Forecasting Air Passengers in a Multiairport Region

David Rubin, Maryland Department of Transportation
Larry N. Fagan,* Golightly International, Inc.

Traditional air passenger forecasting has been done for single airport regions generally by estimating a region's share of a national forecast. This method is inadequate in a multiairport region where forecasts must be distributed geographically. In the Washington-Baltimore region, a method similar to that used in ground transportation planning was applied. A share of the market forecast was developed based on a national forecast by using real yield, per capita disposable income, and government purchases as independent variables. Another forecast was completed based on an on-board origin-destination survey and regional socioeconomic data, including population, government employment, nongovernment employment, and per capita income. The 2 forecasts were then adjusted to provide air passenger trip generation by aviation analysis zone. This forecast was then distributed to other U.S. cities based on an analysis of socioeconomic data for Office of Business Economics analysis areas, including manufacturing product and wholesale and retail trade product. The result was forecasts from aviation analysis zone to other U.S. city for use in distribution.

Forecasts of aviation activity are usually done for individual airports. It is unusual for airport service areas to overlap and even more unusual for an overlap to be considered in the forecasting process. This is especially true for major air carrier airports. The forecasting tradition in aviation planning is to rely on national forecasts and local market shares (1). This process is inadequate in a region with competitive air carrier airports. A process similar to ground transportation planning more adequately produces usable forecasts in a multiairport environment.

Forecasts in a multiairport environment must be geographically distributed. The generated air trips have to be split between the airports, and distance from the airport is a major factor in determining airport choice. The other major factor, airline schedules, must also be considered in developing air passenger distributions (2, 3, 4). Even in regions where all airports are operated by the same agency, as a market phenomenon, they compete. The airports in the Baltimore-Washington region are operated by different agencies. The Maryland Department of Transportation operates Baltimore-Washington International (BWI), and the Federal Aviation Administration operates Washington National (DCA) and Dulles International (IAD).

DCA, because it is convenient to downtown Washington, attracts as many planes as can use it. Quotas have been placed on hourly operations, and limits have been placed on types and destinations of aircraft. DCA serves as a connecting hub; about 1.5 million connecting passengers use the airport annually. It also attracts passengers from the entire Baltimore-Washington region, including 4.5 million people who, when asked, would prefer to be at BWI or IAD. DCA has no long-haul flights to such markets as Kansas City, Dallas, Houston, Denver, Las Vegas, Los Angeles, San Francisco, and Seattle. Long-haul traffic is split between BWI and IAD. In most markets, IAD attracts a larger percentage of the long-haul traffic. Frequency and quality of service play a great role in the individual's airport choice because most passengers are not greatly inconvenienced by either airport.

The Maryland Department of Transportation, after acquiring BWI, changed the name of the airport from Friendship International to emphasize that it serves the Baltimore-Washington region. An aggressive promotion of the facility has been started, including advertisement, access service improvements, marketing to the airlines, and development of an aviation system plan for the state. For the first time, the market is being considered in its entirety (5).

To forecast air passenger trips in the Baltimore-Washington region, a 2-pronged approach was used. Macroforecasts were developed based on national forecasts and historic market shares and were distributed to destination cities. At the same time, microforecasts were used in which forecasts were developed for 72 aviation analysis zones based on the relationship between trip generation and socioeconomic factors in the region.

A national macroforecast was developed for the Maryland Aviation System Plan as one tool to be used in reviewing Baltimore-Washington regional participation for its existing and projected share of the 50-state U.S. domestic passenger historical series and forecast. Spe-

Publication of this paper sponsored by Special Committee on Air Transport Activities of the Transportation Research Board.

*Mr. Fagan was with R. Dixon Speas Associates when this research was performed.

cifically, this forecast consists of data reflecting the projected level of revenue passenger-kilometers, passenger originations, and passenger enplanements on domestic trunk and regional carriers. The basic model used for development of this forecast is an equation developed through a Hildreth-Lu multivariate technique.

GENERAL APPROACH FOR MODEL DEVELOPMENT

The least squares fit to a forecast equation with 3 independent variables will be of the form

$$Y(t) = A + BX_1(t) + CX_2(t) + DX_3(t) + e(t) \quad (1)$$

where

A, B, C, D = coefficients to be determined and
e(t) = model estimation term (which, one would hope, is small).

It may turn out that

$$e(t) = f[e(t-1)] \quad (2)$$

which means that the size of the error term is a function of the error at the previous period of time. This is called autocorrelation and is not desirable. A consequence of autocorrelation if the least squares estimation technique is applied is that the various statistical measures (t-test and F-test) conventionally used to indicate the goodness of fit of a model are no longer valid.

An approach to the problem of autocorrelation is the Hildreth-Lu technique. It assumes that the functional relationship between the 2 error terms is of the form

$$e(t) = Ke(t-1) + g(t) \quad (3)$$

where

K = a constant smaller than 1 and
g(t) = estimation error term.

This technique also provides a remedy in the case of first order autocorrelation. The Hildreth-Lu solution makes use of lagged variables, and the equation contains not only the expected independent variables $X_1(t)$, $X_2(t)$, and $X_3(t)$ but also the variables $X_1(t-1)$, $x_2(t-1)$, $X_3(t-1)$, and $Y(t-1)$.

In more specific terms, the Hildreth-Lu technique as used in this analysis contains 3 independent variables: (a) real yield, (b) per capita disposable income, and (c) government purchases of goods and services expressed in constant 1958 dollars. The dependent variable in the model is per capita revenue-passenger-kilometers.

The Hildreth-Lu equation, which is based on a 19-year history from 1955 to 1973, is

$$\begin{aligned}1.6Y(t) = &-170.68 + [0.379\,45 \times X_1(t)] - [0.270\,74 \times X_2(t-1)] \\ &+ [3.2354 \times X_2(t)] - [2.3085 \times X_2(t-1)] \\ &- [29.665 \times X_3(t)] + [21.167 \times X_3(t-1)] \\ &+ [0.713\,51 \times Y(t-1) \times 1.6] \end{aligned} \quad (4)$$

In this equation the coefficient of determination = 0.9606, the Durbin-Watson statistic = 1.5007, and $F(3,14) = 113.91$. The t-statistics are as follows:

$X_1(t)$ = 6.8109 per capita disposable personal income,
$X_2(t)$ = 3.5251 government purchases of goods and services, and
$X_3(t)$ = 1.4126 yield.

The Hildreth-Lu equation in the form of equation 4 is the final product of a series of multivariate analyses in which as many as 10 variables were tested. The 3-variable equations are the result of these statistical tests and the desire to select not only a statistically reliable but also a logical model.

After arriving at a forecast of per capita revenue-passenger-kilometers for the forecast period, applying a population forecast of persons 16 years old and older and a trip length projection to obtain a forecast of originating passengers was simple. Application of a projected connection factor then yielded a forecast of passenger enplanements. The projections of both trip length and connection factor were developed by carefully examining past trends and by making projections of these trends through 1995.

Model results in the form of high, median, and low forecasts of originations and enplanements are given in Table 1. Applicable growth rates are given in Table 2. The median and high forecasts have been adopted as reasonable for assessing future traffic levels. The low forecast range is viewed as a pessimistic planning minimum. The high-low forecast ranges relative to U.S. originations and enplanements are shown in Figure 1. (The data in Tables 1 and 2 and Figure 1 are for 50-state scheduled service for air carriers that are on Civil Aeronautics Board certified routes.)

INDUSTRY FORECASTS

Exploring other industry forecasts is often useful in assessing the reasonableness of different forecasts developed at the same time. To this end, we reviewed 4 industry enplanement forecasts and compared them with the results of the aforementioned (Speas) model. Such a calendar year comparison is given in Table 3.

Table 4 gives the relationship of Baltimore-Washington and total U.S. trip originations. From 1960 to 1973, the study area percentage of U.S. trip originations has remained fairly constant except for certain short-term fluctuations in the mid 1960s. Since 1960, the region has experienced an average annual compound rate of growth of about 9 percent. From 1970 to 1973, however, this rate has slowed to approximately 5.3 percent/year. After an extensive review of forecasts for expected population and other Baltimore-Washington regional socioeconomic indicators, a multivariate technique was applied to arrive at estimates of future Baltimore-Washington trip origins as a percentage of total U.S. trip origins. These forecast data are also given in Table 4. Based on the past history of certificated air carrier traffic and regional and national economic trends, the regional percentage of trip originations is expected to show a modest increase. However, the absolute levels of passenger demand (originations) are expected to be substantial as indicated in Table 4.

MICROFORECAST OF REGIONAL AIR CARRIER PASSENGER TRAFFIC

The distribution of forecast air travel passenger originations by planning district relies on analysis of past travel behavior and the socioeconomic environment in which that travel took place. Available data in survey form were used that described the extent to which the public has used available air transport services. As part of the Maryland Aviation System Plan, an on-board survey of passengers was conducted for November 7 to 14, 1973; November 28 to December 13, 1973; and January 9 to 23, 1974. The on-board surveys were conducted on selected flights de-

Table 1. U.S. domestic airline passenger data.

Type of Data	Year	Range	Per Capita Revenue-Passenger-Kilometers	Population ≥16 Years	50-State Revenue-Passenger-Kilometers	Trip Length (km)	Originating Passengers (millions)	Connection Factor	Passenger Enplanements (millions)
Actual	1970		1209.8	142.8	173	1600	108	1.386	149
	1971		1212.2	145.3	176	1628	108	1.401	151
	1972		1322.5	147.7	195	1638	119	1.437	171
	1973		1388.7	150.3	209	1650[a]	126[a]	1.405	177
Forecast	1975	Low	1440	155.3	223	1673	134	1.40	187
		Median	1540	155.3	239	1673	143	1.40	200
		High	1640	155.3	255	1673	152	1.40	213
	1980	Low	1812	167.0	291	1731	175	1.40	245
		Median	1955	167.0	326	1731	189	1.40	264
		High	2098	167.0	350	1731	202	1.40	283
	1985	Low	2256	175.4	396	1793	221	1.40	309
		Median	2467	175.4	433	1793	241	1.40	334
		High	2677	175.4	469	1793	262	1.40	367
	1990	Low	2708	184.4	499	1857	269	1.40	377
		Median	3017	184.4	556	1857	300	1.40	419
		High	3325	184.4	613	1857	330	1.40	462
	1995	Low	3195	194.4	621	1923	323	1.40	452
		Median	3648	194.4	709	1923	369	1.40	516
		High	4101	194.4	797	1923	415	1.40	580

Note: 1 km = 0.6 miles.

[a]Estimated.

Table 2. Growth rate percentages for Table 1 data.

Time Period	Per Capita Revenue-Passenger-Kilometers	Population ≥16 Years	50-State Revenue-Passenger-Kilometers	Trip Length (km)	Originating Passengers (millions)	Connection Factor	Passenger Enplanements (millions)
1970 to 1975	6.3	1.7	8.0	0.9	7.1	0.2	7.3
1975 to 1980	5.1	1.5	6.6	0.7	5.9	—	5.9
1980 to 1985	5.0	1.0	6.0	0.7	5.3	—	5.3
1985 to 1990	4.4	1.0	5.5	0.7	4.8	—	4.8
1990 to 1995	4.2	1.2	5.0	0.7	4.7	0.1	4.7

Note: 1 km = 0.6 miles.

Figure 1. High and low forecasts for U.S. originations and enplanements.

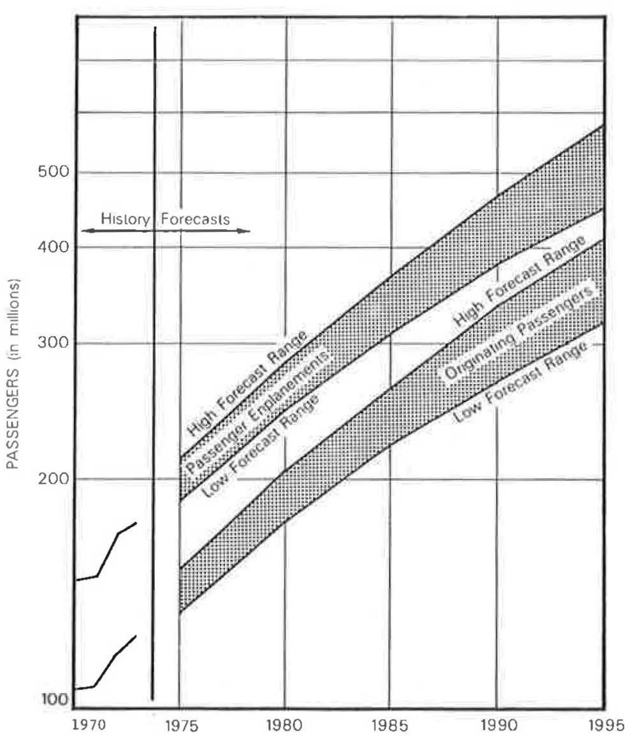

Table 3. 1974 industry forecasts for trunk and regional domestic scheduled service.

	Enplaned Passenger Forecasts (millions)					
					Speas	
Year	ATA	Boeing	General Electric	Lockheed	Median	High
1975	194.3	194.6	224.8	183.8	200.1	213.1
1980	254.6	261.1	275.6	254.7	264.1	283.3
1985	320.8	333.3	319.4	343.2	333.8	366.5
1990	394.5	410.0	—	—	419.4	462.3
1995	479.8	493.7	—	—	516.2	580.4
2000	567.7	579.9	—	—	—	—

Table 4. Relation of Baltimore-Washington and total U.S. trip originations.

Type of Data	Year	Range	United States	Study Area	Percent of United States
Actual	1960		38 868 000	1 755 090	4.52
	1961		40 143 000	1 950 990	4.86
	1962		42 754 000	2 179 970	5.10
	1963		49 047 000	3 165 740	6.45
	1964		55 697 000	2 880 380	5.17
	1965		65 593 000	3 367 420	5.13
	1966		75 069 000	3 791 650	5.05
	1967		88 435 000	4 545 330	5.14
	1968		103 746 000	4 885 130	4.17
	1969		111 697 000	5 317 500	4.76
	1970		107 952 000	5 045 800	4.67
	1971		108 267 000	4 938 820	4.60
	1972		119 267 000	5 385 170	4.52
	1973		127 474 000	5 884 641	4.62
Forecast	1985	Median	241 250 000	11 411 000	4.73
		High	261 790 000	12 383 000	4.73
	1995	Median	368 750 000	17 810 000	4.83
		High	414 580 000	20 024 000	4.83

parting the 3 commercial airports in the Baltimore-Washington region (6). The demand for air passenger service is strongly and directly affected by the socioeconomic environment for air travel as well as by the costs and quality of available transportation services. Nationality, population size, level and mix of employment, and level of per capita income have been found to have a significant impact on the volume of air travel and its rate of growth. Similarly, the transportation variables of time, cost, and frequency are known to have a direct and important influence on passenger traffic development. Consequently, reliable estimates of air travel by planning district require an understanding of the relationships between demand and the factors that influence demand.

By using survey results and a history of the socioeconomic environment in which travel took place, we developed a history of and forecast originations by planning district. For each of the 72 analysis zones, factored originations were segregated by type of trip (government business, nongovernment business, and all other). Three cross-sectional models were then developed to forecast independently the 3 types of trips on a zone basis together with the socioeconomic indicators relevant for each zone. The obtainable and usable variables were government employment, nongovernment employment, population, per capita income, and income product. The results from modeling each of the 3 types of trips for zones 1 through 72 in the aggregate proved to be unsatisfactory. For example, one conclusion in the case of zone-generated government trips was that the travel-generating potential of federal employees is vastly different from that of state employees in both frequency of travel and characteristics of mode of travel. Under the circumstances, therefore, further segregation of the aviation analysis zones into additional regions was found to be more satisfactory for logical and statistical results as well as for providing usable future system planning inputs. The 3 subregions identified were the city of Baltimore (zones 1 to 9), the remainder of the Baltimore metropolitan area (zones 10 to 28), and the Washington metropolitan area (zones 29 to 72).

The equations developed for the microforecast effort are indicative of the 3 subregions and for the 3 types of trips. For example, the equation that was used to generate government-related business trips in a zone of the Baltimore region is

$$\text{Trips from zone (1)} = -633.6 + 0.549\,36 \times \text{government employment (1)} + 0.060\,09 \times \text{population} + 1.1678 \times \text{per capita income} \quad (5)$$

In this equation, the coefficient of correlation = 0.80, the Durbin-Watson statistic = 1.4, and the F value = 4.9.

In general, the equations developed for the Washington area zones tended to be better predictors than those for the Baltimore area zones. After calibration and an add-on adjustment for external zones 73 to 78, the summation of the microforecast modeling effort revealed a slightly higher absolute number of originations for the forecast period than the median and high ranges of the macroforecast modeling effort:

Year	Macroforecast Median	Macroforecast High	Baseline Microforecast
1985	11 411 000	12 383 000	13 594 000
1995	17 810 000	20 024 000	22 094 000

The high range of the macroforecast was selected as being the most reasonable planning range inasmuch as microforecasts often have a tendency to overestimate the true generating potential of an area. This hypothesis proved to be true in this instance because the microforecast model tended to overestimate the base year, 1973, before normalization. Therefore, the microforecast was scaled to the high macroforecast control totals, preserving the zone distribution by type of trip.

OTHER U.S. CITY VERSUS BALTIMORE-WASHINGTON REGIONAL GENERATED DEMAND

The basic external demand unit considered in the regional study of originations was the number of round-trip passengers between the Baltimore-Washington region and another U.S. city. The forecast of city-pair demand was developed first from the region (not zone) to destination because previous studies have found that attempting to develop a forecast of originations from an aviation analysis area by type of trip or trip purpose to a specific external destination is an extremely hazardous exercise. Little reliability can be achieved, and the statistical results are often suspect. Therefore, the zone forecasts by type of trip and the regional forecasts to other U.S. city destination were developed independently.

For all destinations receiving nonstop service in 1971, 1972, or 1973 and any other destinations having an equal or greater number of passengers as any city receiving nonstop service in 1973, the absolute number of inbound plus outbound originations was determined based on the Civil Aeronautics Board 1973 Origination-Destination Ticket Sample. These "qualifying markets" were then grouped to the applicable Office of Business Economics (OBE) analysis area. The distance from the Baltimore-Washington region to each external OBE destination was determined. In addition, the following variables were compared for each external OBE area and the Baltimore-Washington region: per capita income, employment, manufacturing, and wholesale and retail trade. In the initial analysis, grouping markets by unit of distance block was determined to be essential to improvement of the modeling results. For example, the general form of the equation used in forecasting other U.S. markets 1600 or more km (1000 or more miles) from the region was

$$L(1) = -11\,274 + 0.088\,275\,6 \times M(1) + 0.486\,57 \times N(1) - 5.4730 \times P(1)(1.6) \quad (6)$$

where

L = round-trip originations from the region to the OBE external destination,
M = wholesale and retail trade product,
N = manufacturing product, and
P = distance in kilometers.

In this exercise, the coefficient of correlation = 0.995, the Durbin-Watson statistic = 1.6, and the F-value = 301.8. The results of the modeling effort reflect no adjustment for nonstop gateway flows. The results are rank-ordered by unit of distance block classification.

For those markets that did not fulfill the aforementioned criteria to be considered as qualifying markets, a concerted effort was made to redistribute these flows over the qualifying markets. The next step in this procedural analysis was to allocate 1973, 1985, and 1995 traffic levels over all nonstop sectors. All traffic must flow over a nonstop sector because all operations in a mode carry some portion of all traffic to the first nonstop destination. Future nonstop services were hypothesized as feasible when future traffic average daily round-trip flows attained a level of 100 passengers.

The final step was to merge all adjusted city-pair data with the region's aviation analysis zones. Trip purpose for each OBE city was gleaned from the survey. This information, in the form of proportions, was used with the trip-purpose detail developed from the micro forecast.

In the future, the existing trip-purpose proportions of traffic to each other U.S. city as derived from the survey may change. One must keep in mind that the survey, in addition to being a one-time application, is subject to the bias of seasonality. Therefore, although the survey proportions by type of trip for each other U.S. city are recognized as being usable inputs, they are secondary to the primary proportional derivation, the microforecast.

CONCLUSIONS

The forecasts developed were adequate for use in the modeling of air passenger distribution. They appeared reasonable on inspection and intuitively correct. They had their limitations largely because change was being forecast based on a one-time measurement. Time series data were available because a similar survey had been done in 1968, but the status of those data and their incompatibility made time-series analysis a costly and time-consuming process that could not be done within budget and time constraints.

Some of the equations appear counterintuitive. Signs on coefficients are not what would be expected, and relevant factors do not show an expected strong correlation. Part of this problem is a result of the lack of time series. Another is the result of data representing other factors. For example, as population goes up, air travel is expected to go up. Especially in the city of Baltimore, population becomes a surrogate variable for density, and, as density increases, air trips decrease.

This method is applicable in other multiairport environments, including areas in which a second airport is being considered for the future. With greater national application, a better understanding of trip-generation propensity would be developed.

REFERENCES

1. R. W. Pulling and H. J. Guth. Forecasting Traffic for Airport Development. In Airport Economic Planning (G. P. Howard, ed.), MIT Press, Cambridge, 1974.
2. Airport System Study. Association of Bay Area Governments, Berkeley, Calif., final plan, 1972.
3. J. Augustinus. An Air Passenger Distribution Model for the New York-New Jersey Area. In Airport Economic Planning (G. P. Howard, ed.), MIT Press, Cambridge, 1974.
4. A Study for the Development of Stewart Airport. Metropolitan Transportation Authority, New York, Phase 1 Rept., Vol. 2, 1973.
5. The Air Passenger Distribution. Maryland Aviation System Plan, Maryland Department of Transportation, Baltimore-Washington International Airport, Working Paper 5, 1975.
6. Air Traffic Activity Survey and Forecast Methodology. Maryland Aviation System Plan, Maryland Department of Transportation, Baltimore-Washington International Airport, Working Paper 2, Dec. 1974.