Statistical Extreme Value Analysis of the SFO Taxiway Centerline Deviations for 747 Aircraft

Prepared Under a

Cooperative Research Development Agreement Between

The Federal Aviation Administration (FAA) and The Boeing Company

Statistical Extreme Value Analysis of the SFO Taxiway Deviations for 747 Aircraft[1]

1 Introduction

This is a summary of the risk analysis of taxiway centerline deviations of 747 aircraft at the San Francisco International Airport (SFO). Data were collected from the 75 feet-wide taxiway "F" using four lasers placed along a straight segment of the taxiway within a span of 500 ft. Deviations of aircrafts' nose and main gears from taxiway centerline were measured at the four laser locations, hence producing eight measurements for every taxiing aircraft. Every aircraft was also photographed twice using a camera mounted at Laser 3.

We employ a two-step discrimination process to identify a Boeing 747. First, initial discrimination is based on the landing gear geometry, i.e., the 75.54 feet longitudinal distance between a 747 nose and main gear and outer-to-outer main gear tire width of 41.33 feet. Second, we visually inspect images of the aircrafts identified in the first step for the final discrimination. A total of 293 aircraft were positively identified as 747 for their taxiway centerline deviations to be used in the analysis.

The goal of the study is as follows. First, investigate whether sinusoidal functions can be used to reasonably approximate taxiway paths generated by taxiing aircrafts under the assumption that large aircrafts meander around the taxiway centerline following some sinusoidal path. Secondly, if such functions can be realized, modeling deviations collected by four lasers to mimic the aircraft taxiway path along the entire length of the taxiway. Third, providing a basis for understanding the extreme behavior of centerline deviations resulting from such

[1]Data collected by the FAA. Analysis and report prepared by Ranjan Paul, The Boeing Company, August 26, 2009.

predicting deviations from the second step. The second step requires extrapolating the deviation behavior of the screened 747 to more extreme levels of deviation to understand what could be encountered if a much higher sample was observed. In other words, the requirement is to estimate the probability of an event that is much more extreme than that have already been observed.

The four lasers were placed along a straight segment of the SFO taxiway "F" with separations 227 feet between Laser 1 and Laser 2, 98 feet between Laser 2 and Laser 3 and 175 feet between Laser 3 and Laser 4. A diagram with locations of lasers and taxiway "F" with respect to the SFO airport can be found in Figure 1. The lasers were purposely placed at unequal distances from each other. It is well known that fitting sinusoids to equally-spaced points leads to aliasing in that sinusoids with frequencies differing by an integer multiple of the "Nyquist" frequency, π/δ, have the same values at such equally-spaced locations. However, here the situation is slightly different in that we fit four coupled sinusoids simultaneously.

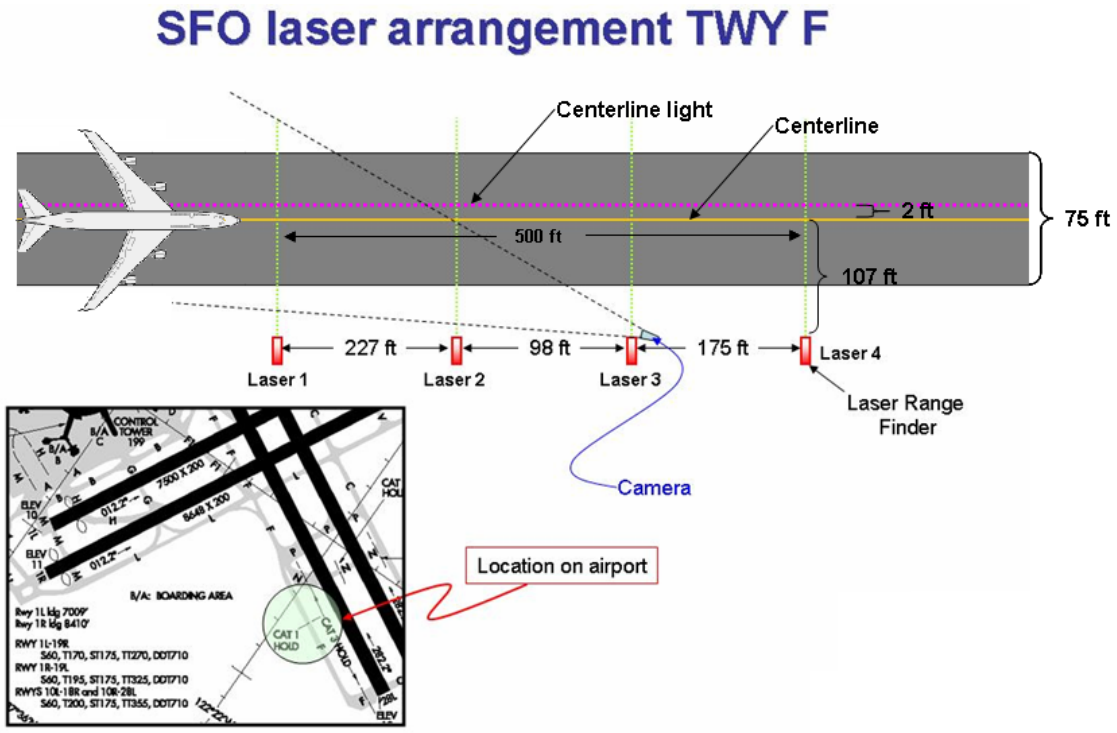

Figure 1: Diagram showing location of taxiway F and lasers.

Taxiway "F" is 75 feet wide and the centerline is right in the middle of the taxiway. The

lasers were located 107 feet away from the taxiway centerline and 109 feet from the centerline light. That means there is a 2 feet separation between centerline and centerline lights. The row of centerline light is on the opposite side of the taxiway centerline from the lasers. In similar previous studies [1,2,3], a maximum of two lasers were used. The purpose of using four lasers is mainly to capture presumed taxiway meandering of large aircraft. It also gives an opportunity to compare results using one single laser with those obtained using all four lasers.

2 Data Collection and Screening

Data were collected from approximately eleven thousand aircraft of which complete sets of 8 measurements were taken from 6956 aircraft. Due to technical problems, measurements from 3 or fewer lasers were obtained from the remaining aircraft. The scope of this study is within the 6956 aircraft for which 55645 measurements were obtained.

We start with a screening process to discriminate 747 which is the aircraft model of interest. First, we apply the specification on lateral distance between nose and main gear of a 747-400 model. Since aircraft do not always taxi at a constant speed and/or perfectly parallel to the taxiway centerline, measurements taken by lasers deviate from the specifications. It has been shown in the previous studies [2,3] that any estimated lateral distance outside the range 18.5 ± 1.5 ft should indicate aircraft other than a 747. Second, we screen estimated wheelbase against 747 specification. Again, due to acceleration/deceleration and angle of travel, any ratio of estimated wheelbase to true wheelbase (76.54 ft) outside 1 ± 0.4 is considered to be excessive [2,3], and hence not corresponding to a 747. This screening for lateral and wheelbase distances resulted in 988 potential 747 aircraft. Detail dimensional specifications of a typical 747-400 can be seen in Figure 2. As mentioned earlier, each taxiing aircraft were photographed twice for further identification. Third, all the $1976 (= 988 \times 2)$ images were visually inspected. This was very helpful in positively identifying 747 because some other aircrafts models (e.g., some series of B777, MD12 and A340) have very similar geometry to that of a 747. Thus we end up with 2344 deviation measurements from 293 positively identified 747 all of which turned out to be east bound traffic. It should be mentioned that airacraft taxiing at night did not contribute to the final 293 because one can not see

anything from the image but the head/flash light of the aircraft. Also, grass grown tall between mowings in front of the camera resulted in some obscured shots where aircraft is either mostly or totally hidden and could not be positively identified.

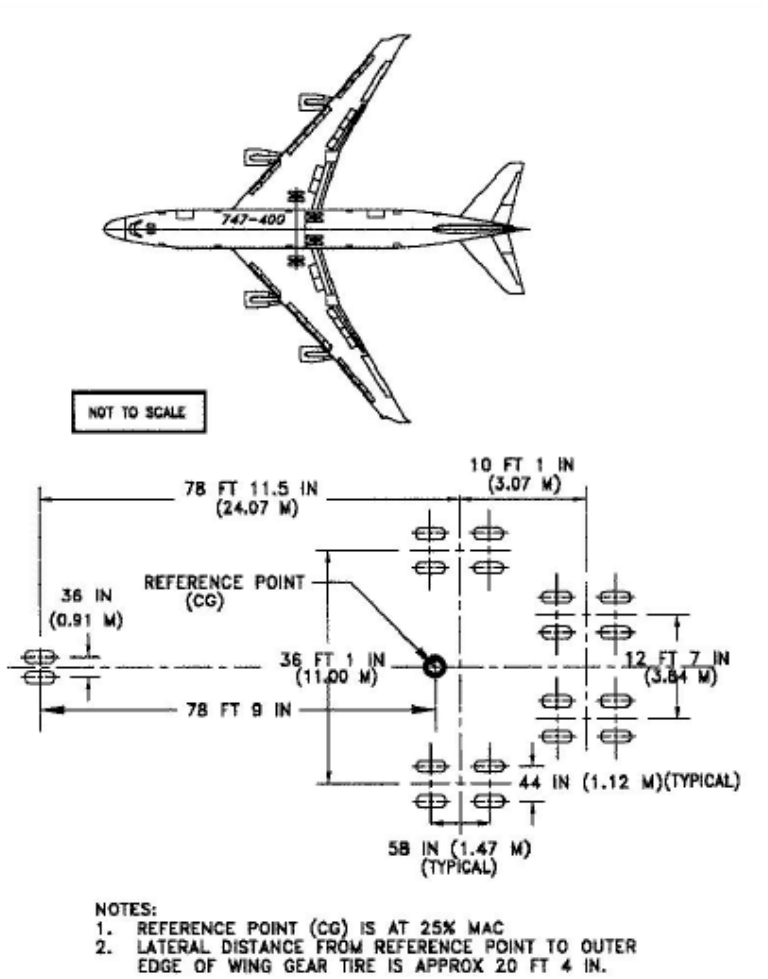

Figure 2: Diagram showing dimensions of a 747.

Data from the 293 747s were put through various consistency checks. They include checking for time trend, plotting deviations of nose and main gear against each other, plotting deviations from the same gear but from two different lasers against each other and checking for outlier. No particular time trend was observed when plotting deviations against calendar time. Also, measurements taken at different lasers or of different gears seem to be consistent. No systematic pattern was observed in these plots. We checked for symmetry of the

distributions of observed deviations at individual lasers and found that they are all skewed on the side of the centerline light (positive deviations). In the next section we describe how we model deviations from four lasers by sinusoidal functions to represent taxiway travel meandering.

3 Modeling

As stated earlier, we will attempt to mimic taxiway travel paths of aircrafts by modeling with sinusoidal functions. Using 747 specifications and distances between lasers, we formulate a system of sinusoidal equations which we explain in the following subsection.

3.1 Fitting a Sinusoidal Path Parallel to Taxiway Centerline

The question was raised to what extent the main gear would follow a sinusoidal path around the taxiway centerline or around a line parallel to the taxiway centerline when the nose gear follows such a path. It was felt that understanding the answer to this might be of some use in relating nose to main gear deviations at a fixed laser location and also may enable the fitting of a full aircraft path to nose and main gear measurements from each of two different laser locations. Such a path would be useful in assessing lengthwise rather than pointwise risks of exceeding the taxiway edge at any point along a straight taxiway segment.

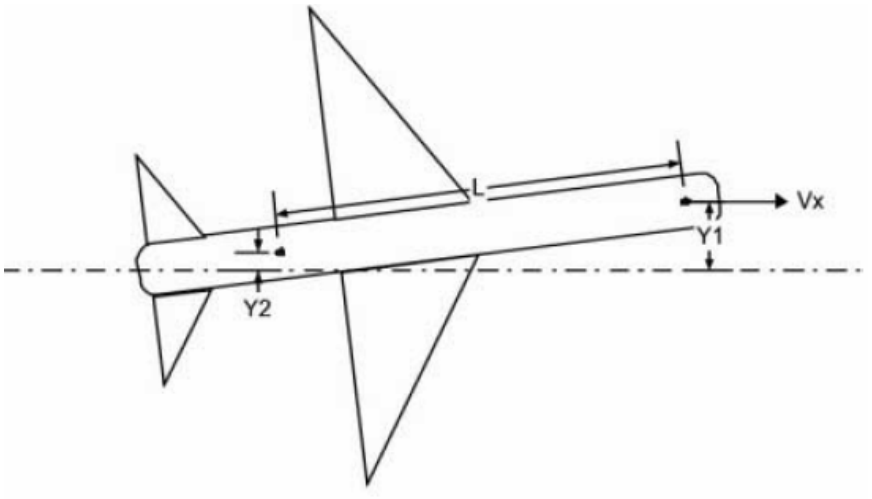

Figure 3: Diagram relating main gear and nose gear deviations.

Using Figure 3 we denote by $Y_1 = Y_{ng}$ the lateral deviation of the nose gear and $Y_2 = Y_{mg}$ the lateral deviation of the main gear. The geometry of a 747 and the measurement of eight aircraft centerline deviations from the taxiway centerline (nose and main gear location at four different lasers) offers the possibility of capturing more than just the pointwise deviations from the taxiway centerline provided we are willing to assume that, to a rough approximation, each such path can be seen as a sinusoidal path with longitudinal axis parallel to the taxiway centerline, i.e., we model the deviation of the aircraft centerline from the taxiway centerline as follows at the nose and main gear locations, respectively:

$$Y_{ng}(t) = \Delta + A\sin(2\pi t/T + \theta)$$

and

$$Y_{mg}(t) = \Delta + A^*\sin(2\pi t/T + \theta + \psi),$$

where Δ is the lateral offset of the sinusoidal path relative to the taxiway centerline, A is the amplitude of the nose gear sinusoidal path as measured from the sinusoid centerline, A^* ($A^* = f(A, L', \psi, T)$) is the amplitude of the main gear sinusoidal path as measured from the sinusoid centerline, and T is the period for one oscillation of the nose gear. The distance between nose gear axle and first main gear axle is denoted by L' and parameters θ and ψ are related to phase shift. These parameters can be better described diagrammatically in Figure 3.1. Thus $A\sin(\theta)$ is the sinusoidal deviation from the sinusoid centerline at time t = 0. The four quantities A, θ, T, and Δ vary from aircraft to aircraft. Since we have eight measurements for the centerline deviations we thus should be able to find values for these parameters that give the best fit and one may then treat such best fit sinusoidal paths as the actual paths of nose and main gear. We can then determine how far out each such main gear path will reach, i.e., we can obtain $A^* + \Delta$ when $\Delta \geq 0$ or $-A^* + \Delta$ when $\Delta \leq 0$. This would give us a maximal deviation along the whole length of the straight taxiway. So far the pointwise measurements at the lasers would typically not give us the deviations at the amplitude of the sinusoidal path. We now set up the relationships that need to be fitted. We denote by D_{12} the distance between Laser 1 and Laser 2 ($D_{12} = 227$ ft), D_{23} the distance

between Laser 2 and Laser 3 ($D_{23} = 98$ ft) and D_{34} the distance between Laser 3 and Laser 4 ($D_{34} = 175$ ft). Thus the distance between Laser 1 and Laser 4 was 500 ft.

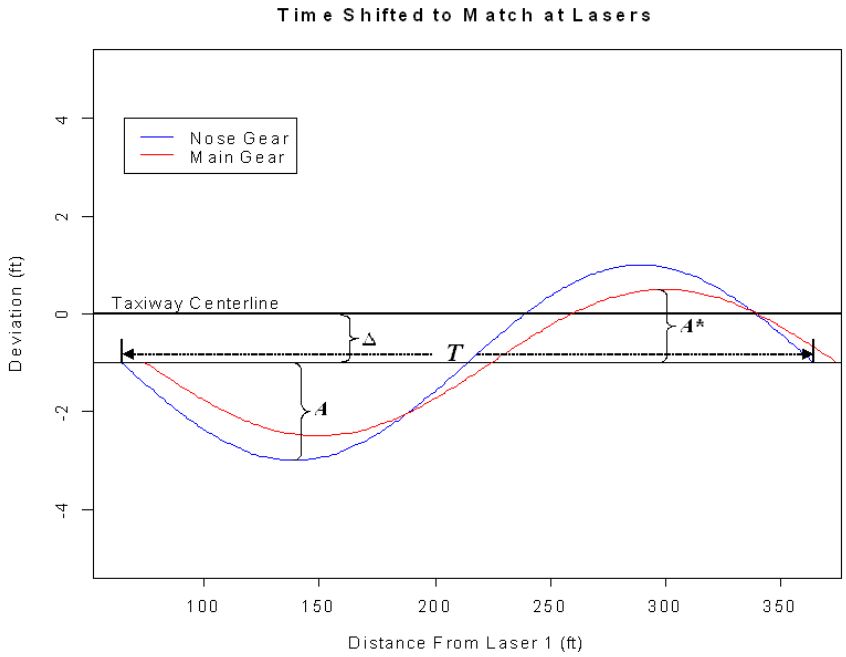

Figure 4: Illustration of nose and main gear sinusoidal paths.

Let us denote by t_i^{ng} the time when nose gear deviation is registered by Laser i ($i = 1, 2, 3, 4$). Similarly, let t_i^{mg} be the time when main gear deviation is registered by Laser i ($i = 1, 2, 3, 4$). Then time $t_1^{mg} = L^{'}/V_x$, where V_x is the velocity of the aircraft along the taxiway centerline and $L^{'}$ is the distance between nose gear axle and first main gear axle (along the aircraft centerline).

For the purpose of simpler reference we treat the time t_1^{ng} when the nose gear deviation $y_{ng}(t_1^{ng})$ is registered by the first laser as $t_1^{ng} = 0$. Then the time t_1^{mg} when the main gear deviation $y_{mg}(t_1^{mg})$ is registered at the first laser is $t_1^{mg} = L^{'}/V_x$, the time t_2^{ng} when the nose gear deviation $y_{ng}(t_2^{ng})$ is registered at the second laser is $t_3 = D_{12}/V_x$, and the time t_2^{mg} when the main gear deviation $y_{mg}(t_2^{ng})$ is registered at the second laser is $t_4 = (D_{12} + L^{'})/V_x$ and so on. In order to fit sinusoidal paths to the observed deviations from four lasers, one needs to find A, θ, T, and Δ such that the fitted paths pass the observed deviations as close as possible. We estimate these parameters such that the quantity

$$\sum_{i=1}^{4}\left[\hat{Y}_{ng}(t_i^{ng}) - y_{ng}(t_i^{ng})\right]^2 + \sum_{i=1}^{4}\left[\hat{Y}_{mg}(t_i^{mg}) - y_{mg}(t_i^{mg})\right]^2 \quad (1)$$

is minimized under two constraints, where $\hat{Y}(t)$ is the fitted value of the sinusoidal at t. The solution obtained is often referred to as the "Least Squares" solution. We discuss the constraints in the next section. Here

$$\begin{aligned}
Y_{ng}(t_1^{ng}) &= \Delta + A\sin(\theta),\\
Y_{mg}(t_1^{mg}) &= \Delta + A^*\sin(2\pi L'/V_xT + \theta + \psi),\\
Y_{ng}(t_2^{ng}) &= \Delta + A\sin(2\pi D_{12}/V_xT + \theta),\\
Y_{mg}(t_2^{mg}) &= \Delta + A^*\sin(2\pi(L' + D_{12})/V_xT + \theta + \psi),\\
Y_{ng}(t_3^{ng}) &= \Delta + A\sin(2\pi(D_{12} + D_{23})/V_xT + \theta),\\
Y_{mg}(t_3^{mg}) &= \Delta + A^*\sin(2\pi(L' + D_{12} + D_{23})/V_xT + \theta + \psi),\\
Y_{ng}(t_4^{ng}) &= \Delta + A\sin(2\pi(D_{12} + D_{23} + D_{34})/V_xT + \theta),\\
Y_{mg}(t_4^{mg}) &= \Delta + A^*\sin(2\pi(L' + D_{12} + D_{23} + D_{34})/V_xT + \theta + \psi).
\end{aligned}$$

In the system of equations above, we can see that V_x always appears in conjunction with T as V_xT. Therefore, we reparameterize the system of equations with $\kappa = V_xT$ which is the distance the nose gear travels in order to complete one complete sinusoidal oscillation, called the wavelength. Thus knowledge of V_x becomes unnecessary in finding the set of unknown parameters $\{A, \theta, \kappa, \Delta\}$. The two constraints under which (1) is minimized are as follows. Firstly, the fitted sinusoidal paths are not allowed to cross the taxiway edge. This restriction is used for the practical reason that no observed deviation of airplane centerline were beyond 10 ft from the taxiway centerline. Secondly, κ is bounded below by 100 ft, i.e., wavelengths smaller than 100 ft are not allowed. This seems to be reasonable because it is not practical for a 747 with 76.5 ft wheelbase to complete a full sinusoidal oscillation in less than 100 ft.

Here we briefly the discuss results of fitting sinusoidal paths to aircraft centerline deviations. Estimated κ for most of the paths are found to be between 200 and 500 feet. This implies us that most of the fitted paths complete a full sinusoidal wavelength within 200 to 500

feet. None of the estimated κ are even close to the constraint lower bound of 100 feet. The distribution of estimated Δ is skewed to the positive side and between -2 and 6 ft in most cases. The skewness to the positive side implies that for most aircraft, the center of the fitted sinusoidal paths are further away from the taxiway centerline when they were centered on the side of the centerline light. We have seen similar skewness when we examined the distribution of the pointwise deviations. Estimated amplitudes A are found to be mostly under 4 ft in absolute value.

3.2 Extracting Extremes From Fitted Sinusoidal Paths

In order to use the fitted sinusoidal paths to perform extreme value analysis, we need to extract extremes from the fitted paths generated by the main gear of the surveyed aircraft. These extracted extremes should represent maximum deviations along the whole length of a straight taxiway. The idea is to sample maximum fitted-deviations on each side of the taxiway centerline. For different estimated values of A^* and Δ, we use the following scheme to extract maximum deviation X_i:

$$X_i = \begin{cases} \Delta \pm A^*, & \text{if } \Delta + A^* > 0 \text{ and } \Delta - A^* < 0, \\ \Delta + A^*, & \text{if } \Delta > 0 \text{ and } \Delta - A^* \geq 0, \\ \Delta - A^*, & \text{if } \Delta < 0 \text{ and } \Delta + A^* \leq 0. \end{cases}$$

Following the above equations, we use both minimum and maximum from the fitted sinusoidal path if the fitted path crosses the taxiway centerline. We use minimum if the fitted path is entirely on the opposite side from the centerline light. We use maximum if the fitted path is entirely on the same side as the centerline light. Any deviation on the same side as the centerline light is then called "Positive" deviation and any deviation on the opposite side from the centerline light is called "Negative" deviation.

We compare the distribution of the fitted deviations to that of the observed deviations. The two distributions look similar in having heavier tails on the right (on the side of the centerline light). We observe that the centers (using medians) of both distributions are not at the centerline of the taxiway. For observed deviations, the median of the distribution is at

1.55 ft while the median of the fitted deviations is at 2.05 ft. This indicates a possible bias due to the centerline light which is 2 ft away from the taxiway centerline. One may call it as "centerline light bias". We correct for this bias by shifting the distributions by subtracting their corresponding medians before performing any risk assessment in later sections. One needs to back adjust a given threshold by corresponding median when estimating the risk.

It is obvious that fitting sinusoidal functions does not necessarily inflate the maximum observed deviation. It can also deflate the observed maximum deviation depending on the least squares solution. It is important to understand that it is the shape of the distribution generated by the fitted deviations rather than particular inflation/deflation which play a vital role in risk estimation.

3.3 Extreme Value Analysis

Extreme value analysis deals with the stochastic behavior of a process at unusually large or small levels. This requires estimating the probability of events that are more extreme than any that have been observed in the sample. Extreme value theory provides a framework that enables such extrapolation is derived from asymptotic argument.

Suppose $X_1, X_2, ..., X_n$ is a sequence of independent random variables having a common distribution function F. The extreme value model focuses on the statistical behavior of $M_n = \max\{X_1, X_2, ..., X_n\}$. The distribution of M_n can be derived exactly for all values of n as

$$Pr\{M_n \leq x\} = Pr\{X_1 \leq x,\ X_2 \leq x, ..., X_n \leq x\} = F^n(x).$$

It is easy to see that the distribution of M_n degenerates to a point mass ($\rightarrow 0$) as $n \rightarrow \infty$. To cope with this difficulty, M_n is renormalized by

$$M_n^* = \frac{M_n - b_n}{a_n},$$

where $\{a_n\}$ ($a_n > 0$) and $\{b_n\}$ are sequences of constants. For appropriately chosen $\{a_n\}$ and $\{b_n\}$, the distribution of M_n^* stabilizes, hence avoiding the difficulties that arise with degenaration of M_n. The distribution of M_n^* is approximated by a limiting distribution for appropriately chosen $\{a_n\}$ and $\{b_n\}$.

3.4 Extreme Value Asymptotic Theorem

If there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that

$$Pr\left\{\frac{M_n - b_n}{a_n} \leq x\right\} \rightarrow G(x)$$

for a non-degenerated distribution function G, then G is a member of the generalized extreme value (GEV) family

$$G(x) = exp\left\{-\left[1 + c\frac{(x-\mu)}{\sigma}\right]^{-1/c}\right\},$$

on support $\{x:\ 1 + c(x-\mu)/\sigma > 0\}$, where $-\infty < \mu < \infty,\ \sigma > 0$ and $-\infty < c < \infty$.

We note here that the above domain condition $1+c(x-\mu)/\sigma > 0$ for the limiting distribution $G(x)$ implies that its upper range is unlimited when $c = 0$ and that it is limited by $x < \mu-\sigma/c$ when $c < 0$. In the latter case one therefore models a hard upper limit for the behavior of $(M_n - b_n)/a_n$, and thus presumably for M_n since one uses the approximation for some reasonably large n for fixed a_n and b_n. However, a_n and b_n could be quite large so that for practical purposes this upper bound could still be quite far off and appear as infinite. In any case, we thought it is worthwhile to point out this special status of $c < 0$, in particular with respect to the current application of 747 deviations from the taxiway centerline because of the somewhat limiting effect of a taxiway edge when it is perceived by the pilot. After having fixed on the proper number k of extremes we then ran the EXTRAP program on the respective number of extremes in the adjusted deviation sample. The resulting extrapolated deviation thresholds were then tabulated for various specified risk levels.

The primary concern is to assess the risk of exceeding thresholds beyond the observed range. Conversely one can ask for the deviation thresholds that are exceeded with various specified small risks, say $10^{-5}, 10^{-6}, 10^{-7}$. Methodology for this was developed in [4] and other methods are discussed in [5]. In essence all of these methods deal with the k most extreme observations in either tail of a given random sample, $X_1, X_2, ..., X_n$. A major question is how large to choose k. Too large a k will introduce influences from the center of the data that have little to do with the extreme behavior of such deviations and thus could lead to bias. Too small a k will leave us open to the typically strong fluctuations in the tail of the

data and will thus result in too much uncertainty in the risk extrapolations. Although [4] proposed a method for choosing k we will not use it here. Instead we take a close look at the tail behavior of the data and make a judgment call on the proper choice of k.

To explain the connection between the distributional behavior of M_n and that of the k largest sample values we digress briefly. Originally the above distributional limit result was used as follows. View the sample $X_1, X_2, ..., X_n$ as a collection of k subsamples of size m where $km = n$. If the original n cannot be cleanly subdivided into k subsamples of size m one may have to discard some observations. If m is large enough one may regard the k maxima from the k subsamples as having roughly the above limiting distribution. Thus one could model the sample of k maxima by that distribution and estimate its three parameters by the method of maximum likelihood. If k is small one deals with the issues of greater variability/uncertainty accompanying small samples. If m is not large enough then the approximating distribution G may not be a good approximation yet and thus there will be some bias which persists no matter how large k is for fixed m. Given that n is fixed and we have to work within the constraint $n = km$ we are stuck between the trade-off of bias and variability.

This view and approach through subsamples has a disconcerting aspect. Namely, how should the allocation of the original sample of size n into k subsamples of size m be done? Furthermore, different allocations would lead to different analysis results. Given that one wants to treat the k maxima from these k subsamples as a random sample (independent) one would have to insist on a random subdivision of $X_1, X_2, ..., X_n$ into k subsamples. This would render the final analysis results subject to extraneous randomness, i.e., different analysts would arrive at different answers by using the same method but not the same random number generator and/or random seed.

Typically the k maxima so obtained will come fairly close to the k largest observations $Y_1 \geq Y_2 \geq ... \geq Y_k$ in the original sample of size n. It is likely but not guaranteed that a random subdivision of $X_1, X_2, ..., X_n$ into k subsamples will allocate one of those $Y_1 \geq Y_2 \geq ... \geq Y_k$ to each of these k subsamples. When that does not happen one uses a lesser extreme in the analysis and one does not make the most efficient use of the extreme data in the original sample. This concludes our digression.

In determining the proper tail depth k we employ two approaches. In the first we compute estimates of the extreme value index c for various values of k and examine when, in terms of k, the variation of these estimates around some level transitions into a drifting off behavior. A second diagnostic examines whether an expected consequence of the assumed extreme value limiting assumption does show up in the data to a reasonable tail depth.

3.5 Determination of Tail Depth k

We use the modified Hill estimator [4] to estimate the extreme value index c for varying k largest sample values. We examine the estimates of c to see when it starts to deteriorate. This gives us an indication of how large a k one should use. Determining this point of deterioration is somewhat subtle. Typically the extreme value index estimates will fluctuate strongly when based on small values of k. As k gets larger these estimates settle down to a more stable value until they start to drift off to a different level. Sometimes this first settling down range of k is short and may be difficult to distinguish from the wild fluctuation and the drifting off behavior. A preferred way of viewing these estimates as a function of k is to plot k on a logarithmic scale and inspect the graph visually.

Secondly, we examine a plot of the median values of exceedence over a threshold u for varying values of u. This provides us a diagnostic for the extreme value limiting assumption as discussed in [5]. The idea is to check whether for large values of u the median excess over u shows an approximate linear relationship as a function of u.

3.6 Nonparametric Extrapolation

The primary objective is to assess the risk of exceeding thresholds beyond the observed range of centerline deviation. This requires a method to extrapolate the tail of the distribution. We use a non-parametric tail extrapolation technique in this study [4]. The technique assumes that the random sample comes from a continuous population which follows one of the three extreme value distribution types. Let $X_1, X_2, ..., X_n$ be a random sample from continuous

distribution $F(x)$ and $Y_1 \geq Y_2 \geq ... \geq Y_k$ be its ordered statistics. For the pth quantile x_p

$$\gamma = Pr\{Y_i \geq x_p\} = I_{1-p}(i, n-i+1),$$

where $I()$ is the incomplete beta function. Then for each (i, γ), there is a unique and exact solution to determine $p = p_i = p_{i,\gamma,n}$ [6,7]. Thus Y_i can serve as an exact $100\gamma\%$ upper confidence bound for x_p. For a normal population, $(Y_i, h(p_i))$ should follow a linear relationship with $h(p_i)$ as the log-odds transformation of p_i.

The three forms of extreme value distribution type can be expressed simultaneously as

$$H_{c,\lambda,\delta}(x) = exp\left\{-\left[1 + c\frac{(x-\mu)}{\sigma}\right]^{-1/c}\right\},$$

on support $\{x:\ 1 + c(x-\mu)/\sigma > 0\}$, where $-\infty < \mu < \infty,\ \sigma > 0$ and $-\infty < c < \infty$. For $p \approx 1$, the pth quantile x_p can be approximated by solving $H_{c,\lambda,\delta}(x) = p^n$ as

$$\begin{aligned} x_p &= \delta\frac{(-n\ln(p))^{-c}-1}{c} + \lambda, \qquad \text{for } c \neq 0 \\ &= \lambda + \delta(-\ln(-n\ln(p))), \quad \text{for } c = 0. \end{aligned}$$

Choosing $\gamma = 0.5,\ Y_1 \geq Y_2 \geq ... \geq Y_k$ serve as median unbiased estimates of $x_{p_1}, x_{p_2}, ..., x_{p_k}$ for appropriately chosen k [4]. This can also be used for finding the estimated exceedence risk or upper confidence bound for the estimated risk via the following formula

$$p = 1 - \exp\left\{-\frac{1}{n}\left[1 + c\left(\frac{y-\lambda}{\delta}\right)\right]^{-1/c}\right\},$$

where y is the specific threshold and p is the risk estimate or upper bound for the risk of exceeding the threshold y. Estimates of parameters $\lambda,\ \delta$ and c are obtained using a weighted least squares fitting of the linear relationship.

Here we illustrate how one should use estimated parameters and the formula above to estimate risk or thresholds. Suppose we want to estimate the risk of aircraft longitudinal centerline deviation with a threshold of 10 ft on the positive side using the fitted sinusoidal deviations. In the formula for estimating p above, one should use $y = 10 - 2.05$. For the

same scenario but for the negative deviations, use $y = 10 + 2.05$. Here 2.05 ft is the median of the sinusoidal fitted deviations. Similarly, for a given extreme risk level p, the median should be back adjusted to x_p by adding or subtracting 2.05 depending on the side of the deviation in question. For the side of the centerline light (positive deviation), one adds the corresponding median to x_p but subtracts for the opposite side (negative deviation) of the taxiway centerline.

4 Results and Discussion

In this section we tabulate and discuss estimates of risk obtained using extreme value analysis. We apply the extreme value modeling and the nonparametric tail expansion techniques to different data sets and compare their results. Seven different data scopes have been utilized for the comparative analysis:

- Maximum deviations from the fitted sinusoidal paths using data from all four lasers,
- Pointwise deviations from all four lasers and
- Four data sets containing pointwise deviations from four individual lasers.

Estimated values of the parameters for the extreme value analysis and nonparametric tail expansion are tabulated in Table 1. All estimation is done using median adjusted deviations. Numbers in parentheses are those corresponding to the nonparametric 95% upper bound.

Table 2 shows threshold (quantile) estimates by exceedence risk for a 747 longitudinal centerline deviation from the taxiway centerline. Numbers in parentheses show 95% upper confidence bounds for the threshold estimates.

From the earlier exploratory analysis and the risk analysis results in Table 2 it is evident that the two sides of the taxiway centerline demonstrate different characteristics in terms of deviation risk. Using deviation data from all four lasers (observed or sinusoidal-fitted), the side of the taxiway centerline with the centerline light is exposed to higher risk than the other side of the centerline. There may be other explanations but one possibility is that the pilot is trying to avoid hitting the the centerline lights with the nose gear tire or the inner tire of the rear column of the main gear. The vertical distance between the inner edges of

the inner tires of the rear main gear column is about 11.6 ft. The outer edge of the nose gear tire is about 1.4 ft away from the aircraft longitudinal centerline. Thus the distance between the outer edge of the main gear tire and inner edge of a inner tire of the rear main gear column is about 6.9 ft.

Table 1: Extreme value parameter estimates for longitudinal aircraft centerline deviations. Numbers in parentheses correspond to 95% upper confidence bounds.

Data	Positive Deviation (Centerline Light Side)				Negative Deviation (Laser Side)			
	$\hat{k}$	$\hat{\lambda}$	$\hat{\delta}$	$\hat{c}$	$\hat{k}$	$\hat{\lambda}$	$\hat{\delta}$	$\hat{c}$
Extremes from Fitted Sinusoidal (Median=2.05 ft)	13	7.504 (9.342)	1.038 (1.249)	-0.252	23	4.574 (4.99)	0.249 (0.270)	-0.356
	n=174				n=174			
Pointwise Deviations from All 4 Lasers (Median=1.55 ft)	9	8.867 (11.410)	1.430 (1.856)	-0.169	38	4.207 (4.643)	0.331 (0.394)	-0.115
	n=591				n=581			
Pointwise Deviations from Laser 1 only (Median=1.38 ft)	17	5.040 (6.316)	0.803 (1.020)	-0.111	13	3.466 (4.096)	0.450 (0.635)	0.094
	n=147				n=146			
Pointwise Deviations from Laser 2 only (Median=1.22 ft)	8	5.148 (6.279)	0.624 (0.697)	-0.374	18	3.555 (4.050)	0.315 (0.391)	-0.144
	n=147				n=146			
Pointwise Deviations from Laser 3 only (Median=1.68 ft)	12	5.750 (7.046)	0.835 (1.104)	-0.053	13	3.811 (5.596)	0.549 (0.746)	0.028
	n=147				n=146			
Pointwise Deviations from Laser 4 only (Median=1.95 ft)	10	8.537 (11.270)	1.532 (1.806)	-0.289	20	4.024 (4.808)	0.500 (0.604)	-0.174
	n=146				n=147			

Since the centerline light is only 2 ft away from the taxiway centerline, the pilot may be trying to avoid hitting the centerline lights with the nose gear tires or trying to keep the

nose gear tire on the side of the centerline light to avoid hitting with the centerline light assembly with inner tire of the rear main gear column. In other words, pilots may be trying to keep the centerline light in between the two inner tires of the rear main gear assembly without hitting the centerline light fixtures by the nose gear tires.

But this does not explain why the distribution of deviations are skewed to right to the side with the centerline light. Looking at the aircraft taxiing paths there seems to be an overall shift starting at Laser 3 towards the side of the centerline light. Some aircraft veered quite far ($\approx$ 10 ft) while passing Laser 4. This may not be surprising to a subject-matter expert but we can only provide some statistical explanation. This may be occurring naturally as a random phenomenon. Or since this is a sample of only 293 aircrafts, this may be a sampling bias. Whatever the reason may be, we have to use the sample in hand and recognize that the results are obtained using a sample of only 293 aircraft.

Using the fitted positive deviation the estimated threshold of 9.32 ft corresponds to a 10^{-7} risk which translates into a $9.32 + 20.665 = 29.985$ ft deviation of the outer edge of the outer main gear tire on the side with the centerline light. This is still considerably smaller than the half width of a 75 ft wide taxiway. The 20.665 ft distance corresponds to the distance between aircraft longitudinal centerline and outer edge of the outer main gear tire of a 747-400. Other older models of 747 use tires of width 19 ft or smaller. The estimated threshold at risk level 10^{-7} is $7.31 + 20.665 = 27.97$ ft on the opposite side from the centerline light which is even smaller than that for the same side of the centerline light.

The threshold estimates using fitted and observed deviations from all four lasers show higher risk of deviation for the side with the centerline lights. Similar results are also seen using observed deviations from Laser 4. Evidently Laser 4 has some level of influence in the skewness of the right tail when combined data from all four lasers are used. We mentioned earlier that lasers were placed in a mid-section of the straight segment of the taxiway "F". The excess veering of some aircraft at Laser 4 (which contribute to the skewness to the right) may be just a random phenomenon unless there is a turn right after passing Laser 4 which does not seem to be the case.

Threshold estimates obtained using sinusoidal-fitted deviations are somewhere in the middle of the results obtained using individual lasers. This should not be surprising because there

is an averaging effect when using the fitted deviation. The least squares solution for the sinusoidal fit minimizes sum of squares of the vertical differences between observed deviations and the fitted curve. That means that the estimate of $|\Delta + A|$ does not always exceed the corresponding observed maximum deviation.

5 Conclusion

In this study, first, we attempted to mimic taxiway travel paths of a 747 aircraft along a straight segment of the taxiway "F" at the San Francisco International airport by sinusoidal functions. A system of sinusoidal equations were fitted using deviation data collected by four lasers placed in a span of 500 ft. We have used the structural geometry of a 747-400 aircraft to impose appropriate conditions for paths generated by the nose and main landing gears. We have obtained a reasonably good fit to the sinusoidal functions using the least squares estimation technique. The fitted sinusoidal functions show that most of the 747 travel a single wave length along a sinusoidal path within 500 ft which validates the appropriateness of placing four lasers within a span of 500 ft.

The fitted deviations as well as observed deviations are used to perform taxiway deviation risk estimation using statistical extreme value analysis and subsequent nonparametric tail extrapolation. Results show that there is a higher risk of deviation on the side of the taxiway centerline with the centerline lights compared to the other side when using fitted or all observed deviations from all four lasers. Analysis of individual lasers showed mixed results in this respect. At the risk level 10^{-7}, none of the point estimates (positive or negative) of the thresholds showed any evidence of the outer edge of the outer main gear tire crossing the edge mark of the taxiway.

The purpose of the risk analysis using sinusoidal-fitted deviation paths is to be able to claim that the estimated exceedence risks are lengthwise along the straight stretch of SFO taxiway "F". We have shown than there is an averaging effect in risk estimation using the fitted deviations. One extension of this study would be to constrain the fitted deviations to be equal or greater than the observed deviations.

One important question is whether these results can be generalized to other similar or different taxiways at other airports or other aircraft model? Although in a previous study [3] it has been shown that the data from two airports show very similar characteristics, one need to look at data from many more taxiways at various locations in order to be able to generalize any inference. As for expecting similar results for different aircraft types, one should expect different steering response behaviors for different-sized aircrafts.

The limitations of the extreme value analysis needs to be recognized when using these results. The extreme value models are developed using asymptotic assumptions and idealized circumstances. Our risk analysis via sinusoidal fitting is based on a sample of $1172 \times 2 = 2344$ deviations recorded at nose and main gears. They are divided into positive and negative deviations and analyzed separately due to the different tail behaviors. On the other hand, we only used deviations (1172) for the main gear when analyzing observed data. They were also divided by the sides of the taxiway. Additional caution is necessary concerning the risk extrapolations. There may be additional extreme value behavior that has not yet been realized which may exhibit more extreme characteristics. We have not taken into account the uncertainties in the choice of κ and the estimate of extreme value index c which arise from inherent sampling variation while computing confidence intervals of the estimated risks.

Acknowledgement

Portions of this work was accomplished with the support of funds from FAA Contract DTFAWA-05-P-00226 from The Office of Flight Standards Branch, AFS-450. Special thanks belongs to George Legarreta, Ryan King, and Jim Patterson of the FAA; and to Stephen Jones, Kaz Konya, and Jerry Robinson of The Boeing Company for their feedback and support of this study.

References

[1] Booker, A. (1995), "Statistical Analysis of Aircraft Deviations from Taxiway Centerline," Boeing Information & Support Services.

[2] Scholz, F. W. (2003), "Statistical Extreme Value Analysis of ANC Taxiway Centerline Deviations for 747 Aircraft", FAA-Boeing Cooperative research report.

[3] Scholz, F. W. (2003), "Statistical Extreme Value Analysis of JFK Taxiway Centerline Deviations for 747 Aircraft", FAA-Boeing Cooperative research report.

[4] Scholz, F.W. (1995), "Nonparametric Tail Extrapolation", ISSTECH-95-014, ISS Technology, Boeing Information.

[5] Coles, S. (2001), "An Introduction to Statistical Modeling of Extreme Values", Springer-Verlag, London.

[6] Hoaglin, D.C. (1983), "Letter values: A set of selected order statistics." in Understanding Robust and Exploratory Data Analysis, eds. D.C. Hoaglin, F. Mosteller, and J.W. Tukey, John Wiley & Sons, New York, 33-57.

[7] Filliben, J.J. (1975). "The probability plot correlation coefficient test for normality", Technometrics 17, 111-117.

Table 2: Threshold (ft) by Exceedence Risk for 747 longitudinal centerline deviations from taxiway centerline. Numbers in parenthesis correspond to 95% upper confidence bounds.

Data	Exceedence Risk									
	Positive Deviation (Centerline Light Side)					Negative Deviation (Laser Side)				
	1e-7	1e-6	1e-5	1e-4	1e-3	1e-7	1e-6	1e-5	1e-4	1e-3
Extremes from Fitted Sinusoidal	9.32 (11.94)	9.11 (11.69)	8.74 (11.25)	8.09 (10.46)	6.92 (9.06)	7.31 (7.79)	7.29 (7.77)	7.25 (7.72)	7.16 (7.62)	6.95 (7.39)
Pointwise Deviations from All 4 Lasers	14.15 (18.72)	13.37 (17.71)	12.23 (16.23)	10.53 (14.03)	8.04 (10.79)	7.69 (8.50)	7.41 (8.16)	7.04 (7.72)	6.55 (7.15)	5.93 (6.40)
Pointwise Deviations from Laser 1 only	8.79 (11.45)	8.17 (10.67)	7.39 (9.67)	6.39 (8.37)	5.05 (6.70)	13.70 (17.98)	11.04 (14.23)	8.90 (11.21)	7.18 (8.78)	5.80 (6.82)
Pointwise Deviations from Laser 2 only	5.57 (6.89)	5.54 (6.85)	5.45 (6.76)	5.25 (6.54)	4.78 (6.01)	6.52 (7.45)	6.35 (7.23)	6.11 (6.93)	5.77 (6.52)	5.30 (5.94)
Pointwise Deviations from Laser 3 only	11.09 (14.67)	9.96 (13.17)	8.68 (11.47)	7.23 (9.55)	5.59 (7.38)	12.67 (16.03)	11.00 (13.75)	9.43 (11.62)	7.96 (9.62)	6.58 (7.75)
Pointwise Deviations from Laser 4 only	11.67 (15.31)	11.47 (15.07)	11.08 (14.61)	10.32 (13.72)	8.85 (11.98)	8.44 (9.73)	8.23 (9.49)	7.93 (9.12)	7.47 (8.57)	6.79 (7.74)