

AIRPORT SYSTEMS PLANNING, DESIGN AND MANAGEMENT

Assignment 1 Forecasting Exercise

Prepared by Richard de Neufville assisted by Jay Falker

Learning Objectives:

- To provide a realistic experience of the practical issues in forecasting of airport traffic;
- To demonstrate how any forecasting of traffic is built upon a series of assumptions, built on more and more assumptions -- a conceptual "house of cards"; and
- Thus to demonstrate that all traffic forecasts should be treated with caution and that, as a practical matter, planners should assume that the actual levels will be within a broad band away from the estimate.

Nature of Exercise:

Your general task is to develop forecasts from the same data available to the professionals involved in the 1997 Master Planning process for Los Angeles. You will encounter the same issues, and will be able to appreciate from your own experience the practical impossibilities of ever developing an accurate forecast.

The data are available in a Microsoft Excel ® file so that you can easily carry out statistical analyses, graph the data, and thus focus on the conceptual issues. The data are extracted from the Draft Chapter III Forecasts of Aviation Demand Report, dated February 26, 1996, prepared for the Los Angeles Department of Airports by Landrum and Brown consultants as part of their work on the Master Plan for the Los Angeles International Airport (LAX).

The data are basically of two sorts:

- historical records of traffic and demographic features of the Los Angeles region, and
- forecasts of future demographic features, needed as inputs for the equations used to create forecasts of traffic.

You will note that both the historical records and the forecasts come in several versions. In the case of Los Angeles, as indeed practically everywhere, different agencies measure and estimate the same factors in different ways, and come to different values. The Excel ® file, which you may download from the course web page, describes these data in more detail.

Your Tasks:

You are to proceed through each of the main assumptions involved in a forecasting exercise, thinking about each of them critically in the manner requested below. The first 4 are conceptual, the next involve using the computer to analyze the data.

Assumption 1: Specify the variables to be used to "explain" the levels of traffic. The LAX Master Plan focused on four explanatory demographic variables:

- the total population of the region,
- the level of employment,
- the per-capita income, and
- the "yield" of airlines for air travel (that is, the average revenue per unit distance, which can be taken as a measure of the relative price of airline tickets).

You are to discuss whether these are the best variables to use. What other factors might be important influences on the level of traffic in an area? Why are the ones selected best? Why should you use both population and employment, when these two variables are generally almost directly proportional to each other (as you can test)? (They can thus substitute for each other, and their coefficients change correspondingly.)

Assumption 2: Specify the form of the variables used. The LAX Master plan used the absolute levels of the “explanatory” variables in the analysis. Alternatively, as is often done in economic studies, they could have used a variety of forms:

- logarithmic, on the grounds that people’s sensitivity to economic factors is typically non-linear (e.g., the response of people to changes in price is generally represented by a demand curve);
- differences, on the grounds that persons respond not simply to the overall level of the price of an item they could buy, but to its price relative to other goods;
- segmented, on the grounds that the analysis should focus not on the total population of an area, but on the number of persons in the segments of the population that could afford to fly.

You should think about and discuss the choice of the form of the variables used to “explain” the level of traffic. What might be the best choice in this regard?

Assumption 3: Specify the form of the equation used in the statistical analysis. The LAX Master Plan fitted a linear equation to the data, whereas they could have used various curves. Among other possibilities, it might be more reasonable to fit an exponential curve of the form e^{rt} which corresponds to something growing at a compound rate of $r\%$ per unit time.

You should think about and discuss the choice of the equation used to develop a model connecting the “explanatory” variables with the item to be forecast. As for the discussion of Assumption 2, think about what might be the best choice.

Assumption 4: The number of periods to be considered in the analysis. The LAX Master Plan used data going back to 1975 (20 years).

Discuss the wisdom of this choice, noting that the circumstances prevalent in previous years are possibly quite different from what they are today. Specifically in this case, the airlines in the United States were deregulated in 1978, leading the airlines to offer different kinds of fares and services. More recently, travel throughout North America and indeed across the Pacific has been made much more competitive. Furthermore, more recent years are always probably more indicative of the immediate future than much earlier years.

Assumption 5: Choice of the formula to be used for forecasting. The LAX Master Plan used a specific formula as identified in the Excel ® file. You are to compare this equation with several alternatives that you will calibrate statistically, using the “REMI” data for the demographic characteristics as used by the LAX Master Plan. In each case consider the R^2 as a first order measure of the goodness of the fit of the model with the historical record.

Your formula should forecast LAX Domestic Passengers (total O&D traffic). The alternatives you should compare are:

1. A forecast based on the expression used by the LAX Master Plan;
2. A forecast based on the same as (1), but without the employment variable;
3. A forecast expressing traffic as a function of yield to a power;
4. A forecast based on the correlation between the level of traffic and e^{rt} ;
5. A forecast based on the correlation between traffic and any irrelevant variable for which you can find data for every fifth year (wine production in France, population of Chile, tons of steel produced worldwide, -- amuse yourself) [In a previous year, the best correlation in the class was with the prison population of the State of Oregon!]; and
6. Whatever better formula you can devise or prefer.

Discuss these comparisons and the process which you would pick the formula to use to forecast traffic for Los Angeles.

Assumption 6: Choice of estimates used for the “explanatory variables”. The LAX Master Plan selected the “REMI” estimates as the preferred input for their formula for predicting future traffic. However they could have used any of the 3 others they had available (as you do in the Excel ® file). Alternatively, they could have used combinations of these data, such as the Maximum, Average, or Minimum values.

Calculate the forecasts derived from these different possibilities, both for the formula used by the LAX Master Plan, and for your preferred formula. Plot the results, and note the range of forecasts in 20 years.

Bring this exercise to closure by discussing what you feel this experience has taught you about the validity of forecasts, and the confidence planners should place in them.