



# A Data Driven Approach for Characterization of Ramp Area Push Back and Ramp-Taxi Processes

William J. Coupe \* and Dejan Milutinović †

*Jack Baskin School of Engineering, University of California, Santa Cruz, CA 95064, USA*

Waqar Malik ‡

*University of California, Santa Cruz, NASA Ames Research Center, Moffett Field, CA 94035, USA*

Yoon Jung §

*NASA Ames Research Center, Moffett Field, CA 94035, USA*

The purpose of this research is to better understand and characterize the time duration distributions between the push back, stop and taxi events within the ramp area transit time. The ramp area transit time is the time between push back and arrival at the taxiway spot. Specifically, the goodness-of-fit of log-normal and gamma distributions are assessed to different processes within ramp area transit time data that was collected at the Charlotte Douglas airport over a three day period. In this paper, we also assess the goodness-of-fit of two-dimensional parametric distributions to sampled two-dimensional ramp area conflict distribution data. Better understanding of both the distributions related to the different processes within the ramp area transit time and the conflict distributions are critical to the execution of safe and efficient ramp area operations.

## I. Introduction

Air traffic demand continues to grow and forces operations of the National Air Space (NAS) to keep pace with increased capacity. A key concept in air traffic flow management is balancing demand and capacity to maintain safety while minimizing delays. Accuracy of the wheels-off time estimate affects the quality of traffic demand estimates. Inaccurate traffic demand estimates can result in either inadequate or unneeded flow restrictions [1]. A large source of error in predicting the demand results from the prediction errors of the taxi times [2], which also plays a role in the prediction of the wheels-off time. Wheels-off time uncertainty is caused by uncertainty in gate departure time, ramp area transit time, and taxi-out time. The ramp area transit time is the time between push back from the gate and arrival at the taxiway spot, which is the ramp area exit point.

The purpose of this paper is to better understand time distributions of different processes within the ramp area transit time. This type of analysis is important because the uncertainty within the distributions should be accounted for in order to execute safe and efficient ramp area operations. Mean value estimates for the different processes can be used to schedule aircraft, however, many aircraft will deviate from the mean. This can force aircraft to slow down or stop along the route to avoid a loss of separation.

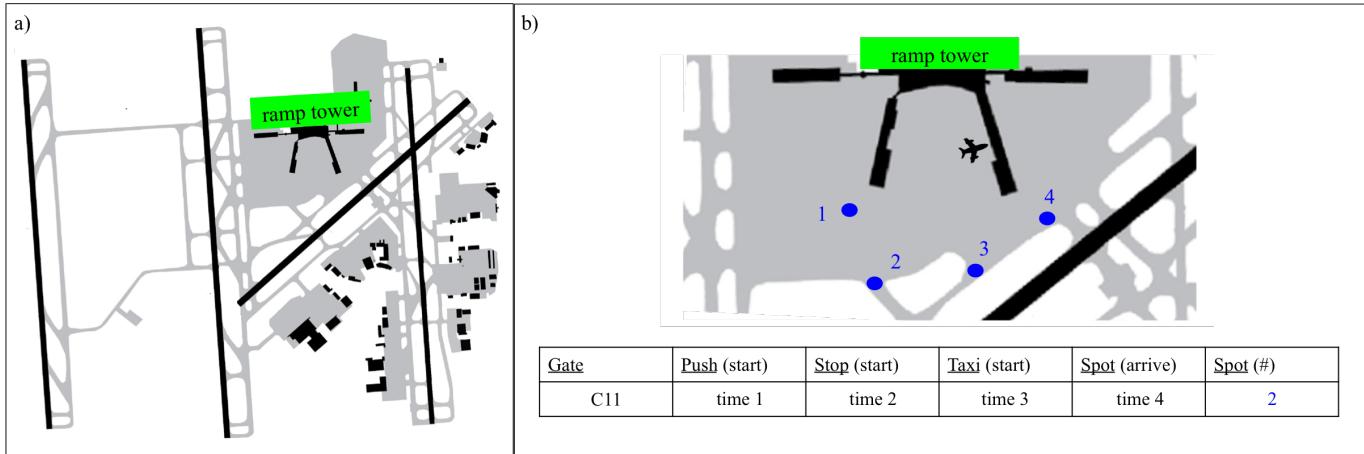
In contrast, we propose to use the full distribution for the different processes within the ramp area. This provides for schedules where aircraft can travel along their route *unimpeded* in the presence of other aircraft and trajectory uncertainties [3,4]. A recent study [5] estimated as much as 18% of fuel consumption during taxi operations was due to stop-and-go activity. The study also concluded that under the assumption of 15 knots or greater speed for all unimpeded aircraft, there is the potential to reduce overall fuel consumption on the surface by at least 21%.

\*AIAA Graduate Student Member, Graduate Student, Computer Engineering Department, UC Santa Cruz.

†Associate Professor, Computer Engineering Department, UC Santa Cruz

‡Research Scientist, University Affiliated Research Center, MS 210-8, Moffett Field, CA 94035.

§AIAA senior member and Aerospace Engineer, NASA Ames Research Center, MS 210-6, Moffett Field, CA 94035.



**Figure 1.** a) CLT airport surface. b) Zoomed in view of CLT south sector and illustration of the experiment set up. Data was collected by observer located in the ramp tower.

In this paper we analyze ramp area transit time data that was collected at the Charlotte Douglas (CLT) airport over a three day period. Using the collected data, we also infer conflict distributions representing combinations of push back times, for aircraft at different gates, that lead to a conflict in their sampled trajectory data. Understanding the conflict distribution is important because it plays a critical role in computing aircraft push back time windows that ensure conflict free trajectories [6] within the ramp area.

Unlike maneuvers on taxiways and runways, aircraft maneuvers within the ramp area are typically not confined to well-defined trajectories. The shape and timing of the trajectories are influenced by the pilot and are stochastic in nature. Most previous research focused on taxi-out times and did not analyze ramp area transit time. The taxi-out times were modeled using a Erlang and log-normal distribution [7–11] and the goodness-of-fit was also assessed [12, 13].

In our previous work [3, 4] based on a scaled-down robot experiment, we collected aircraft trajectory data related to the ramp area transit time. We used data from a robot experiment because trajectory data are not readily available mainly due to the lack of surveillance equipment in the ramp area. Moreover, investments in collecting ramp area trajectory data are unlikely unless the usefulness of the data in increasing airport efficiency is illustrated. We then modeled ramp area trajectories as a stochastic process with three discrete states: push back, stop, and taxi. The time spent in each discrete state was modeled as a gamma distribution and fitted to the robot experiment data.

In this paper, we explore not only the gamma distribution, but also a log-normal distribution to the collected airport operational data for the push back, stop and taxi processes. Since the Erlang distribution is a special case of the gamma distribution we do not consider it here. The goodness-of-fit of the distribution is assessed using hypothesis testing with multiple statistical tests.

We then sample a large number of ramp area trajectories using the fitted parametric distributions as input to a stochastic model of ramp area trajectories. The sampled trajectories are used to estimate conflict distributions defined by the time separation of two aircraft at the taxiway spot. These two-dimensional conflict distributions are analyzed and the goodness-of-fit of a multivariate Gaussian, Gaussian Copula, and t-Copula are considered.

This paper is organized as follows. In Section II, we describe the experiment that was performed at the CLT airport and report the raw data that was collected. Then in Section III we assess the goodness-of-fit of the log-normal and gamma distributions to the collected data. In Section IV we use the fitted distributions to sample ramp area trajectories and conflict distributions. Next, in Section V we analyze the goodness-of-fit of a multivariate Gaussian, Gaussian Copula, and t-Copula. Lastly, in Section VI we provide an overview of the work that was done and concluding remarks.

## II. Data Collection Methodology and Raw Collected Data

In this paper, we analyze data that was collected within the CLT south sector ramp area between August 23-25, 2015. The layout of the CLT airport is shown in Fig. 1a and a zoomed in view of the south sector ramp area is illustrated in Fig. 1b. The ramp tower is colored in red and provides the south sector ramp controllers with a view down the center alley.

The experimental data was collected by an observer located in the ramp tower. The observer had access to the controller radio frequency and was able to monitor communication between pilots and the ramp controller. This allowed for the observer to distinguish among trajectories that are allowed to proceed unimpeded and trajectories that are held by controllers for various reasons.

The goal of the experiment was to collect unimpeded trajectory data between the gate and various spots within the ramp area. Although both departure and arrival trajectory data were collected, in this paper we focus on the departure data. For a single departure trajectory, six pieces of data were collected. 1) We collected the gate number from which the aircraft trajectory begins. 2) We recorded the time that the aircraft push back is initiated. 3) We recorded the time that the stop phase of the trajectory is initiated. 4) We recorded the time that the taxi phase of the trajectory is initiated. 5) We recorded the time that the trajectory arrives at the spot. 6) We recorded the spot number that the trajectory arrived at. The four spots are indicated in Fig. 1b.

After the data was collected over the three day period, the data was processed which provides the time that each trajectory spends in discrete states push back, stop, and taxi. Figure 2 shows an example of the processed data that was collected. In the figure, each sub figure illustrates the histogram of data for different gates as well as different days. Within each subfigure there are 3 histograms which illustrate the histogram for the push back maneuver, the stop maneuver and the taxi maneuver from top to bottom respectively. Data from all gates is shown in the first column, data from the middle gates B6-B12 and C7-C13 is shown in the second column and data from the back gates B2-B4 and C3-C5 is shown in the second column. Data that was collected over all three days is shown in the first row, data collected on the first day is shown in the second row, data collected on the second day is shown in the third row and data collected on the third day is shown in the fourth row. In Fig. 3 we analyze the collected data from all gates over all days colored in red and in Fig. 4 we analyze the collected data from the middle gates over all the days colored in magenta.

## III. Statistical Testing of Collected 1-D Time Distributions

In this Section and Section V we consider the problem of comparing samples from two probability distributions  $f$  and  $g$  defined on the domain  $\mathcal{X}$  by proposing the null hypothesis  $H_0 : f(x) = g(x)$ , for every  $x \in \mathcal{X}$ . Statistical tests are implemented to test against the alternative hypothesis  $H_1 : f(x) \neq g(x)$ , for some  $x \in \mathcal{X}$ . We use the statistical tests to compare the collected data to known parametric distributions. If the collected data and the parametric distributions are a good fit within a 95% confidence interval then we will accept the null hypothesis, else we reject the null hypothesis.

### *Kolmogorov-Smirnov 1-Dimensional Test:*

The Kolmogorov-Smirnov(KS) [14,15] test is a 1-dimensional distribution free test. We analyze the empirical distribution CDF  $\mathbb{F}_n$  against the CDF of the parametric distribution  $\mathbb{G}$ . The KS statistic for empirical distribution  $\mathbb{F}_n$  with  $n$  iid observations

$$D_n = \sup_x |\mathbb{F}_n(x) - \mathbb{G}(x)| \quad (1)$$

By the Glivenko-Cantelli theorem [16], if the sample  $\mathbb{F}_n(x)$  comes from the distribution  $\mathbb{G}(x)$ , then  $D_n$  converges to 0 almost surely in the limit when  $n$  goes to infinity.

### *Kernel Method 2 Sample Test:*

The kernel method [17] is a non-parametric 2 sample test. The test statistic is the distance between the means of the two samples mapped into a reproducing kernel Hilbert space (RKHS). The statistical test is implemented based on a large deviation bound for the test statistic.

Given observations  $X := \{x_1, x_2, \dots, x_m\}$  and  $Y := \{y_1, y_2, \dots, y_n\}$  drawn independently and identically distributed (i.i.d.) from  $f$  and  $g$  respectively, let  $\mathcal{F}$  be a class of functions  $\hat{f} : \mathcal{X} \rightarrow \mathbb{R}$  and let  $f, g, X, Y$  be defined as above. We define the maximum mean discrepancy (MMD) and its empirical estimate as

$$\text{MMD}[\mathcal{F}, f, g] := \sup_{\hat{f} \in \mathcal{F}} \left( E_{x \sim f}[\hat{f}(x)] - E_{y \sim g}[\hat{f}(y)] \right) \quad (2)$$

$$\text{MMD}[\mathcal{F}, X, Y] := \sup_{\hat{f} \in \mathcal{F}} \left( \frac{1}{m} \sum_{i=1}^m f(x_i) - \frac{1}{n} \sum_{i=1}^n f(y_i) \right) \quad (3)$$

Let  $\mathcal{F}$  be a unit ball in a universal RKHS  $\mathcal{H}$  defined on the compact metric space  $\mathcal{X}$  with associated kernel function  $k(\cdot, \cdot)$ . Then  $\text{MMD}[\mathcal{F}, f, g] = 0$  if and only if  $f = g$ . Since  $E[\hat{f}(x)] = \langle \mu(f), \hat{f} \rangle$ , we may write

$$\text{MMD}[\mathcal{F}, f, g] = \sup_{\|\hat{f}\|_{\mathcal{H}} \leq 1} \langle \mu[f] - \mu[g], \hat{f} \rangle = \|\mu[f] - \mu[g]\|_{\mathcal{H}} \quad (4)$$

Using  $\mu[\mathcal{X}] := \frac{1}{m} \sum_{i=1}^m \phi(x)$  and  $k(x, x') = \langle \phi(x), \phi(x') \rangle$ , an empirical estimate of MMD becomes

$$\text{MMD}[\mathcal{F}, X, Y] = \left[ \frac{1}{m^2} \sum_{i,j=1}^m k(x_i, x_j) - \frac{2}{mn} \sum_{i,j=1}^{m,n} k(x_i, y_j) + \frac{1}{n^2} \sum_{i,j=1}^n k(y_i, y_j) \right] \quad (5)$$

Intuitively we expect  $\text{MMD}[\mathcal{F}, X, Y]$  to be small if  $f = g$ , and the quantity to be large if the distributions are far apart. The statistical test can be carried out using a large deviation bound with the acceptance region

$$\text{MMD}[\mathcal{F}, X, Y] < 2\sqrt{\frac{K}{m}} \left( 1 + \sqrt{\log \alpha^{-1}} \right) \quad (6)$$

where  $\alpha$  is the desired significance level of the test.

### **Quantile-Quantile (Q-Q) plot:**

The quantile-quantile (Q-Q) plot [18], is a visual goodness-of-fit test that can be applied to 1-D distributions. Corresponding to any ordinate value  $p$  there are two quantile values  $q_f(p)$  and  $q_g(p)$ . A Q-Q plot of samples from  $f$  and  $g$  is then just a scatter plot of  $q_f(p)$  versus  $q_g(p)$  for various  $p$ .

If  $f$  and  $g$  are drawn from the same distributions, then the plot of  $f$ -quantiles versus  $g$ -quantiles will be a straight line configuration with slope 1, pointed towards the origin. If  $g$  is a linear function of  $f$  then the corresponding Q-Q plot will still be linear but with possibly changed location and slope. It is this linear invariance property which has made the use of Q-Q plots appealing.

For the case in which the variables have heavy tails, the Q-Q plot tends to emphasize the comparative structure in the tails and to blur the distinctions in the 'middle' where the densities are high. The reason for this is that the quantile is a rapidly changing function of  $p$  when the density is sparse (in the tails) and a slowly changing function of  $p$  where the density is high (in the middle). [18].

### **Analysis of Collected 1-D Time Distributions**

In this section we analyze the data that was collected by evaluating the results of the KS test, the kernel two sample test and the Q-Q plot. Figure 3 and Fig. 4 show the results of the different statistical tests applied to aircraft from all gates and applied to aircraft from the middle Gates B6-B12 and C7-C13 respectively. Each figure contains 3 rows of figures where the top row analyzes the push back data, the middle row analyzes the stop data and the bottom row analyzes the taxi data.

Both Fig. 3 and Fig. 4 contain 3 columns of subfigures. In the first column we show the raw data with a histogram and also show the three different parametric distributions that were fitted to the data including a normal distribution, gamma distribution, and log-normal distribution.

The second column of subfigures represent the analysis of the three 1-D statistical tests which compare the collected data in each row to a gamma distribution. The third column of subfigures represent the analysis of the three 1-D tests which compare the collected data to a log-normal distribution. Within each subfigures in

column 2 and 3 we show the Q-Q plot, followed by the results of the kernel two sample test using confidence level  $\alpha = 0.05$ , followed by the results of the KS test using confidence level  $\alpha = 0.05$ . The statistical tests are run for 100 independently drawn samples of the two distributions.

For example in Fig. 3, the subfigure in row 1 and column 2 contains the results of the three statistical tests applied to the push back data for trajectories originating from all gates. In contrast, in Fig. 4, the subfigure in row 2 and column 2 contains the results of the three statistical tests applied to the stop data for trajectories originating from the middle gates B6-B12 and C7-C13.

The results in Fig. 3 and Fig. 4 appear to illustrate that there is a trade-off between the two distributions that we fit in column 2 (gamma distribution) and column 3 (log-normal distribution). The results of the KS test and kernel 2 sample test estimate that the log-normal distribution is a better fit for the collected data of push back and stop. The Q-Q plots, however, show a different result as the tails of Q-Q plot for the log-normal distribution are more nonlinear than the tails of the Q-Q plot for the gamma distribution. This implies the right tail of the log-normal distribution is heavier than the data that was collected. Although the log-normal distribution might do a good job at fitting the mean and CDF of the distribution, the gamma distribution appears to provide a better fit along the right tail of the distribution. Fitting the tails of the distribution is important because overestimating the density along the right tail of the distribution can impact ramp area throughput as the separation at the taxiway spot should be increased to accommodate for the greater uncertainty.

#### IV. Sampled Ramp Area Trajectories and Sampled Conflict Distributions

We begin this Section by defining the stochastic model of aircraft trajectories. The essence of the idea is to capture data influenced by the human operator and use the observed distributions as input to a stochastic model. We use the model to sample [19, 20] a large number of realistic trajectories. The sampled aircraft trajectories are used to generate a probabilistic measure of conflict within the ramp area. These distributions will be analyzed in Section V.

The discrete states  $i = 0, 1, \dots, 4$  defined as gate, push back, stop, taxi, and spot are the building blocks of our stochastic hybrid automaton model [21]. A single ramp area departure trajectory for aircraft  $i$  is described by our automaton where each discrete state is defined by the continuous time evolution:

For  $q = 0$ (gate),  $q = 2$ (stop),  $q = 4$ (spot):

$$dx^i = 0, \quad dy^i = 0, \quad d\theta^i = 0, \quad dv^i = 0 \quad (7)$$

For  $q = 1$ (push back):

$$dx^i = -v^i \cos(\theta^i)dt, \quad dy^i = -v^i \sin(\theta^i)dt, \quad d\theta^i = u_\theta^i(t)dt, \quad dv^i = u_v^i(t)dt + \sigma_v^i dW_v^i \quad (8)$$

where  $dW_v^i$  is an increment of a unit intensity Wiener process and  $\sigma_v^i$  is a scaling factor for the intensity of the variations in the velocity  $v^i$ . The control  $u_\theta^i(t)$  and  $u_v^i(t)$  defines the evolution of the heading angle and velocity of the deterministic component of the trajectory, respectively. The deterministic component of the trajectory is defined by the path aircraft follow within the NASA Ames Future Flight Central (FFC) [22] simulations of the CLT ramp area. Given that we are interested in sampling unimpeded trajectories, we believe this provides a good representation of the path a pilot would follow in the absence of other aircraft.

For  $q = 3$ (taxi):

$$dx^i = v^i \cos(\theta^i)dt, \quad dy^i = v^i \sin(\theta^i)dt, \quad d\theta^i = u_\theta^i(t)dt + \sigma_\theta^i dW_\theta^i, \quad dv^i = u_v^i(t)dt + \sigma_v^i dW_v^i \quad (9)$$

where  $dW_\theta^i$  is an increment of a unit intensity Wiener process and  $\sigma_\theta^i$  is a scaling factor for the intensity of the variations in the heading angle  $\theta^i$ . The introduction of the terms  $dW_v^i$  and  $dW_\theta^i$  in the equations transforms the otherwise deterministic system of equations into a stochastic process.

We use the data that was captured during the three day experiment at CLT to provide a distribution over the continuous interval of time that we could spend in any given discrete state. We use these distributions of time to fit parametric distributions for push back, stop and taxi. Sampling from these distributions provides the time each trajectory should spend within each discrete state. We then sample a large number of realistic trajectories, see Fig. 5. In the first column we illustrate the evolution of trajectories in time and

---

**Algorithm 1** Conflict Distribution: Aircraft  $i$  vs Aircraft  $j$ 

---

```

Assume  $t_i = 0$ 
Set N = 1,000
for  $t_j = -200:1:200$  do
  for  $k = 1:N$  do
    • Randomly sample aircraft  $i$  and  $j$  from their respective family of trajectories.
    • Measure the spatial proximity of the aircraft along the route and provide a conflict flag
      if aircraft lose spatial separation.
  end for
  • Return conflict ratio for the relative schedule  $t_j - t_i$ 
end for
• Return conflict ratio for all relative schedules at a resolution of 1 second.

```

---

in the second column we show the distribution of the different discrete states push back, stop and taxi. The distributions of push back, stop and taxi are color coded to represent the relationship to the middle and front gate distributions of Fig. 2, Fig. 3, and Fig. 4. The sample distributions of stop and taxi were accepted by the K-S test and the kernel 2 sample test when analyzed for goodness-of-fit. The sample distribution of push back were accepted by the K-S test and rejected for the kernel 2 sample test.

After sampling trajectories, we estimate the probability density function for trajectory duration of aircraft  $i$  in the absence of any other aircraft in the ramp area. We refer to this type of distribution as *natural* since the aircraft is unimpeded. We are interested in computing push back windows for aircraft  $i$  such that the aircraft arrives at the terminal node at the scheduled time  $t_i$ . Therefore, we enforce a terminal condition in time for the sampled trajectories and this generates a distribution for the push back time. In addition, enforcing this terminal condition in time provides us with a set of departing trajectories that all enter the FAA controlled taxiway at the same time.

Using the family of trajectories defined by the natural distributions of aircraft  $i$  and  $j$ , we generate a measure of conflict defined by the ratio of conflicting trajectories to total trajectories. We compute the measure of conflict by fixing the terminal time of aircraft  $i$  in time such that  $t_i = 0$ . Next we fix the terminal time of aircraft  $j$  in time, e.g.  $t_j = -200$ . Given the relative schedule defined by  $t_j - t_i$ , there exists a family of trajectories for both aircraft  $i$  and  $j$  that push back from their respective gates and taxi to the terminal node as required.

For the relative schedule  $t_j - t_i$ , we sample a single trajectory from the family of trajectories for aircraft  $i$  and  $j$ , measure their spatial proximity along the route, and provide a conflict flag if the aircraft lose spatial separation. If we continue this process of randomly sampling from the family of trajectories with fixed terminal times, we compute a conflict ratio for the relative separation in time at the taxiway spot, see Algorithm 1. The fixed terminal times are considered for every whole second and the estimated conflict distribution provides a measure of conflict at a resolution of 1 second, see Fig. 6a.

## V. Statistical Testing of Sampled Two-Dimensional Conflict Distributions

In this Section we use the hypothesis testing method described in Section III to assess the goodness-of-fit of a multivariate Gaussian, Gaussian Copula, and t-Copula to the sampled distribution of red conflict points seen in Fig. 6b. We begin this section by describing the two new test statistics derived from the Wald-Wolfowitz Runs Test and the Kolmogorov Smirnov Test which we use to analyze the fit of the two-dimensional distributions.

### **Wald-Wolfowitz Two-Dimensional Runs Test:**

The Wald-Wolfowitz runs test [23] is a 1-dimensional distribution free test. Consider two random samples  $F \sim f$  and  $G \sim g$ . A run of a sequence is defined as a maximal non-empty segment of the sequence consisting of adjacent equal elements. For example the ranked list distributed as

$$FFFFGGFF$$

consists of 3 runs, 2 of which are  $F$  and the other  $G$ . Under the null hypothesis  $H_0 : f = g$ , the number of runs in a sequence of  $N$  elements is a random variable whose conditional distribution given  $n$  observations

$F$  and  $m$  observations  $G$  is approximately normal with

$$\mu = \frac{2 * m * n}{N} + 1 \quad (10)$$

$$\sigma^2 = \frac{(\mu - 1)(\mu - 2)}{N - 1} \quad (11)$$

This test can be extended to  $d$ -dimensions with the multivariate runs test proposed by [24] and later shown to be consistent against the general alternative hypothesis  $H_1$  [25]. The multivariate runs test extends the notion of a run to  $d$ -dimensional space by constructing the minimum spanning tree (MST) of the data set  $F \cup G$ . After the MST has been constructed, we eliminate all edges where the vertices of the edge are from different families of the data. We define the multivariate run test statistic

$$R = 1 + \text{Number of cross matches} = \text{Number of disjoint trees} \quad (12)$$

Under the null hypothesis  $H_0$  the expected value of  $R$  and variance can be derived through a combinatorial argument as

$$E[R] = \frac{2mn}{N} + 1 \quad (13)$$

$$Var[R] = \frac{2mn}{N(N-1)} \left[ \frac{2mn-N}{N} + \frac{C-N+2}{(N-2)(N-3)} [N(N-1)-mn+2] \right] \quad (14)$$

where  $C$  is the number of edge pairs that share a common node in the MST.

Define the new statistic

$$W = \frac{R - E[R]}{(Var[R])^{\frac{1}{2}}} \quad (15)$$

Under the null hypothesis  $H_0 : f = g$ , the test statistic  $W$  approaches the standard normal distribution. Similar tests can be run where we use optimal non-bipartite matching [26], matching based on minimum energy [27], or nearest neighbors [28].

### Kolmogorov-Smirnov Two-Dimensional Test:

Extending this test from 1-dimension to  $d$ -dimensions is not straightforward. The KS test requires the definition of a probability function that is independent of the direction of ordering, which is not possible given that there are  $2^d - 1$  ways of defining a CDF in a  $d$ -dimensional space [29]. Furthermore, tests based on binning face the hurdle of "the curse of dimensionality": a high dimensional space is mostly empty and binning tests can only start to be effective when the data sets are very large [30]. To address this, authors [31, 32] defined the statistic independent of any particular ordering by finding the largest difference under all possible orderings.

### Analysis of Goodness-of-Fit for Two-Dimensional Conflict Distributions

Now we analyze the distribution of red conflict points illustrated in Fig. 6b. This type of analysis is important because understanding the conflict distributions help us better understand the risk of conflicts. From a visual perspective, the conflict distributions for some values such as  $t_j - t_i = 0$  or  $t_j - t_i = 20$  seem "regular" and familiar. For some relative schedules, such as  $t_j - t_i = -20$  or  $t_j - t_i = 150$ , the distributions exhibit odd shapes and skew that do not appear so "regular". These difference in distribution could be reasonable given the difference in ratio of conflicts for the two schedules.

In order to assess the goodness-of-fit of the conflict distributions and potential qualitative differences, we select the conflict distributions for  $t_j - t_i = 20$  and  $t_j - t_i = -20$  seen in Fig. 6b. We analyze the goodness-of-fit of a multivariate Gaussian, Gaussian Copula, and t-Copula distribution to the samples. Figure 7 shows the results of the hypothesis testing applied to the two distributions. In the figure, the first row of subfigures are the statistical tests applied to the conflict distribution  $t_j - t_i = -20$  and the second row are the tests applied to the conflict distribution  $t_j - t_i = 20$ . The first column is assessing the fit of the multivariate Gaussian, the second column assessing the fit of the Gaussian Copula, and the third column assessing the fit of the t-Copula.

Within each subfigure the top image shows the distribution of the Wald-Wolfowitz test statistic  $W$  for 100 independent tests. Below the distribution is the binary decision for the Wald-Wolfowitz test to reject the null hypothesis  $H_0$  for a confidence interval  $\alpha = 0.05$ . Below that we have the binary decision for the K-S test to reject the null hypothesis  $H_0$  for a confidence interval  $\alpha = 0.05$ . Lastly we illustrate the kernel threshold and the kernel statistic for the kernel two sample test for with confidence level  $\alpha = 0.05$ .

The second row of subfigures within Fig. 7 analyzes the conflict distributions for  $t_j - t_i = 20$ . As can be seen in the figure, the multivariate Gaussian does not seem to be a good fit for the distribution. The Wald-Wolfowitz test statistic has a mean less than  $-1$  which indicates that we are getting less cross-matches on average than we should if the null hypothesis  $H_0$  were true. For the Gaussian Copula and the t-Copula the test statistic  $W$  mean is approaching 0, indicating that we are getting the expected number of cross-matches if the null hypothesis  $H_0$  were true. The distributions of the test statistic  $W$  seem different for the Gaussian Copula and t-Copula. Given that the test statistic  $W$  should approach the standard normal, this could indicate the Gaussian Copula is a better fit.

Next, consider the first row of subfigures within Fig. 7 which analyze the conflict distributions for  $t_j - t_i = -20$ . As can be seen, the Wald-Wolfowitz test statistics  $W$  is rejected for a large number of test for the relative schedule  $t_j - t_i = -20$  regardless of the distribution that is being fit to the data. This indicates that the conflict distributions are statistically different from the parametric distributions we are considering. This is in contrast to the K-S test statistic and kernel two sample test statistic which are accepted for a large number of tests for the Gaussian Copula and t-Copula distributions, but not the multivariate Gaussian. This indicates that the multivariate Gaussian is not a good fit, but the Gaussian Copula and t-Copula may be a good fit.

To get a better understanding why the Wald-Wolfowitz test statistic  $W$  would reject tests when the K-S and kernel two sample test accept the test, we visually analyze the distributions. The first row of subfigures in Fig. 8 show the sampled conflict points for  $t_j - t_i = -20$  in red and samples from the fitted distributions in blue. Compare the first row of subfigures to the second row of subfigures, which are defined the same way for the conflict distribution  $t_j - t_i = 20$ .

From Fig. 8 it seems like there may be a reason that the Wald-Wolfowitz test rejects the samples for  $t_j - t_i = -20$  but accept samples for  $t_j - t_i = 20$ . For the figures analyzing distributions from  $t_j - t_i = 20$  the concentration of blue and red points seems rather uniform. There are no large concentrations of blue or red points throughout the domain, instead they are equally distributed amongst themselves. In contrast, the figures analyzing the distributions from  $t_j - t_i = -20$  have higher concentrations of blue or red points where individual colonies seem to appear. The Wald-Wolfowitz test statistic  $W$  is able to catch on to this difference in concentration and reject the tests.

Overall, this analysis suggests that the conflict distributions can be qualitatively different depending on the conflict ratio. For conflict distributions that have a high ratio of conflicts the distributions seem more “regular” and this notion was confirmed when we found that the Gaussian Copula and t-Copula distributions are strong candidates to fit the distributions. For conflict distributions that have a low ratio of conflicts the distributions do not seem so “regular” and this notion was confirmed with one of the statistical tests, but was not supported by two other statistical tests.

This contradiction between statistical tests introduces an interesting question. What is the most appropriate metric for our purposes to assess the goodness-of-fit of a two-dimensional distribution? Whereas the K-S test and kernel two sample test can confirm that the mean and CDF are a good fit, we know that the tails may be poorly accounted for. Intuitively, the Wald-Wolfowitz Runs test seems like a more natural candidate in two-dimensions and our results showed the test to distinguish between two distributions more stringently.

Understanding the conflict distribution is important because it plays a critical role in computing aircraft push back time windows that ensure conflict free trajectories [6]. For example, if we underestimate the density along the tails of the conflict distributions this can lead to unexpected conflicts in the ramp area and a greater likelihood of aircraft having to slow down or stop to avoid a loss of separation. Moreover, knowing the conflict distributions are a good fit to a parametric distribution can help us derive confidence intervals related to the push back windows that are ultimately computed.

## VI. Conclusion

In this paper, we analyzed the time duration distributions of different processes within the ramp area transit time. The processes that we consider are push back, stop, and taxi. The analyzed data were collected between August 23-25, 2015 at the Charlotte Douglas (CLT) airport. Specifically, we assessed the goodness-of-fit of a gamma distribution and log-normal distribution to the different processes.

The analysis of data shows that the log-normal and gamma distributions are reasonable fit to push back, stop, and taxi. The analysis also show that there exists a trade-off between the two distributions. Whereas the log-normal may better fit the mean and CDF of the distribution, the gamma distribution provides a better fit along the right tail of the distribution.

Next, we analyzed the goodness-of-fit of a multivariate Gaussian, Gaussian Copula, and t-Copula distribution to the sampled two-dimensional conflict distributions. The analysis showed that the conflict distributions may be qualitatively different depending upon the conflict ratio. For conflict distributions with a high ratio of conflicts, the Gaussian Copula and t-Copula are reasonable fits. For conflict distributions with a low ratio of conflicts, the statistical tests contradict each other and it is not clear if these distributions are a good fit.

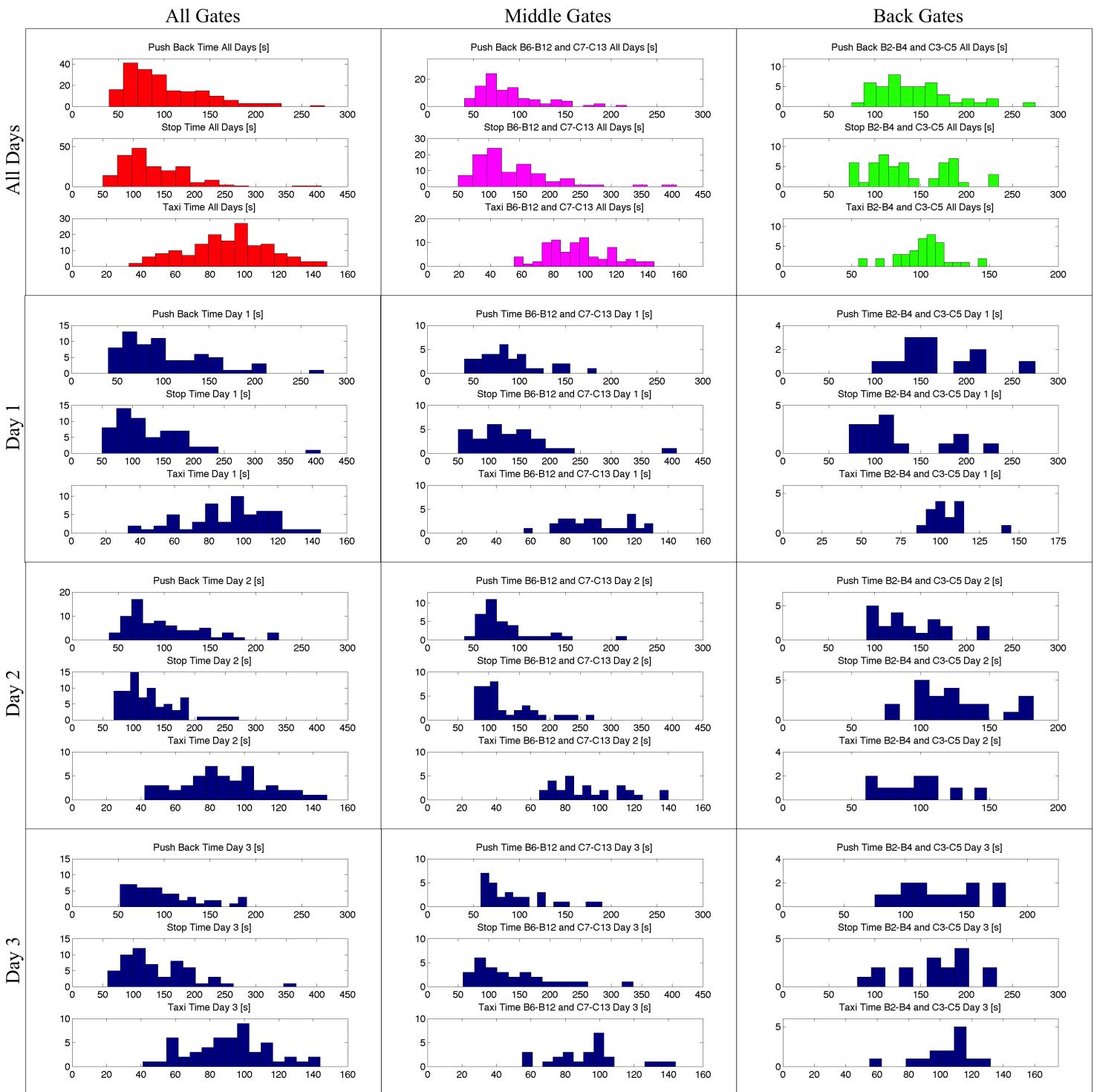
The results show that parametric distributions can be a good fit for the duration of the different processes within the ramp area transit time and the two-dimensional conflict distributions. This is important because the better that we understand these distributions, the more we can anticipate the uncertainty that is intrinsic within them. Furthermore, knowing the conflict distributions are a good fit to a parametric distribution can help us derive confidence intervals related to the push back time windows that are eventually computed. Future work will use more ramp area data to analyze whether the distributions of the different processes within the ramp area are time dependent or stationary processes.

Understanding the uncertainty within the distributions is critical for the safe and efficient execution of ramp area operations. For example, overestimating the right tail of the distributions for the different processes within ramp area transit time can impact throughput as the separation at the taxiway spot should be increased to accommodate for the greater uncertainty. Similarly, if we underestimate tails of the conflict distributions this can lead to conflicts that were not anticipated and a greater likelihood of aircraft having to slow down or stop to avoid a loss of separation. Ultimately, these distributions are important because they play a critical role in computing push back time windows that ensure aircraft can travel unimpeded from their gate to the departure runway queue in the presence of other aircraft and trajectory uncertainties.

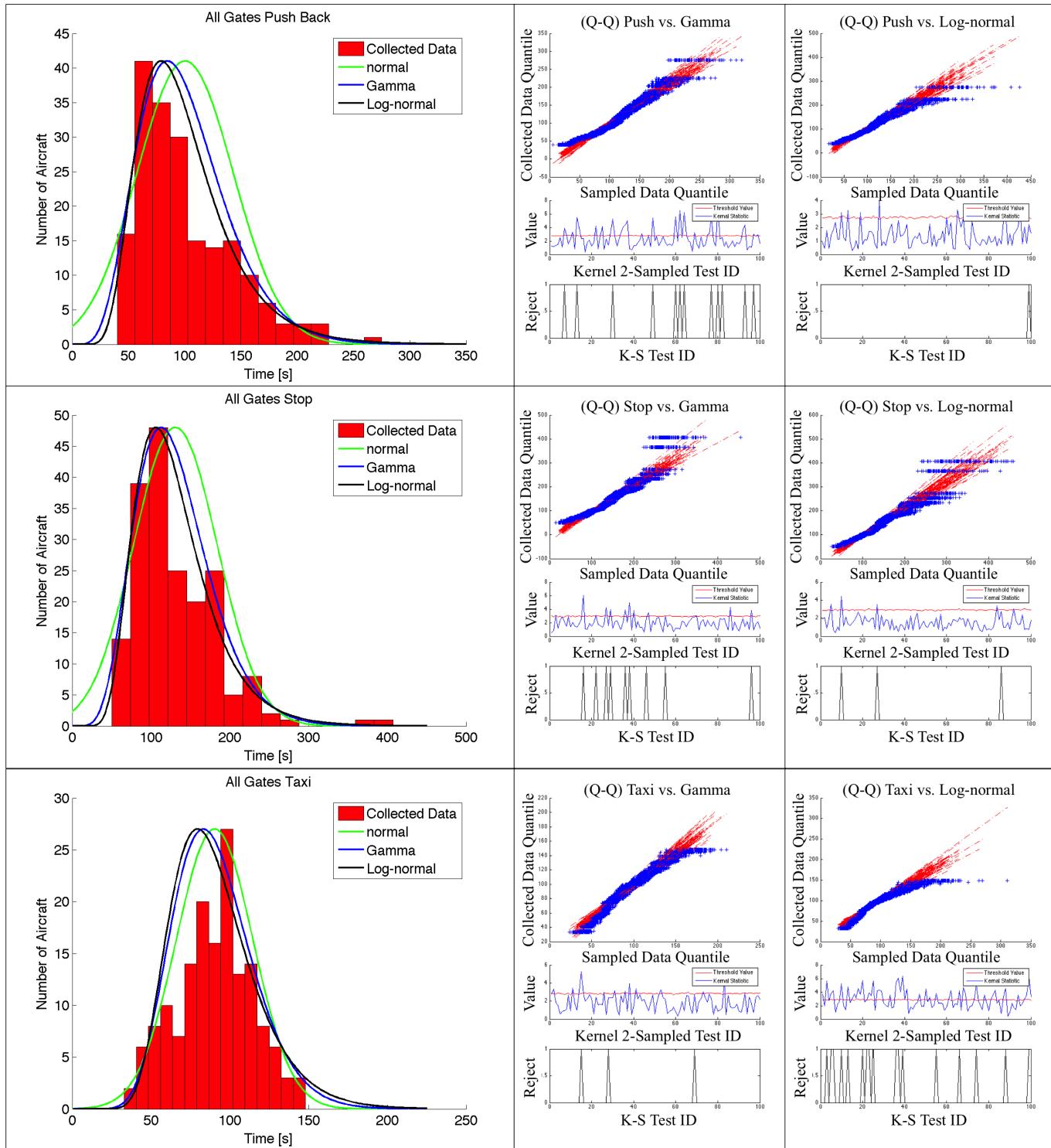
## References

- <sup>1</sup>Chatterji, G. B. and Zheng, Y., "Wheels-Off Time Prediction Using Surface Traffic Metrics," *12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, 2012, p. 5699.
- <sup>2</sup>Srivastava, A., "Improving Departure Taxi Time Predictions Using ASDE-X Surveillance Data," *Digital Avionics Systems Conference (DASC), 2011 IEEE/AIAA 30th*, IEEE, 2011, pp. 2B5-1.
- <sup>3</sup>Coupe, W. J., Milutinović, D., Malik, W., Gupta, G., and Jung, Y., "Robot Experiment Analysis of Airport Ramp Area Time Constraints," *AIAA Guidance, Navigation, and Control Conference (GNC)*, Boston, MA, 2013.
- <sup>4</sup>Coupe, W. J., Milutinović, D., Malik, W., and Jung, Y., "Integration of Uncertain Ramp Area Aircraft Trajectories and Generation of Optimal Taxiway Schedules at Charlotte Douglas (CLT) Airport," *AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*, Dallas, TX, 2015.
- <sup>5</sup>Nikoleris, T., Gupta, G., and Kistler, M., "Detailed Estimation of Fuel Consumption and Emissions During Aircraft Taxi Operations at Dallas Fort Worth International Airport," *Published in the journal Transportation Research Part D: Transport and Environment*, Vol. 16D, Issue 4, June 2011.
- <sup>6</sup>Coupe, W. J., Milutinović, D., Malik, W., and Jung, Y., "Optimization of Push Back Time Windows That Ensure Conflict Free Ramp Area Aircraft Trajectories," *AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*, Dallas, TX, 2015.
- <sup>7</sup>Hansen, M., Nikoleris, T., Lovell, D., Vlachou, K., and Odoni, A., "Use of Queuing Models to Estimate Delay Savings from 4D Trajectory Precision," *Eighth USA/Europe Air Traffic Management Research and Development Seminar*, 2009.
- <sup>8</sup>Hengsbach, G. and Odoni, A. R., "Time Dependent Estimates of Delays and Delay Costs at Major Airports," Tech. rep., Cambridge, Mass.: MIT, Dept. of Aeronautics & Astronautics, Flight Transportation Laboratory, 1975, 1975.
- <sup>9</sup>Simaikakis, I. and Balakrishnan, H., "A Queuing Model of the Airport Departure Process," *Transportation Science*, 2011.
- <sup>10</sup>Anderson, K., Carr, F., Feron, E., and Hall, W., "Analysis and Modeling of Ground Operations at Hub Airports," 2000.
- <sup>11</sup>Khadilkar, H. and Balakrishnan, H., "Network Congestion Control of Airport Surface Operations," *Journal of Guidance, Control, and Dynamics*, Vol. 37, No. 3, 2014, pp. 933–940.
- <sup>12</sup>Simaikakis, I. and Balakrishnan, H., "Probabilistic Modeling of Runway Interdeparture Times," *Journal of Guidance, Control, and Dynamics*, Vol. 37, No. 6, 2014, pp. 2044–2048.

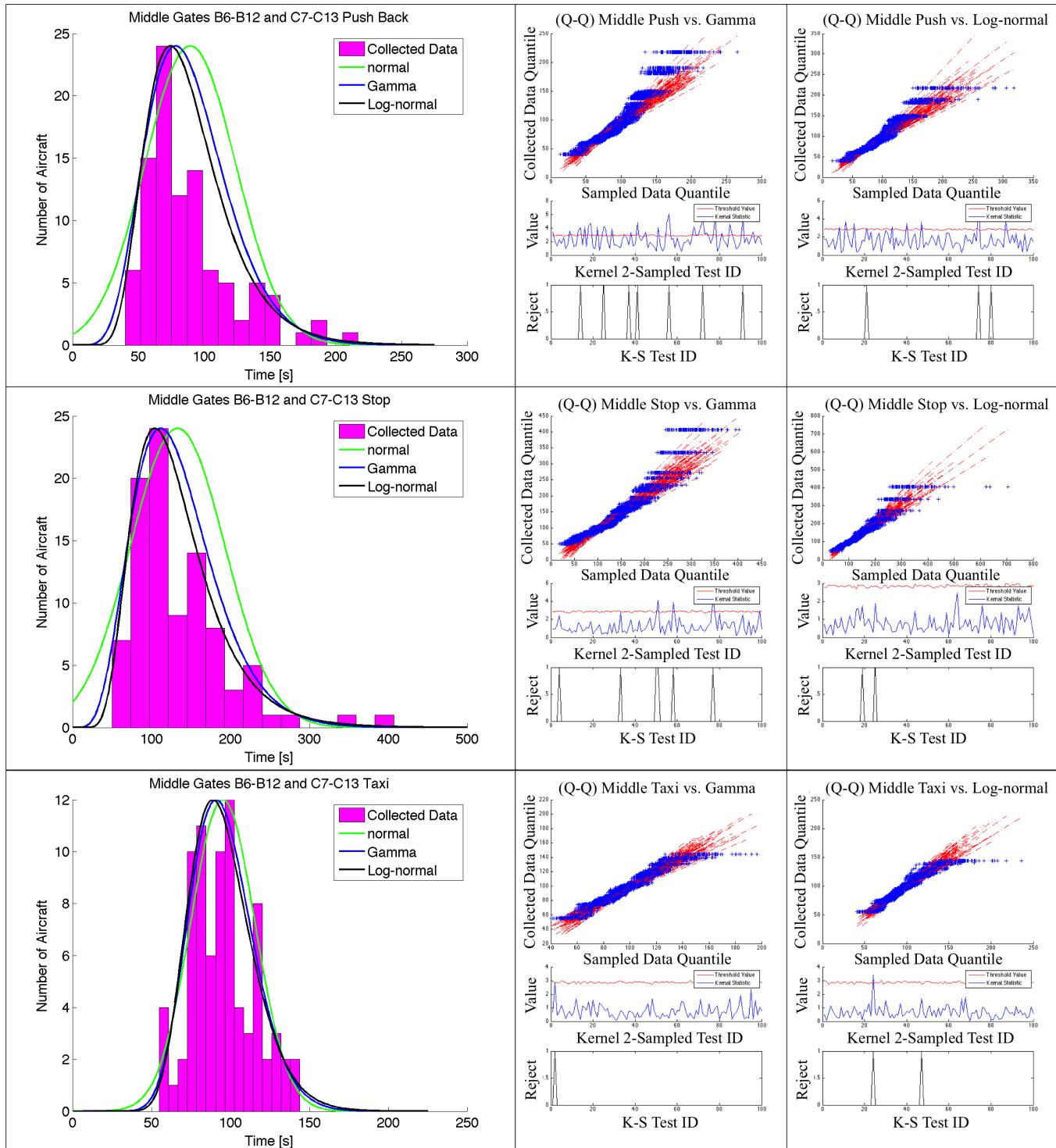
- <sup>13</sup>Simaiakis, I. and Pyrgiotis, N., "An Analytical Queuing Model of Airport Departure Processes for Taxi Out Time Prediction," *AIAA Aviation Technology, Integration and Operations (ATIO) Conference*, 2010.
- <sup>14</sup>Smirnov, N. V., "On the Estimation of the Discrepancy Between Empirical Curves of Distribution for Two Independent Samples," *Bull. Math. Univ. Moscou*, Vol. 2, No. 2, 1939.
- <sup>15</sup>Massey Jr, F. J., "The Kolmogorov-Smirnov Test for Goodness of Fit," *Journal of the American statistical Association*, Vol. 46, No. 253, 1951, pp. 68–78.
- <sup>16</sup>Wellner, J. A. et al., "A Glivenko-Cantelli Theorem and Strong Laws of Large Numbers for Functions of Order Statistics," *The Annals of Statistics*, Vol. 5, No. 3, 1977, pp. 473–480.
- <sup>17</sup>Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., and Smola, A., "A Kernel Two-sample Test," *The Journal of Machine Learning Research*, Vol. 13, No. 1, 2012, pp. 723–773.
- <sup>18</sup>Wilk, M. B. and Gnanadesikan, R., "Probability Plotting Methods for the Analysis of Data," *Biometrika*, Vol. 55, No. 1, 1968, pp. 1–17.
- <sup>19</sup>Milutinović, D., "Utilizing Stochastic Process for Computing Distributions of Large-Size Robot Population Optimal Centralized Control," *10th International Symposium on Distributed Autonomous Robotic Systems (DARS)*, Lausanne, Switzerland, 2010.
- <sup>20</sup>Gillespie, D. T., "Exact Stochastic Simulation of Coupled Chemical Reactions," *The Journal of Physical Chemistry*, Vol. 81, 1977.
- <sup>21</sup>Milutinović, D. and Lima, P., *Cells and Robots: Modeling and Control of Large-Size Agent Populations*, Springer, 2007.
- <sup>22</sup>Dorighi, N. and Sullivan, B., "Future Flight Central: A Revolutionary Air Traffic Control Tower Simulation Facility," *AIAA Modeling and Simulation Technologies Conference and Exhibit*, 2003.
- <sup>23</sup>Wald, A. and Wolfowitz, J., "On a Test Whether Two Samples Are from the Same Population," *The Annals of Mathematical Statistics*, Vol. 11, No. 2, 1940, pp. 147–162.
- <sup>24</sup>Friedman, J. H. and Rafsky, L. C., "Multivariate Generalizations of the Wald-Wolfowitz and Smirnov Two-sample Tests," *The Annals of Statistics*, 1979, pp. 697–717.
- <sup>25</sup>Henze, N. and Penrose, M. D., "On the Multivariate Runs Test," *Annals of statistics*, 1999, pp. 290–298.
- <sup>26</sup>Rosenbaum, P. R., "An Exact Distribution-free Test Comparing Two Multivariate Distributions Based on Adjacency," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 67, No. 4, 2005, pp. 515–530.
- <sup>27</sup>Aslan, B. and Zech, G., "New Test for the Multivariate Two-sample Problem Based on the Concept of Minimum Energy," *Journal of Statistical Computation and Simulation*, Vol. 75, No. 2, 2005, pp. 109–119.
- <sup>28</sup>Schilling, M. F., "Multivariate Two-sample Tests Based on Nearest Neighbors," *Journal of the American Statistical Association*, Vol. 81, No. 395, 1986, pp. 799–806.
- <sup>29</sup>Lopes, R. H., Hobson, P. R., and Reid, I. D., "Computationally Efficient Algorithms for the Two-dimensional Kolmogorov-Smirnov Test," *Journal of Physics: Conference Series*, Vol. 119, IOP Publishing, 2008, p. 042019.
- <sup>30</sup>Scott, D. W., *Multivariate Density Estimation: Theory, Practice, and Visualization*, John Wiley & Sons, 2015.
- <sup>31</sup>Bickel, P. J., "A Distribution Free Version of the Smirnov Two Sample Test in the P-variate Case," *The Annals of Mathematical Statistics*, 1969, pp. 1–23.
- <sup>32</sup>Peacock, J., "Two-dimensional Goodness-of-fit Testing in Astronomy," *Monthly Notices of the Royal Astronomical Society*, Vol. 202, No. 3, 1983, pp. 615–627.



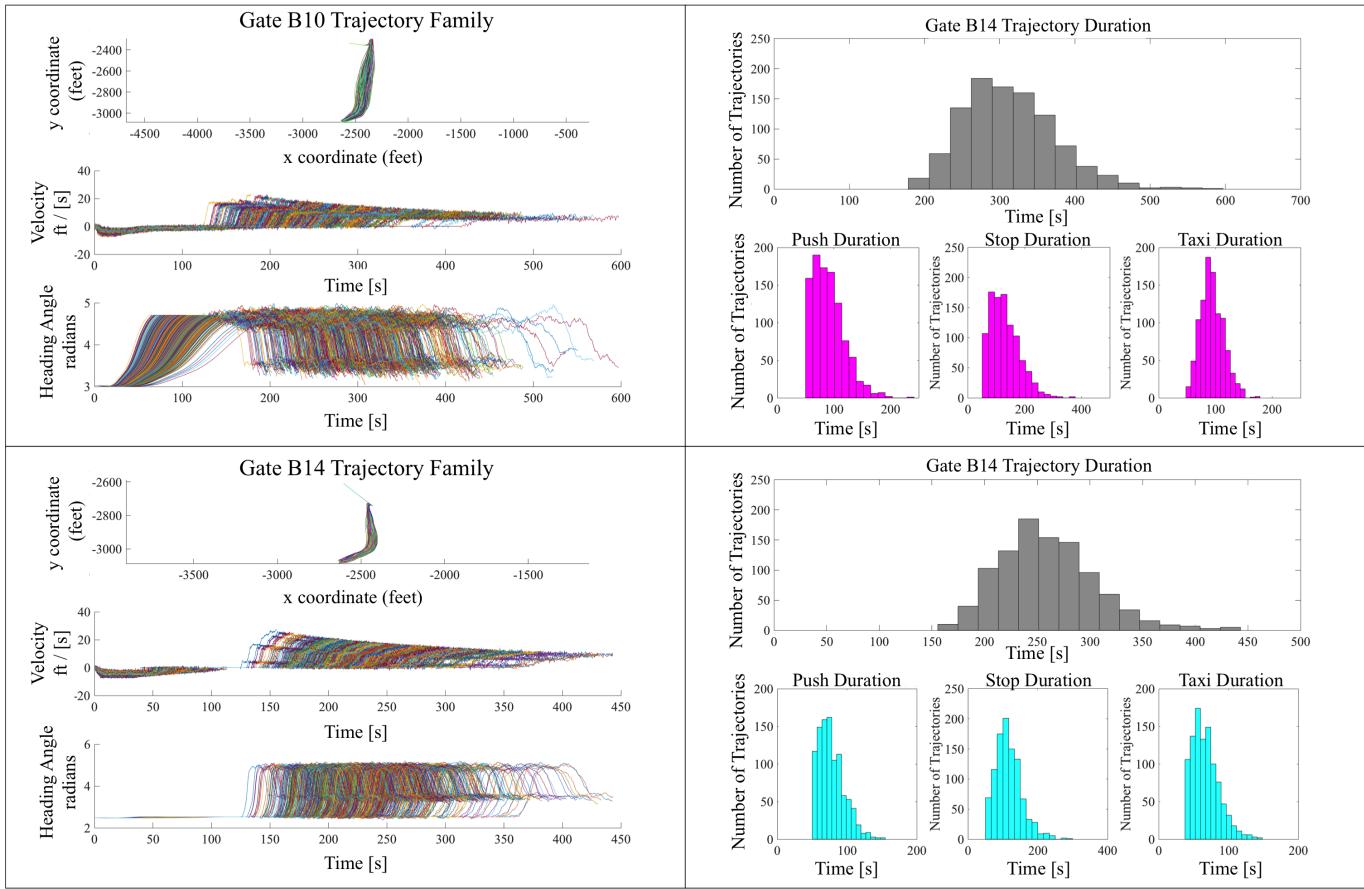
**Figure 2.** The processed data is illustrated using histograms. The *x*-axis represents the time spent in seconds to complete each process and the *y*-axis represents the number of aircraft within each bin. Data from all gates is shown in the first column, data from the middle gates B6-B12 and C7-C13 is shown in the second column and data from the back gates B2-B4 and C3-C5 is shown in the third column. Data that was collected over all three days is shown in the first row, data collected on the first day is shown in the second row, data collected on the second day is shown in the third row and data collected on the third day is shown in the fourth row.



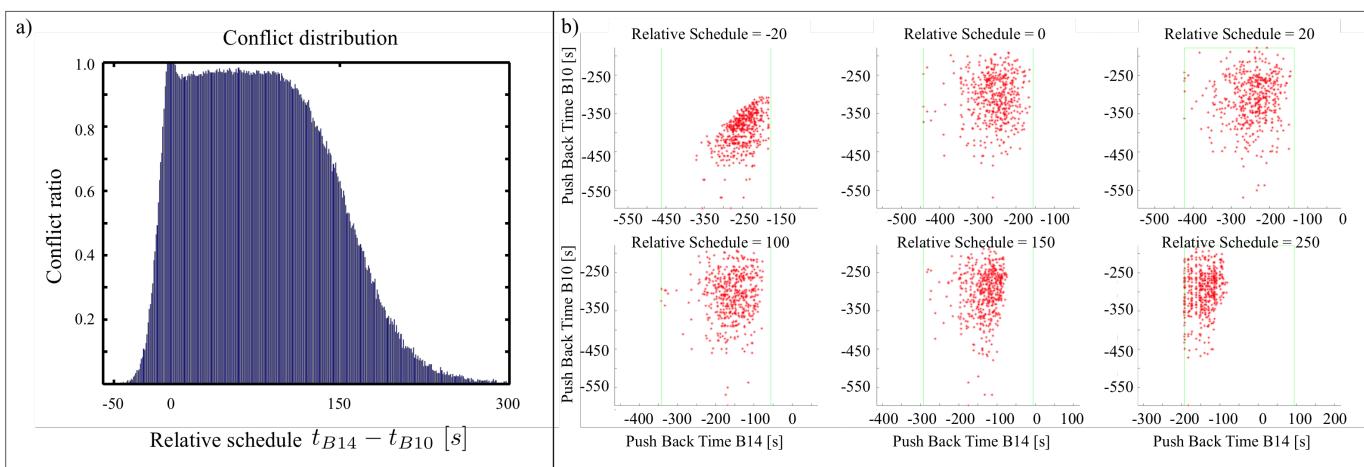
**Figure 3.** Analysis of collected data from all gates over all days. The push back data is in the first row, the stop data the second row, and the taxi data the third row. The first column shows the histogram of data and the fitted distributions, the second column show the results of the three different statistical tests assessing the goodness-of-fit of the gamma distribution to the collected data, and the third column show the results of the three different statistical tests assessing the goodness-of-fit of the log-normal distribution to the collected data.



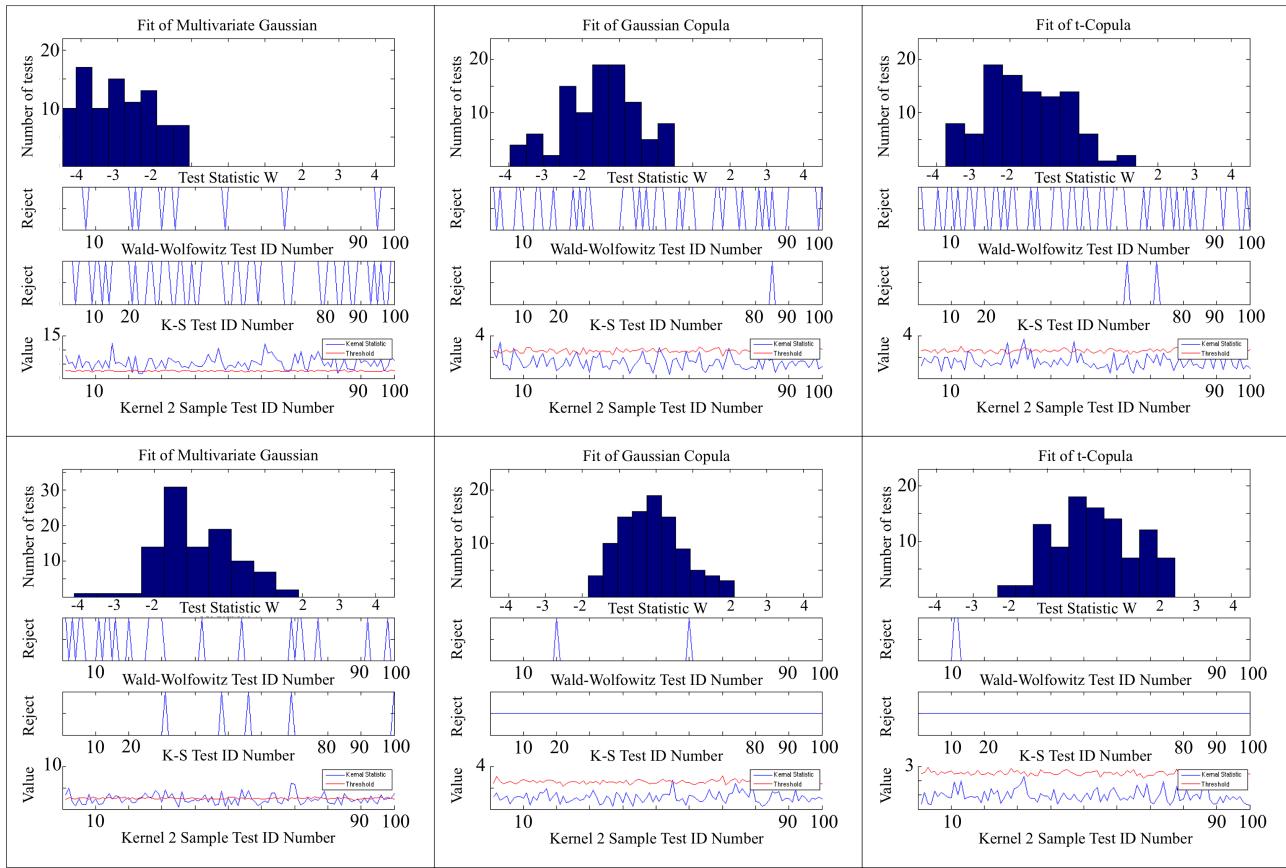
**Figure 4.** Analysis of collected data from middle gates B6-B12 and C7-C13. The push back data is in the first row, the stop data the second row, and the taxi data the third row. The first column shows the histogram of data and the fitted distributions, the second column show the results of the three different statistical tests assessing the goodness-of-fit of the gamma distribution to the collected data, and the third column show the results of the three different statistical tests assessing the goodness-of-fit of the log-normal distribution to the collected data.



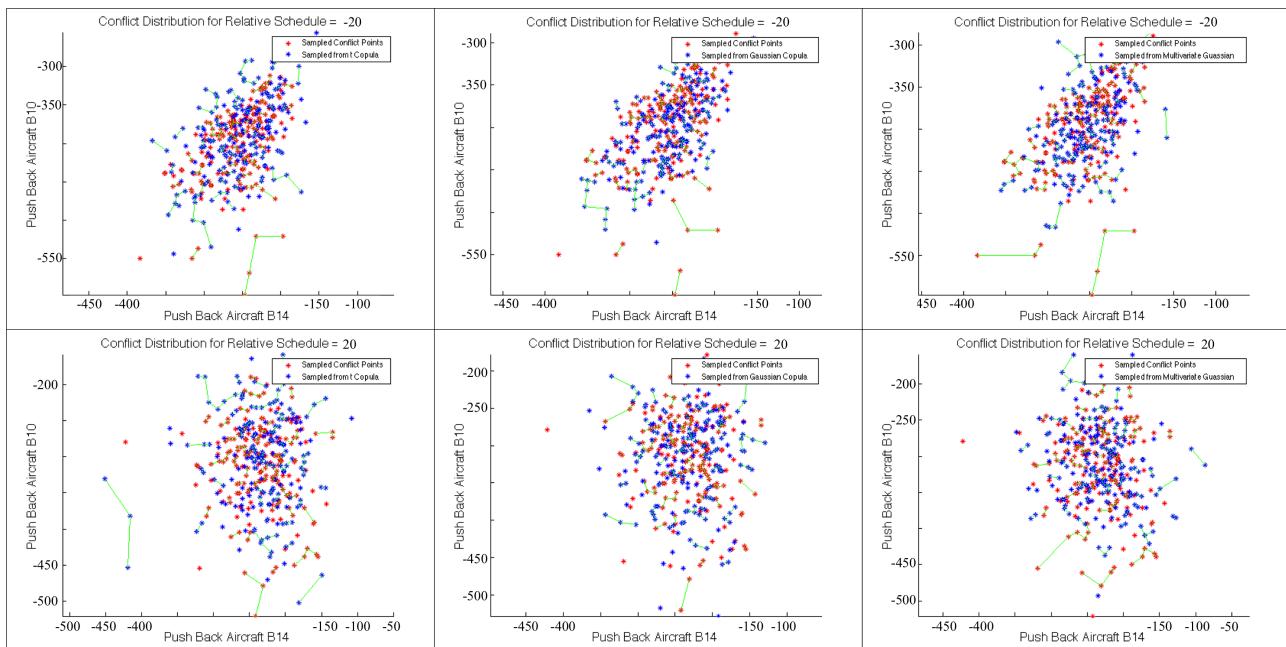
**Figure 5.** Sampled trajectories from the stochastic model that was described in Section 5. The sampled trajectories from gate B10 are shown in the first row and the sampled trajectories from gate B14 are shown in the second row. In the first column we illustrate the evolution of trajectories in time and in the second column we show the distribution of the different discrete states push back, stop and taxi. The distributions of push back, stop and taxi are color coded to represent the relationship to the middle and front gate distributions of Fig. 2, Fig. 3, and Fig. 4. The sample distributions of stop and taxi were accepted by both the K-S test and the kernel 2 sample test when analyzed for goodness-of-fit. The sample distribution of push back were accepted by the K-S test and rejected for the kernel 2 sample test.



**Figure 6.** a) Sampled conflict distribution estimated from the conflict ratio defined by the relative taxiway schedule  $t_{B14} - t_{B10}$ . b) Two-dimensional conflict distributions for different values of the relative schedule ranging from  $t_{B14} - t_{B10} = -20$  to  $t_{B14} - t_{B10} = 250$ . The shape and structure of the conflict distributions appear to be different for different values of the conflict ratio.



**Figure 7. Analysis of two-dimensional conflict distributions.** The first row analyzes the sampled distribution  $t_{B14} - t_{B10} = -20$  and the second row assesses the goodness-of-fit to the sampled distribution  $t_{B14} - t_{B10} = 20$ . The first column assesses the goodness-of-fit of a multi-variate distribution, the second column assesses the goodness-of-fit of a Gaussian Copula, and third column assess the goodness-of-fit of a t-Copula.



**Figure 8. Analysis of the sampled conflict points (red) with the samples from the parametric distribution (blue) and the edges of the MST that do not include cross matches (green).** The first row analyzes the sampled distribution  $t_{B14} - t_{B10} = -20$  and the second row assesses the goodness-of-fit to the sampled distribution  $t_{B14} - t_{B10} = 20$ . The first column analyzes a multi-variate distribution, the second column analyzes a Gaussian Copula, and third column analyzes a t-Copula.