Institute of Transportation Studies
University of California at Berkeley

Delay and Flight Time Normalization Procedures for Major Airports: LAX Case Study

Mark Hansen
Tatjana Bolic

RESEARCH REPORT
UCB-ITS-RR-2001-5

NATIONAL CENTER OF EXCELLENCE FOR
NEXTOR
AVIATION OPERATIONS RESEARCH

June 2001
ISSN 0192 4095

Acknowledgments

This report documents research supported by the National Center of Excellence in Aviation Operations Research (NEXTOR) under FAA Contract DTFA03-97-D-004 and Cooperative Agreement 01-C-UCB-02. The authors would like to thank Dave Knorr, Rich Gutterud, and Ed Meyer of the FAA and Joe Post from the Center for Naval Analyses for their support and encouragement in this project. The opinions expressed herein are those of the authors and may not reflect those of the persons or organizations just cited. The authors are also solely responsible for any errors or omissions in this report.

Abstract

This report presents methodologies for normalizing performance of the National Airspace System (NAS). The purpose of the study is to develop the capability of isolating the performance of NAS enhancements, such as those being made under the Free Flight Phase I program. It is often not possible to observe the effect of such enhancements directly, because of the confounding influences of weather, demand, and conditions elsewhere in the system. The analysis presented here shows how linear and non-linear regression models can be used to statistically remove a large proportion of these confounding effects, increasing the possibility that the effects associated with the enhancement will be detectable.

The particular focus of this study is on arriving flights at Los Angeles International airport, where two FFP1 tools, Traffic Movement Advisor (TMA) and Passive Final Approach Spacing Tool (PFAST), are being deployed. We develop a metric that captures the daily variation in flight times (including departure delay and gate-to-gate time) for LAX arrivals. This metric, which we term the Daily Flight Time Index (DFTI), is a weighted average where the weights reflect the proportions of flights coming from different destinations over the analysis period. We then analyze the day-to-day variation in DFTI, relating it to weather, demand, and average delays at origin airports. Our data set extends over 41 months from January 1997 through May of 2000.

Our approach was to develop a “baseline” model and then compare it with a variety of others. The baseline model contains 9 weather factors (scores from which are generated from applying principal component analysis to 32 underlying weather variables), 2 demand factors, and an origin airport delay variable, in a simple linear form. It explains about 75% of day-to-day variation in DFTI. Origin airport congestion is the most important source of variation, followed by several weather factors relating to temperature, visibility, and wind. Demand is the least important source of variation in DFTI over the time period analyzed. Most of the effects observed are intuitively reasonable: for example, we find that the DFTI decreases with visibility. Some are more mysterious--for example, DFTI is found to decrease with temperature at LAX.

Several other models are estimated and compared to the baseline model. A response surface model that includes quadratic and interaction terms as well as linear ones offers some improvement in fit (adjusted R^2 of 0.82 as compared to 0.74) albeit with a vastly increased number of coefficients. A non-linear model also performs somewhat better. Model performance is quite insensitive to the number of weather factors used. However, models that capture weather by categorizing days rather than employing quantitative weather factors perform somewhat less well.

Finally, TRACON logs from outlier observations were inspected to find reasons that the model predictions were inaccurate for these days. Generally, it was possible to discern explanations on days when the model under-predicted the DFTI. These included facility outages, overly stringent ground delay programs, east flow operation, and, in one case, the closure of half the airport due to a visit by Air Force 1.

TABLE OF CONTENTS:

LIST OF FIGURES:

LIST OF TABLES:

1. Introduction

Free Flight Phase 1 (FFP1) is a Federal Aviation Administration program for improving the performance of the National Airspace System (NAS) through the deployment of advanced technologies for air traffic management. Most of the FFP1 technologies, including the Center-Terminal Automation System (CTAS) and the Surface Movement Advisor (SMA), are oriented to the airport terminal area. The FFP1 program involves deploying these technologies at limited set of the airports featuring high traffic and large amounts of aggregate delay.

In addition to the deployment activities, FFP1 includes a significant evaluation component. In broad terms, the purpose of the evaluation is to determine the impact of the program on NAS performance. Within the FAA performance evaluation framework, dimensions of performance include safety, delay, efficiency, predictability, flexibility, and system productivity. The FFP1 program office has developed a number of specific metrics pertaining to each of these aspects of performance. By tracking these metrics before and after deployment, it is hoped that the impacts of FFP1 can be adduced.

This evaluation effort faces a significant hurdle. As FFP1 is implemented, the world does not stand still. Weather, a major determinant of NAS performance, is ever changing. Demands on the NAS, which have grown steadily over the years, are expected to continue to do so in the future, with significant ramifications on system performance. Indeed it is possible that FFP1 will itself trigger user responses that affect demand. Finally, in addition to weather and demand, there is a plethora of other factors—enhancements to the NAS infrastructure not related to FFP1, facility outages, and so on—that may also cause changes in the performance metrics.

An additional complication derives from the fact that the operations at one airport are affected by conditions at another. Obviously, if a flight departure is delayed by conditions at the flight origin, this will cause an arrival delay at the destination airport, which may in turn cause a second departure delay and arrival delay at an airport further downline. Conversely, a ground delay program may trigger departure delays at an origin airport as a result of congestion at the destination. As a result of these forward and backward propagation effects, performance trends at any one location in the NAS are influenced by changes throughout the system.

Given these circumstances, it is not advisable, and may not be possible, to observe the performance impacts of FFP1 through simple before-and-after comparisons of the performance metrics. Rather, it is necessary to normalize such comparisons so that, to the greatest extent possible, the non-FFP1 influences can be controlled for. Only when this is done can before-and-after comparisons be translated into with/without ones.

While the need for normalization is easy to recognize, the task of normalization may be quite difficult. There is wide day-to-day and hour-to-hour variation in NAS performance. While many of the sources of this variation are understood, there has been relatively little research that attempts to systematically relate performance variation to its underlying causes. For example, we do not know how much of the observed variation in performance results directly from observable differences in weather conditions and fluctuations in demand, and how much is the result of other, less easy to identify, causes. A related question, which has also been little studied, is how to represent and quantify the relationships between performance and the factors influencing it. For example, is it better to treat weather as a set of continuous variables and treat performance as a function of these variables, or to treat weather as a discrete variable by adopting some classification scheme (clear and sunny days, stormy days, etc) and model performance variation in terms of these weather categories? Or, are the relationships so complex and non-linear as to render normalization impossible by either of these methods, or any others?

With such questions in mind, we have undertaken a series of studies in which we statistically analyze NAS performance variation in relation to factors related to weather, demand, and conditions elsewhere in the system. Each of the studies focuses on one airport. The airport studied here is Los Angeles International (LAX). FFP1 technologies that are being implemented at LAX include the two CTAS components: Traffic Management Advisor (TMA), which will be used by the Los Angeles ARTCC, and the Passive Final Approach Spacing Tool (PFAST), which is being implemented at the Southern California TRACON. The immediate purpose of this study, however, is not to assess the benefits from these deployments, but to analyze performance trends at LAX during the pre-deployment period. It is expected that the methodologies developed in this effort may then be used at a later time as part of the FFP1 evaluation.

The essence of our approach is to develop and statistically model a set of daily level performance metrics for LAX. The overarching metric is a weighted average of flight times into LAX, which we term the Daily Flight Time Index (DFTI). This flight time is measured as the interval between the scheduled departure and the actual arrival, and can be decomposed into several components, including the departure delay, the taxi-out time, the airborne time, and the taxi-in time. Our research revolves around observing and analyzing day-to-day variation in the DFTI metric and its components.

The remainder of this paper is organized as follows. Section 2 overviews the methodology employed in this research. Section 3 discusses performance trends, in terms of the DFTI and its components, at LAX. Sections 4 and 5 present our methods of introducing demand and weather into the normalization process, while Section 6 does the same for conditions elsewhere in the NAS. The next several sections present estimation results for models relating performance to weather, demand, and conditions elsewhere in the NAS. In Section 7, we present a baseline linear model of the DFTI. In Section 8, we consider similar models for the individual DFTI components. Section 9 considers a variety of alternative models. Section 10 discusses "outlier" days in which predicted performance was considerably different from what actually occurred. Finally, Section 11 presents conclusions and recommendations.

2. Methodology Overview

This research is concerned with fitting models of the general form:

$$P\_METRIC_t = f(WX_t, DEMAND_t, ORG\_DELAY_t) + \varepsilon_t \quad (1)$$

where:

P_METRIC_t is the value of a performance metric for day t based on flights arriving at the study airport (in this case LAX);

WX_t is a vector of weather variables for day t;

$DEMAND_t$	is a vector of demand variables for day t;
ORG_DELAY_t	is a measure of delay at origin airports;
ε_t	is a stochastic error term.

While a variety of performance metrics could be analyzed using this type of model, in the work presented here we consider only the DFTI metric, and the components of that metric, as discussed above. Details on the calculation of this metric are presented in the next section. As a performance metric, the DFTI is quite similar to average arrival delay. It has two important advantages over that more conventional measure, however. First, it is insensitive to changes in the amount of "padding" built into the flight schedule. As is well known, airlines add extra time in the expectation that their flights will be delayed, and the amount of such padding has generally increased over time. All else equal, the padding increases will cause delays against schedule to go down. By using the DFTI, we avoid this "artificial" effect. Second, use of DFTI permits the decomposition of flight time, and hence delay, changes into the components mentioned above: departure delay, taxi-out time, flight time, and taxi-in time. Since conventional delay metrics require reference to some scheduled time, they can capture only the first of these.

We employ a metric defined at the daily level, rather than some finer time scale, for several reasons. First, as elaborated below, this permits a larger number of flight origins to be included in the averaging. Second, we can treat each day as an independent observation, ignoring interrelationships whereby performance in one time period affects performance in some other period. Such interrelationships cannot be ignored in sub-day time scales, since flights that are unable to land in one time period become additional demand in a subsequent period, often resulting in delays for other flights.

In developing models of the general form given in (1), we recognized that our primary goal is normalization as opposed to testing specific hypotheses or identifying particular causal mechanisms. The ultimate purpose of this research is to isolate the performance impacts of FFP1, not those of weather, demand, or origin airport delay. Our aim is to control for these factors, which is not the same as understanding exactly how and why they influence performance. In later sections, explanations and interpretations for the effects we observe will be offered, but some of these are rather speculative, and their validity is not essential given our emphasis on statistical control. The ways in which such thinking shaped our research effort will become evident in the sections that follow.

3. Computing the Daily Flight Time Index for LAX

We used individual flight data drawn from the Airline Service Quality Performance (ASQP) database to develop a daily time series of average flight times for flights arriving at LAX. As previously explained, we refer to this daily average value as the Daily Flight Time Index-DFTI. The DFTI is a weighted average of individual flight times into LAX. For individual flights we obtain from ASQP the time interval between its *scheduled* departure from its origin gate and its *actual* arrival at its gate at LAX. In addition, we obtain the components of this interval, which include the departure delay, the taxi-out time, the airborne time, and the taxi-in time.

To construct the DFTI (and its components) from the individual flight data, we take a weighted average where the weights reflect the proportion of flights from each origin into LAX over the period of analysis. Since the same weights are used for all the days in the sample, DFTI's are comparable even when the mix of long-haul and short-haul flights changes over time.

We limited the set of origins in the DFTI average in two ways. First, we eliminated airports within 200 miles of LAX. This reduced the influence of correlation between conditions at the origin airport and conditions at LAX. Second, for purposes of computing the average it was necessary that, for every day considered, there must be at least one completed flight from each origin. If an origin had no flights on a given day, one can either exclude the day from the sample or the origin from index. By excluding just a few days, the number of origins that can be included in the DFTI average was greatly increased. In the case of LAX, over the five plus years that we analyzed, we were able to include some 23 origins in the DFTI by eliminating just 4 of 1978 days. The set of origins, and their associated weights, are shown in Table 1. By way of comparison, in order to include 34 origins in the DFTI, it would have been necessary to eliminate 1024 days.

Table 1. Set of Origins and Their Associated Weights.

Origin Airport	Weight (in Percentage)
LAS	12.1
SFO	11.4
PHX	11.1
OAK	9.0
ORD	8.4
SEA	5.9
DFW	5.6
SJC	4.4
PDX	3.8
SMF	3.7
SLC	3.6
HNL	3.5
IAH	3.0
ATL	2.2
TUS	2.2
ABQ	1.8
STL	1.8
MIA	1.7
ELP	1.4
MCO	1.3
PIT	1.3
CVG	1.0

Figures 1, 2, and 3 summarize trends in DFTI for LAX since 1995. Figure 1 simply plots the daily values. It is evident from Figure 1 that DFTI varies substantially from day to day. Most of the time, it is within the range of 140 to 160 minutes. But there are a considerable number of days in which DFTI exceeds 180 minutes, and a few where it goes above 200. Taking 140 minutes as a nominal value for good days with essentially no delayed flights, we see that on very bad days the average delay (measured against the 140 minute standard) can exceed an hour per arrival.

Figure 2 plots the 30-day moving average of the DFTI, with the data points coded by season. A seasonal pattern is evident, with higher values in the fall and winter and the lowest values generally in the summer. Recent years have seen some divergence from this pattern. The springs of the last three years have been considerably worse than previous ones, as was the summer of 1999 compared to those of 1996-1998. Also, during the fall of 1999, the DFTI was stable throughout the quarter instead of climbing toward the end. The early winter of 2000 was also considerably better than average, although high DFTI values returned in the latter part of that quarter. Overall, Figure 2 suggests some trend toward increasing DFTI over the past five years, although that trend is dominated by seasonal variation.

Figure 3 plots the 7-day DFTI moving average, decomposed into three flight time components: time at origin (departure delay plus taxi-out time), airborne time, and taxi-in time. While airborne time is clearly the largest component of DFTI, it is evident that "spikes" in the DFTI are normally accompanied by similar "spikes" in time at origin, suggesting that the latter is the largest source of DFTI variation. To investigate this question more directly, we analyzed the variance in DFTI, using the identity:

$$\begin{aligned} VAR(DFTI) &= VAR(TAO) + VAR(ABT) + VAR(TIT) + \\ &2 \cdot COV(TAO, ABT) + 2 \cdot COV(TAO, TIT) + 2 \cdot (ABT, TIT) \end{aligned} \quad (2)$$

Where:

DFTI is the daily flight time index;

TAO is the time at origin;

ABT is the airborne time;

Figure 1. Daily DFTI Values for LAX

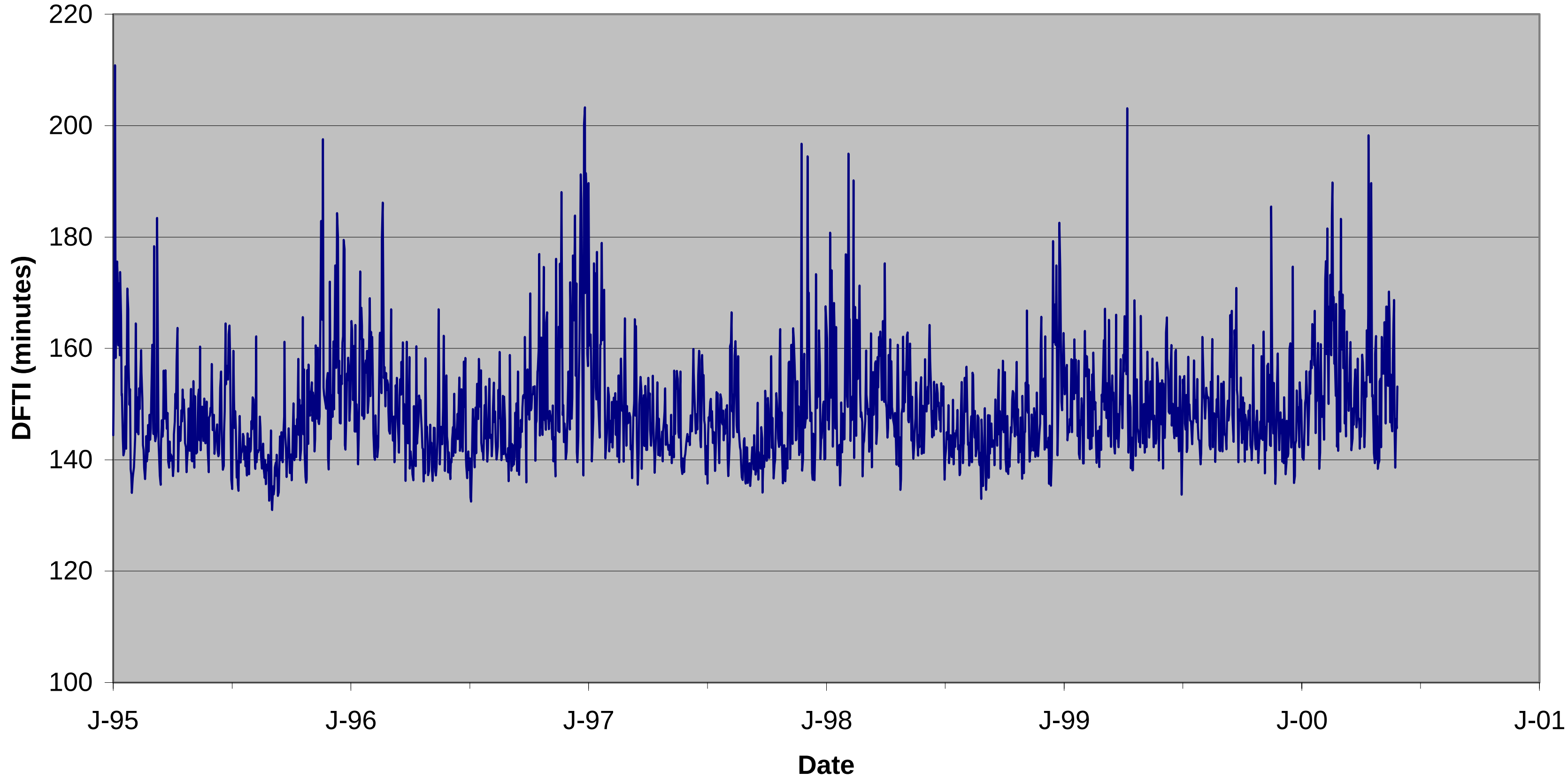


Figure 2. 30-Day Moving Average of DFTi, by Season

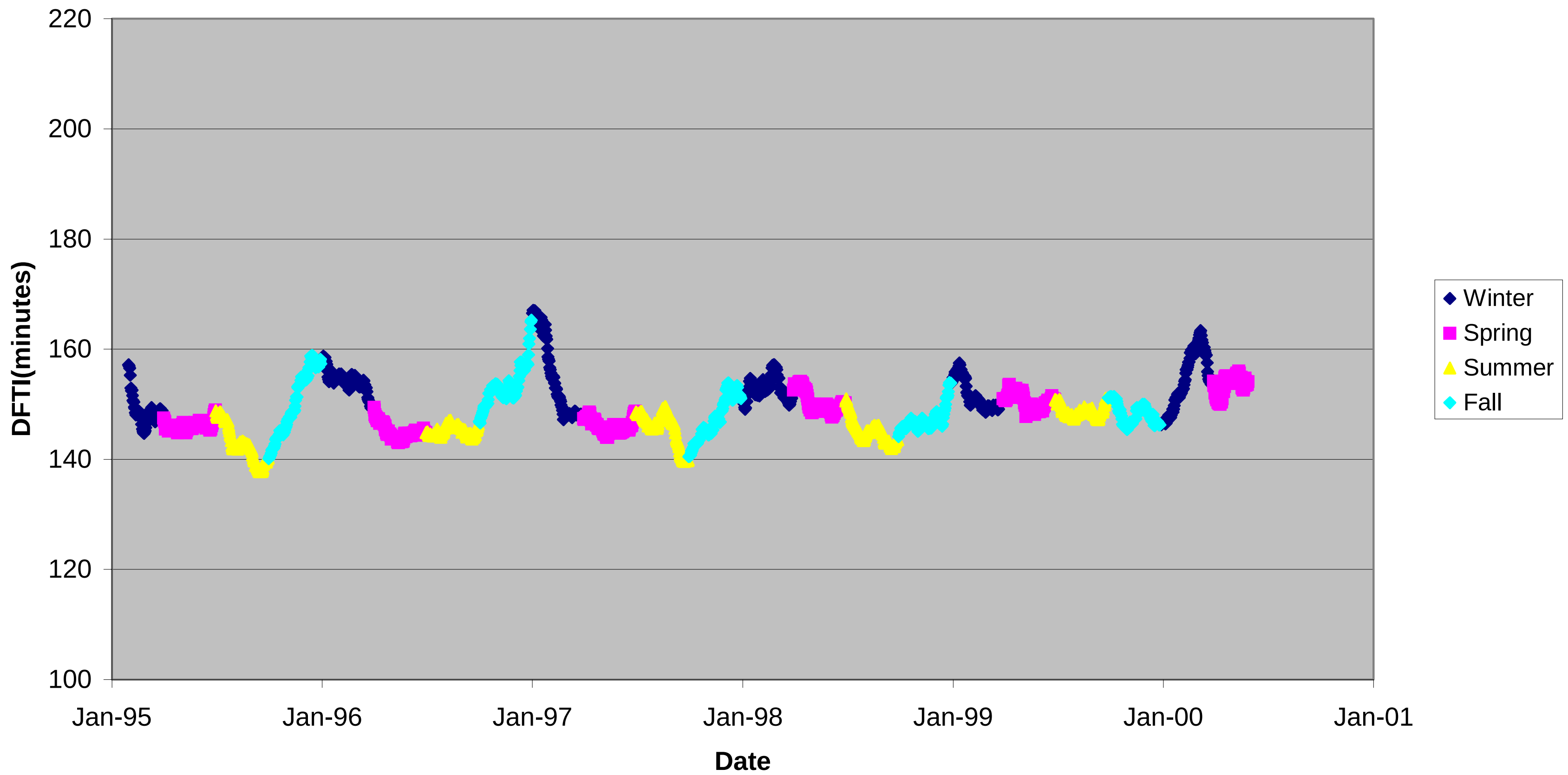
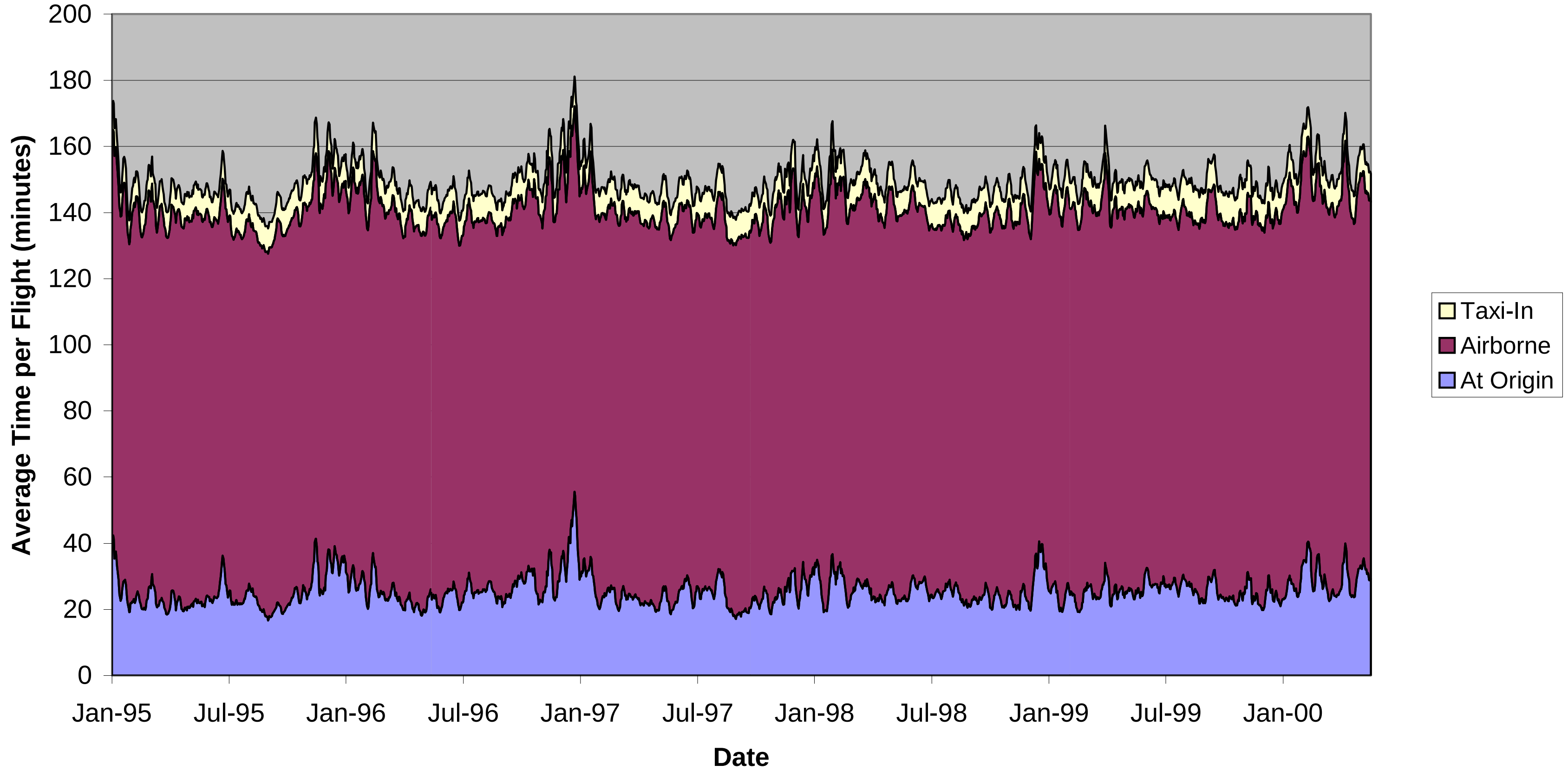


Figure 3. 7-day DFTI Moving Average, Three Components



TIT is the taxi-in time:

$VAR(\cdot)$ is the variance;

$COV(\cdot)$ is the covariance.

Figure 4 shows the decomposition of $VAR(DFTI)$ into these various components. As suggested above, $VAR(TAO)$ is the largest source of DFTI variance. The second greatest source, in most years, is the $COV(TAO, ABT)$ term, while $VAR(ABT)$ is the third, and *COV(TAO,TIT)* fourth. Thus variation in time at origin contributes to variation in DFTI not only directly, but also through its co-variation with the other flight time components.

The significant co-variation between time at origin and airborne time may be explained in several ways. First, air traffic management will cause flights to be held at their origins as a result of congestion either en route or in the LAX terminal area. In effect, this creates a link between time at origin and airborne time that is captured by the covariance. Second, despite the 200 mile requirement there is some inherent correlation between weather conditions at the various origin airports—particularly the nearer ones, which figure prominently in the DFTI average—and LAX.

4. Weather Normalization Variables

4.1 Qualitative Discussion

The first part of the section will cover some qualitative explanations of weather influence on delays at the LAX, and second part of the section will cover development of the weather variables used for DFTI metric modeling.

LAX airport has two sets of parallel runways (Figure 5). Outboard runways are used for arrivals. Inboard runways are used for departures, but they can be used for departures and arrivals simultaneously. Simultaneous use of the runways depends both on the weather situation and on the number of departures. If the demand for departures is high, arrivals will not be allowed on the inboard runways, regardless of weather. That usually happens around 8 am and noon.

Figure 4. Variance Decomposition for DFTI

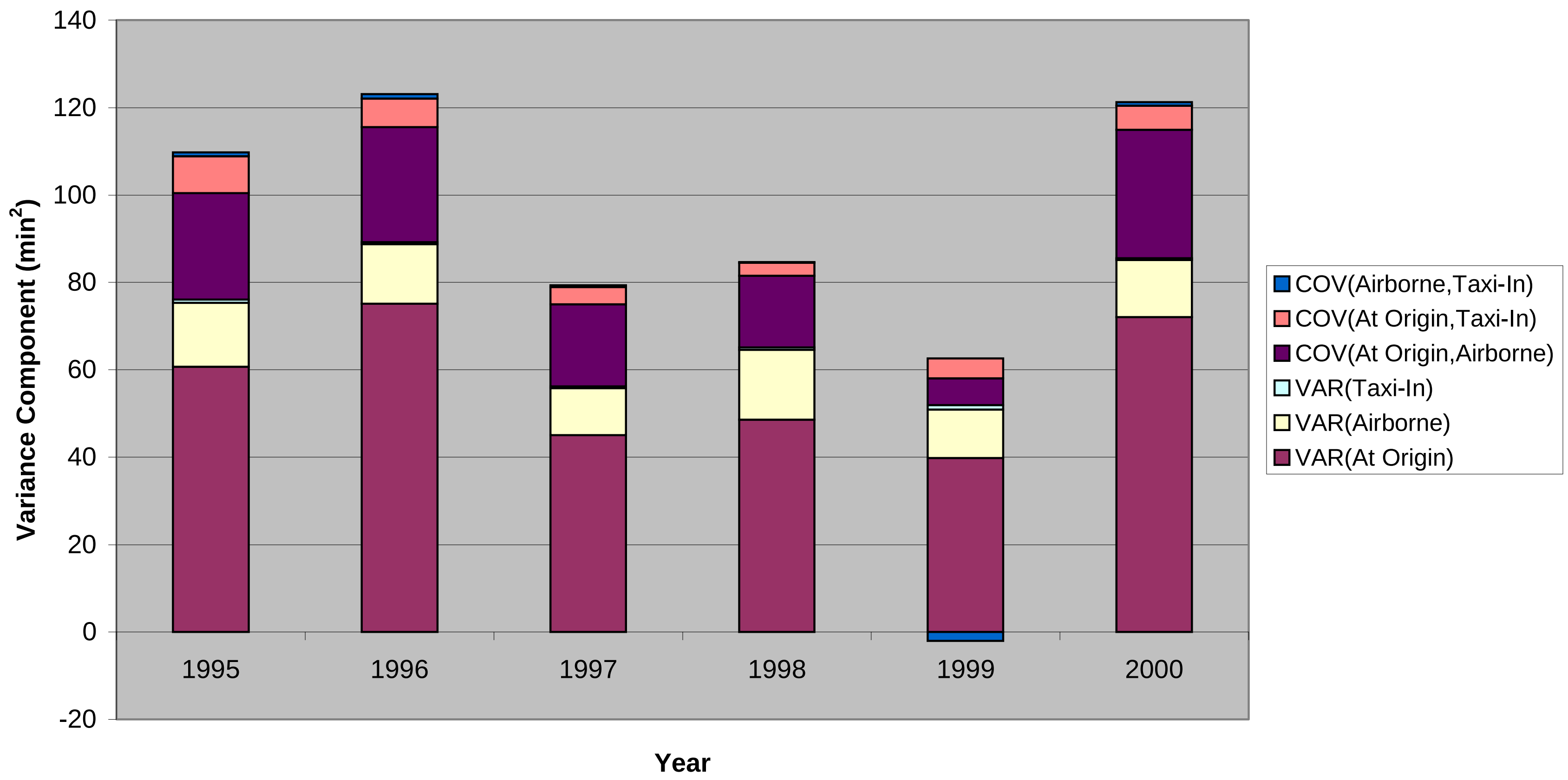
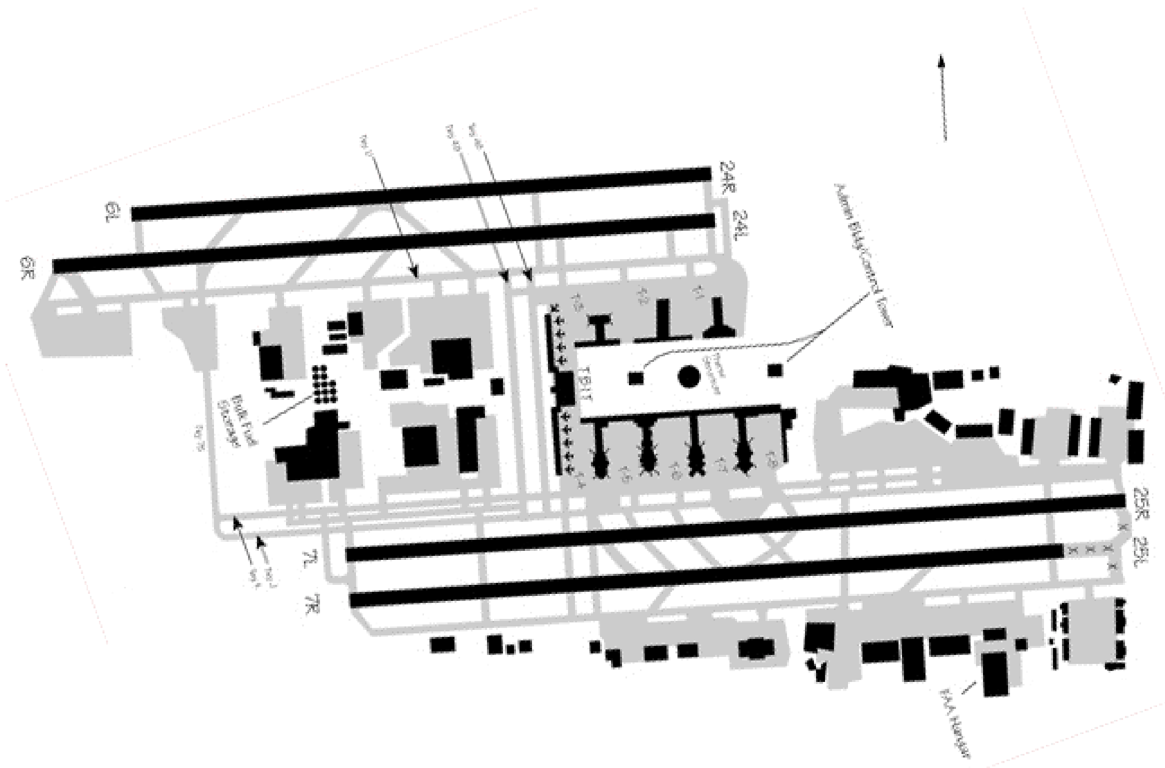


Figure 5. LAX Airport Configuration



The number of aircraft that can land during one hour is referred to as Airport Acceptance Rate (AAR), and it mostly depends on weather (and sometimes on departure demand too). From the Terminal Management Unit (Southern California TRACON) we learned that depending on the weather situation Airport Acceptance Rates for LAX are the following:

- 84 for VFR (Visual Flight Rules) operations when all four runways are in use,
- 72 for VFR operations when less than four runways are in use and no wind present,
- 68 for IFR (Instrument Flight Rules) operations using two runways (lower visibility and/or high winds), sometimes with visual finals,
- 62/60 for IFR operations when there is a low cloud ceiling present,
- It can be even lower than that, depending on the weather situation.

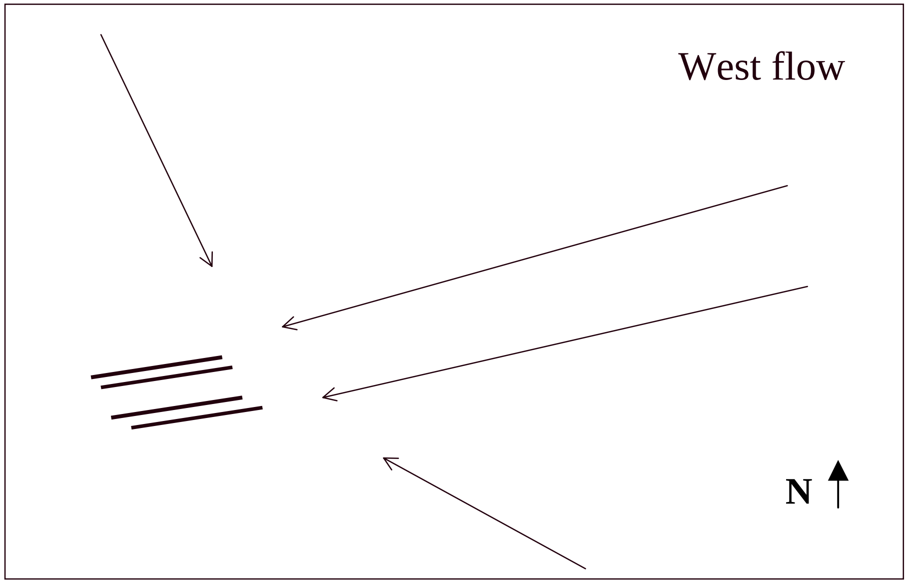
Delays occur whenever the demand is greater than the airport capacity. In the case of arrivals, delays happen whenever the demand is greater than the Airport Acceptance Rate. If the demand is greater than the capacity, aircraft have to wait somewhere until the airport is able to serve them. Delays can be imposed to the aircraft at the origin airport, en-route or in the terminal area (holding). As airlines usually schedule for fair weather any decrease in AAR caused by weather (and other factors) can lead to delays.

Several weather factors influence operations in Terminal Area. In the following text we will focus on the following:

- Winds,
- Clouds,
- Precipitation,
- Haze.

Winds. Due to winds over the year, arrivals at LAX use the west flow 90-95% of the time (Figure 6). Sometimes, AAR can decrease if the tail wind component is strong. When the tail wind component is strong, pilots cannot decrease their aircraft's speed

Figure 6. West Flow at LAX



according to the approach procedure, so air traffic controllers have to separate them more than usual. AAR decreases with greater separation between aircraft.

Operations change to east flow when the wind speed and angle do not allow for the west flow operations. During the day east flow is in effect until the winds change. During the night (midnight till 6.30 am) LAX is running operations in the east flow due to a noise abatement procedure. AAR rate goes down whenever the east flow is used. The fact that AAR decreases while utilizing east flow is not directly connected to the weather situation. Rather, the lower rates derive from the lack of familiarity of both pilots and air traffic controllers with the east flow procedures and the absence of high-speed runway exits for the east flow traffic (see Figure 5).

Clouds. Operations at LAX are VFR if the cloud ceiling is higher than 5000 ft. VFR operations allow for use of inboard runways for arrivals. If the cloud ceiling goes bellow 5000 or 4000 ft, operations are IFR, but some aircraft still can get permission for visual final approach. With cloud ceiling lower than 3000 or 2000 ft operations are IFR only. It is already described in the previous text how the VFR or IFR operations influence Airport Acceptance Rates.

From Terminal Management Unit we found out that sometimes AAR decreases when the cloud ceiling is between 8000-12000 ft west of the airport, along the approach routes. We do not see direct connection between the lowering of the AARs and clouds in that area. It may be that it influences the AAR when the demand is high and when controllers try to sequence the aircraft further away from the airport.

Precipitation. Precipitation slows down the air traffic as well. The visibility is lower when precipitation is present and that slows down the pilots. The runway occupancy time increases with precipitation. When runways are wet it takes longer for aircraft to slow down, so the occupancy time increases. Both visibility and runway wetness lowers the AAR.

Haze. Haze is a weather phenomenon that is specific for LAX basin. The haze is elevated polluted air. Sometimes, even when there are no clouds, the haze layer forces controllers to use IFR procedures.

4.2 Normalization Procedure

In this part, we develop the variables that will be used to normalize for weather at LAX. As revealed above weather can affect NAS performance in general, and our DFTI metric in particular, in a variety of ways. Performance may therefore be related to a wide range of meteorological parameters. In analyzing these relationships on a daily level, we must also consider the fact that performance effects are interrelated with the daily patterns of demand, as we already showed in previous text. Our task is to capture the effects of weather in a manner consistent with these complexities, and yet to do so in a manner that is economical enough to allow for meaningful statistical analysis.

To meet these aims we used principal components analysis, a type of factor analysis, to develop a set of daily metrics that characterize weather conditions at LAX. The data underlying these metrics was obtained from the CODAS weather database, which provides hourly data on temperature, wind, cloud ceiling, visibility, precipitation, and mode of operation (VFR versus IFR). For the factor analysis, we used the CODAS data to develop summary information for four daily periods: early morning (0-600), morning (600-1200), afternoon (1200-1800), and evening (1800-2400). For each of these periods, we calculated eight measures. These included:

- Average temperature
- Average visibility
- Average wind speed
- Total precipitation
- Proportion of time with VFR operation
- Proportion of time cloud ceiling was 3000 ft or under (“Low” cloud ceiling)
- Proportion of time cloud ceiling was over 3000 and under 8000 ft (“Medium” cloud ceiling)
- Proportion of time could ceiling was over 8000 and under 10000 ft (“High” cloud ceiling)

We had four daily observations for each of these eight variables, or a total of 32 variables per day. In the subsequent analysis, each variable is converted into a standardized variable (zero mean, unit variance). The factor analysis procedure was then used to collapse the 32 standardized variables to a smaller number of factors, which are also constructed so that they have zero mean and unit variance. Each of the factors is a linear combination of the original 32 variables. The factors are constructed so that the first accounts for the largest possible amount of the variation in the 32 variables, the second accounts for the largest possible amount of variation unaccounted for by the first factor, and so on. While 32 factors are needed to fully capture the variation in 32 variables, a much smaller number of factors will generally account for most of the variation. This is particularly true when the variables are highly intercorrelated, as they are in this case.

Figure 7 summarizes results of the factor analysis. The first factor (Number 1) accounts for over than 20 percent of the variation in the original 32 variables, and over 50 percent of the variation is explained by four factors. Nine of the 32 factors explain more than 1/32 of the variation. This is a useful threshold for determining how many factors to retain, since a “factor” constructed from just one of the 32 variables would do just this well if the variables were completely uncorrelated. Thus, on the basis of this criterion, we retained nine variables for subsequent analysis. Together, these nine variables account for 73 percent of the variation in the weather data. Figure 8 shows the proportion of variation of each weather variable that is accounted for by the nine factors. The factors capture the majority of the variation for virtually every factor, and over 70 percent of the variation for most. Visibility, VFR, and low ceiling variables are particularly well captured, while for precipitation, winds, and medium/high ceilings the factors do somewhat less well.

Table 2 shows the correlations between the nine factors and the 32 underlying weather variables. The first factor is highly correlated with visibility and VFR conditions, and negatively correlated with a low cloud ceiling. The second factor has high negative correlations with temperature, and moderate positive correlations with morning winds, medium cloud ceiling, and high ceilings. High wind and temperature conditions throughout the day are associated with the third factor. Table 3 provides summary of qualitative interpretations of each of the nine factors.

In considering these factors, it is important to recognize that, by construction, they are mutually orthogonal. That is, any factor is uncorrelated with any other factor. Thus, for

Figure 7. Variation Explained by LAX Weather Factors

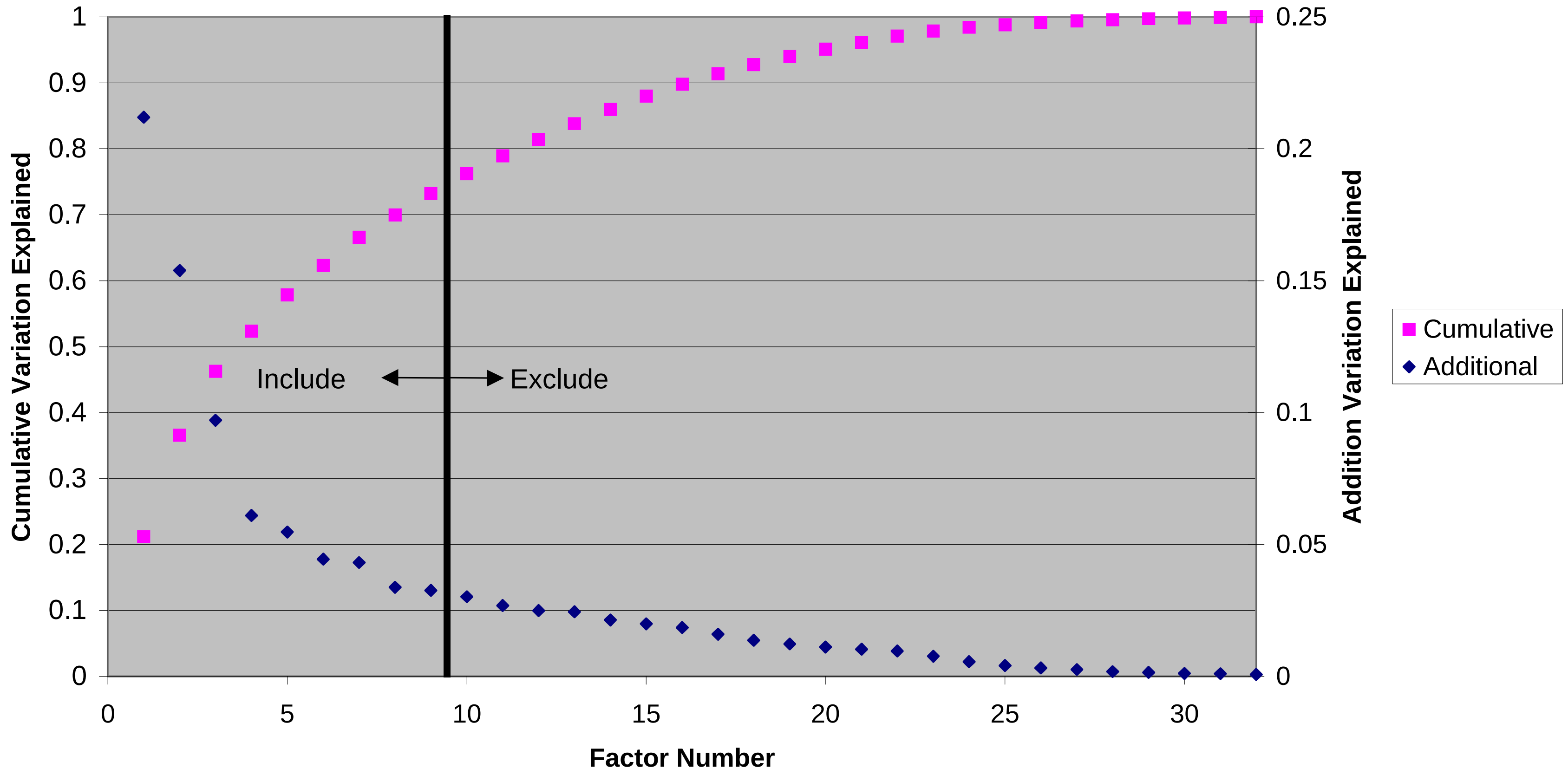


Figure 8. Proportion of Weather Variables Explained by Factors

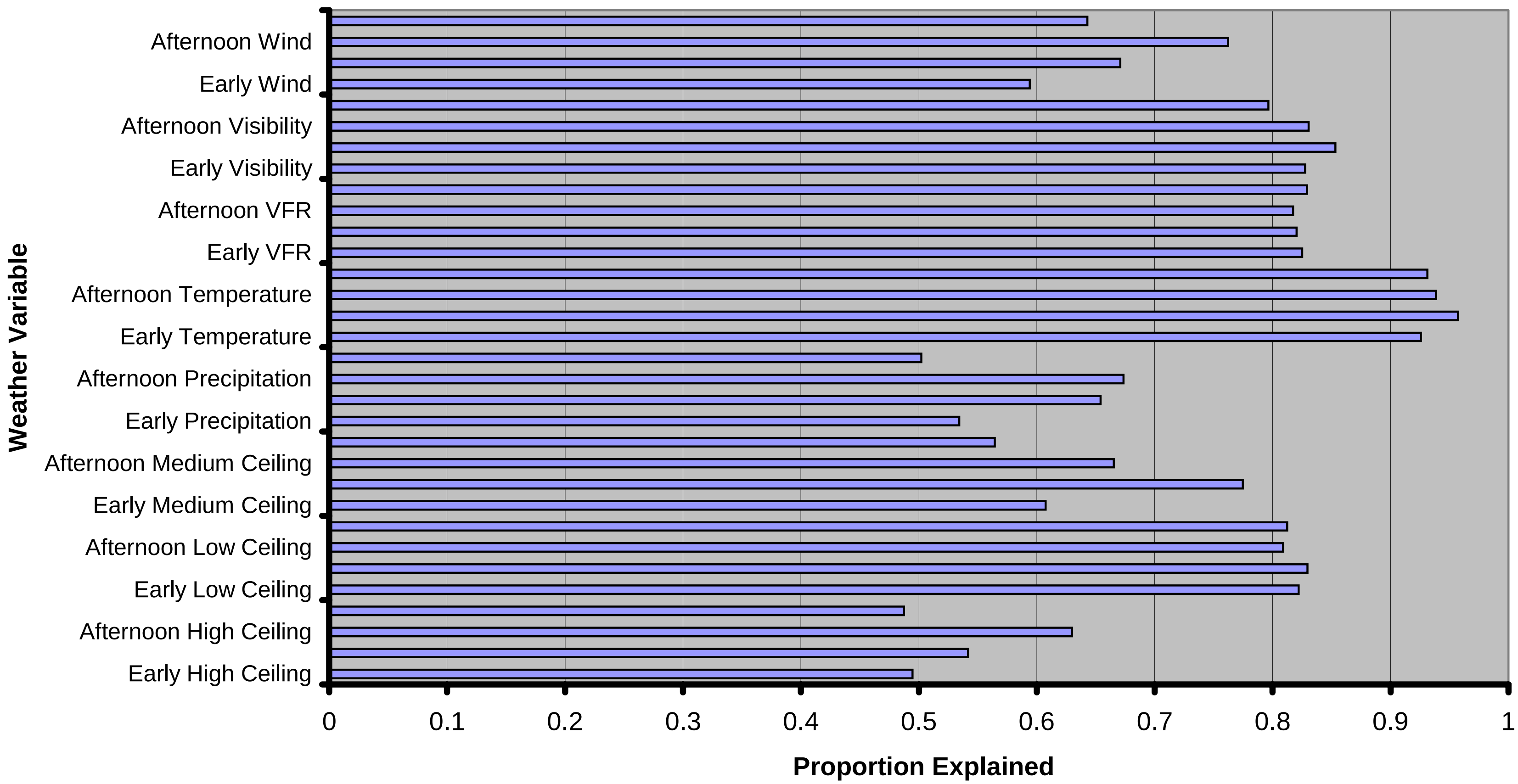


Table 2. Correlations Between the Factors and Weather Variables

		FACTOR								
Variable	Day Time	1	2	3	4	5	6	7	8	9
Wind	Early am	0.12	0.48	0.31	-0.27	0.22	-0.00	-0.05	0.36	-0.02
	Late am	0.08	0.40	0.53	-0.20	0.35	0.05	0.01	0.06	-0.22
	Afternoon	0.14	-0.02	0.50	-0.37	0.47	-0.04	0.14	-0.19	-0.28
	Evening	0.08	0.28	0.52	-0.25	0.38	-0.10	0.07	-0.25	-0.08
Temperature	Early am	-0.26	-0.61	0.65	0.21	-0.07	0.09	-0.00	-0.00	-0.06
	Late am	-0.12	-0.76	0.53	0.26	-0.05	0.08	-0.00	-0.01	-0.08
	Afternoon	0.02	-0.83	0.43	0.22	-0.05	0.08	0.00	0.04	0.01
	Evening	-0.08	-0.78	0.49	0.24	-0.08	0.12	0.02	0.00	0.03
Visual Operations	Early am	0.74	-0.04	-0.27	0.28	0.234	0.06	-0.24	-0.02	-0.05
	Late am	0.78	-0.18	-0.18	0.19	0.13	-0.09	-0.12	0.09	-0.24
	Afternoon	0.72	-0.25	-0.04	-0.14	-0.18	0.06	0.40	0.06	-0.10
	Evening	0.76	-0.01	-0.10	-0.08	-0.07	0.39	0.19	-0.19	0.04
Visibility	Early am	0.70	0.01	0.27	0.05	0.10	-0.23	-0.37	0.02	0.26
	Late am	0.76	-0.01	0.27	-0.02	0.05	-0.30	-0.21	0.15	0.22
	Afternoon	0.73	-0.08	0.32	-0.19	-0.13	-0.19	0.02	0.12	0.29
	Evening	0.68	0.02	0.30	-0.26	-0.10	0.02	-0.02	-0.05	0.41
Precipitation	Early am	-0.01	0.28	0.15	-0.11	0.02	0.30	0.27	0.51	-0.00
	Late am	-0.06	0.40	0.18	0.10	0.16	0.62	-0.11	0.00	0.18
	Afternoon	-0.08	0.40	0.16	0.36	0.27	0.46	-0.16	-0.10	0.19
	Evening	0.01	0.35	0.10	0.55	0.12	-0.08	0.17	0.12	0.05
Low Ceiling	Early am	-0.74	-0.01	0.26	-0.28	-0.17	-0.11	0.25	0.05	0.15
	Late am	-0.77	0.12	0.16	-0.19	-0.02	0.05	0.13	-0.06	0.36
	Afternoon	-0.72	0.19	0.05	0.12	0.22	-0.10	-0.40	-0.08	0.12
	Evening	-0.76	-0.04	0.07	0.07	0.13	-0.39	-0.17	0.16	-0.04
Medium Ceiling	Early am	0.13	0.48	0.32	-0.01	-0.33	0.05	-0.20	-0.08	-0.31
	Late am	0.08	0.47	0.39	0.00	-0.50	-0.04	-0.21	-0.09	-0.30
	Afternoon	0.00	0.57	0.23	0.13	-0.51	0.04	-0.07	0.03	-0.04
	Evening	0.01	0.52	0.27	0.18	-0.39	-0.00	-0.08	0.14	0.04
High Ceiling	Early am	0.06	0.38	0.12	0.50	0.21	-0.09	0.18	0.00	-0.02
	Late am	0.10	0.32	0.10	0.39	0.12	-0.21	0.43	0.17	-0.04
	Afternoon	0.15	0.37	0.11	0.37	0.00	-0.30	0.40	-0.22	0.16
	Evening	0.11	0.30	0.11	0.05	-0.12	-0.10	0.14	-0.57	0.07

Table 3. LAX Weather Factor Interpretations, Principal Components

FACTOR	Interpretation
1	High visibility and absence of low ceiling throughout day.
2	Cold temperatures, precipitation, and medium-to-high cloud ceiling throughout day. High winds except in afternoon.
3	High winds and temperatures throughout day.
4	High cloud ceiling until evening, afternoon precipitation, low winds and warm temperatures throughout day.
5	High winds and absence of a medium cloud ceiling throughout day.
6	Daytime precipitation, low visibility except in evening, VFR operations in evening.
7	High ceiling during daytime. Low visibility during early part of day. VFR operations during afternoon.
8	Precipitation and high winds during early part of day, lack of a high cloud ceiling during latter part of day.
9	High visibility, particularly in the evening. IFR operations and low cloud ceiling in late morning. Low winds during daytime hours.

example, a day with a high score for Factor 1 may also have a high score on Factor 6. Since, during the daytime, these factors are, respectively, positively and negatively correlated with visibility, a day with such a set of factor scores would (subject to its scores on the other factors) be expected to have average visibility.

To facilitate interpretation of the weather factors, we rotated them. The objective of factor rotation is to create factors that are highly correlated (either positively or negatively) with some variables while having low correlation with others. Various rotation procedures have been developed; for the weather data we chose the promax procedure. This is an oblique rotation method. Unlike the original factors, those generated by oblique rotations may be correlated with each other. The correlation will be limited since the rotated factors must span the same variable space as the original ones.

Table 4 shows the correlations between the rotated factors and the original 32 weather variables. As intended, the new factors tend to have very high correlations with certain variables and low correlations with others. For example, Factor 1 corresponds essentially to high temperature. Table 5 provides qualitative interpretations of all nine factors. The relative brevity of the descriptions is an indicator of the value of factor rotation.

5. Demand Normalization Variables

As in previous studies, (Hansen and Wei, Multivariate Analysis of the Impacts of NAS Investments: A Case Study of A Major Capacity Expansion at Dallas-Fort Worth Airport) we use the concept of hypothetical deterministic delay (HDD) to capture the intensity of demand at LAX. The HDD concept is illustrated using a representative daily schedule into LAX, as depicted in the queuing diagram in Figure 9. The *scheduled* curve in this figure represents cumulative scheduled arrivals. It is derived from the CODAS daily schedule data, which records scheduled arrivals, based on the Official Airline Guide, for 15-minute intervals. One can see that on the day depicted, 1175 arrivals were scheduled into LAX, with peak demand levels (shown as steep portions of the cumulative curve) in the late morning, mid-afternoon, and late evening. The curve as drawn continues into the next day, but only arrivals scheduled for the first day are considered, thus the curve becomes flat after midnight of the first day.

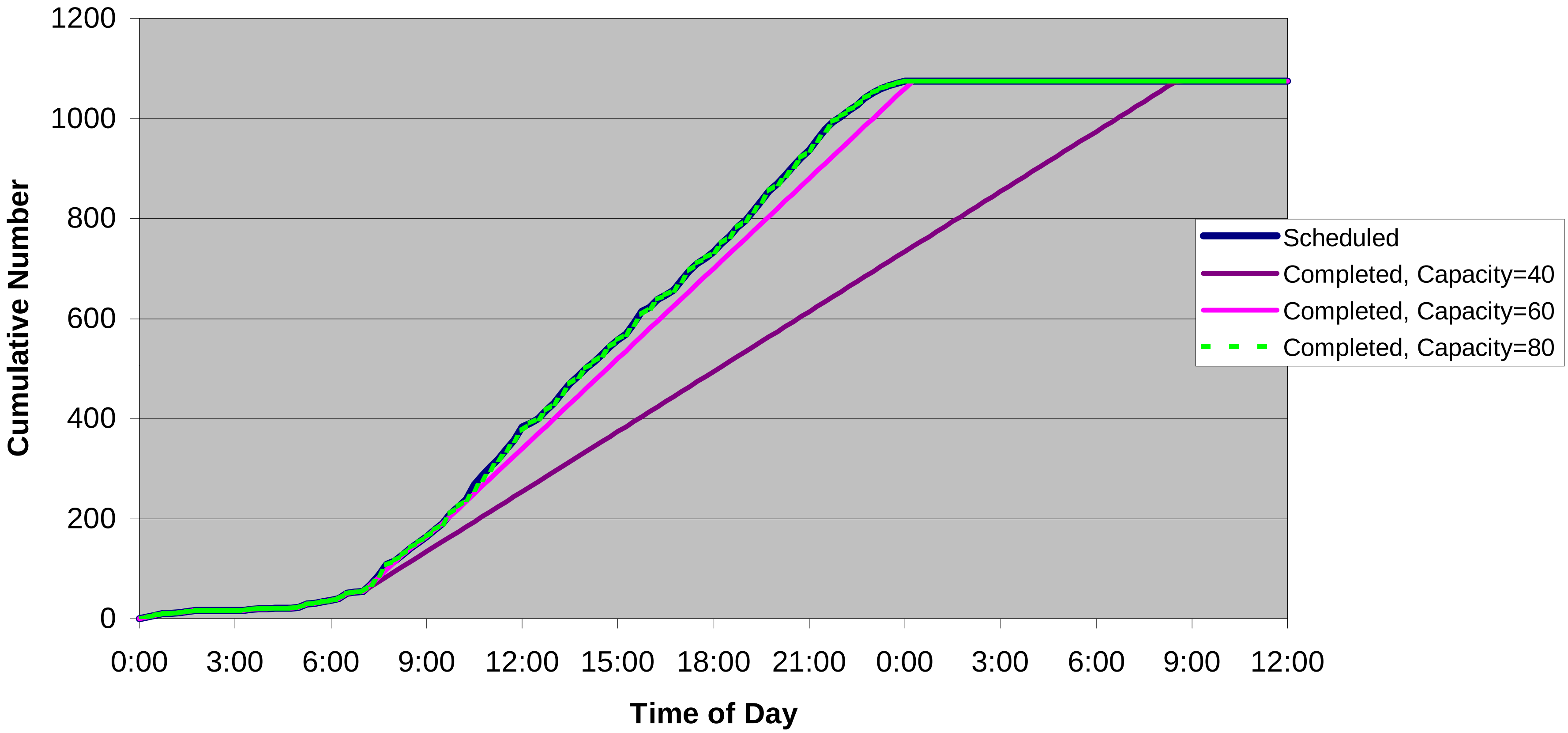
Table 4. Correlations Between the Rotated Factors and Original Weather Variables
(Promax method)

		FACTOR								
VARIABLE		1	2	3	4	5	6	7	8	9
Wind	Early am	-0.32	-0.06	-0.02	0.31	0.34	0.51	0.18	0.22	0.48
	Late am	-0.12	-0.06	-0.03	0.20	0.34	0.76	0.21	0.26	0.26
	Afternoon	0.14	0.00	0.11	0.17	-0.04	0.83	-0.04	-0.04	0.01
	Evening	-0.06	-0.12	-0.01	0.25	0.19	0.77	0.16	0.17	-0.08
Temperature	Early am	0.92	-0.21	-0.18	-0.06	-0.09	0.10	-0.14	-0.13	0.01
	Late am	0.98	-0.03	-0.06	-0.02	-0.23	0.01	-0.19	-0.23	-0.04
	Afternoon	0.95	0.08	0.07	0.08	-0.33	-0.08	-0.24	-0.29	-0.02
	Evening	0.96	-0.03	0.00	0.02	-0.27	-0.06	-0.21	-0.22	-0.04
Visual Operations	Early am	-0.13	0.87	0.40	0.41	-0.15	-0.08	0.08	0.09	-0.12
	Late am	0.01	0.87	0.48	0.41	-0.12	-0.01	0.06	-0.22	-0.01
	Afternoon	0.08	0.43	0.84	0.38	-0.10	0.02	-0.00	-0.36	0.06
	Evening	-0.12	0.49	0.85	0.40	-0.01	0.05	-0.03	0.14	-0.14
Visibility	Early am	0.02	0.55	0.26	0.85	0.11	0.20	0.08	0.02	-0.09
	Late am	0.01	0.53	0.37	0.88	0.11	0.21	0.12	-0.14	0.04
	Afternoon	0.07	0.32	0.56	0.86	0.11	0.20	0.04	-0.20	0.05
	Evening	-0.02	0.22	0.58	0.82	0.13	0.22	-0.03	0.05	-0.07
Precipitation	Early am	-0.15	-0.18	0.14	0.00	0.20	0.15	0.20	0.20	0.62
	Late am	-0.18	-0.12	-0.01	-0.02	0.24	0.16	0.10	0.78	0.14
	Afternoon	-0.14	-0.01	-0.16	-0.03	0.19	0.12	0.28	0.80	-0.03
	Evening	-0.11	0.06	-0.12	0.02	0.20	0.00	0.67	0.27	0.08
Low Ceiling	Early am	0.15	-0.87	-0.42	-0.36	0.05	0.06	-0.07	-0.11	0.13
	Late am	0.02	-0.86	-0.50	-0.36	-0.01	0.01	-0.04	0.21	0.01
	Afternoon	-0.04	-0.42	-0.85	-0.36	0.03	0.00	-0.02	0.32	-0.07
	Evening	0.14	-0.47	-0.84	-0.42	-0.07	-0.04	0.01	-0.16	0.12
Medium Ceiling	Early am	-0.20	0.02	0.05	0.15	0.74	0.28	0.14	0.20	0.01
	Late am	-0.14	-0.08	0.02	0.16	0.86	0.22	0.14	0.12	-0.02
	Afternoon	-0.27	-0.19	0.01	0.11	0.78	-0.00	0.31	0.27	0.06
	Evening	-0.22	-0.15	-0.05	0.18	0.69	0.02	0.36	0.27	0.14
High Ceiling	Early am	-0.15	0.11	-0.08	0.03	0.19	0.13	0.66	0.28	-0.01
	Late am	-0.14	0.03	0.06	0.05	0.16	0.11	0.71	0.03	0.15
	Afternoon	-0.20	-0.03	0.12	0.18	0.20	0.09	0.71	0.08	-0.26
	Evening	-0.20	-0.09	0.16	0.12	0.26	0.17	0.24	0.16	-0.53

Table 5. LAX Weather Factor Interpretations, Rotated Factors

FACTOR	Interpretation
1	Warm temperatures throughout day.
2	VFR operations and absence of low cloud ceiling in the morning.
3	VFR operations and absence of low cloud ceiling in the afternoon.
4	High visibility throughout day.
5	Medium cloud ceiling throughout day.
6	High winds throughout day.
7	High cloud ceiling throughout day; evening precipitation.
8	Precipitation in late morning and afternoon.
9	Precipitation in early morning.

Figure 9. Hypothetical Queuing Diagram for LAX

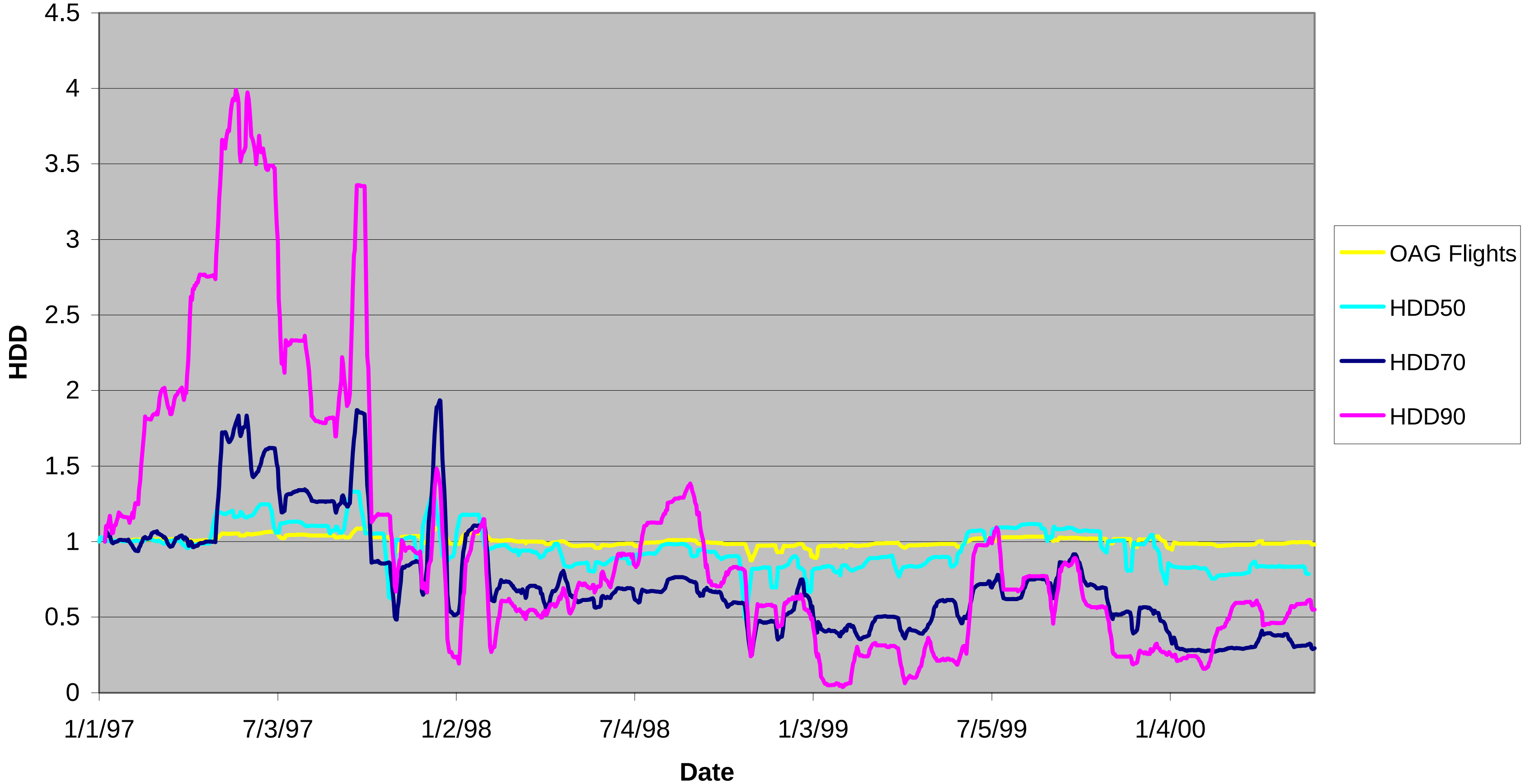


The other curves in Figure 9 are *hypothetical* cumulative curves for completed arrivals assuming capacity levels of 40, 60, and 80 arrivals per hour. The curves are constructed so that their slope never exceeds the associated capacity level, and so their value never exceeds the curve for scheduled arrivals. When a given completed arrival curve overlaps the scheduled curve, the airport is able to handle demand without delaying planes, and throughput is demand-limited. When a completed arrival curve deviates from the scheduled curve, it is because the capacity does not allow the airport to keep up with demand. Thus, in these situations, the throughput rate is capacity limited. One can see that if capacity is 80, the airport can almost always keep up with demand, with demand exceeding capacity slightly only for a brief time around 10 am. In the other cases, the demand exceeds capacity for most of the day. For a capacity of 60, there is a queue beginning around 10 am that does not clear until midnight, and when capacity is 40, the queue grows steadily throughout the day, and does not clear until well into the next day.

The area between the *scheduled* curve and a given *completed* curve represents the total time difference between when arrivals are scheduled to occur and when they can occur, given the capacity assumed in the completed curve. In short, this area measures cumulative delay from serving a given demand with a given capacity. When divided by the number of arrivals, this becomes an average delay. We refer to this average as the *hypothetical deterministic delay* (HDD), and employ it as a parameter characterizing the "stress" that a given scheduled arrival curve places on the airport. By using different assumed capacities, a family of such parameters can be constructed. Together, this family of parameters provides a thorough characterization of the demand placed on the airport by the arrival schedule. When capacities well below demand levels are used, the HDD increases linearly with the number of flights. HDD's for higher capacities capture the incidence, intensity, and duration of rush periods and are only weakly correlated with overall demand levels.

Figure 10 plots trends in values of three HDD parameters as well as total scheduled arrivals since January of 1997. The values plotted are seven day moving averages, normalized so that the average for the first week of January 1997 is 1.0. Total scheduled arrivals have changed relatively little—it is almost always within 10 percent of the

Figure 10. Trends in Values of HDD Parameters and Scheduled Arrivals since 1997.



baseline value. The HDD parameters vary more dramatically, with the magnitude of the fluctuations correlated with the assumed HDD capacity. For example, in the summer of 1997, the HDD50, HDD70, and HDD90 values were respectively 4, 1.7, and 1.2 times their January values. Since the summer of 1997, the HDD's have generally declined, with HDD70 and HDD90 less than half their baseline values for most of the period since the beginning of 1999. This implies that while the total scheduled operations at LAX have remained quite steady, the degree of peaking in the schedule has declined.

To help visualize the variation in scheduled demand implied by these trends, cumulative scheduled arrivals for three days - June 24, 1997, January 17, 1999, and June 29, 1999 - are plotted in Figure 11. The first date was a high demand day with HDD values and total scheduled flights well above baseline. The second was a low demand day with regard to both HDD and total flights. The third day had a fairly large number of flights, but a relatively smoothed out schedule so that HDD values at higher capacities are quite low. (Table 6 compares the various HDD values for the three days.) Comparing the first and second plot it is clear that the onset of significant flight activity occurs about 30 minutes earlier on the high demand day, and the late morning rush, late afternoon, and late evening rush periods on this day are far more intense. The cumulative difference in flights is about 20 percent, but because of the differences in rush hour intensities, HDD differences at higher capacity levels are much more pronounced. Comparing the June days in 1997 and 1999, we see that total flights are nearly identical until the late evening, when there is a considerably stronger arrival rush for the 1997 day. But even during the earlier parts of the day, there is considerably stronger peaking in 1997, particularly in the late morning, which again drives up the HDD values.

Figure 11. Cumulative Scheduled Arrivals

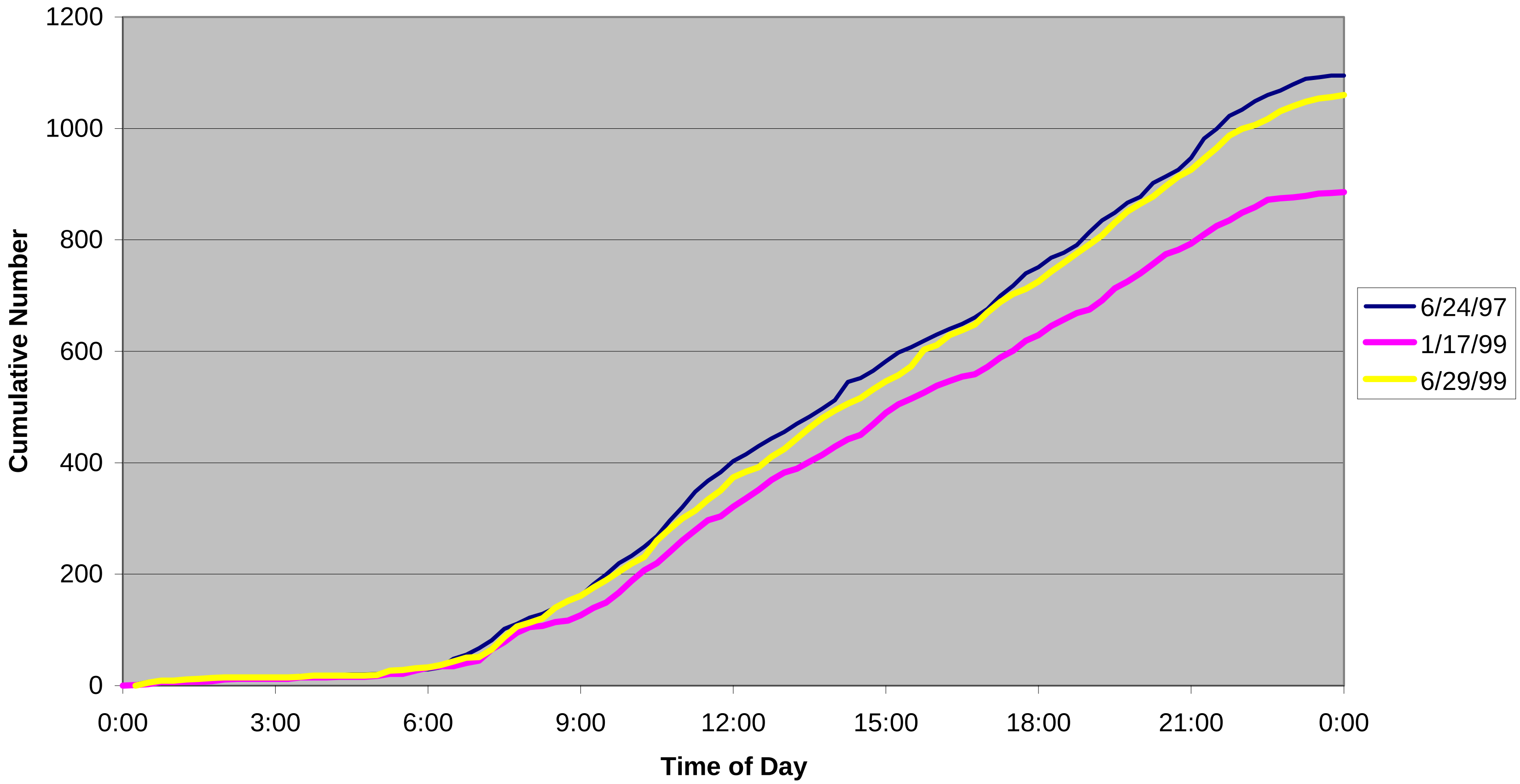


Table 6. Comparison of HDD Values for Three Representative Days

DAY	HDD50	HDD60	HDD70	HDD80	HDD90	HDD100	HDD110	HDD120
6/24/97	124.55	44.62	6.86	2.64	1.07	0.40	0.16	0.09
1/17/99	52.88	6.96	0.87	0.06	0.00	0.00	0.00	0.00
6/29/99	111.90	28.71	3.11	0.85	0.29	0.11	0.04	0.00

Because the capacity assumed in computing the HDD parameter can take on any value, it is natural to ask how many HDD parameters are required to satisfactorily capture daily variation in demand at LAX. As with weather, we used factor analysis to answer this question. In particular, we computed the HDD for capacity values ranging from 10 to 120, and then applied factor analysis to these 12 variables. The results are summarized in Table 7. Just two factors capture 93 percent of the variation in the 12 variables. Without rotation, one of the factors is positively correlated with all 12 variables. This can be interpreted as a measure of overall demand. The second factor is positively correlated with high capacity HDD's and negatively correlated with low capacity HDD's. This is known as a "contrast factor" and captures situations when high capacity HDD's are large relative to low capacity HDD's. When the factors are rotated (in this case, using the varimax procedure, which preserves orthogonality), the first factor is highly correlated with low capacity HDD's, and the second factor with high capacity HDD's.

Table 7. Initial and Rotated Demand Factors Based on HDD

	Initial Factors		Rotated Factors	
	FACTOR1	FACTOR2	FACTOR1	FACTOR2
HDD10	0.86	-0.46	0.95	0.22
HDD20	0.89	-0.42	0.95	0.27
HDD30	0.88	-0.46	0.96	0.24
HDD40	0.87	-0.46	0.96	0.23
HDD50	0.91	-0.37	0.93	0.32
HDD60	0.92	-0.20	0.83	0.45
HDD70	0.92	0.10	0.62	0.68
HDD80	0.89	0.34	0.45	0.84
HDD90	0.86	0.46	0.34	0.91
HDD100	0.82	0.53	0.27	0.94
HDD110	0.78	0.57	0.21	0.94
HDD120	0.70	0.59	0.14	0.91

6. Origin Airport Delay Normalization

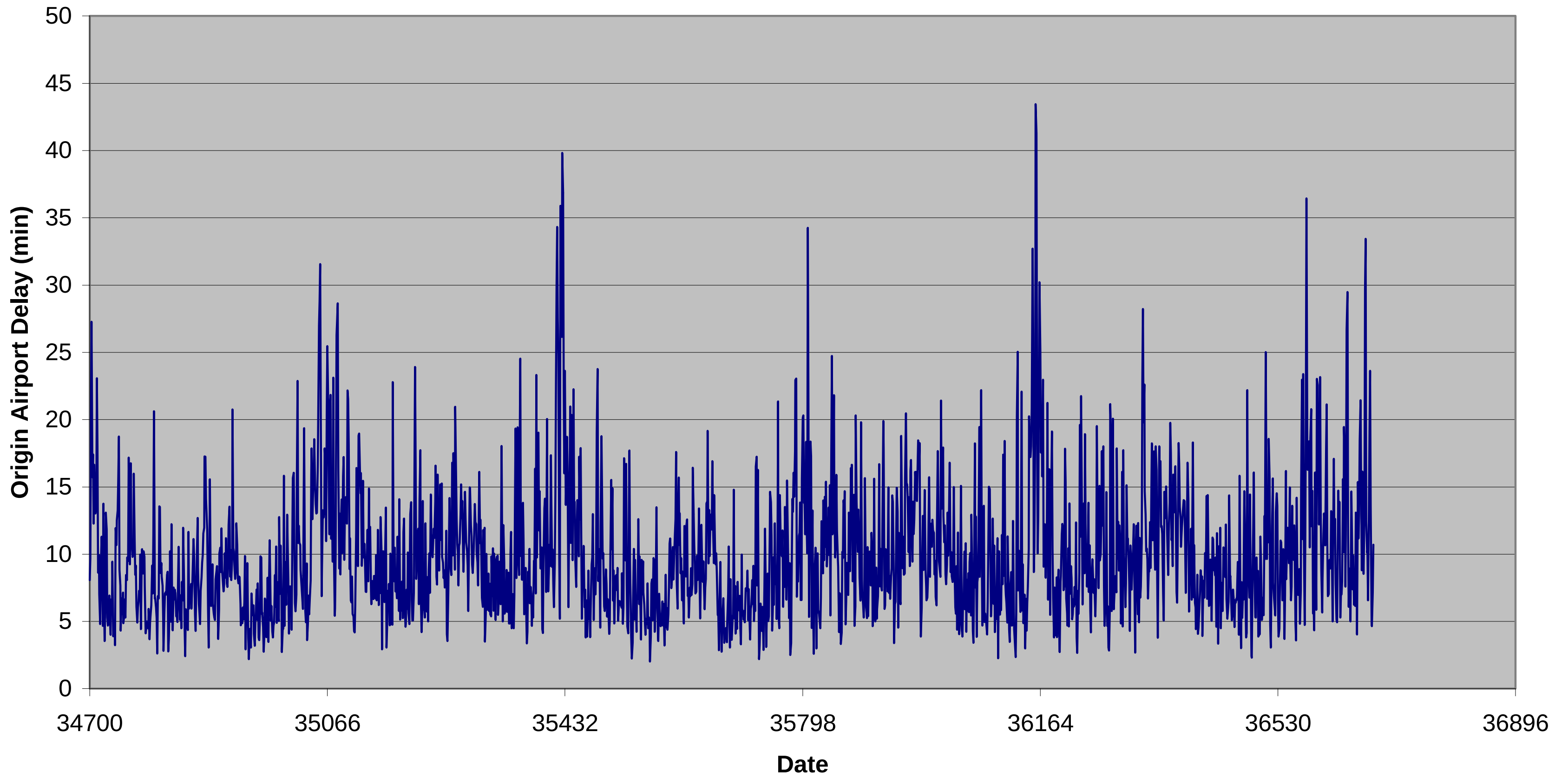
The third and final normalization variable is delay at the origins of flights bound for LAX. Obviously, if an origin airport has severe congestion that is delaying outbound flights, arrivals from that origin into LAX are likely to be effected, even when conditions at LAX are ideal. We seek a daily-level variable that accounts for this effect.

At first glance, the answer may seem obvious: since we compose our DFTI from flight level data that include the departure delay, why not simply subtract out the departure delay before computing the index? The reason this does not suffice is that some departure delays are the result of conditions at the destination rather than at the origin. This is obviously the case when a ground-hold program is put into effect. But there are a number of other mechanisms through which this linkage can occur as well. For example, the Southern California TRACON may issue a call for release (CFR) for an origin airport under which departure clearances for LAX-bound flights are issued only after consultation with the TRACON. It is also not unusual for dispatchers of individual airlines to hold flights based on their own assessments of conditions at the destination. In light of all this, it is clearly inappropriate to assume that LAX is “blameless” for departure delays at up-line points.

Short of a case-by-case review, there is no foolproof way to apportion the responsibility for any given departure delay so that the LAX contribution can be removed. However, for normalization purposes, we propose an origin departure delay index analogous to the DFTI. For every airport included in the DFTI average (see Table 1) we compute, on a daily basis, the average departure delay for all flights not bound for the LAX area (including LAX itself and all airports within 200 miles of it). Then, we compute a weighted average across all the DFTI airports, using the same weights used in the DFTI itself. In effect, this average measures what the departure delay of an average flight to LAX would be if that flight was actually bound for another destination.

Figure 12 plots the daily origin delay metric for LAX since beginning in January 1995. Under the best conditions, the metric is less than 5 minutes, and for the most part it remains under 15 minutes. On the worst days, it can spike as high as 30 or 40 minutes. These days are generally in the winter months, particularly January. As shown in Figure

Figure 12. Origin Airport Delay Time Series for LAX



13, which is a 30-day moving average of the same data, there is some evidence of an upward trend in metric beginning in January 1998. After this period, the average never dips below 7.5 minutes, a value undercut fairly often in the earlier period.

7. Regression Modeling of DFTI at LAX

As explained earlier, we used the weather and demand factors as explanatory variables to model the variation of the DFTI, and its components, for LAX. Equation (1) above is the general form of the model, but before estimation can proceed it is necessary to specify the model further. Our approach to specification is to begin with the linear form:

$$DFTI_t = \alpha + \sum_i \omega_i \cdot WX_{it} + \sum_j \partial_j \cdot DMD_{jt} + \theta \cdot ODEL_t + \varepsilon_t \quad (3)$$

where:

WX_{it} is the value of weather factor i on day t

DMD_{jt} is the value of demand factor j on day t

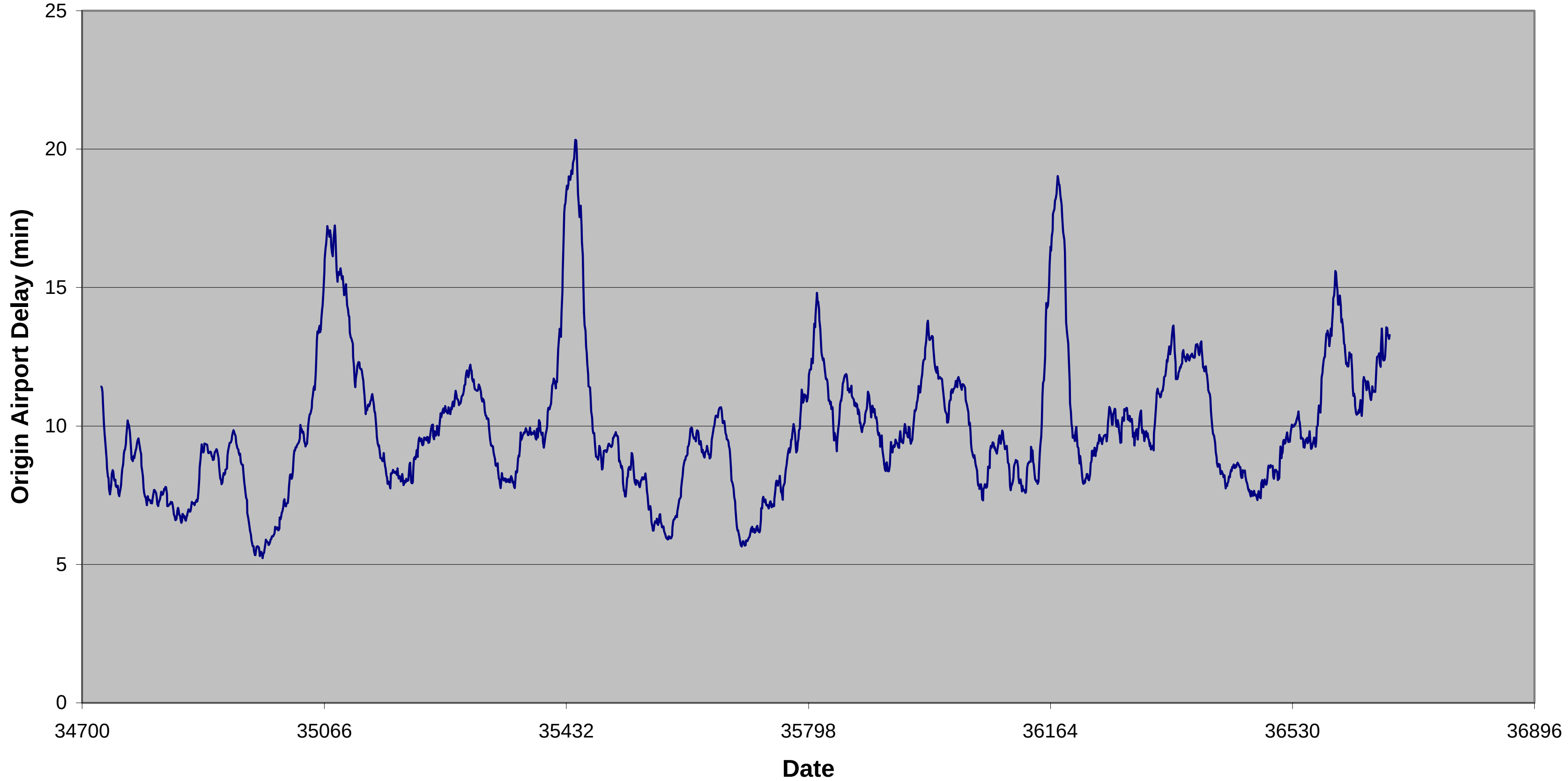
$ODEL_t$ is the measure of average departure delay for LAX origin airports

ε_t is a stochastic error term

The weather and demand factors are the rotated ones described earlier, and the average departure delay is the variable discussed in the previous section.

We adopt equation (3) as our baseline form because it is relatively simple and easy to estimate, and because estimation results are easy to interpret. As the baseline, it serves as the “hub” for other models that differ from it in various ways. We also use this specification to model the DFTI components time-at-origin (TOA), airborne time (ABT), and taxi-in time (TIT). Use of the linear form means that the coefficients in (3) will be sums of the corresponding coefficients in the components models. We subsequently try a number of alternative model specifications. These include a quadratic response surface model, a non-linear model, models in which weather is represented by more or fewer factors, and finally a non-parametric model

Figure 13. Origin Airport Delay 30-Day Moving Average for LAX



To complete specification of (3) it is necessary to make assumptions about the error term, ε_t. Early estimation results for LAX, as well as experience with similar models for other airports, reveal that the errors are heteroscedastic, with greater errors generally occurring on days when predicted DFTI values are high. The phenomenon is illustrated in Figure 14, where the predicted values for DFTI are those obtained from an ordinary least squares (OLS) regression of (3). OLS estimation on heteroscedastic data yields estimates that are unbiased, but inefficient. Moreover, the standard errors on the coefficients are biased, making it difficult to judge their statistical significance. To remedy this problem, we ran a feasible generalized least squares (FGLS) procedure. Under FGLS, we employ a prediction of squared residual as a weight on each observation. We obtained the prediction by regressing the absolute value of the OLS residuals against the explanatory variables included in (3). This estimation procedure yields results that are unbiased and asymptotically efficient.

The estimation data set includes daily observations from the beginning of 1997 - when CODAS became available - through May 2000. Table 8 contains the estimation results from the FGLS procedure. All but one coefficient estimate are significant at the 0.05 level, and all but two at the 0.001 level. The adjusted R^2 of 0.74 implies that the model explains about three quarters of the DFTI variation occurring in the data set. As just explained, the accuracy of the model predictions varies, but overall the standard error of a prediction is under 5 minutes.

We now consider the performance effects of weather, demand, and origin airport delay as they are revealed in this model. From the coefficient on *ODEL* we learn that an additional minute in the expected origin departure delay adds about 1.1 minutes to the expected DFTI at LAX. This coefficient is certainly of reasonable magnitude—if the expected origin delay increases by a certain amount, one would expect the DFTI to follow suit. It is of some interest that that estimate is slightly, but statistically significantly, above 1. This suggests either that there is some correlation between the conditions that generate high average departure delays and other factors that cause DFTI to increase, or that the origin delays themselves cause increased delays of other kinds. For example, it may be that

Figure 14. Observed and Predicted Values for DFTI

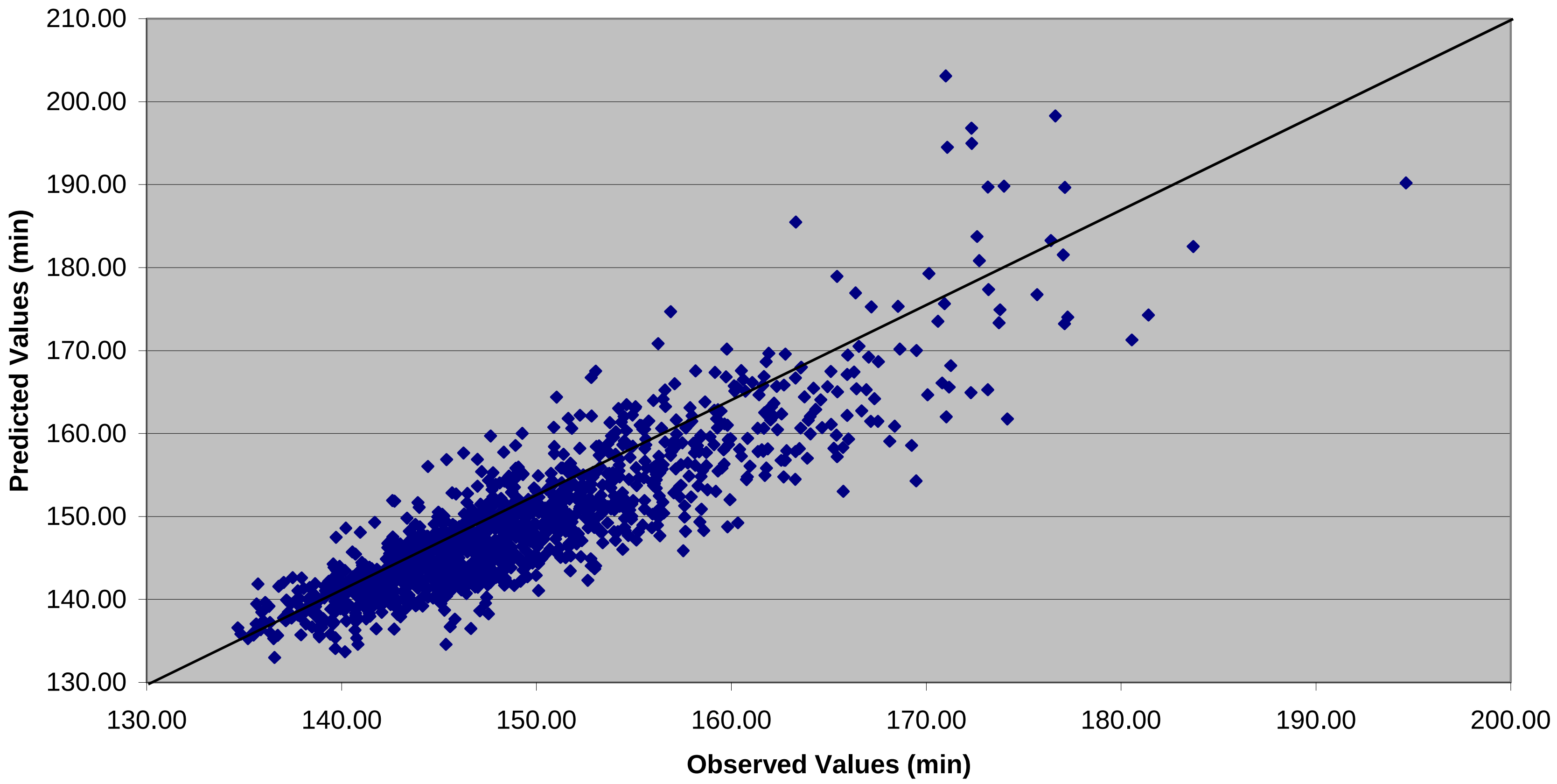


Table 8. Estimation Results From the FGLS Procedure, DFTI Values

Variable	Description	Estimate	T - statistic	P - value
INTERCEPT	Intercept	138.055	567.065	0.0001
ODEL	Origin airport departure delay	1.128	44.351	0.0001
WX_1	Warm daily temperatures	-1.357	-12.101	0.0001
WX_2	VFR ops, no low cloud ceiling in the morning	-0.988	-7.116	0.0001
WX_3	VFR ops, no low cloud ceiling in the afternoon	-1.123	-7.583	0.0001
WX_4	High visibility throughout day	-0.449	-3.575	0.0004
WX_5	Medium cloud ceiling throughout day	1.440	10.555	0.0001
WX_6	High winds throughout the day	0.512	4.531	0.0001
WX_7	High cloud ceiling throughout day	0.911	4.172	0.0001
WX_8	Precipitation in late morning and afternoon	1.871	8.324	0.0001
WX_9	Precipitation in early morning	-0.379	-2.614	0.0091
DMD_1	Peak demand	0.075	0.725	0.4685
DMD_2	Base demand	0.440	4.574	0.0001
ADJUSTED R^2			0.743	

departure delays disrupt normal demand patterns in the LAX terminal area and thus force additional metering and delays to LAX-bound flights.

All of the weather factors are statistically significant at the 0.01 level or better. Four of them, Factors 5, 6, 7, and 8, are positively related to DFTI. Since the factors are standardized (mean=0, standard deviation=1) variables, the coefficient on Factor 8 implies that a 1 standard deviation increase in this variable causes DFTI to increase about 1.9 minutes. Factor 8, midday precipitation, has the strongest effect. In addition to causing wet runways and lower deceleration rates, precipitation is probably an indicator of adverse wind and visibility conditions that are not fully captured by the other weather factors. Factor 5, medium cloud ceiling, also has a strong positive effect on DFTI, with a coefficient of 1.4. The influence of this factor probably derives from its impact on the ability to use the in-board runways for arrivals. Factors 7 and 6, high cloud ceiling and high winds, have weaker effects, with coefficients of 0.9 and 0.5 respectively. The latter result is certainly to be expected. High winds may force the airport to operate in east flow or cause controllers to keep larger separations in order to assure that minimums are not violated. It is less obvious why high cloud ceiling would have an impact. Generally, we may presume that the presence of such a ceiling is associated with some visibility condition, such as haze, that is not reflected in the other weather variables.

The other five weather factors are negatively associated with DFTI. Of these, the strongest impact belongs to Factor 1, temperature, for which an increase of 1 standard deviation causes a 1.4-minute DFTI reduction. The strength of this effect is somewhat puzzling. One possibility is that the high temperature is negatively correlated with fog. Another is that it is an indicator of clear weather conditions with the earth's surface in full sunlight. Finally, the impacts may derive from effects of temperature on aerodynamic performance. For example, cruise speeds are positive related to temperature because of its effect on Mach number. The other factors with negative coefficients include Factor 3, VFR operations and absence of low cloud ceiling in afternoon, Factor 2, VFR operations and absence of cloud ceiling in morning, Factor 4, high visibility throughout day, and Factor 9, precipitation in early morning. Explanations for the first three of these are obvious. As to the last, it is hard to imagine how early morning precipitation could have any direct, negative effect on flight times. The most likely explanation is that this effect occurs in the context of Factor 8, high daytime precipitation. If Factor 8 is high, then there is likely to be high precipitation in the period from 6 am to noon. If Factor 9 is also high, then this precipitation is more likely occurring during the early part of this period, when it has less impact on operations.

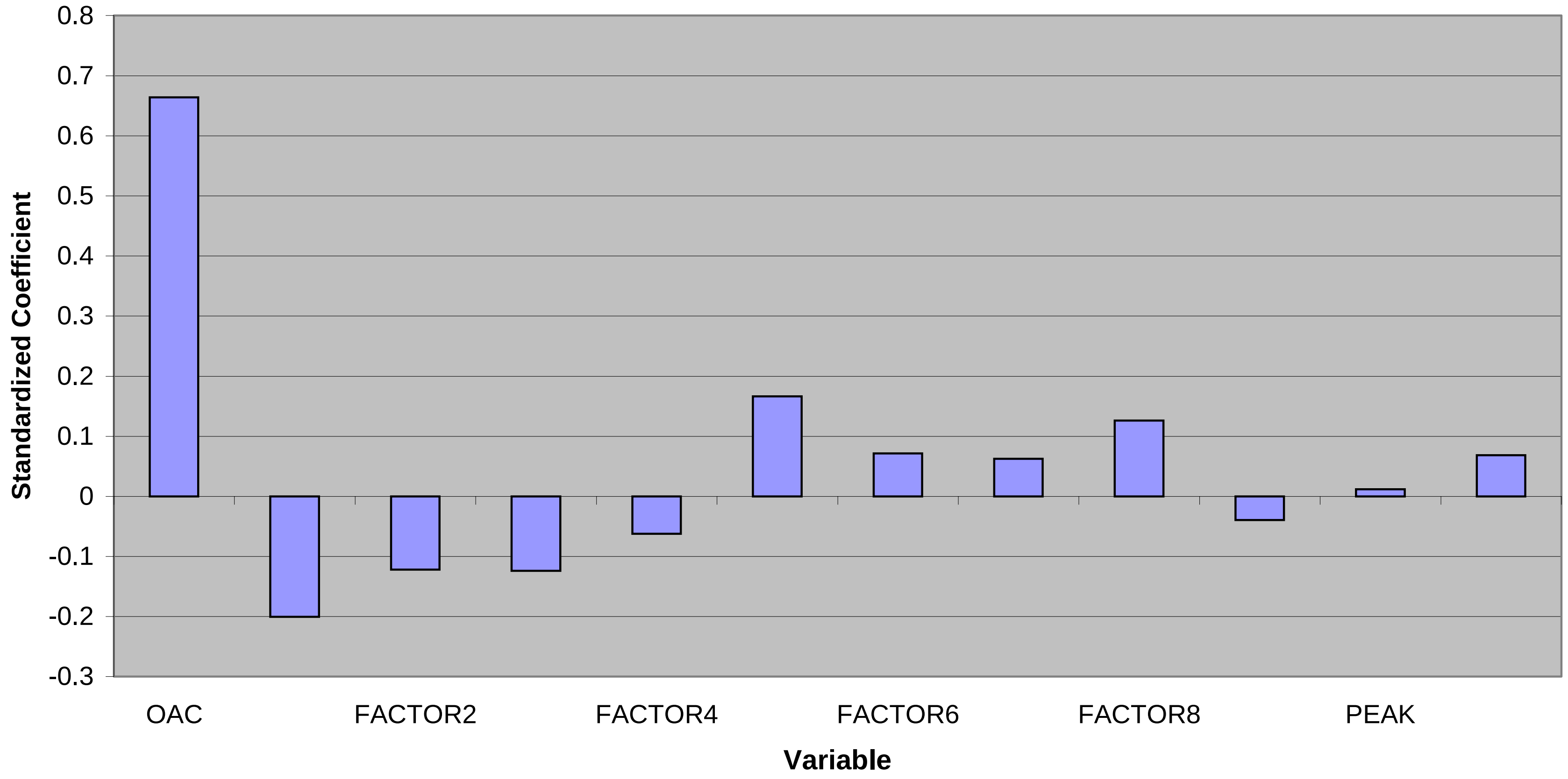
Estimates for the demand factors are both positive, but only the base demand factor (correlated with low capacity HDDs) is statistically significant. This implies that the DFTI is affected by the volume of operations throughout the day, not just during periods of peak demand. This is not surprising in light of the rather even temporal pattern of demand at LAX. The impacts of the demand factors will be further elucidated when estimation results for the DFTI component models are discussed, in the next section.

The above results reveal the sensitivity of the DFTI to the various explanatory variables. It is also interesting to compare the contributions of the explanators to DFTI variation. These contributions depend not just on sensitivity, but also on the degree of variation in the explanatory variable. To assess them we compute standardized regression coefficients, which relate the change in the dependent variable to the change in independent variables when both are measured in terms of standard deviations. That is, a standardized coefficient of 0.5 implies that a 1 standard deviation change in the independent variable leads to a 0.5 standard deviation change in the dependent variable. The standardized coefficients for the FGLS model are plotted in Figure 15. The standardized coefficient for origin airport delay is by far the largest, with a value over 0.6. The magnitudes of the weather factor coefficients vary between 0.04 and 0.2. The demand variables have been the least important sources of DFTI variation over the analysis period.

8. Regression Modeling of DFTI Components

As discussed earlier, the DFTI can be decomposed into four components: departure delay, taxi-out time, airborne time, and taxi-in time. It is therefore possible to similarly decompose the regression results obtained above. In other words, a regression coefficient in the DFTI model is the sum of regression coefficients for identically specified models

Figure 15. Standardized Coefficients for the FGLS Model



of the DFTI components. By estimating the component models, we obtain additional insight concerning how the various explanatory variables influence performance.

Table 9 summarizes the regression results for the DFTI components. The three components considered are time at origin (departure delay plus taxi-out time), airborne time, and taxi-in time. The models were estimated using the weights used in the total DFTI model. This may not provide the most efficient estimates, since the different components may have different patterns of heteroscedasticity, but it preserves the identity between the DFTI regression coefficients and the sums of the coefficients for the DFTI components.

Not surprisingly, the time at origin component is most strongly influenced by average origin departure delay. In addition, seven of the nine weather factors are significant at the 5 percent level - five of these at the 1 percent level. Factors 8 and 5, precipitation and medium cloud cover, are the most important positive correlates with time at origin. Factors 2 and 3, morning and afternoon VFR conditions and absence of low cloud cover, are the factors that decrease time at origin. It should be emphasized that these weather conditions pertain to LAX, not the origin airport. Presumably, air traffic management procedures create the linkage between LAX weather and time at origin, through ground holds, ground stops, and other actions. This is also the case for the demand variables, of which only the base demand is statistically significant.

All nine weather factors have statistically significant impacts on airborne time. Factors 8 and 5 again have the largest positive coefficients, with Factors 6 and 7, high cloud ceiling and high winds, close behind. High temperature has by far the largest negative coefficient. Since this factor does not affect time-at-origin, it appears that the mechanism involved is not one that is governed by air traffic management actions. One possibility is that surface temperatures at LAX correlate with the upper air temperatures, which in turn affect the speed of sound, and hence the airspeed equivalent of a given Mach number. Further research is required to determine whether this or some other mechanism is at work. The morning and afternoon visibility factors (2 and 3) and (again somewhat mysteriously) early morning precipitation are also associated with lower airborne times. Of the two demand factors, only the one associated with peak demand levels is

Table 9. Estimation Results From the FGLS Procedure, Values for DFTI Components

	Time-at-origin		Airborne time		Taxi-in time	
Variable	Estimate	P - value	Estimate	P - value	Estimate	P - value
INTERCEPT	14.588	0.0001	115.594	0.0001	7.874	0.0001
ODEL	1.099	0.0001	-0.012	0.4621	0.041	0.0001
WX_1	-0.065	0.4011	-1.474	0.0001	0.182	0.0001
WX_2	-0.722	0.0001	-0.233	0.0100	-0.033	0.2290
WX_3	-0.669	0.0001	-0.348	0.0003	-0.105	0.0003
WX_4	-0.201	0.0198	-0.186	0.0232	-0.062	0.0125
WX_5	0.599	0.0001	0.846	0.0001	-0.005	0.8567
WX_6	0.154	0.0480	0.428	0.0001	-0.069	0.0021
WX_7	0.372	0.0132	0.503	0.0004	0.036	0.3995
WX_8	0.897	0.0001	0.796	0.0001	0.179	0.0001
WX_9	-0.060	0.5485	-0.316	0.0008	-0.003	0.9158
DMD_1	0.034	0.6366	0.234	0.0005	-0.193	0.0001
DMD_2	0.260	0.0001	0.060	0.3367	0.120	0.0001
ADJUSTED R^2	0.804		0.427		0.213	

significant. This contrasts with the result for DFTI as a whole. As will be explained below, this is because the positive (and intuitively reasonable) effect of the peak demand on airborne time is offset by its negative impact on taxi-in time. Finally, and as expected, average origin departure delay is found to have essentially no effect on airborne time.

The third component, taxi-in time, does not vary much with the weather factors, although some effects are statistically significant. The temperature and precipitation factors (1 and 8) are both positively related to taxi-in time. While the latter effect is understandable, the former is more curious. One possible interpretation is that high temperatures are associated with high landing speeds (since aircraft stall speeds depend on air density), which in turn increase runway occupancy times. The good visibility factors (2 and 3) also are negatively related to taxi-in times, as is the high winds factor (6). A possible explanation of the latter is that with high winds, aircraft must land at lower ground speeds and thus exit the runways sooner. Among the demand factors, peak demand is negatively associated with taxi-in times while for base demand the association is positive. A plausible explanation for the former is that, when peak demands are high, there is more use of the in-board runways for arrivals. In-board arrivals have shorter taxi times because the distances to the terminal are shorter and no runway crossings are required. Finally, delay at the origin airport has a small but statistically significant effect on taxi-in time. While there is no obvious way to account for this result, it may be related to the reasons, discussed above, why the coefficient on origin delay in the overall DFTI model is slightly greater than 1. Whatever the uncertainties in explaining the results of the taxi-in time model are, the magnitudes of the coefficients are all rather small, so the phenomena at work exert little influence on overall flight times and arrival delays.

9. Alternative Model Specifications

The model presented above is the simplest and most straightforward version that incorporates the effects of the nine weather factors, two demand factors, and origin airport delay. By the same token, the baseline model is highly restrictive and embodies strong assumptions about the underlying performance relationships. Moreover, the choice of nine weather factors in the baseline model is somewhat arbitrary, suggesting that alternatives should be tried. In this section we investigate other model specifications that either have different mathematical forms, or include different numbers of weather factors

9.1 A Response Surface Model

This model extends the previous model by estimating a quadratic "response surface" for DFTI. The form of the model is:

$$DFTI_t = \alpha + \sum_i \beta_i X_{it} + \sum_i \gamma_i X_{it}^2 + \sum_i \sum_{j>i} \lambda_{ij} X_i X_j + \varepsilon_t \qquad (4)$$

where the *X* variables include all the independent variables in equation (3). One can see that the previous model is a restricted version of this one, with the γ_i and λ_{ij} forced to zero. By removing these restrictions, we can investigate non-linear effects on the individual explanatory variables as well as interaction effects between variables. The "price" of these capabilities is a drastically increased set of parameters, from 13 to 91.

The estimation procedure again involved FGLS in order to account for heteroscedasticity in the error term. Estimation results from the procedure appear in Table 10. Estimates for the first order terms are somewhat different than those from the baseline model, although the signs are maintained except for the peak demand factor, which is now negative and significant. Five of the 12 quadratic terms (whose estimates appear on the diagonal of Table 10) are significant at the 0.05 level (bold numbers). These include the weather factors for medium cloud ceiling, high winds, all-day precipitation, and early morning precipitation (Factors 5,6,8, and 9 respectively) and the peak demand factor. When interpreting the signs of these terms, it is important to remember that all variables are measured as deviations from their means so the effect of a quadratic term is the same whether the variable is a certain amount above or below average. For example, in the case of Factor 5, the positive linear coefficient combined with the negative quadratic coefficient implies that DFTI increases with this factor when it is at its mean value, but that the rate of increase diminishes as Factor 5 increases from below average to above average.

Of the 66 interaction terms, whose estimates appear on the diagonal in Table 10, 15 are significant at the 0.05 level. Two of the stronger interactions are those for Factor 8 (all-day precipitation) combined with Factor 5 (medium cloud ceiling leading to evening precipitation) and Factor 6 (high winds). Both of these interactions are positive, meaning that (for example) all-day precipitation combined with high winds has a greater impact on DFTI than either of these conditions separately. Presumably, these interactions capture the difference between a steady rain, or a clear blustery day, and a major storm.

The interactions involving the base demand variable are also of interest. They reveal that when there is high demand DFTI is more sensitive to visibility (as captured by Factors 2 and 3) and to a high cloud ceiling leading to evening precipitation (Factor 7). Since each of these weather conditions likely to affect airport capacity, it is not surprising that their impacts are strongest when demand is greatest.

Given the large number of insignificant quadratic and interaction terms, it is natural to ask whether the added predictive power of the response surface model can justify its greater complexity. To quantify the statistical improvement, one can compare the coefficient of determination (R^2) for the linear model, to one with both linear and quadratic terms, and to the full model that also includes the interaction terms. Adding the 12 quadratic terms causes the R^2 to increase by about 0.01 from 0.76 to 0.77, while the 66 interaction terms increase the R^2 to about 0.82. This may seem like a modest return. Statistically, however, it is significant. F-tests of the hypotheses that all the quadratic terms are zero and that all of the interaction terms are zero reveal that both must be rejected even at a 0.0001 level. Intuitively, it should not be surprising that the restrictions of the simple linear model are not born out by empirical observation.

Table 10. Estimation Results for the Response Surface Model

	First Order	Second Order FACTOR: ODEL	WX_1	WX_2	WX_3	WX_4	WX_5	WX_6	WX_7	WX_8	WX_9	DMD_1	DMD_2
ODEL	**1.161**	-0.006	-0.001	0.021	0.013	0.020	**0.104**	0.017	-0.032	**0.243**	-0.003	**-0.078**	0.003
WX_1	**-1.639**		-0.186	**0.442**	0.309	**0.439**	0.016	-0.278	0.287	-0.241	0.103	-0.285	0.210
WX_2	**-1.250**			-0.199	0.078	0.244	-0.055	-0.106	**-0.570**	-0.103	**-0.488**	0.044	**-0.468**
WX_3	**-1.005**				0.140	-0.056	0.255	-0.264	-0.296	-0.172	-0.128	0.278	**-0.340**
WX_4	-0.061					0.031	-0.036	**0.402**	0.187	-0.517	0.049	-0.176	-0.085
WX_5	**2.464**						**-0.432**	-0.125	0.016	**0.717**	0.212	**0.321**	0.221
WX_6	**0.594**							**-0.160**	-0.040	**1.050**	0.140	-0.060	-0.066
WX_7	**0.886**								0.057	-0.161	**0.524**	-0.455	**0.591**
WX_8	**1.011**									**-0.346**	-0.027	-0.062	-0.052
WX_9	**-0.987**										**-0.131**	0.221	-0.185
DMD_1	**-0.762**											**0.293**	-0.077
DMD_2	**0.782**												-0.022
ADJUSTED R^2		0.820											

Bold numbers are significant at 0.05 percent level.

9.2 A Non-linear Model

Both the baseline and the response surface models are linear in parameters. Although (as the latter model illustrates) such models can be used to investigate non-linear effects, it is also useful to try functional forms that require non-linear estimation. An example of such a form is:

$$DFTI_t = \alpha + \beta \cdot ODEL_t + \lambda \exp(\sum_i \gamma_i X_{it}) + \varepsilon_t \qquad (5)$$

In this equation $\alpha + \beta \cdot ODEL_t$ is the minimum value for DFTI given the conditions at the origin airport. The exponential term is additional flight time resulting from delays. Here, we assume that this extra time is an exponential function of the various independent variables (including *ODEL*) specified in equation (3). This expression is non-linear in the $\gamma' s$, and therefore cannot be estimated using linear regression. Rather, non-linear least squares must be used.

Initial estimation results revealed that this model, like the previous ones, is subject to heteroscedasticity in the error terms. The approach taken for overcoming this problem is similar to that for the previous models. We estimated the model first using non-linear least squares, then regressed the absolute values of the residuals against the independent variables, and finally estimated a transformed version of the original model in which both sides are divided by the expected error.

Estimation results appear in Table 11. The minimum DFTI in the absence of delay at the origin airport and ideal conditions at LAX is about 128 minutes. The estimated value of β is approximately 1. All of the weather factors as well as the base demand factor are statistically significant, and their signs are consistent with those obtained from the baseline model. Their relative importance, however, is somewhat different. Factor 1, temperature, has the largest negative coefficient, while Factor 5, medium cloud ceiling, has the largest positive one.

Again, a natural question is whether the non-linear model performs better than the baseline linear model. Comparing the non-linear model with a linear one having the same explanatory variables, we find that the former has slightly greater explanatory power,

Table 11. Estimation Results for Non-Linear Model

	Parameter	Description	Estimate	Asymptotic Standard Error
Linear variables	α	Intercept	**128.086**	1.824
	β	Origin delay (ODEL) coefficient	**0.958**	0.006
Exponential variables	λ	Exponential coefficient	**9.323**	1.944
	γ_1	Weather factor 1 coefficient	**-0.146**	0.034
	γ_2	Weather factor 2 coefficient	**-0.099**	0.023
	γ_3	Weather factor 3 coefficient	**-0.094**	0.020
	γ_4	Weather factor 4 coefficient	**-0.048**	0.014
	γ_5	Weather factor 5 coefficient	**0.118**	0.024
	γ_6	Weather factor 6 coefficient	**0.056**	0.015
	γ_7	Weather factor 7 coefficient	**0.054**	0.014
	γ_8	Weather factor 8 coefficient	**0.062**	0.008
	γ_9	Weather factor 9 coefficient	**-0.020**	0.009
	γ_{10}	Base demand factor coefficient	**0.058**	0.016
	γ_{11}	ODEL coefficient	**0.016**	0.093

even when we take into account the fact that it has two more coefficients. The specification of the non-linear model is also somewhat more intuitive, since it features a base flight time component combined with a delay component that is always positive. These advantages must be weighed against the greater challenges of non-linear estimation, and the less desirable statistical properties of non-linear estimators.

9.3 Models With Different Numbers of Weather Factors

The baseline model contains nine weather factors, as explained above. This choice for the number of factors was based on the principal that each factor should explain as much variation in the 32 weather variables as a single one of these variables would if the variables were mutually orthogonal. While this is a widely used criterion, it has no real theoretical justification. It is therefore interesting to investigate how changing the number of factors affects the performance and estimation results of the baseline model.

To explore this, we repeated the factor extraction, factor rotation, and model estimation procedures employed for the baseline model, except that we specified a priori the number of factors that would be extracted. Models with 3, 6 and 12 weather factors were estimated. As shown in Figure 5, they respectively account for 46, 62, and 81 percent of the variation in the original 32-variable data set. Interpretations for the factors, after an oblique promax rotation, are shown in Table 12.

Regression results for the 3, 6, and 12-factor DFTI models appear in Table 13. There are two main findings, which at first seem contradictory. First, the performance of the models is roughly the same. While increasing the number of weather factors (beyond three) certainly provides a more complete depiction of the weather at LAX on a particular day, this increased fidelity does not contribute to our ability to explain day-to-day variation in DFTI. Second, the factors are almost all highly significant, even in the case of the 12-factor model. In other words, when we use a highly detailed depiction of weather each “detail” (with one exception in the 12-factor model) has a significant effect on DFTI.

How can these findings be reconciled? The key point is that the factors in the 3-factor model are, to a good approximation, aggregates of the factors of the other models. Insofar as these factors being aggregated have about the same effect on DFTI, there is no loss in

Table 12. Interpretation of the Factors, After an Oblique Promax Rotation

FACTOR	Three Factor Model	Six Factor Model	Twelve Factor Model
1	High visibility and absence of low cloud ceiling.	High visibility and absence of low cloud ceiling.	High temperature.
2	High winds, precipitation, and medium/high cloud ceiling.	High temperature.	Morning visibility.
3	High temperature.	Medium cloud ceiling.	All-day visibility.
4		High winds and visibility.	High winds.
5		High cloud ceiling in day with evening precipitation.	Low cloud ceiling and visibility.
6		Daytime precipitation.	High cloud ceiling in day with evening precipitation.
7			Absence of low cloud cover and evening visibility.
8			High morning winds and medium cloud ceiling in afternoon and evening.
9			Medium morning cloud ceiling.
10			Daytime precipitation.
11			Early morning precipitation and high winds.
12			High cloud ceiling late in day.

model fit from the aggregation. Inspection of the estimation results for the 6- and 12-factor models in Table 13 reveals coefficient estimates for many of the factors are indeed approximately equal.

Examining the estimation results for the 3-factor model, we find that the estimates on origin airport delay, peak demand, and base demand are much the same as in the other models. Of the weather factors, the coefficients on the high visibility and high temperature factors are negative, while that on the winds/precipitation factor is strongly positive. All three estimates have low standard errors and are highly significant.

9.4 Models With Non-Parametric Weather Effects

All of the models treated so far have assumed that weather effects can be decomposed into the effects of a few continuous variables derived from various meteorological parameters. An alternative approach is to posit that there are a few discrete categories of days, and that DFTI relationships should be developed separately for each day category. Formally, this implies that the DFTI should be modeled as $DFTI_t = f_{d(t)}(\vec{X}_t)$ where *d(t)* is a discrete-valued function indicating the category to which day *t* belongs. Furthermore, we assume that $d(t) = g(\vec{X}_t)$, i.e. that the day category is a function of the *X* variables.

Given the last assumption, the models presented here, explain performance in terms of the same weather, demand, and origin airport delay variables employed in the earlier models. Econometrically, however, the forms presented here are more flexible because they permit variation in DFTI to be studied without imposing an analytic function on the data. To achieve this flexibility, however, one must lump days with similar, but nonetheless somewhat different, conditions and treat them as though they come from a single homogeneous category.

The first order of business in order to develop these types of models is to identify day categories. We employed cluster analysis for this purpose. Cluster analysis treats a set of variables as a set of coordinates in an n-dimensional space, and then finds sets of observations that are nearby each other in this space. Numerous clustering algorithms have been developed. We chose Ward's minimum variance method. In this method, every observation is initially in its own cluster. In each iteration, a pair of clusters is

Table 13. Regression Results for 3,6, and 12-factor DFTI Models

	12-Factor Model		6-Factor Model		3-Factor Model	
Variable	Estimate	P - value	Estimate	P - value	Estimate	P - value
INTERCEPT	137.999	0.0001	138.045	0.0001	138.167	0.0001
ODEL	1.131	0.0001	1.127	0.0001	1.108	0.0001
DMD_1	0.159	0.0762	0.117	0.2023	0.139	0.0854
DMD_2	0.446	0.0001	0.429	0.0001	0.429	0.0001
WX_1	-1.390	0.0001	-1.909	0.0001	-1.960	0.0001
WX_2	-0.967	0.0001	-1.310	0.0001	2.843	0.0001
WX_3	-0.543	0.0001	1.853	0.0001	-1.363	0.0001
WX_4	0.621	0.0001	0.678	0.0001		
WX_5	1.202	0.0001	1.408	0.0001		
WX_6	0.860	0.0001	1.527	0.0001		
WX_7	-0.004	0.9756				
WX_8	1.350	0.0001				
WX_9	0.841	0.0001				
WX_{10}	1.344	0.0001				
WX_{11}	-0.355	0.0469				
WX_{12}	0.279	0.0249				
ADJUSTED R^2	0.749		0.742		0.753	

merged into one. The pair chosen is the one that would result in the least increase in within-cluster variance. The iteration continues until all observations have been joined into one cluster. By inspecting the changes the within cluster variance after each iteration, it is often possible to identify natural clusters for which within-cluster variance is small but will increase significantly if additional clustering is carried out.

As coordinates, we used the weather factors based on both the 9-factor representation that is employed in the baseline model and the 3-factor representation that was discussed in the previous section. In addition to capturing most of the variation in the original weather data, these factors are standardized and thus form a coordinate system that is independent of the units in which the original data were measured. Demand and origin airport delay variables are not considered in the clustering process, but rather retained continuous variables.

Figure 16 illustrates the clusters obtained from the 3-factor weather representation. Nine clusters were identified; each is plotted in Figure 16 with a different symbol. For example, the squares in Figure 16 are days with below-average temperatures and visibility, and somewhat below-average storm activity. The pyramids are days with average temperatures, visibility somewhat above average, and somewhat above-average storm activity. Table 14 summarizes the mean factor scores of each cluster.

Two models were estimated using the clusters. In the first, and simpler, model, has the form:

$$DFTI_t = \alpha_0 + \alpha_{d(t)} + \sum_i \beta_i X_{it} + \varepsilon_i$$

Where *d(t)* is the cluster to which day t belongs. In this model, the effect of the cluster is on the intercept term only. In the second model, the slope coefficients are also allowed to vary according to the cluster. That is:

$$DFTI_t = \alpha_0 + \alpha_{d(t)} + \sum_i (\beta_i + \beta_{id(t)}) X_{it} + \varepsilon_i$$

Figure 16. Illustration of Cluster Analysis for LAX

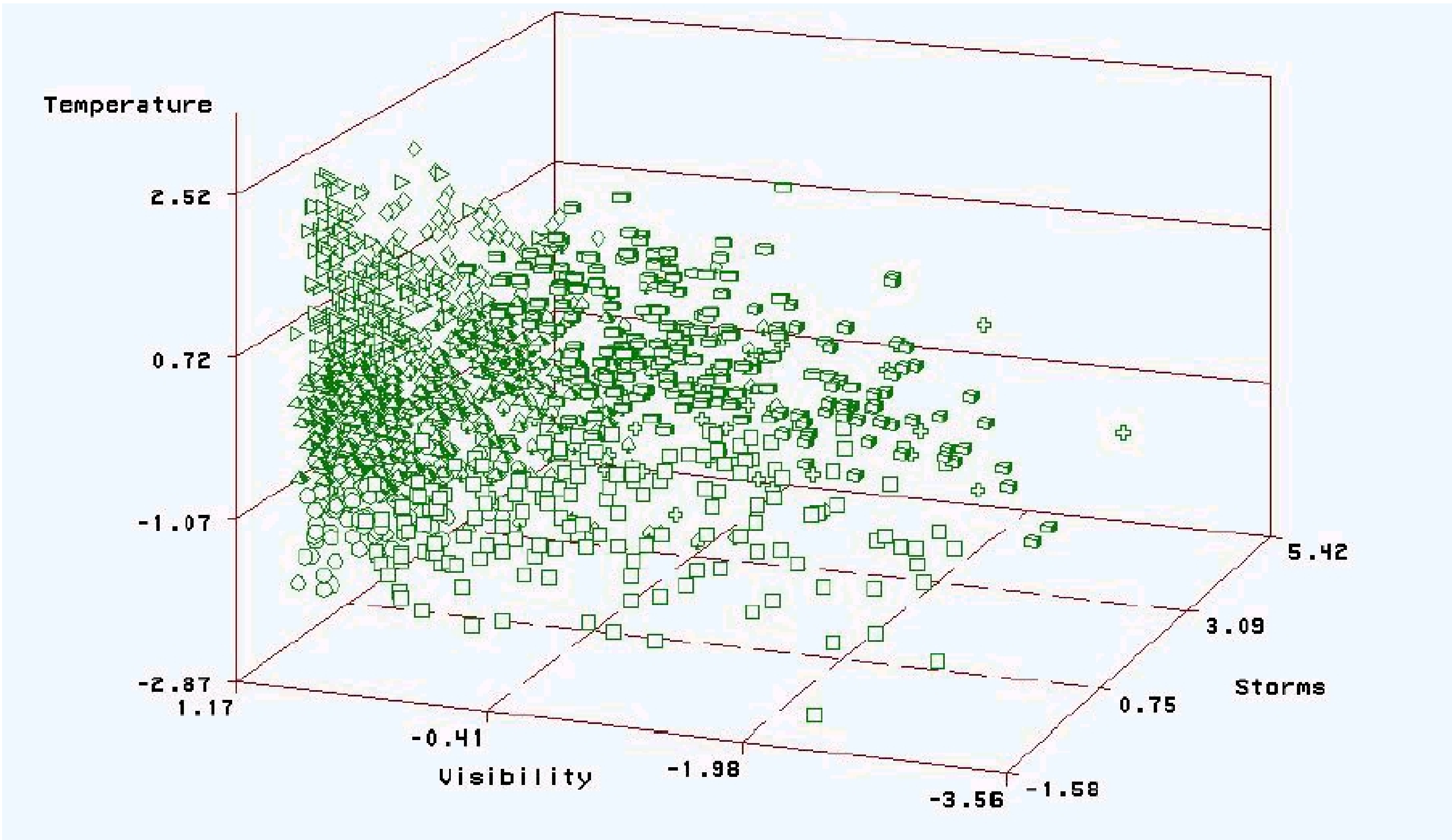


Table 14. Mean Factor Scores of Clusters for 3-factor Weather Representation

CLUSTER	NUMBER OF DAYS	FACTOR 1 Visibility	FACTOR 2 Storm	FACTOR 3 Temperature
1	176	0.762	-0.245	-0.963
2	235	0.374	-0.467	0.027
3	182	0.697	-0.521	1.189
4	174	-0.874	-0.642	-0.980
5	171	0.416	1.417	-0.540
6	62	0.604	1.080	1.170
7	158	-0.961	-0.148	0.956
8	52	-2.486	-0.426	0.475
9	31	-0.789	3.086	-1.250

Estimation again employed a FGLS procedure involving an initial estimation using OLS, a second step in which the OLS residual absolute values are regressed, and a final step in which observations are weighted according to the predicted residual.

Estimation results for the 3-factor model appear in Table 15. Most of the cluster terms are highly significant. Days in Cluster 9 have, by far, the largest DFTI values, some 23 minutes greater than Cluster 1, which serves as the reference in this analysis. From Table 14 we see that cluster 9 features high storm activity and low visibility. The only cluster with DFTI's significantly lower than Cluster 1 is Cluster 3, days with generally higher temperatures than those in Cluster 1. The non-cluster variables, base demand and origin airport delay, are highly significant as in the previous models. The R^2 of 0.70 is less than what was obtained in the models presented in previous sections.

When the second version of the model is used, significant interactions between the cluster and the origin airport delay and demand variables are found. For example, the effect of origin airport delay appears to be less for days in Cluster 7 (high temperature, low visibility). In contrast, DFTI appears to be more sensitive to demand in Cluster 8 days, which are characterized by very low visibility. Clusters 7 and 9, which also feature below-average visibility, exhibit a similar tendency. While clearly significant, the presence of these interactions does not greatly improve the fit of the model, the R^2 of 0.71 is only slightly higher than that obtained for the model without interactions.

The same procedure was employed using the 9-factor characterization of weather. In this case, the cluster analysis identified 12 groupings, mean factor values of which are summarized in Table 16. FGLS estimation results for using the simple fixed effect cluster model appear in Table 17. All the cluster fixed effects, which are again estimated using Cluster 1 as a baseline, are highly significant. The clusters with the highest fixed effects are 9, 10, 11, and 12. What these clusters have in common is a negative factor score on Factor 1, which corresponds to below-average temperatures. The only cluster effect that is negative is that for cluster 3, which is characterized by above average temperatures. This model has an adjusted R^2 of 0.72, somewhat higher than the previous cluster models but still below what was obtained in the earlier models.

Table 15. Estimation Results for the 3-factor Model

	MODEL 1		MODEL 2	
VARIABLE	Estimate	P-value	Estimate	P-value
INTERCEPT	136.375	0.0001	136.516	0.0001
CLUSTER 2	-0.234	0.5702	-0.416	0.6315
CLUSTER 3	-2.954	0.0001	-3.199	0.0002
CLUSTER 4	3.286	0.0001	4.186	0.0001
CLUSTER 5	5.066	0.0001	2.745	0.0261
CLUSTER 6	1.595	0.0051	0.803	0.5086
CLUSTER 7	-0.012	0.9800	2.210	0.0449
CLUSTER 8	4.455	0.0001	3.366	0.1434
CLUSTER 9	22.923	0.0001	14.099	0.0159
ODEL	1.151	0.0001	1.144	0.0001
ODELCLUSTER 2			0.014	0.8541
ODELCLUSTER 3			0.026	0.7405
ODELCLUSTER 4			-0.086	0.3169
ODELCLUSTER 5			0.214	0.0315
ODELCLUSTER 6			0.080	0.4886
ODELCLUSTER 7			-0.266	0.0079
ODELCLUSTER 8			0.047	0.8234
ODELCLUSTER 9			0.717	0.1125
DMD_2	0.453	0.0004	-0.0571	0.0493
DMD_2CLUSTER 2			1.155	0.0048
DMD_2CLUSTER 3			0.853	0.0370
DMD_2CLUSTER 4			1.331	0.0070
DMD_2CLUSTER 5			1.620	0.0041
DMD_2CLUSTER 6			0.883	0.0413
DMD_2CLUSTER 7			2.428	0.0001
DMD_2CLUSTER 8			4.320	0.0002
DMD_2CLUSTER 9			2.429	0.2888
ADJUSTED R^2	0.700		0.708	

Table 16. Mean Factor Scores of Clusters for 9-factor Weather Representation

CLUSTER	NUMBER OF DAYS	FACTOR 1 Warm daily temperatures	FACTOR 2 VFR ops, no low cloud ceiling in the morning	FACTOR 3 VFR ops, no low cloud ceiling in the afternoon	FACTOR 4 High visibility throughout day
1	183	-0.635	0.787	0.379	0.551
2	106	-0.704	0.655	0.423	0.902
3	249	1.028	0.609	0.492	0.354
4	137	-0.244	0.304	-0.015	-1.138
5	195	0.560	-1.222	-0.215	-0.445
6	153	-0.277	-0.039	0.038	0.288
7	100	0.080	-1.411	-2.394	-1.391
8	63	-0.650	-0.136	0.776	0.786
9	26	-1.188	-0.042	-0.977	-0.094
10	14	-1.184	-0.786	1.235	0.345
11	11	-1.120	0.544	0.202	0.270
12	4	-1.442	-1.646	-0.497	-0.481
	FACTOR 5 Medium cloud ceiling throughout day	FACTOR 6 High winds throughout the day	FACTOR 7 High cloud ceiling throughout day	FACTOR 8 Precipitation in late morning and afternoon	FACTOR 9 Precipitation in early morning
	-0.531	-0.837	-0.232	-0.125	0.117
	-0.298	1.544	-0.066	-0.076	0.260
	-0.550	-0.085	-0.301	-0.136	-0.114
	-0.477	-0.637	-0.241	-0.172	-0.011
	-0.095	0.062	-0.101	-0.346	0.191
	1.646	0.242	-0.105	-0.014	0.030
	-0.429	-0.307	-0.230	-0.019	-0.042
	0.787	0.484	1.120	0.085	-2.077
	1.842	0.760	1.495	3.994	0.141
	1.798	1.582	1.784	0.255	4.787
	0.981	-0.014	7.082	0.277	-0.413
	1.819	1.853	0.956	11.028	2.011

Table 17. Estimation Results for the 9-factor Model

VARIABLE	MODEL 1 Estimate	MODEL 1 P-value	MODEL 2 Estimate	MODEL 2 P-value
INTERCEPT	135.331	0.0001	136.095	0.0001
CLUSTER 2	1.582	0.0009	1.844	0.0793
CLUSTER 3	-1.199	0.0008	-2.509	0.0004
CLUSTER 4	1.608	0.0005	2.381	0.0086
CLUSTER 5	1.825	0.0001	1.428	0.1811
CLUSTER 6	5.620	0.0001	2.299	0.0454
CLUSTER 7	6.154	0.0001	4.667	0.0025
CLUSTER 8	5.031	0.0001	1.172	0.4070
CLUSTER 9	21.095	0.0001	3.713	0.5637
CLUSTER 10	10.058	0.0002	7.056	0.2785
CLUSTER 11	11.789	0.0001	2.817	0.7760
CLUSTER 12	28.012	0.0001	22.166	0.0583
ODEL	1.153	0.0001	1.075	0.0001
ODEL CLUSTER 2			-0.016	0.8555
ODEL CLUSTER 3			0.140	0.0310
ODEL CLUSTER 4			-0.088	0.2895
ODEL CLUSTER 5			0.028	0.7692
ODEL CLUSTER 6			0.324	0.0007
ODEL CLUSTER 7			0.126	0.3368
ODEL CLUSTER 8			0.360	0.0022
ODEL CLUSTER 9			1.523	0.0036
ODEL CLUSTER 10			0.358	0.4665
ODEL CLUSTER 11			0.756	0.3734
ODEL CLUSTER 12			0.589	0.6014
DMD_2	0.230	0.0586	-0.215	0.3783
DMD_2CLUSTER 2			0.389	0.3199
DMD_2CLUSTER 3			0.404	0.2327
DMD_2CLUSTER 4			0.508	0.1926
DMD_2CLUSTER 5			1.328	0.0086
DMD_2CLUSTER 6			0.500	0.4161
DMD_2CLUSTER 7			2.449	0.0039
DMD_2CLUSTER 8			0.657	0.2800
DMD_2CLUSTER 9			-1.180	0.6256
DMD_2CLUSTER 10			4.949	0.1369
DMD_2CLUSTER 11			9.887	0.0129
DMD_2CLUSTER 12			-6.535	0.2597
ADJUSTED R^2	0.721		0.732	

When the cluster model is expanded to include interactions between the cluster effects, origin airport delay, and demand, most of the interaction effects are statistically insignificant. The three major exceptions on the origin delay side are clusters 6, 8, and 9. DFTI's for days in these clusters are all more sensitive to origin airport delay than is the case for cluster 1. The major commonality of these factors is a score for Factor 2, afternoon visibility, which is fairly close to the mean, and well below the Factor 2 score for cluster 1, days in which have unusually good afternoon visibility. Significant interactions occur for clusters 5, 7, and 11. The latter cluster, in particular, appears to feature an unusually strong impact of demand on DFTI. Its major distinguishing characteristic is a very high incidence of early morning precipitation (Factor 9). As with the other non-parametric models, however, the fit of this one is somewhat less than what was obtained from the earlier models. These results suggest that loss in information from lumping days together in discrete categories outweighs the greater flexibility of non-parametric models, at least for the categorization method employed here.

10. Outliers

The results presented above suggest that 70-80% of the day-to-day variation in DFTI at LAX can be explained using relatively simple multivariate models that take into account the effects of weather, demand, and delay at origin airports. What about the remaining 20-30%? To identify some of the factors contained in this residual variation, we identified a set of "outlier" days for which model predictions (using the baseline linear model) substantially overestimated or underestimated the DFTI. For these days, we reviewed Traffic Management Unit (TMU) logs provided to us by the Southern California TRACON. These logs identify key events affecting traffic conditions in the TRACON, and actions taken by TMU, often in conjunction with other ATC facilities, in response. They also provide hourly counts of scheduled and actual arrivals at LAX. In addition to reviewing the written records, we visited the TRACON and spoke with controllers, traffic managers, and supervisors about the factors that affect delays at LAX.

At the LAX airport, main arriving traffic flow direction is to the west. East flow is employed from midnight to 6:30 am local time, due to a noise abatement procedure. The Airport Acceptance Rate depends on the weather situation (IFR or VFR rules) as well as on the fleet mix. If the share of heavy aircraft and B-757s is greater than 30%, that can reduce the Airport Acceptance Rate.

Table 18 summarizes the results of our investigation for days where there was a considerable difference between predicted and observed DFTI values, as well as the reasons for these discrepancies. The days considered "good" are with higher predicted than observed values for DFTI. One common feature of these days is that, hour-by-hour, the demand was consistently less than the Airport Acceptance Rate (AAR). While demand, as well as factors that influence the Airport Acceptance Rate, are considered in

our models, we are not able in our methodology to precisely distinguish days that feature such a favorably hourly pattern.

To illustrate, one of the "good" days was 2/17/98. The airport ran visuals throughout that day, with no delays reported. Inboard runways were used for arrivals too. Neither the demand nor the actual number of arrivals at LAX exceeded the given Airport Acceptance Rate at any time during the day. Thus operations ran smoothly throughout the day even with one of the runways closed for 20 minutes due to runway checks.

When we investigated “bad” days, for which the predicted values were less than observed, we found a broader set of explanatory factors. Some of the reasons for higher observed values are east traffic at the airport, or changes from east to west traffic and vice-versa. In a few cases one of the runways was closed for some time, which can also lead to a higher delays. Also in 6 out of 8 cases delay programs were in effect, so maybe that contributed to the delays as well. In one of the cases air traffic control had problems with equipment, which also may have led to a higher values of delay.

One of these "bad" days was 12/05/97 with a DFTI 23 minutes greater than predicted. ATCSCC directed an all airport ground delay program including props for the entire day, which resulted in peaks and valleys of traffic throughout the day. The program did not deliver enough aircraft to meet AAR and delayed many flights unnecessarily. During one period of time (1600-1800) the traffic was extremely light as a result of the program. Flight schedules were backed up that much, that by late afternoon aircraft were missing their Estimated Departure Controlled Times (EDCT) from LAX. At first, the TRACON was releasing these aircraft as soon as possible, but over time the number of aircraft

Table 18. Summary for Outliers

Month	Day	Year	Observed	Predicted	Residual	Reason
2	17	1998	154.284	169.487	-15.203	VFR whole day, demand less than AAR
1	4	1998	161.753	174.163	-12.410	No apparent reason
9	25	1997	145.880	157.538	-11.658	No apparent reason, demand was less than AAR
9	4	1998	149.236	160.328	-11.092	West traffic before 6:30
6	16	1998	148.768	159.806	-11.038	IFR whole day, actual arrivals less than AAR
2	7	1998	158.560	169.253	-10.693	Limited visuals, demand less than AAR
4	28	1998	136.470	146.628	-10.158	VFR whole day, demand less than AAR
1	21	2000	164.383	151.037	13.346	IFR whole day, EDCT, RWY 25L closed for 3.5 h
9	26	1999	170.834	156.237	14.597	IFR whole day, EDCT invoked
12	22	1999	174.703	156.880	17.823	Few E/W changes, RWY 25R closed for 2 h
4	17	2000	198.286	176.626	21.660	East Traffic, erratic winds, EDCT
11	19	1999	185.485	163.310	22.175	IFR whole day, EDCT, RWY 25R closed 1 h
2	6	1998	194.986	172.332	22.654	West to East change during rush, program invoked, WX
12	5	1997	194.491	171.082	23.409	East traffic, EDCT program, backed out
11	26	1997	196.798	172.325	24.473	Storm during the day, east then west traffic, equipment problems

missing their EDCT times grew to a point where the TRACON had to invoke internal call for release (CFR) program to avoid sector saturation. The TMU feels that the CFR program successfully minimized holding and balanced the delivery of aircraft and produced the better actual arrival rate. On top of ground hold program, LAX was running east flow throughout the day.

During 1/21/00, the operations were IFR throughout the day shift, returning to visual operations in the evening. Airforce 1 arrived at the LAX at start of the late morning rush period. There was also a ground delay program for BOS, EWR, CFR LAX, and PHX. For about 20 minutes during the day shift an internal ground stop for LAX was invoked. From these facts, it seems that the reason for higher values in the observed DFTI is the ground delay program.

For the 12/22/99, there were no ground delay programs invoked. However there were multiple runway changes during the day due to wind shifts, which made very complex dayshift. Every time that LAX changes to the east traffic the Airport Acceptance Rate goes down. On top of the runway changes, LAX had very high surface traffic congestion, causing additional arrival and departure delays.

The final “bad” day considered was 4/17/00. This featured IFR east traffic, unusually low AARs of between 54 and 58, and a ground hold program imposed on flights originating from the western states (a so called 10-W program). There is evidence that the program delivered too few aircraft in one period of time and too many in another, which led to an internal ground stop. After that time, visuals began to work and excess inventory was exhausted fairly rapidly.

11. Conclusions

The need to assess and isolate the impacts of FFP1 on NAS performance has motivated this analysis of day-to-day variation in average flight times (as measured by the DFTI) into LAX. We find that the majority of the variation can be statistically accounted for in terms of average delays at the origin airport, weather at LAX, and demand at LAX as measured by the peaking characteristics of the flight schedule. While all of these factors are statistically significant, there is a clear hierarchy, with origin airport delay the most important, and demand the least. This reinforces the point that any individual airport is part of a larger system and that delays are strongly influenced by non-local factors.

We estimated a variety of models explaining day-to-day variation in DFTI. As we move from a simple and restrictive model that is strictly linear in the explanatory variables to more sophisticated models with quadratic interactions and non-linear forms, we realize a modest gain in statistical performance. Conversely, there is some degradation when non-parametric models based on days categorized by weather conditions are used. While these differences are of some interest, it appears that the content of the model - what explanatory variables it includes - is more important than the form.

Similarly, but perhaps more surprisingly, we find that model fits are quite insensitive to the level of detail at which local weather is included in the model. Linear models with three, six, nine, and twelve weather factors all yield roughly the same level of performance. In light of this, the “principal of parsimony” suggests the three-factor model as a candidate for further study and refinement in future research.

The old saw that “correlation is not causation” must be born in mind when interpreting the results of this research. Many of the statistical results can be understood in terms of well-known mechanisms in which cause and effect are clear. For example, it is obvious that departure delays at origin airports will cause arrival delays at LAX. Other results are more mysterious. For example, why is it that the surface temperature at LAX is correlated with the DFTI? Does this derive from a visibility effect, a wind effect, or is the effect actually due to temperature? Likewise, how can we account for the inverse relation between peak demand and taxi-in time? Is it really from increased landing on the in-board runways, or is there some other reason?

While these are puzzling questions, their answers are not of central importance given the objective of normalization. It is precisely because we don’t understand all the mechanisms at work that statistical normalization is a valuable tool in gauging the impact of FFP1 technologies.