes and emergency returns to airfield after takeoff, a departure becomes an approach only after the aircraft is established downwind or, in the case of a straight in approach, is established on final
- A missed approach becomes a departure once a controlled climb is established.
- Accident location will be the point of initial impact or touchdown
- Accidents involving inflight collision with an obstruction will be included but will be noted as such
- Pilot control will be recorded so that separate plots can be made, if desired, of those that had no control and those that were able to choose a landing spot
- Helicopter accidents will not be included due to our inability to link them to a specific runway, airport, or even an improved landing site.

The following, then, is the process to determine one datapoint:

- Read NTSB computer records ('minibriefs') to obtain list of accidents that fit criteria
- Request full Factual Accident Report (on microfiche) of selected accidents from NTSB
- Search all documents in microfiche record for information on exact accident locations
- Where necessary, use additional resources (e.g., telephone cross-directories, local street maps, USGS Quadrangle maps, etc.) to determine bearing and distance information
- Add record to accident database.

Following the above procedure, 12,700 accident 'minibriefs' were reviewed resulting in a request for full Factual Reports on the 2,633 accidents that fit our profile. Due to problems with missing microfiche at the NTSB, only 2,450 accident files were received. A search of these files resulted in 400 datapoints. Appendix B contains a list of all variables included in the accident database and the sources of information for each variable as well as an example of the full information contained in one database record.

3.0 Characteristics of Aircraft Accidents

Table 1 provides a list of the of the average values of the variables contained in the 400 accidents that make up the database. These can be distinguished in terms of a number of characteristics including: accidents were almost equally divided between arrivals (47.5%) and departures (52.5%); the majority of approach accidents occurred during visual approaches (67.89%); approximately 65% of accidents took place during the day ; and 74.5% of all flights were conducted under visual flight rules (VFR).

One of the more important accident parameters is whether the accident involved a loss of pilot control or not. The reason is that under controlled conditions a pilot will utilize open space to land the aircraft which would, in principle, reduce the number of injuries and amount of damage. We found that more than half, 58.5%, of the accidents occurred under conditions of the pilot having no control and 30.5% with the pilot having some control. In the remaining 11% of the accidents, degree of control could not be determined.

Inflight collisions occurred in 166 cases or 34% of all accidents in the database. In roughly half of these, the collision was a major factor in the accident. Collisions include contact with trees, wires, fences, and buildings but not other aircraft.

Table 1

Average Values of Select Variables in Accident Database

Category		Number*	Percentage
Accidents		400	
Arrivals		190	47.5
Visual Approaches		129	67.89
Precision Approaches		45	23.68
Non-Precision Approaches		15	7.89
Departures		210	52.5
Time			
	Day	262	65.5
	Night	119	29.75
	Dusk	17	4.25
	Dawn	2	0.5
Flight Rules			
	VFR	298	74.5
	IFR	102	25.5
Pilot Control			
	None	234	58.5
	Some	122	30.5
	No Information	44	11
Inflight Collision		166	41.5
Factor?			
	Yes	87	52.4
	No	79	47.6
Number Of Engines			
	Single	290	72.5
	Twin	110	27.5
Landing Pattern			
	Left	182	87.5
	Right	26	12.5
Aircraft Damage			
	Destroyed	299	74.75
	Substantial	97	24.25
Accidents With Onboard Fatalities		235	58.75
Accidents With Ground Fatalities		4	1.0
Accidents With Onboard Serious Injuries		109	27.25
Accidents With Ground Serious Injuries		5	1.25
Average Runway Length		4935	

*Numbers in each category may not add up to 400 due to missing data in some files

In the vast majority of accidents, injuries, fatalities and damage are associated with the aircraft rather than with people and property on the ground. In 75% of the accidents the aircraft was destroyed and over half (58.75%) of the accidents involved onboard fatalities.[2] Ground fatalities or serious injuries occurred in only 2.25% of the accidents.

Three important parameters contained in the data set are runway length, and both accident bearing and accident distance measured from the appropriate runway end. Runway length can be crucial in determining whether an accident occurs and, given its occurrence, whether or not it fits the criteria for the database. A departure accident which occurs on takeoff from a 2,000 foot runway and is located on or near the extended centerline 3,000 feet from the departure end would quite possibly not even have been an incident had the runway been 6,000 feet in length. The aircraft could have safely put back down on the runway and taxied or been towed back to the ramp. The average runway length at airports where the accidents occurred was 4,935 feet with a standard deviation of 2,205 feet. Sixty-eight percent of the sample falls within a range of plus or minus one standard deviation from the [arithmetic] mean. Exhibit 1 illustrates the distribution of runway lengths in the sample.

The importance of the average accident bearing and distance from the approach or departure end of the runway is in establishing the spread, or area of greater or less accident occurrence. The average relative bearing of arrival accidents from the approach end of the runway was 187 °, or 7° left of centerline, and had a standard deviation of 56°. A spread of plus or minus one standard deviation from the mean covers an area from 243° to 131° relative for a total of 112°. The mean accident distance measured along the Y axis from the runway threshold was 3,391 feet with a standard deviation 5,600 feet. Thus a range of plus or minus one standard deviation from the average extends from 8,991 feet before to 2,209 feet past the runway threshold and contains 78% of the approach accidents.

Departures, as one would expect, have a quite different set of statistics. Measured from the departure end of the accident runway, the average bearing of an accident was 354° relative, or 6° left of centerline, with a standard deviation of 75°. Thus the spread of departure accidents is greater than that for approaches. The mean departure accident distance measured along the Y axis was 1,294 feet with a standard deviation of 3,780 feet. If one moves the point of measurement for departure accidents from the departure end of the accident runway to the beginning point of the takeoff roll, the mean accident distance is 5,750 feet with a standard deviation of 4,498 feet and the mean angle changes to 360° with a standard deviation of 22°.

---

[2] In contrast, the NTSB reported in its "Annual Review of Aircraft Accident Data, U.S. General Aviation" that for calendar year 1989 only 390 or 20% of the 1.984 fixed wing aircraft accidents involved fatalities. The NTSB data, however, include accidents that occured on the runway (e.g., wheels-up landings) that tend to be less severe than the off- runway accidents that were the focus of this study.

Exhibits 1 through 5 provide a graphic representation of the variables just discussed.

Exhibit 1

Distribution of Runway Lengths in Database

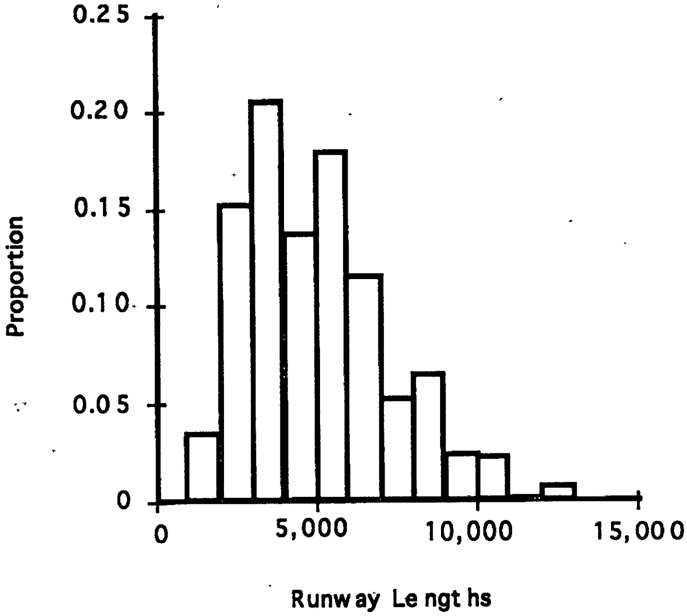


Exhibit 2

Cumulative Distribution of Runway Lengths in the Database

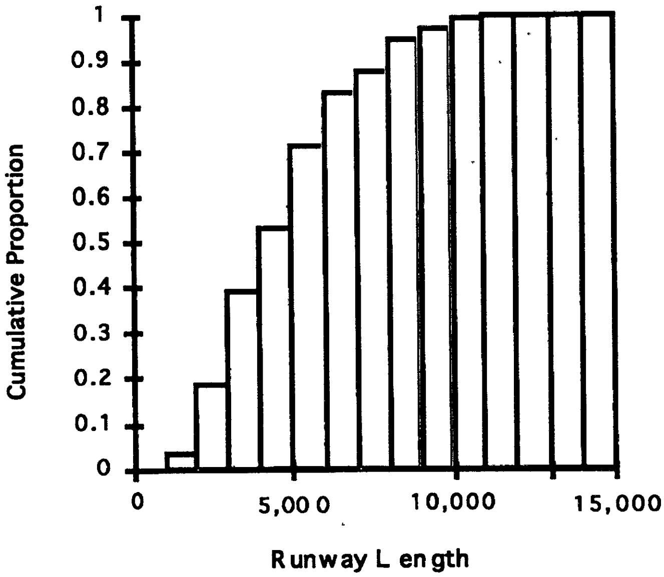


Exhibit 1 shows the distribution of runway lengths contained in the sample. There is clearly a wide dispersion about the mean. Approximately 71% of the runways were less than 6,000 feet long with the majority of the remainder in the 6,000 to 9,000 foot range. As discussed earlier, runway length is an important factor in that it can materially affect the severity of an accident and thus the accident's inclusion in our database. This, in turn, will affect the accident probability location contours (as illustrated in the next section) and the safety-risk model.

The striking differences between the location and dispersion of arrival and departure accidents are illustrated in Exhibits 3 through 8. Exhibits 3 and 4 show the distribution of accident locations in terms of distance from the appropriate runway threshold for arrival and departure accidents respectively. This distance is measured along the runway centerline. Exhibits 5 and 6, which provide the cumulative distributions of distances for arrival and departure accidents, emphasize the difference between the arrivals and departures, with the cumulative percentage of departure accidents markedly trailing arrivals until a distance of approximately 7,000 feet is reached. By 15,000 feet, 91% of the departure accidents have occurred compared to 82% of the arrival accidents.

Exhibit 3

Distribution of Arrival Accident Distances Measured From Landing Threshold

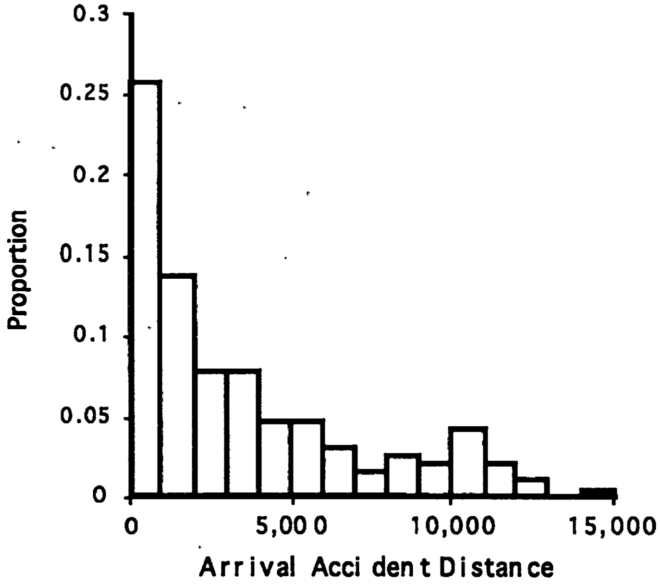


Exhibit 4

Distribution of Departure Accident Distances Measured From Takeoff Roll Beginning Point

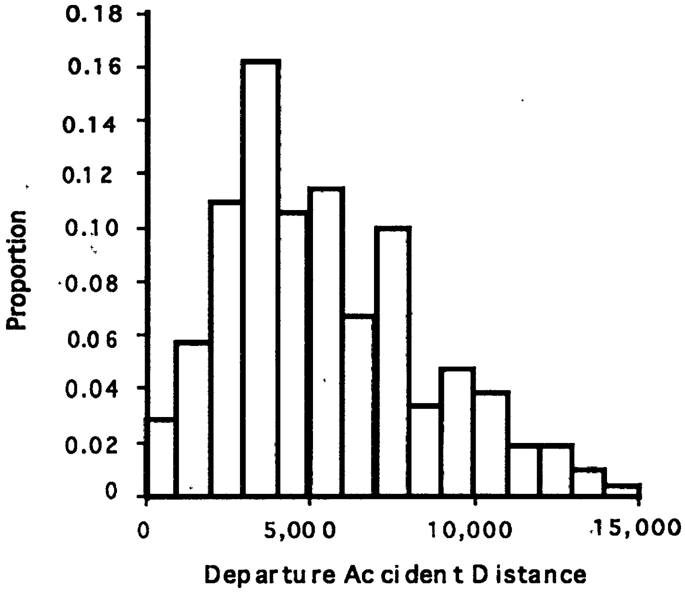


Exhibit 5

Cumulative Proportion of Accident Location for Arrival Accidents

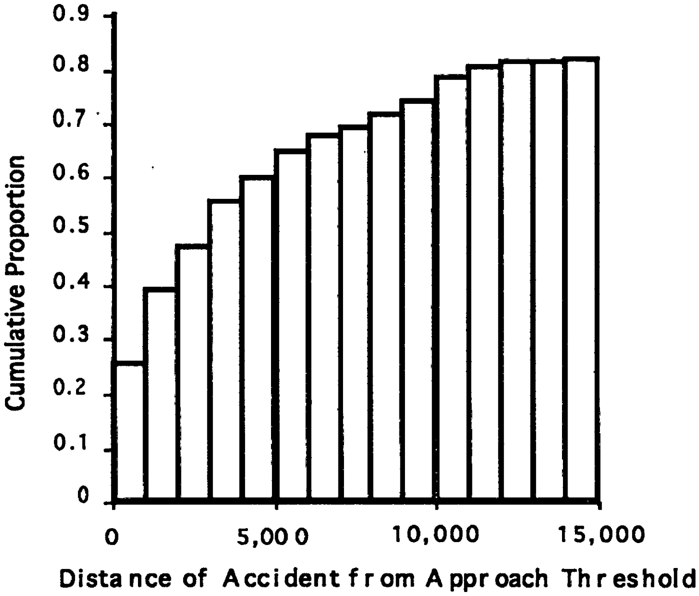
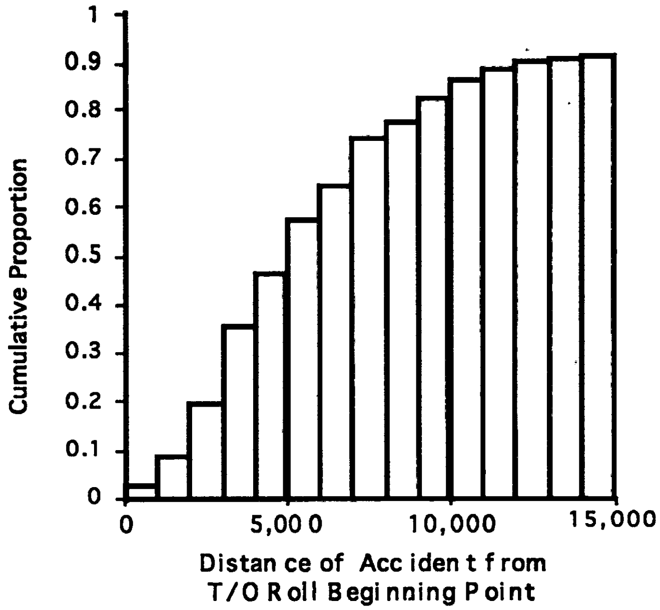


Exhibit 6

Cumulative Proportion of Accident Location for Departure Accidents



The "Y-distance" is only one dimension of the location, however. The distance away from the centerline is also important. This dispersion from the centerline is captured by the measured relative bearing of the accident from the runway approach or departure threshold as appropriate. These are illustrated for arrival and departure accidents, respectively in Exhibits 7 and 8. For arrival accidents 72% are within an arc 30° on each side of the extended runway centerline. The distribution of the remainder is asymmetric with the larger proportion being to the left. Departure accidents are substantially more dispersed with only 48% of the accidents occurring within a 30° arc on each side of the extended centerline with the remainder evenly distributed on both sides.

The exhibits illustrate some of the problems of assessing the data in only one dimension. In the following section we develop accident location contours which provide a two dimensional picture of the distribution of accidents over space.

Exhibit 7

Distribution of Arrival Accident Bearings
(Measured from Runway Approach Threshold)

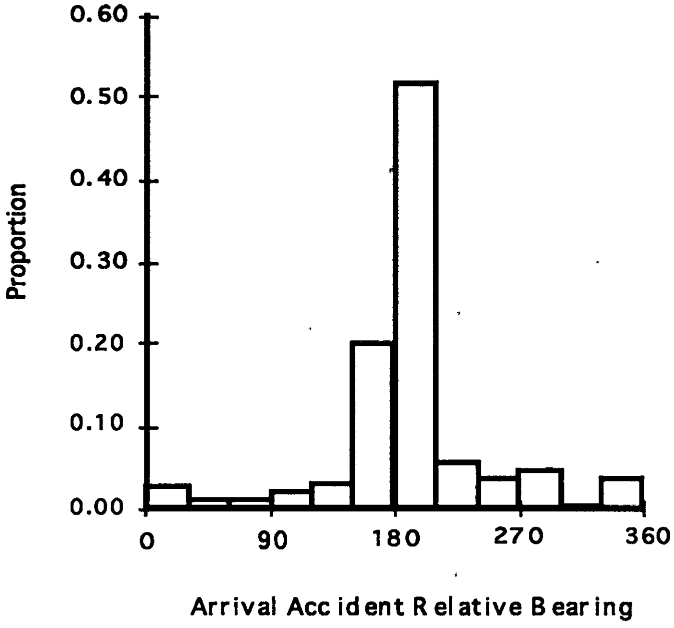
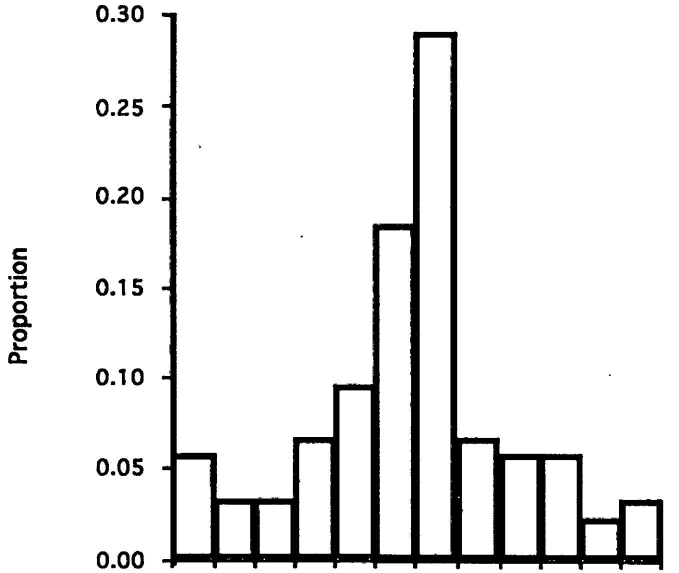


Exhibit 8

Distribution of Departure Accident Bearings
(Measured From Runway Departure End)

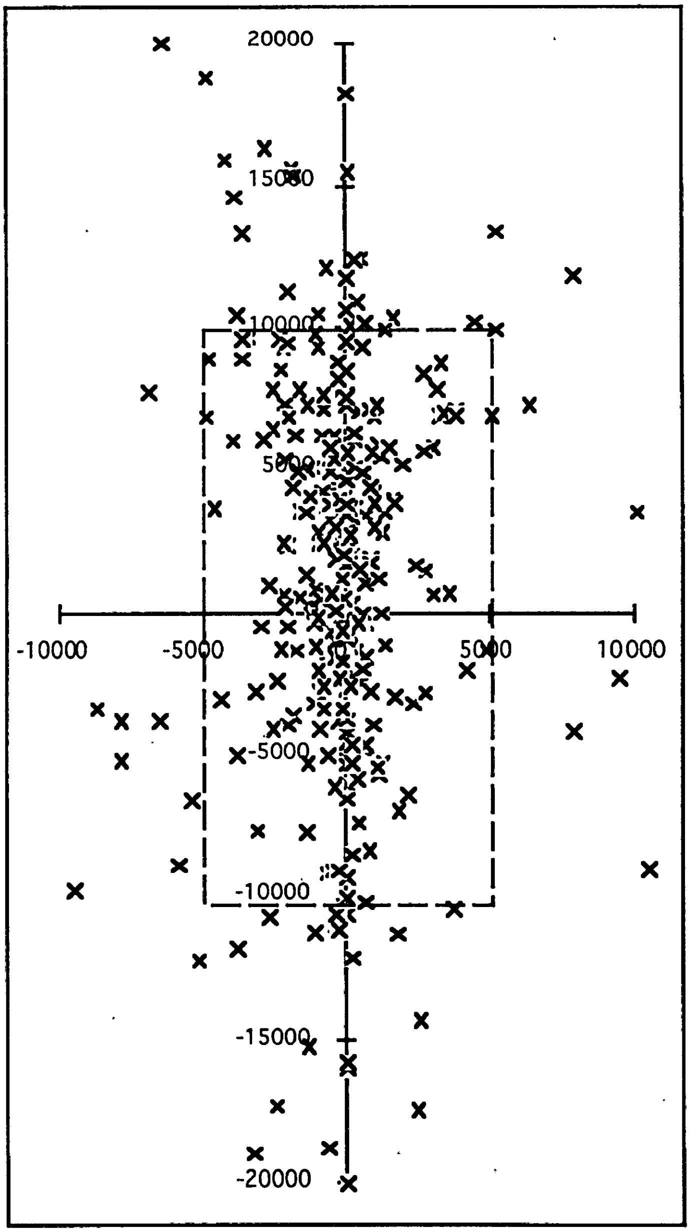


4.0 Accident Contours

Almost all (97.5%) of the accidents in the database lie in an area 40,000 feet long by 20,000 feet wide. As shown in Exhibit 9, most of these points (82%) are within a 20,000 foot by 10,000 foot area. For the sake of clarity and scale, only this smaller area is used for depicting accident scatterplots and contours.

Exhibit 9

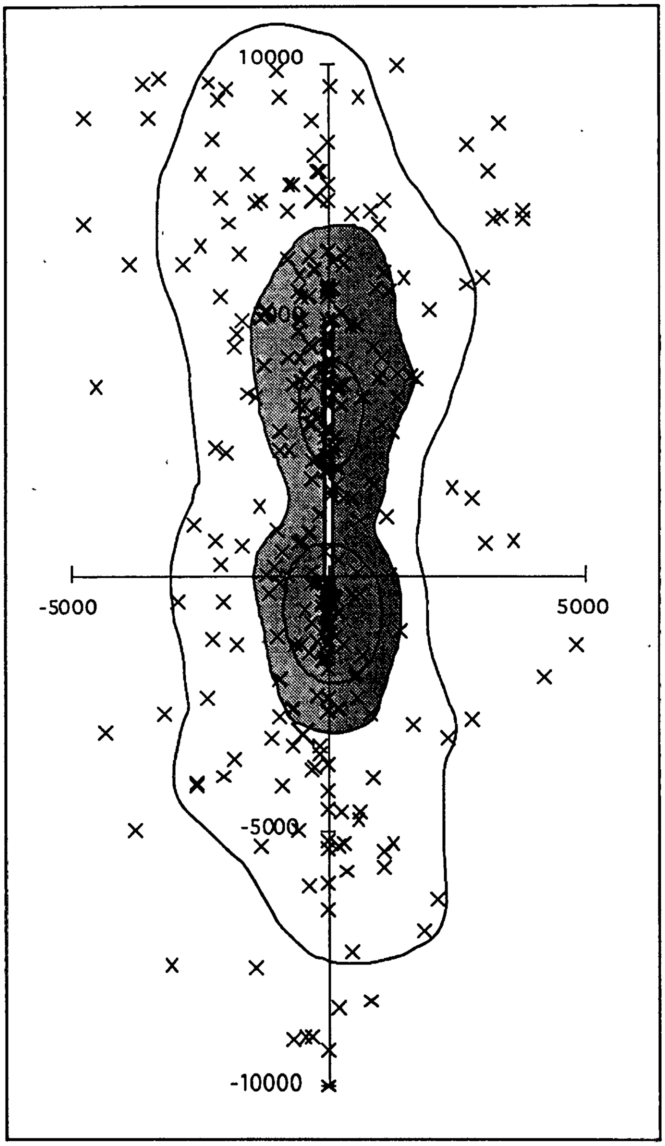
Scatter Plot of All Accidents



The accident contours in Exhibit 10 provide a means of visualizing the geographic areas that contain 25%, 50%, and 75% of the 400 accidents. The shaded area, for example, includes the 25% and 50% contour lines and thus covers 50% of the accident sites. Included in the drawing is a 4,935 foot runway (the mean runway length in the database sample). The contour lines are created by smoothing the data and then joining cells that have the same number of accidents. A cell is approximately 300 feet by 300 feet. (See Exhibit 11)

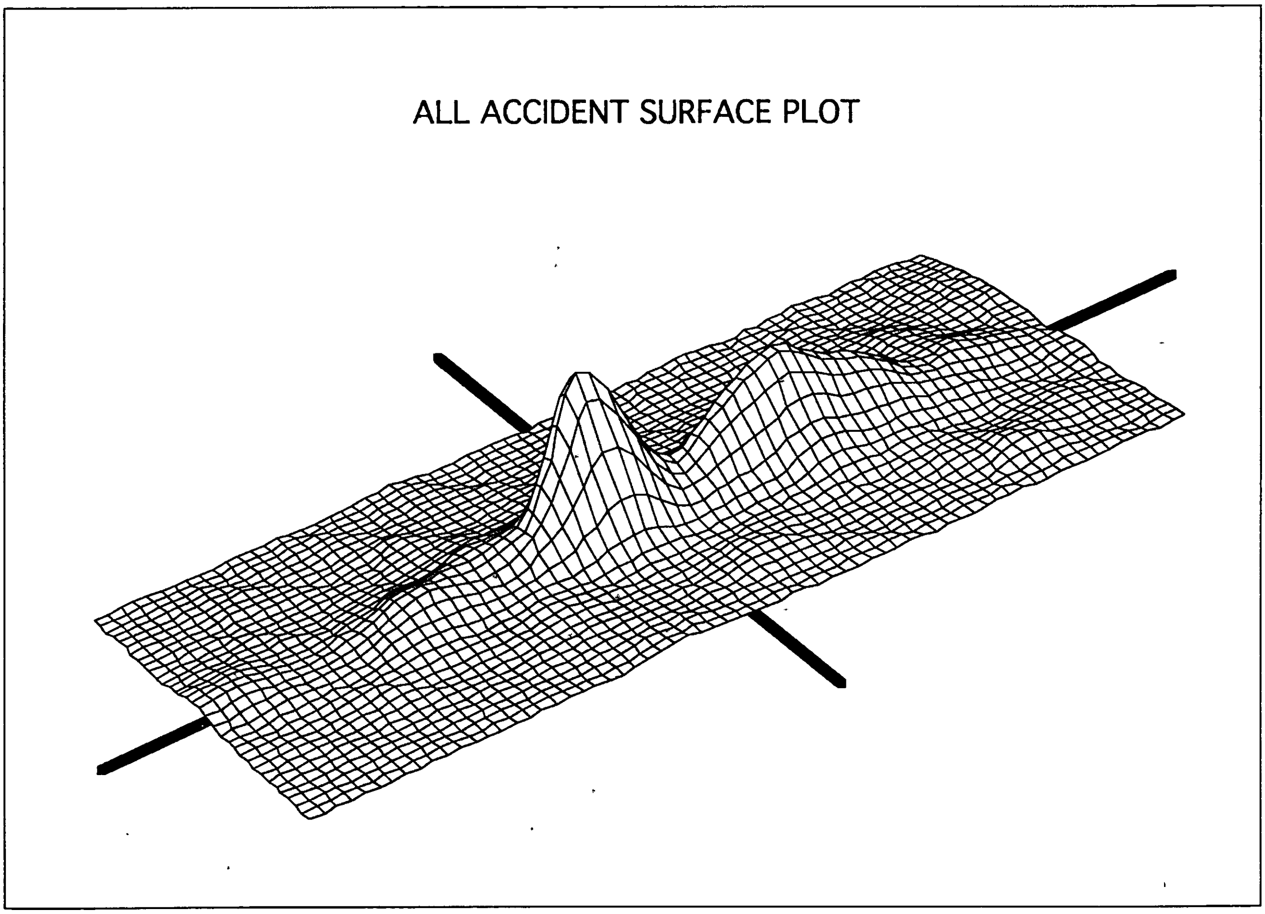
Exhibit 10

Contour Plots of Accident Frequency: 25 Percent Contours



Perhaps the best way to obtain an understanding for the spatial distribution of accidents is to view it in a 'three dimensional' format where the number of accidents in a particular grid area (in this case 300' by 300') are summed to achieve a feeling of greater height in areas where more accidents occur. This view also illustrates a continuous rather than the step function shown in the contour plot. The result is shown in Exhibit 11 where the runway approach threshold is at the intersection of the two axes, with the runway extending up and to the right.

Exhibit 11



Finally, Exhibits 12, 13, and 14 provide views of the accident locations by category, covering arrivals, departures, and accidents with no pilot control, respectively.

Exhibit 12

Scatter Plot of Arrival Accidents

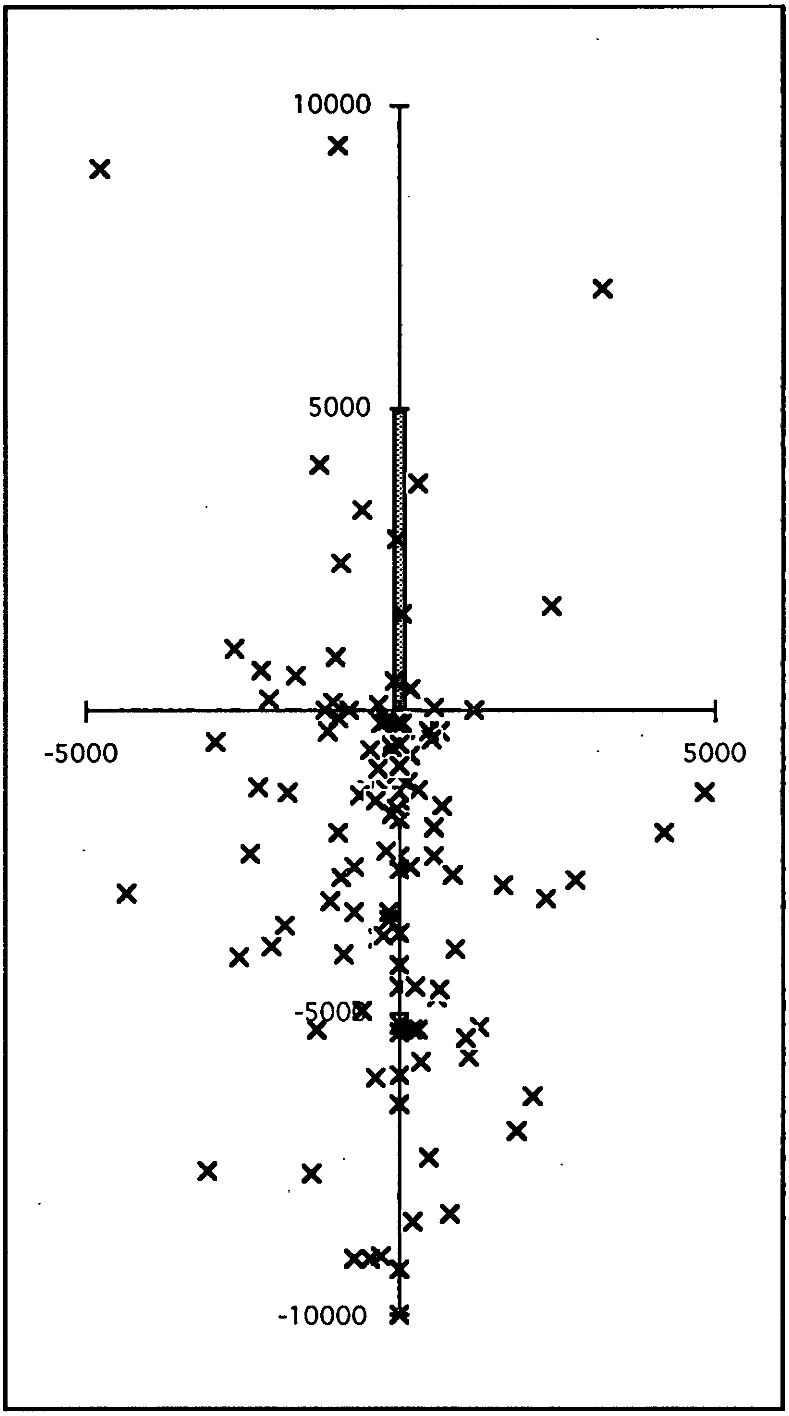


Exhibit 13

Scatter Plot of Departure Accidents

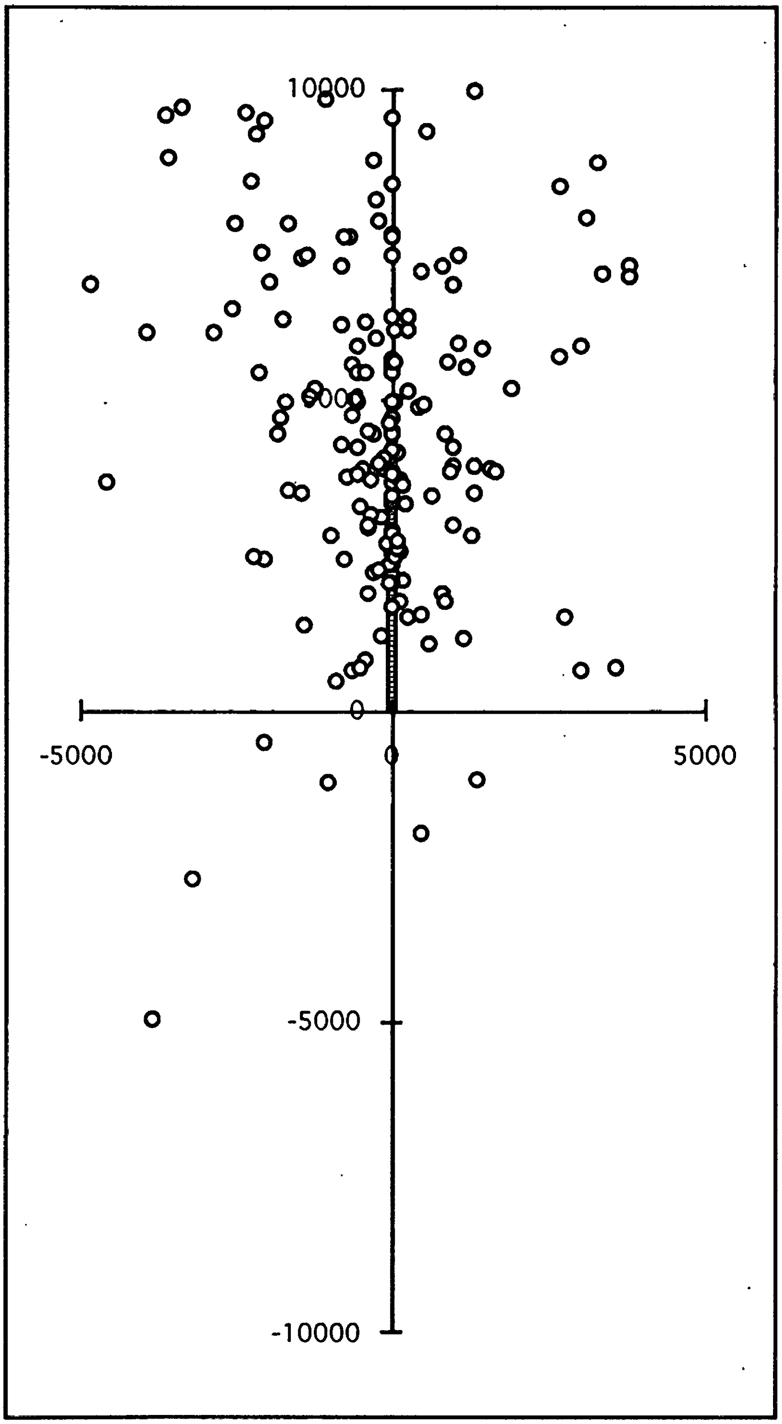
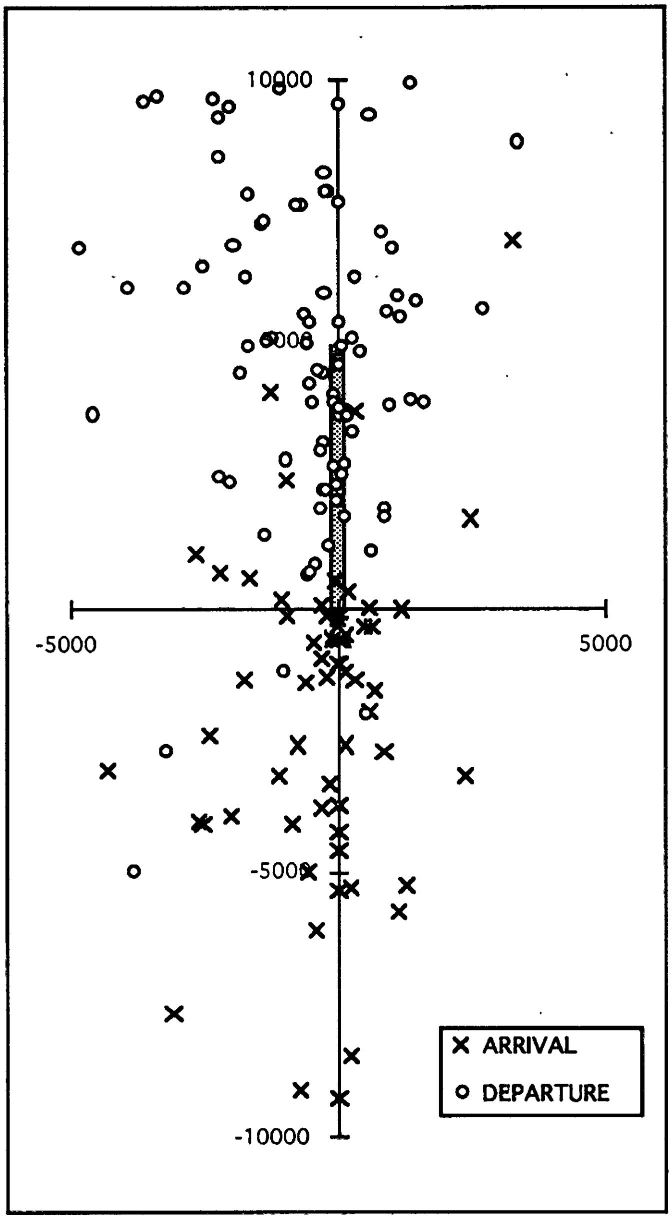


Exhibit 14

Scatter Plot of Accidents For Which There Was No Pilot Control and No Inflight Collision



5.0 Land Use Planning and Modeling Accident Location Probabilities

Current land use planning in the area surrounding airports is based on some notion of what the desirable level of safety and economic loss is for both aircraft occupants and people on the ground. Many might point to the record of few lives lost and paucity of damage to structures on the ground due to aircraft accidents as a sign of the value of prevailing planning guidelines. What is often forgotten, however, is that restricting land use or density for a particular parcel of land imposes economic costs directly on both those who own the land, because they have fewer opportunities available to them, and to the community in general, since those land uses which have been restricted in one area will place pressure on land in other areas. In planning the use of land it is important to consider both the costs and benefits of restricting land to one use rather than another. Under present land use guidelines, for example, it may be that the costs of saving a life or reducing property damage are far in excess of their value.[3]

A natural extension to the construction of this accident database is to develop a risk assessment model that can be used both in land use planning and in assessing the costs or benefits to safety of airport management and investment decisions. As we described in the introduction, a decision to use land in one way rather than another carries with it both costs and benefits. Restricting development may reduce the probability of damage to property and injuries to persons in the event of an accident but, at the same time, impose a cost since land may not be used in its most productive manner. As one moves out from the runway the probability of an accident occurring at a particular location decreases in some non-linear way. If a safety zone of a particular shape is to be prescribed, it is important that the incremental gain in safety be balanced against the incremental costs or forgone utility or productivity of land use. For example, if the safety zone were to be extended an additional 1,000 feet from the end of the runway, how many accidents would be included in that area? How does this relationship change as we move further away from the runway? At some point adding land to the safety zone will have a small or negligible impact on accident location probabilities and hence on improved safety.

A risk assessment model is a composite of two models: an accident location probability model, which provides the spatial probability distribution of an accident occurring; and a land use model, which translates the accident probabilities into dollars as measured by the cost of an accident occurring at a particular location. The costs are composed of four broad elements: loss of life or injury to occupants of the aircraft; damage to aircraft; loss of life or injury to people on the ground; and damage to structures on the ground.

The accident probability model is a joint conditional probability. This means that the probability of an accident must first be established and then, given that an accident occurs, what is the probability that it occurs at a given location in relation to the runway. The first probability is a function of the number of the aircraft movements, type and age

---

[3] For those who would argue that human life has an infinite value, simply look at legal decisions or automobile safety improvement decisions to see that finite values have been placed on saving a human life.

characteristics of the aircraft, type and age characteristics of the pilot, environmental conditions, and type of flight. The probability of a specific location for an accident, given an accident occurs, is based on the demonstrated accident rate over space. This rate may be established using the sample of accident data and the contours described in the previous section.

Restating the accident probability model as an economic risk model requires a translation into a dollar metric. An accident can result in bodily injury to both the occupants of the aircraft and people on the ground. Likewise, damage can occur to the aircraft and structures on the ground. As land is used in alternative ways it can increase or decrease the extent of damage. If, for example, all areas are left clear, there will be no damage to people or buildings on the ground. There is also a greater likelihood that, given some control, a pilot will be able to minimize damage to aircraft and occupants. The main costs, therefore, would be related to the lost opportunity of developing the land.

We have made the distinction between controlled and uncontrolled accidents. A controlled accident will have a different, and expectedly lower, cost than an uncontrolled one. In planning land use in the vicinity of an airport, it is important to keep this distinction in mind. If, for example, the majority of off airport landings are 'controlled' [4], it would call for optimizing land use to maximize the expected return to land including the cost of safety on the cost side of the ledger. However, if the majority of accidents are 'uncontrolled', land use planning is more difficult since the expected costs will have either a random distribution or an accident distribution as represented by our accident database. The point is that, in establishing the distribution of accidents spatially, a distinction should be made between controlled and uncontrolled accidents since they will have different location/spatial distributions.

In order to undertake a risk assessment, additional data on movements by airport as well as other characteristics described above are required. Nonetheless, the accident data which is described in this report can be useful in developing as first approximation of the shape of safety zones about runways and in evaluating the current FAA rules regarding safety zones. One need only examine the incremental gain in the inclusion of accidents within the safety zone as the size and shape changes and compare this against the incremental increase in the amount of land which will have a restricted use with its inclusion in the safety zone. If one wanted to translate this into a dollar measure, assumed values for loss of life or injury as well as for aircraft and ground structures can

---

[4] While this was not the case in our sample, with only 30.5 % of the accidents occuring with some pilot control, there may be enough successful off airport forced landings which result in incidents rather than accidents (by NTSB definition) to make this true.

be utilized. Similarly, an assumed value for land in alternative uses may also be used. The use of this data and approach will provide a sounder analytical basis for land use planning in areas adjacent to airport runways.

Appendix A

The Econometrics of Accident Distributions

In this Appendix we provide an overview of an econometric approach to the spatial and temporal distribution of accidents. In it we consider what variables are important in explaining the number of accidents over time and space. One issue to be considered is that accidents take on only nonnegative values and this will affect the specification of the econometric model. The rationale for this modeling effort is that current rules regarding land use planning around airports make the implicit assumption that these rules or guidelines should be applied ubiquitously. This implies that all airports are alike and that the incremental gain from the application of a rule or guideline is the same at every airport. Rather than have this as a maintained hypothesis, we argue that this proposition should be tested. If it is the case that airports fit into particular families based on certain airport and traffic characteristics, it may be the case that some guidelines are more effective with some families of airports than with others.

As a first step consider that for a given airport, the probability of an accident satisfies the following conditions: the probability of an accident is proportional to the geographic area being considered and the length of time; the instantaneous probability of an accident is constant over time and space; the probability of having more than one accident in a given time period is small and at a given point in space even smaller; and accident occurrences are independent. Letting v(f, f+df) be the number of accidents occurring at a given point in space over the interval from f to f+df, it is reasonable to presume that

$$p[v(f,f+df) = 0] = 1-\lambda df + \Theta\ (df)$$

and that

$$p[v(f, f+df) = 1] = \lambda df + \Theta\ (df)$$

therefore,

$$p[v, f+df) \geq 2\ ] = \Theta\ d(f) \text{ when } df \rightarrow \Theta$$

where λ is a parameter and $\Theta\ d(f)$ is for higher order terms assumed to be negligible.

The Poisson distribution satisfies the conditions set out above and will be used as a first choice to represent the distribution of accidents over space. The probability that an airport will be faced (involved) with v accidents over area f is:

$$p\left({}^{v}\!/_{f}, \lambda\right) = \frac{e^{-\lambda f} (\lambda f)^{v}}{v!}$$

where λ is both the mean and variance of the data. Using a Poisson distribution to represent the observed distribution of accidents for a group of airports assumes that all airports have the same probability of being involved in one or more accidents. One could argue that such an assumption is highly restrictive. As an alternative, assume that λ varies between airports and that for a given airport the distribution of accidents follows a Poisson distribution. It will, therefore, be [mathematically] convenient to assume that λ is distributed as f(λ),

$$f(\lambda) = \frac{\lambda^{a-1} \tau^{a} e^{-\tau}}{\Gamma(a)}, \tau > 0, a > 0$$

which is a Gamma density distribution with mean a/τ, variance a/τ^2, and $\Gamma(a)$ is the Gamma function. From the distribution of λ, the average probability of v accidents at an airport over area f becomes a negative binomial distribution:

$$P\left({}^{v}\!/_{f}, a, m\right) = \frac{\Gamma(f+a)}{v!\,\Gamma(a)} \left[\frac{a}{a+mf}\right]^{a} \left[\frac{mf}{a+mf}\right]^{v}$$

where the mean is m=a/t and the variance is $(a/\tau)\,(1+1/\tau)$. If the distribution of accidents is best represented by a Poisson distribution, it would suggest that airports are homogeneous insofar as their respective probabilities of accident are concerned. This implies, among other things, that differential policies with respect to land use or specific applications of regulatory policy based on past experience would have no empirical support.

The models, as set out thus far, do not permit the identification of those factors which would explain the distribution of accidents. For planning purposes it would seem suitable to determine which airport characteristics are significant. In addition, heterogeneity in the airport's probabilities of accident may make the basic Poisson model more attractive. Following Dionne et al. (1992)[5] it is possible to examine the Poisson and the negative binomial models with heterogeneous airports. In both cases, our maintained hypothesis is that the probability of an accident depends upon the airport and traffic characteristics: in the Poisson model, the assumption is that the distribution of

---

[5] Dionne, G., R. Gagne and C.Vanasse (1992), "A Statistical Analysis of Airline Accidents in Canada 1976-1987" Working Paper CRT-811. Center for Transportation research, University of Montreal

accidents about an airport is a one parameter distribution (hence, the mean and variance are equal), while in the case of the negative binomial, allowance is made for an overdispersion (the variance is larger than the mean) in the distribution by assuming that the airport characteristics leave some differences between the unexplained individual airport probabilities and that the unexplained portion follows a Gamma distribution. The practical problem is that the data may not be rich enough to pin the Poisson distribution with heterogeneous firms.

Let A_i be the number of accidents about an airport in a given time period. Assume for the moment that the assumptions of the one parameter Poisson model hold and furthermore that the function providing the linkage to the parameter λ_i to the characteristics is

$$\lambda_i = \exp(x_i\beta)$$

where β is a vector of weights to be estimated. Therefore, the probability that airport i will be faced with v accidents is

$$p(A_i = {}^{v}\!/_{x_i}) = \frac{e^{-\exp(x_i\beta)} \exp(x_i\beta)^{v}}{v!}$$

The mean and variance of A_i are both $\lambda_i=\exp(x_i\beta)$ with the form of exp(•) ensuring that λ_i is nonnegative. The likelihood function of the data (A_i,x_i) is given by

$$L(A;\beta) = \frac{\prod_{i=1}^{n} e^{-\exp(x_i\beta)} \exp(x_i\beta)^{A_i}}{A_i!}$$

In this model using the Poisson distribution there are a number of restrictions. First, the Poisson model is based on the assumption that successive accidents are independent. Therefore, the probability that an airport will experience more than one accident in a given location or within a given time period is considered independent of whether an accident occurred earlier in time or space. A second restrictive assumption is that for a given airport the conditional mean and variance of the distribution are equal. The consequence of this is that the variance may be underestimated and hence yield individual variable significance levels that are too high (the t-statistic). The Poisson model may also generate too small probabilities for the number of accidents to be two or more. The strong link between mean and variance should be broken to rectify some of these problems. Using a Gamma distribution for the individual λ_i, for example, would allow one to generate a two parameter distribution of accidents. Under the Poisson distribution, despite the λ_i being different from airport to airport, the conditional mean and conditional variance of each individual Poisson distribution are both equal to $\lambda_i=\exp(x_i\beta)$.

A method of introducing greater flexibility into the model is to consider that the vector of characteristics x_i is not sufficient to capture the whole difference between airports and to assume that the additional unobserved characteristics can be modeled as a random addition μ_i to x_i; that is,

$$\lambda_i = \exp(x_i\beta + \mu_i)$$

with a probability density function $h(\mu_i)$. The marginal probability that the ith airport will have v accidents becomes:

$$\int P[v / f = 1, x_i, \mu_i]\, h(\mu_i)\, d(\mu_i)$$

$$= \int \frac{e^{-(x_i\beta)+\mu_i} \exp((x_i\beta)+\mu_i)^v}{v!}\, h(\mu_i)\, d(\mu_i)$$

which as Dionne et al (1992) note, is the Compound Poisson distribution; changing the specification of $h(\mu_i)$ generates different cases.

A form to be considered, and use in other safety research, is the negative binomial distribution which can be obtained by writing $\lambda_i = \exp(x_i\beta + \mu_i) = \exp(x_i\beta)\varepsilon$ and assume $\varepsilon_i=\exp(\mu_i)$ follows a gamma density of

$$f(\varepsilon) = \frac{\varepsilon^{a-1}\, e^{-a\varepsilon}\, a^a}{\Gamma(a)}$$

with a mean equal to 1 and a variance of 1/a. To obtain the distribution of accidents, integrate the expression

$$P(A_i = v/x_i) = \frac{\Gamma(v+a)}{v!\Gamma(a)} \left(\frac{\exp(x_i\beta)}{a}\right)^v \left(1 + \frac{\exp(x_i\beta)}{a}\right)^{-(v+a)}$$

which is the negative binomial distribution with parameters a and $\exp(x_i\beta)$; the mean is $E(A_i) = \exp(x_i\beta)$ and the variance is $V(A_i) = \exp(x_i\beta)[1+\exp(x_i\beta/a] = E(\lambda_i)[1+E(\lambda_i)VAR(\varepsilon)]$. As a result ,the variance ends up being a strictly convex function of the mean with the variance/mean ratio being a linear function of the mean.

The rationale for this approach is that in accident analysis undertaken in other modes of transportation, the negative binomial model with a regression component yields a reasonable approximation of the true distribution of accidents. Different statistical tests rejected the Poisson distribution and the negative binomial model without individual characteristics. The practical importance of this issue for land use planning is whether a 'generic' model can be used for airports or whether each airport must be considered in a different class. Acceptance of the Poisson model would lead to acceptance of the generic model. Using a common modeling approach, as appears to be true currently, has implicitly assumed acceptance of the Poisson distribution.

Before this empirical investigation can be undertaken additional data will be required to supplement the current database. In particular, aircraft movement data by volume, type, and flight characteristics across airports is needed.

Appendix B

ACCIDENT DATABASE

The accident database fields, sources and, where necessary, explanations are as follows:

Accident date - NTSB Factual Report
NTSB file number - NTSB Factual Report
Airport name - NTSB Factual Report
Airport location (city and state) - NTSB Factual Report
Aircraft manufacturer - NTSB Factual Report
Aircraft model - NTSB Factual Report
Aircraft weight - NTSB Factual Report
Number of engines - NTSB Factual Report
Arrival/departure - NTSB Factual Report
Takeoff roll beginning point - assumed to be at the runway approach end unless other location given in Factual Report
Type of approach (visual/precision/non-precision) - NTSB Factual Report
Time of accident - NTSB Factual Report
Weather conditions - NTSB Factual Report
VFR/IFR - NTSB Factual Report
Weather a factor? -yes if weather listed as a causative factor by NTSB
Day/night - NTSB Factual Report
Duty runway - NTSB Factual Report
Runway type - NTSB Factual Report
Duty runway magnetic heading - runways are named according to their magnetic heading rounded to the nearest ten degrees. Since most points are plotted on USGS Quadrangle or city maps, bearing and distance are determined <u>relative</u> to the duty runway so exact runway heading is not required and the rounded heading is listed in the database. Where the exact heading is readily available, or if the determination of an accident location requires the extra time necessary to find the exact heading, the field 'Actual?' will so specify.
Actual? - yes/no
Duty runway length - NTSB Factual Report
Duty runway width - NTSB Factual Report

ILS available - Type of instrument approach available for accident runway. Source of information is Approach Plates

Airport landing pattern (left/right) - Airport Facility Directory. If the accident report shows accident aircraft was flying a pattern different than that which is in the directory, the pattern actually flown will be given.

Accident bearing from arrival/departure threshold - NTSB Factual Report

Accident relative bearing from arrival/departure threshold - calculated using RB = accident bearing - runway heading.
If RB < 0, add 360

Accident distance from arrival/departure threshold - NTSB Factual Report

Distance from centerline (X) - calculated using
$X = (\text{accident distance})*(\sin((\text{relative bearing}))$

Distance from landing threshold (Y) - calculated using
$Y = (\text{accident distance})*(\cos(\text{relative bearing}))$
if accident is departure, add runway length

Pilot control (some/none/unknown) - NTSB Factual Report

Swath length - NTSB Factual Report

Swath direction - NTSB Factual Report

Inflight collision with obstruction - NTSB Factual Report

Factor? - Yes/no. Did the inflight collision cause the accident or did it have a material effect on where the aircraft came down? NTSB Factual Report.

Injuries onboard aircraft - NTSB Factual Report

Injuries on ground - NTSB Factual Report

Damage to aircraft - NTSB Factual Report

Damage on ground - NTSB Factual Report

SAMPLE DATABASE PAGE

DATE	NTSB FILE #	AIRPORT NAME	AIRPORT LOCATION
7/1/83	FTW83FA310	ENID-WOODRING (WDG)	ENID ,OK

AIRCRAFT TYPE

A/C MFG BEECH

A/C MODEL A-36

A\C WEIGHT 3600

 # OF ENGINES 1 **REG#** N23684

FLIGHT INFORMATION

ARR\DEP ARR RTN TO A/P

T/O ROLL BEG POINT (Dist from Approachend) NA

TYPE OF APPROACH Visual/ILS (Precision/non-precision) VISUAL

TIME 1807

AIRPORT CONDITIONS

WEATHER CLEAR
10 MILES
180@18

VFR/IFR VFR

WEATHER A FACTOR ?

DAY/NIGHT DAY

RUNWAY INFORMATION

RUNWAY 30 **TYPE** ASPHALT

RWY HDG 300 **ACTUAL** NO

LENGTH 5511

WIDTH 100

ILS AVAIL NONE

PATTERN (LEFT/RIGHT) LEFT **FAA TWR** YES

ACCIDENT INFORMATION

BEARING AND DISTANCE FROM CENTER OF RUNWAY THRESHOLD

BEARING 120

RELATIVE BEARING 180

DISTANCE 1026

X Y COORDINATES

X COORDINATE 0

Y COORDINATE -1026

PILOT CONTROL (Some, None, No Information)	SWATH LENGTH (ft)	DIRECTION
NONE	70	300

INFLIGHT COLLISION WITH OBSTRUCTION (TYPE) NONE FACTOR? NO

INJURIES

Onboard Aircraft		On Ground	
Fatal	2	Fatal	0
Serious	0	Serious	0
Minor	0	Minor	0

DAMAGE

AIRCRAFT	ON GROUND
DESTROYED	NONE

NOTE A/C ESTABLISHED ON LEFT BASE LEG. CRASH OCCURRED AFTER TURN TO FINAL. NORMAL PATTERN FOR RWY 30 IS RIGHT