Stochastic Properties of Enroute Air Traffic—An Empirical Investigation

William J. Dunlay Jr.*
The University of Texas at Austin, Austin, Texas
and
Robert Horonjeff†
University of California, Berkeley, Calif.

This paper presents results of a recent research project into the nature of enroute air traffic. Characteristics of air traffic on a single airway and of traffic passing through an air traffic control (ATC) sector are considered. Flight-strip data for 26 days from the Bradford-High sector were used to investigate these characteristics. Airline flights operating according to a seven-day-per-week schedule comprised a large percentage of the air traffic data studied. It is shown that, despite this underlying daily schedule, air traffic on a single airway and flight level, and even the traffic passing through an ATC sector as a whole, during short-time intervals varies greatly from day to day and can, in fact, be approximated by a Poisson process. Empirical investigations are made of the extent of daily repetition of scheduled airline flights passing through an ATC sector in a given time period. In addition, the distribution of aircraft arrival times at an enroute fix and stochastic properties of successive aircraft separations in an air traffic stream are examined. It is argued that, for short time intervals, the epochs at which aircraft pass a point on an airway are, approximately, the points of a Poisson process. Theoretical and experimental evidence for this argument is presented. Findings of this research are applicable to airspace capacity analysis and ATC system design.

Introduction

THE findings presented in this paper stem from FAA sponsored research on estimating air traffic control (ATC) conflicts.[1,2] It is felt that many of the air traffic properties discovered during that research may be of interest to others doing research in the area of ATC, or airport and airspace capacity. In addition, designers of the national airspace system should find these results of use in their work.

Data

Flight-progress strips for 26 days, March 1 to March 10 and March 15 to March 30, 1973, from the Bradford-High sector of the Chicago Air Route Traffic Control Center (ARTCC) were the source of data for investigating enroute air traffic characteristics. A sketch of the Bradford-High sector is shown in Fig. 1.[2] The times at which a particular flight was expected to pass various enroute fixes along its route were taken directly from its filed flight plan and recorded to the nearest minute on flight-progress strips. When a conflict between aircraft occurs, the usual corrective action prescribed by the controller is a radar vector. The time penalty to a vectored aircraft is almost certainly less than 1 min. Therefore, the fix times shown on flight strips are assumed to be unaffected by control instructions. In addition, since aircraft on airways are relatively free to pass each other (with the air of radar vectors from controllers), flight strip data also constitute a reasonable representation of aircraft flow rates on airways. Flight-strip fix times and actual observed fix times have been shown to compare favorably.[1]

Received May 5, 1975; revision received Sept. 8, 1975. The material in this paper was part of a doctoral dissertation prepared by W.J. Dunlay Jr., at the University of California, Berkeley.[1] This research was supported by the U.S. Federal Aviation Administration under Contract No. DOT-FA-72-WA-2827.

Index categories: Air Transportation Systems; Navigation, Control, and Guidance Theory.

*Assistant Professor, Department of Civil Engineering.

†Professor and Research Engineer, Institute of Transportation and Traffic Engineering. Member AIAA.

Daily Repetition of Flights

From an analysis of traffic passing through the Bradford-High sector it was found that, on the average, 93% of aircraft were on scheduled airline flights listed in the Official Airline Guide, practically all of which were on a seven-day-per-week schedule. The remaining flights were military. Only a few of the high-altitude flights were private aircraft.

Because most high-altitude flights are air carriers operating according to a seven-day-per-week schedule, it was not unreasonable to suspect that approximately the same set of flights passes a point on an airway or through a sector during a specified time period every day. However, it is common for some aircraft to be routed over a different set of airways on a particular day due to varying wind conditions; in addition, a flight may be so late that it falls outside a particular time period, or there may be a substantial number of nonscheduled aircraft.

In order to investigate this question of daily repetitiveness, one can collect data on the flights passing through a sector in a particular time period of, say N days. One way to measure daily repetitiveness from such data is to sort the individual flight numbers into groups according to the number of days ($n = 1, 2, \ldots, N$) that they appeared in the sector. The percent of flights that occurred on n days can be computed for all values of "n." These percentages can be represented graphically by a histogram, as shown in Fig. 2.

Histograms constructed from the Bradford-High flight strip data are shown in Fig. 3 for two different lengths of time interval. Note that there is very little difference in repetitiveness between the 1-hr and the 2½-hr intervals. In both cases, more than 60% of flights occurred on only one of the 26 days. No flights occurred on all 26 days. As one would expect, there is slightly more repetition when one considers a 2½ hr interval than for a 1-hr interval.

In an attempt to explain the surprising lack of repetition in the daily makeup of the sector's traffic, two categories of airline flights were distinguished: level flights and flights climbing or descending from/to a nearly airport (mainly O'Hare or Midway). Flights in the first category are called "overs," and those in the second, "locals." The repetitiveness of these

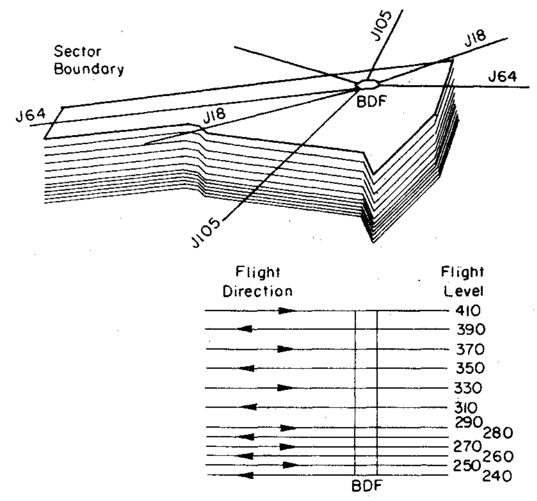

Fig. 1 Typical high-altitude, enroute air traffic control sector.

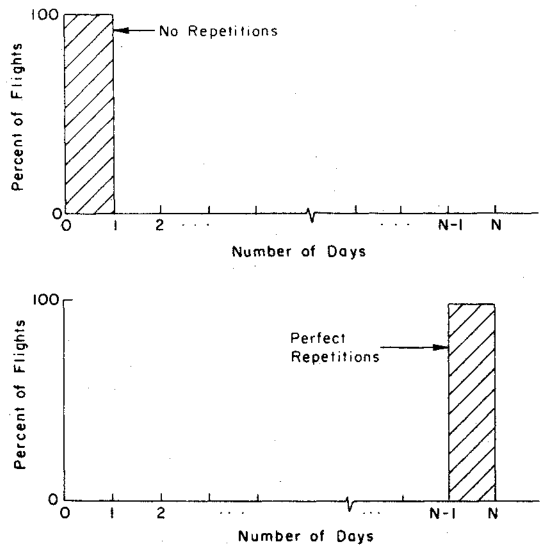

Fig. 2 Histogram for the number of days flights pass through a sector.

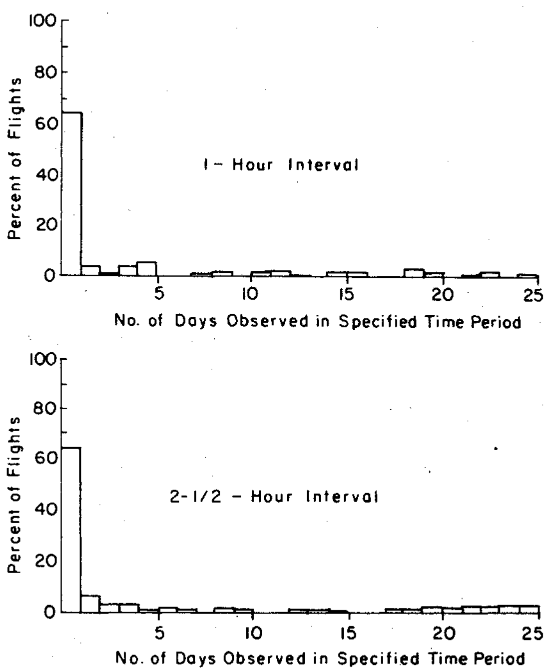

Fig. 3 Repetitiveness for different lengths of time intervals.

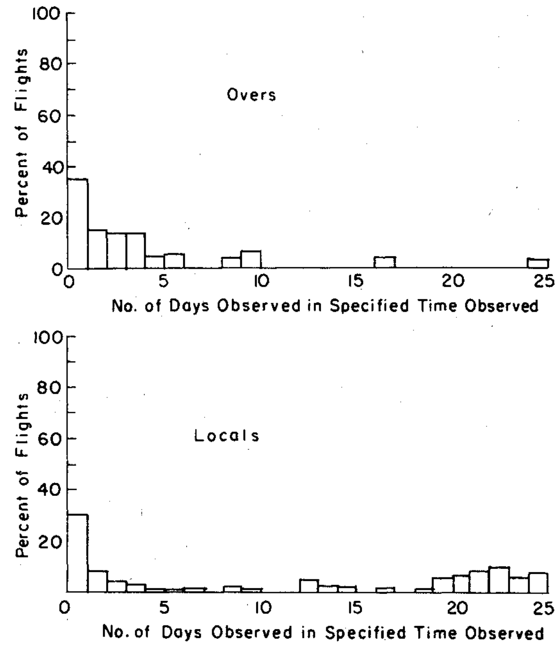

Fig. 4 Repetitiveness of different types of flight.

two types of airline flights in a 2½ hr period on 26 days is represented by Fig. 4. It is clear from this figure that locals are far more repetitious than overs. However, there is surprisingly little repetition even for locals, when one considers that more than half of the local flights passed through this particular sector on fewer than 13 of the 26 days. Air traffic on a single airway is even less repetitive than the traffic passing through a sector as a whole.

Successive Aircraft Spacings

Scatter diagrams of successive aircraft spacings, i.e., spacing i vs spacing $i+1$ for all values of i, are shown in Figs. 5a and 5b for airways J18 and SID1 (ORD), respectively. These data are shown arranged in two-way contingency tables in Table 1, and results of Karl Pearson chi-square tests of

Table 1 Two-day contingency tables and Karl Pearson chi-square tests of independence of successive separations

Successive separations on airway J18, FL330, 26 days
($i+1$)st separation, naut miles

		<75	75-175	>175	total
ith separation	<75	6	7	6	19
	75-175	6	9	5	20
	>175	9	4	8	21
	total	21	20	19	60

$x^2 = 3.29, f=4, P_l = 0.51$; Do not reject independence hypothesis

Successive separations on airway SID1, FL240, 26 days
($i+1$)st separation, naut miles

		<30	30-90	90-180	>180	total
ith separation	<30	12	9	9	6	36
	30-90	12	15	15	11	53
	90-180	12	17	11	8	48
	>180	10	11	15	9	45
	total	46	52	50	34	182

$x^2 = 3.29, f=9, P_l = 0.92$. Do not reject independence hypothesis

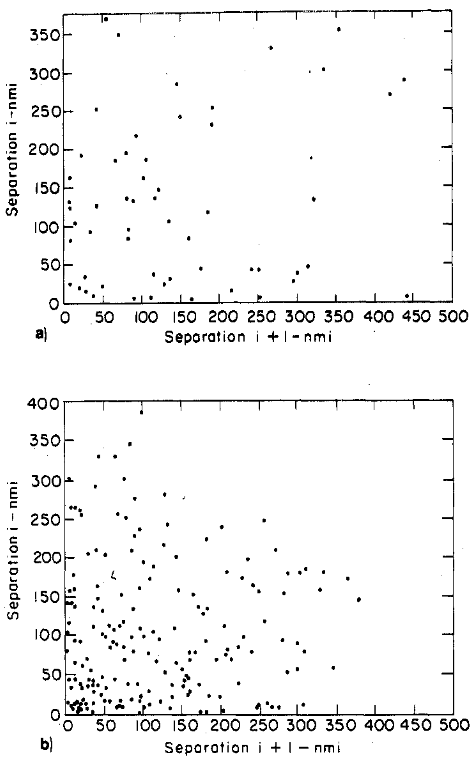

Fig. 5 Scatter diagrams for successive aircraft separations: a) successive separations on airway J18, FL330; b) successive separation on airway SID1, FL240.

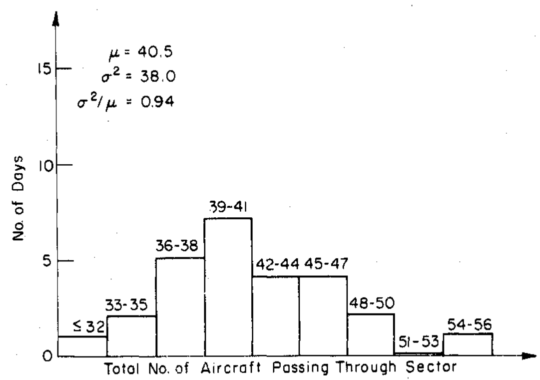

Fig. 6 Distribution of the total no. of aircraft passing through the Bradford-High sector in a 2½-hr period.

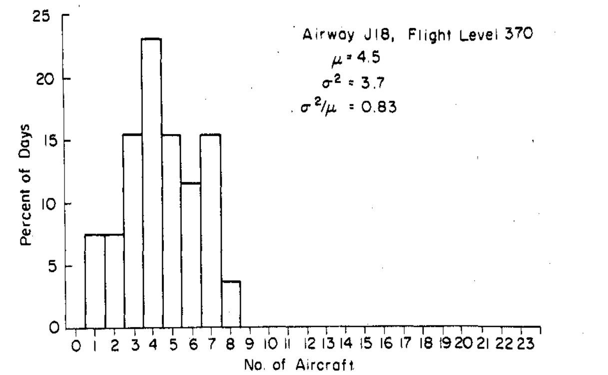

Fig. 7 Distribution of the number of aircraft on a single airway in a 2½ hr period.

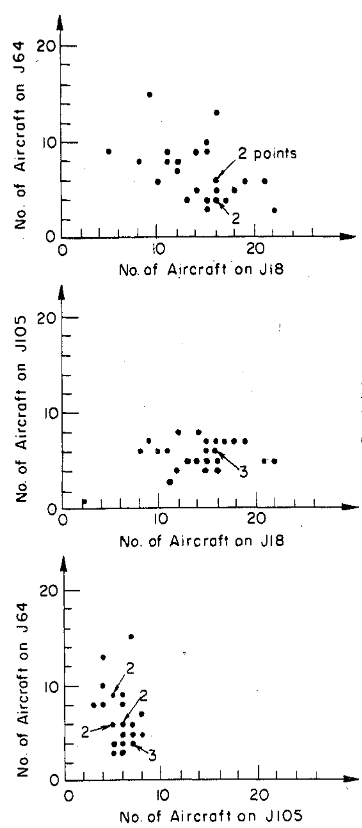

Fig. 8 Scatter diagrams for the no. of aircraft on different airways—all flight levels.

stochastic independence are also shown in that table. From these data, there is no reason to believe that successive aircraft spacings are not independent.

Aircraft Flow Rates

This section is concerned with two related topics: 1) day-to-day variations in aircraft flow rates during a specified time interval, and 2) possible dependence among the flow rates on different routes.

Day-to-Day Variation

From Bradford-High flight-strip data for the same 2½ hr period on 26 days, the number of aircraft passing selected points in the airspace was analyzed. The results of this analysis are displayed as histograms in Figs. 6 and 7 for the sector as a whole (including all airways and flight levels) and a single airway and flight level, respectively. It is interesting to note that there is considerable day-to-day variation in the number of aircraft passing a particular point on an airway, and even through the sector as a whole. Two reasons why one might expect such variation can be proposed: 1) the lack of day-to-day repetition in the flights passing through a sector, and 2) the influence of the daily flight planning process in which the flight level and route of a particular flight depend on such factors as wind forecasts, and payload, and the particular aircraft being used.

Dependence Among Different Flow Rates

Because flights are planned to make the best of prevailing winds and weather conditions, and because there is a finite population of aircraft which would ever pass through a particular sector under any circumstance, it is not unreasonable to suspect that the number of aircraft on one airway or flight level may depend inversely on the number of aircraft on other airways or flight levels. In order to investigate the extent of such dependence, three airways in the Bradford-High sector, which seemed most likely to serve as alternative routes for one another (based on a study of day-to-day switching of routes by flights), were selected for analysis. Scatter diagrams for 26 paired observations of the number of aircraft during a particular time period on one airway vs the number on a second

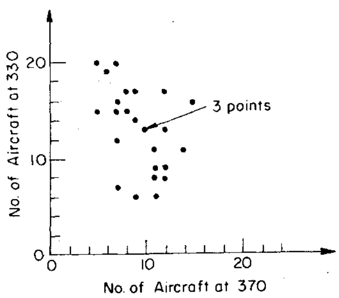

Fig. 9 Scatter diagrams for the no. of aircraft at different flight levels—all airways.

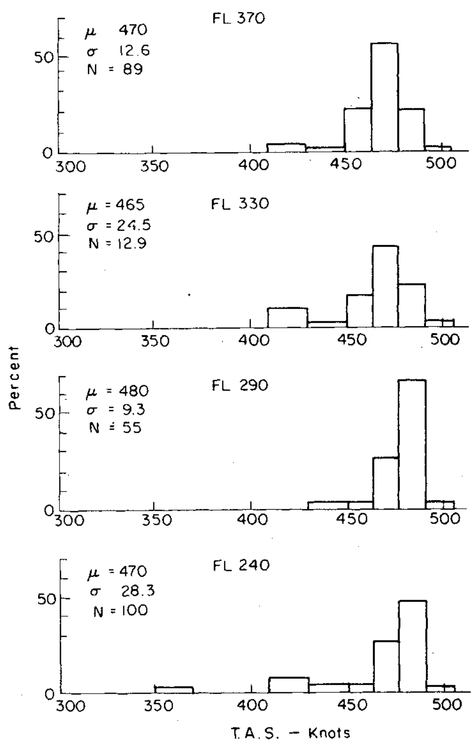

Fig. 10 Histograms for true airspeed at various flight levels.

airway are shown in Fig. 8 for the three possible pair-wise combinations of airways J64, J18, and J105. Figure 9 is a scatter diagram for the number of aircraft at flight level 330 vs the number at 370. Based on an evaluation of the scatter points in Figs. 8 and 9, it is concluded that the dependence among the numbers of aircraft at different points in the Bradford-High airspace is not very strong. However, these graphs do suggest that this type of possible dependence should be investigated in future research.

Distribution of True Air Speeds

Histograms of aircraft velocities for individual flight levels are shown in Fig. 10. These graphs were constructed using true air speed (TAS) data from 26 days of Bradford-High flight strips. The histograms of Fig. 10 were used to define the discrete speed-class probability distribution for different flight levels shown in Table 2. These speed distributions have application in the design and analysis of the ATC system. From Fig. 10 and Table 2 it is seen that the means and standard deviations of true air speeds differ only slightly among different flight levels.

Distribution of Individual Flight Arrival Times

In this section, the nature of the day-to-day variation in the clock times at which individual flights arrive at an enroute fix is investigated. As a notational convenience, such variation, will be referred to as a "Lateness distribution." From 26 days of Bradford-High flight strips for a 2½-hr period, a set of flights which passed through the sector on most of those days was selected. The lateness distributions for these flights have been plotted in the form of histograms in Fig. 11. Also shown on this figure are the sample mean, standard deviation, and sample size for the flight arrival times. As is typical of most scheduled events, these lateness distributions generally are skewed to the right. This is a consequence of flights that leave their origins quite late relative to schedule departure times.

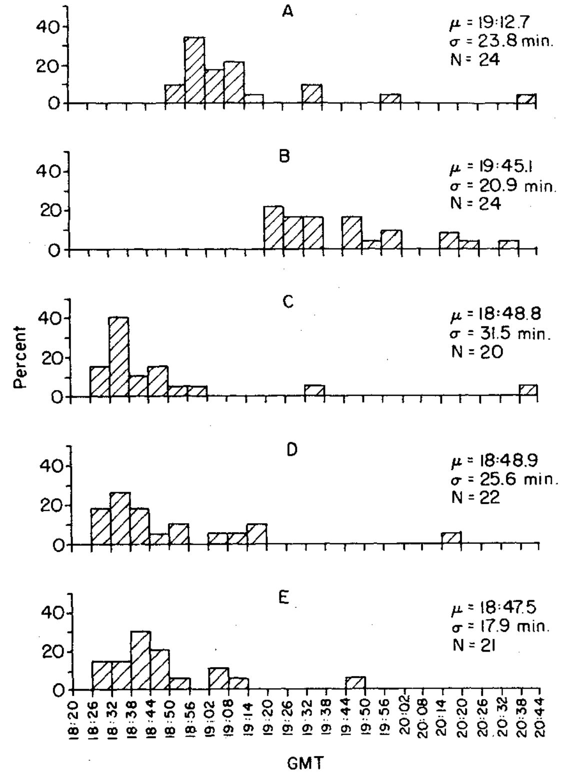

Fig. 11 Lateness distributions for individual scheduled flights.

Poisson Hypothesis

Air traffic can be represented by a stochastic point process, as illustrated in Fig. 12. The points plotted along the time axis are the epochs at which aircraft pass a particular location on an airway. It is important to realize that any graphical depiction of a stochastic point process such as Fig. 12 represents only one realization of the process which, in this research, means the process observed on a single day. The main stochastic properties of the point process, e.g., the distribution of the number of events in an interval or of the elapsed times between events, are elicited by repeated observation of the point process over a large number of days. Thus, it is assumed that different days represent independent and identical conditions for observing the stochastic process.

Table 2 Probability distributions for true air speeds

	Entry flight level							
	370		330		290		240	
i	$v_i^{(a)}$	$Pr\{V=v_i\}$	v_i	$Pr\{V=v_i\}$	v_i	$Pr\{V=v_i\}$	v_i	$Pr\{V=v_i\}$
1	420	0.023	420	0.116	420	0.000	420	0.111
2	440	0.011	440	0.023	440	0.018	440	0.056
3	457	0.202	457	0.171	457	0.018	457	0.056
4	471	0.550	471	0.426	471	0.273	471	0.078
5	458	0.202	485	0.232	485	0.673	485	0.481
6	499	0.011	499	0.031	499	0.018	499	0.019

[a] Speeds expressed in knots.

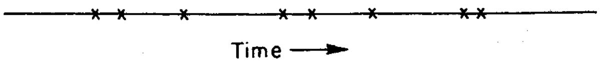

Fig. 12 Stochastic point process representation of air traffic.

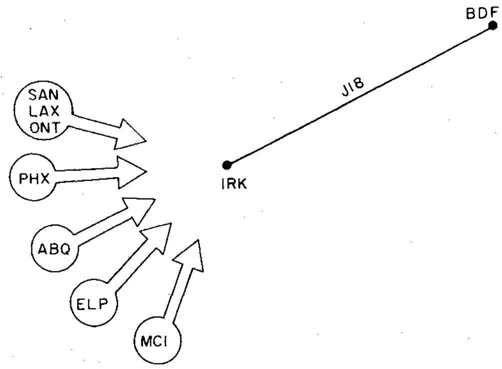

Fig. 13 Possible origins for the traffic on a segment of high-altitude jet route J18.

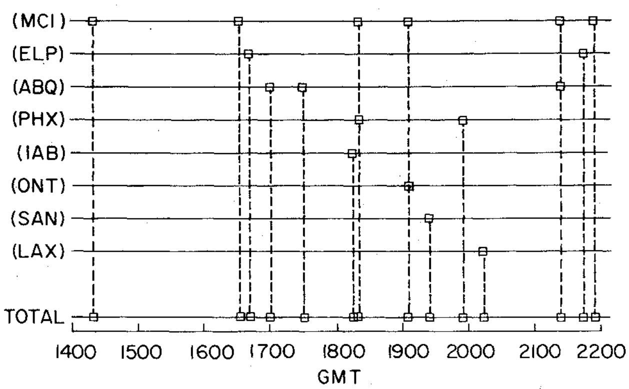

Fig. 14 Superposition of point processes for individual independent origins—one realization. Airway: J18; entry FL = 330; date: 3-2-73; sector: BDF-high.

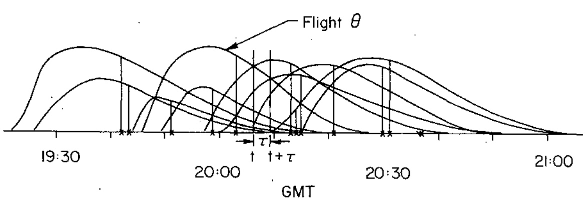

Fig. 15 Lateness distribution of flights on airway J18, flight level 330.

Most enroute airway segments handle flights from many different origins. For example, the segment of airway J18 shown in Fig. 13 handled flights from the seven different origins indicated by the arrows during an 8-hr period on a particular day. On other days, flights from still more origins not shown in Fig. 13 may show up on J18. That is to say, the total population of origins that might contribute to the traffic on a particular airway segment is almost certainly larger than the number of different origins observed on any one day.

The times at which aircraft from each potential origin pass a point on the airway may themselves be represented by a stochastic point process (called a component process), as shown in Fig. 14. In addition, the total or "pooled" process formed by superposing the component processes onto a single time axis is shown at the bottom of the figure. Note that the process one actually observes is the pooled process.

It seems reasonable to assume that the component processes are mutually stochastically independent, since they are generated by conditions at different airports, which are remote from each other. Under this assumption, the local properties of the pooled process are very nearly those of a Poisson process.[4] There are various limit theorems in the theory of stochastic point processes which state that the pooled process tends to the Poisson process as the number of component processes become large.[5]

Not all enroute airways handle flights from many different origins. For example, one airway studied in this research handled aircraft coming from O'Hare International Airport, exclusively. The aforementioned argument for a Poisson process obviously does not apply to such an airway. A second argument for a Poisson process given in the following can be based on the lateness distributions investigated in Fig. 11. Along the time axis in Fig. 15 are shown lateness distributions for individual flights on a single airway (constructed by drawing smooth curves approximately through the tops of individual histograms similar to those in Fig. 11) superimposed on the same time axis. Smooth curves were drawn because there is no reason to assume that the idealized lateness distributions are discontinuous. The areas under the lateness curves of Fig. 15 differ because the histograms from which they were constructed were plotted in such a way that the enclosed area was equal to the percent of days a flight used this particular airway and flight level.

The probability that a particular flight falls in a short time interval, say $(t, t+\tau]$, is equal to the area under that flight's lateness distribution between t and $t+\tau]$, $k = 1,...,n$, where n is the total number of lateness curves that overlap in the interval. Let N_k be a random variable, which takes on the value 1 if flight k is in $(t, t+\tau]$, and zero otherwise; i.e., $Pr\{N_k = 1\} = P_k$ and $Pr\{N_k = 0\} = 1 - P_k$, $\forall k$. Then the total number of flights that fall in the interval is

$$N(t, t+\tau) = \sum_{k=1}^{n} N_k$$

which is a random variable with expectation

$$E\{N(t, t+\tau)\} = \sum_{k=1}^{n} P_k$$

It is clear from Fig. 15 that the P_k are all small as compared with one. It is assumed that the N_k are stochastically independent, i.e., the probability that a particular flight falls in $(t, t+\tau]$ is independent of the probability that any other flight falls in that interval. This is almost certainly the case for short intervals.

For reasonably large n, N_k independent and $P_k << 1$, $k = 1,...,n$, is it a standard result of probability theory that the sum $N(t, t+\tau)$ can be approximately satisfactorily by a Poisson distribution with parameter

$$\lambda = \sum_{k=1}^{n} P_k$$

In particular, $N(t, t+\tau)$ has in the limit as $n \to \infty$ and the largest $P_k \to 0$ (such that the sum λ remains constant) a Poisson distribution (e.g., Ref. 6).

Because the precise value of τ has not been specified, the previous argument also holds for any component of $(t, t+\tau]$. Also, the numbers of events in nonoverlapping components of this interval are stochastically independent and approximately Poisson distributed. Therefore, the events in the interval are approximately those of a Poisson process. Note that the preceding argument for a Poisson process applies best to busy airways, where the number of flights whose lateness distributions overlap in a short time interval is large.

Test of the Poisson Hypothesis

In order to test the hypothesis that air traffic can be represented by a stochastic point process, which locally has

Table 3 Significance probabilities for chi-square goodness-of-fit tests

	Length of time interval, min					
Route/FL[a]	3	6	15	30	75	150
J64/370	0.65	0.07	0.03	0.15	0.17	0.05
J64/330	0.52	0.39	0.70	0.19	0.17	0.22
J18/370[b]	0.68	0.53	0.25	0.25	0.18	0.97
J18/330[b]	0.97	0.70	0.55	0.21	0.15	0.20
J18/290	0.17	0.05	0.64	0.06	0.01	0.50
J105/370	...	...	...	...	0.37	0.54
J105/330	0.15	0.50	0.32	0.31	0.35	0.04
J105/290	...	...	...	...	0.61	0.79
DR1/330	...	...	...	...	0.21	0.94
S1D1/240[b]	0.87	0.39	0.20	0.13	0.23	0.06

[a]Flight level $= FL$. [b]Three busiest airways—account for 75% of all aircraft.

the properties of a Poisson process, it is sufficient to test whether the distribution of the number of points that fall in short time intervals, say of length τ, is Poisson. Consider a relatively long time interval Υ over which the point process on a particular airway is assumed to be stationary. Divide that time interval into $n_I = \Upsilon/\tau$ smaller intervals, each of length τ. Let $X_i^{(j)}$ be the number of points that fall in the ith interval on day j, and assume that data from J different days have been pooled together. The $X_i^{(j)}$ are assumed to represent $n_I J$ independent observations of the same random variable X, the number of points in a time interval of length τ on a particular airway.

The hypothesis to be tested is that, for small τ, X has a Poisson distribution with parameter

$$\hat{\lambda}_\tau = \frac{\sum_{j=1}^{J} \sum_{i=1}^{n_I} X_i^{(j)}}{n_I J}$$

The point estimate $\hat{\lambda}_\tau$ is known to be the unique, minimum-variance unbiased estimator of the parameter in a Poisson distribution. Chi-Square goodness-of-fit tests were used to test this hypothesis (with $J = 26$ days, $\Upsilon = 150$ min, and various values of τ) for as many individual airways as the data would permit, subject to the constraint that the following conditions of a reasonable test be satisfied: 1) expected cell frequencies of at least 3.0, and 2) at least one degree of freedom, for which this test means at least three cells.‡ Details of these tests are shown in Ref. 1. The test results in terms of significance probabilities are summarized in Table 3. The corresponding values of chi square and number of degrees of freedom are summarized in the Appendix.

Recall that significance probability, usually denoted by P_I, is the probability of obtaining a value of chi square as large or larger than the one calculated in the test, given that the hypothesis tested is true. In other words, P_I is the approximate probability that one would make a type I error by rejecting the hypothesis tested if it were true. The significance probabilities in Table 3 can be viewed as rough relative measures of how well the data fit a Poisson distribution. The tests that one would reject at, say, a 10% signfiicance level are indicated by underlines in Table 3. Note that, even for time intervals as long as 150 min, the hypotheses that the number of events in those time intervals are Poisson distributed random variables are not often rejected. It is worth noting that, for most of the airways, the Poisson distribution seems to fit better (as measured by higher significance probabilities) for the shorter time intervals.

By definition of significance probability, one would reject a true hypothesis 10% of the time at the 10% significance level. Since Table 3 contains 48 tests of the Poisson hypothesis, one would expect to reject about five of these tests at the 10% significance level if the Poisson hypothesis were correct. The actual number of rejections at this significance level is eight, as indicated by the underlines in the table. However, for the three busiest airways, only one out of 18 (5.6%) tests was rejected.

Based on the previous test results, the following assumptions concerning the nature of the stochastic point process for airway traffic are made: 1) over very short time intervals, the process has the properties of a Poisson process; and 2) the number of events over longer time intervals can be approximated satisfactorily by a Poisson distribution.

Conclusions

The air traffic properties described in this paper should prove useful to those concerned with the analysis of aircraft flow rates and air traffic control. The following are the main conclusions of this paper: 1) the characteristics of enroute air traffic passing through a small volume of airspace are quite variable from day to day, in spite of the underlying airline schedule; this property should be considered in future airspace and air traffic research; 2) hypothesis tests indicate that the stochastic point process for airway traffic is approximately a Poisson process over short time periods; and 3) tests also indicate that the number of aircraft on an airway or passing through an ATC sector in relatively long time periods is approximately Poisson distributed.

References

[1]Dunlay, W.J., Horonjeff, R., and Kanafani, A., *Models for Estimating the Number of Conflicts Perceived by Air Traffic Controllers,* The Institute of Transportation and Traffic Engineering, University of California, Berkeley, Special Report for FAA, Contract DOT FA-72-WA-2827.

[2]Dunlay, W.J. and Horonjeff, R., "Applications of Human Factors Data to Estimating Air Traffic Control Conflicts," *Transportation Research Journal,* Vol. 8, 1974, pp. 205-217.

[3]Ratner, R.S., Williams, J.O., Glaser, M.B., Stuntz, S.E., and Couluris, G.J., *The Air Traffic Controller's Contribution to ATC System Capacity in Manual and Automated Environments – Volumes I and II,* Stanford Research Institute for U.S. Federal Aviation Administration, Rept. FAA-RD-72-63, 1972.

[4]Cox, D.R. and Lewis, P.A.W., *Statistical Analysis of a Series of Events,* Methuen, London, 1966, pp. 210-213.

[5]Khintchine, A.J., *Mathematical Methods in the Theory of Queueing,* Griffin, London, 1960, pp. 49-56.

[6]Feller, W., *An Introduction to Probability Theory and Its Applications,* Vol. I, 3rd ed., Wiley, New York, 1968, p. 282.

[7]Brownlee, K.A., *Statistical Theory and Methodology,* 2nd ed., Wiley, New York, 1965, p. 217.

‡The precise value of this minimum expected cell frequency has been the subject of debate – see, for example, Ref. 7.