Building a Stochastic Terminal Airspace Capacity Forecast from Convective Weather Forecasts

Diana Michalek and Hamsa Balakrishnan

Abstract

Air traffic delays have been increasing over the last several years, and a significant amount of the delay has been due to convective weather activity that renders portions of airspace unsuitable for flying. Though models exist to optimize flight routes and minimize delay in the face of reduced capacity, these models require a forecast of capacity as input. Unfortunately no such reliable forecast exists. In this paper we study the problem of creating a stochastic model of airspace capacity and present a route-based approach to forecasting capacity in the terminal area. This approach uses deterministic weather forecast data to synthesize dynamic routes through airspace, and correlates features of these routes with blockage in the true weather that materializes. This results in a model for route robustness, which can be mapped into a probabilistic capacity forecast.

1 Introduction

The global commercial aviation industry generated $485 billion dollars in 2007, after a steady increase in revenue over the last decade. In the United States alone, this activity included 660 million passenger enplanements and the transport of over 11 million tons of freight [1, 2]. This increase in demand for air travel has coincided with increased delays in the National Airspace System (NAS). According to Department of Transportation estimates, the annual cost of these delays is over $3 billion in costs to airlines alone [3]. What's more, 76.9% of the delays in 2007 were weather-related, an increase over previous years [4]. This situation will only get worse, as demand for airspace is projected to climb further in the coming years, and it will become increasingly important to operate the airspace system efficiently even in the face of storms [5].

Researchers have been working on the challenge of minimizing delays for several decades, mostly in the field of Air Traffic Flow Management (ATFM). ATFM is the process of making strategic decisions a few hours ahead of the time of operations, to balance the demand for aircraft operations with the capacity of the NAS. The capacity of airspace is affected by the presence of hazardous weather, as aircraft must avoid unstable regions and are often forced to deviate from their planned trajectories. Historically, ATFM models have taken expected capacity to be a known input, and have only focused on optimizing routing decisions based on this input. Under clear weather conditions, these expected capacities are fairly stable and tend to reflect reality. However, under stormy weather conditions, capacity is highly variable, and the use of estimated capacity for planning is unrealistic. More recently researchers have developed stochastic models for TFM, which take as input a probabilistic weather forecast or a probability distribution for capacity.

The missing piece has been that a realistic model of airspace capacity under hazardous weather conditions does not exist. Such a model is critical for decreasing weather-related delays, yet it has been a challenging problem due to the chaotic nature of weather and to the unique requirements of ATFM forecasts. More concretely, forecasts for ATFM must be reasonably precise and fine-grained. Knowing that there is a 30% chance of rain in Boston today, for instance, does not help to determine if there will be a route open from the east into Boston Logan Airport at 5PM, or if flights should incur delay on the ground and avoid entrance into Boston between 5 and 9PM.

In this paper I present the first steps towards creating a stochastic model of terminal capacity, by using currently available weather forecasts. We focus on terminal airspace, or airspace near an airport, since it is the bottleneck of aircraft operations. We consider a scenario in which short-term routing decisions must be made based on available forecast data. We find that the quality of a forecast can be measured in terms of how likely it is for a trajectory through forecast weather to be blocked by true weather conditions. This leads to a model for route robustness, which correlates features of a particular trajectory with blockage. We present the steps taken in creating this model, as well results using weather forecast data for Hartfield-Jackson Atlanta International Airport (ATL) terminal area from the summer of 2007. In addition, we show how route robustness can be mapped into a probabilistic forecast of terminal airspace capacity. Although this work focuses on the terminal area, the methodology developed can be modified to model en-route airspace.

The structure of this paper is as follows. Section 2 gives an overview of relevant TFM models and existing weather forecast research and operational products. Section 3 introduces the Lincoln Lab convective weather forecast, which is used in this work. Section 4 describes the issues with trying to model capacity by studying errors in weather forecasts at individual pixels. Section 5 introduces a trajectory-based approach to modeling capacity, and contains the main contribution of this paper. Finally, Section 6 contains conclusions and future work.

2 Related Work

In this section we present Traffic Flow Management models whose objectives include minimizing delays. These models take as input airspace capacity or weather conditions. In addition, I summarize what is available today in terms of forecasts of airspace capacity and weather. This will highlight the gap between TFM input requirements and the state of weather forecasting.

2.1 ATFM Models

The earliest work in TFM involves large-scale integer programming models which determine how to route a set of aircraft from their planned departure locations to their planned destinations while minimizing the cost of delays. Capacity is the major system constraint, limiting how many aircraft can fly through a region of airspace at once. Models involving both a static set of capacities for each sector of airspace, as well as multiple capacity scenarios, each associated with a probability of occurrence were developed [6, 7].

More recently there has been work on algorithms to efficiently synthesize routes through regions of airspace hit by convective weather. This work takes fine-grained and time-varying weather forecast data as static weather input, and focuses on synthesizing short and flyable routes which don't get too close to regions of airspace hit by weather [11, 12]. However, the weather forecasts are treated as ground truth, and routes are not evaluated against actual weather scenarios.

Finally, the Route Availability Planning Tool (RAPT) product uses Lincoln Lab convective weather forecasts to model jet route blockage deterministically. The product is used operationally to help controllers determine if aircraft can take off, though the product is not used as part of automated ATFM [10].

2.2 Capacity Estimation

Some recent work has begun to look at the problem of creating stochastic models of capacity. [9] considers the problem of estimating the capacity of a sector of en-route airspace by computing a theoretical capacity given weather in the region. This is done through the application of continuous maximum flow theory. However, this work relies on static weather forecasts and does not incorporate uncertainty intervals or any measure of forecast accuracy. This research is taken a step further in [8], which considers weather forecasts accompanied by a region of uncertainty. However, the uncertainty profiles are randomly generated

We expand on this line of research by using maximum flow to aid in identifying bottlenecks in regions of airspace, and helping to synthesize routes which are then validated against observed weather.

2.3 Weather Forecasts

There are several aviation weather forecasts available for the United States. In general, weather forecasts take the form of a grid, where each grid cell, or pixel, corresponds to a square section of 2-dimensional airspace. Each pixel contains a value indicating the severity level of weather at that point.

MIT Lincoln Laboratory's Convective Weather Forecast product is a state-of-the-art 0-2 hour forecast, used throughout the United States to aid air traffic control [14]. The forecast is static, meaning that each pixel contains one deterministic value indicated weather severity, with no information about the likelihood that the forecast is correct, or a distribution over possible severities. Specific details about the forecast are provided in Section 3.

Due to randomness in the weather and the resulting inaccuracy of weather forecasts, creating a plan for routes is not realistic using static forecasts alone. Indeed, flying through a region of airspace that turns out to be stormy would compromise safety. This has lead to research into developing probabilistic weather forecasts for aviation. NCWF-2 is one such forecast developed by the National Center for Atmospheric Research, which at each pixel gives a probability p that the pixel will contain convective weather. Initial validation of the forecasts show that these values of p have significant errors associated with them, and tend to be large overestimates of true values [15].

Another probabilistic weather forecast that has been proposed is based on polygons [13]. In this model, a weather cell is represented by a polygon, and the probability that weather will occur at a point in the polygon decreases with increased distance from the center. This structure is then used to estimate flight lengths and deviation delays. However, the model is not validated against behavior of weather, and for the polygons to have much meaning, they might have to be very large and hence not useful for fine-grained ATFM.

3 Lincoln Lab Convective Weather Forecast

In order to create a realistic model of terminal area capacity, it is necessary to start with some available weather forecast. In this section we present the details of Lincoln Lab's Convective Weather Forecast (CWF), which we use in modeling terminal capacity.

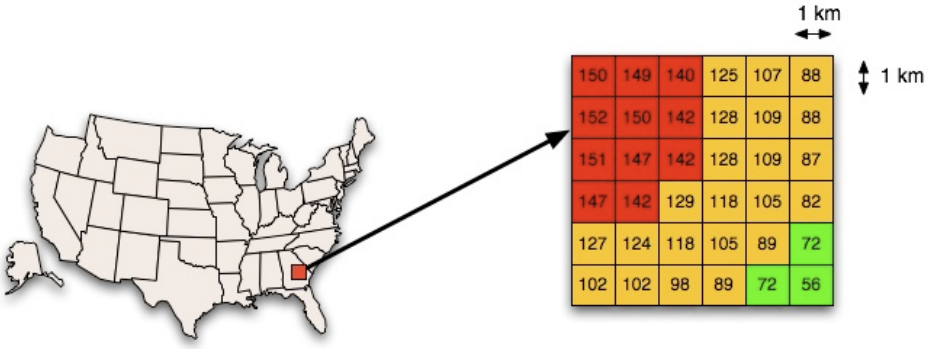

Figure 1: Sample Lincoln Lab Convective Weather Forecast near ATL.

The 0-2 hour CWF consists of a grid of 1km x 1km pixels covering a region of United States airspace [14]. The region will contain the entire continental United States as of 2008, though we only use data from Atlanta's Hartsfield-Jackson Airport (ATL) for the purposes of this work. Each pixel contains a value of Vertically Integrated Liquid (VIL), indicated by an integer value in the range $[0, 255]$. Figure 1 shows a sample forecast for ATL. A VIL value above a certain threshold w corresponds to weather of severity level 3 or higher, which is commonly considered hazardous weather that pilots typically will not fly through. A

forecast has a horizon every 5 minutes between 5 and 120 minutes, and is updated every 5 minutes. In other words, at time T_0, forecasts are available for time $T_0 + 5, T_0 + 10, T_0 + 15, \ldots, T_0 + 120$. The forecast data is accompanied by observed VIL values for the same region of airspace at that time, allowing for validation of the quality of the forecast.

It is important to note that the Lincoln Lab forecast is static, as it contains a single VIL value at each point in space, with no accompanying probability distribution or error estimate. This static forecast is useful in giving a general idea of what weather will look like, and is used in various decision support tools by air traffic controllers and airlines. However, no large-scale historical validation of accuracy has been performed. Lincoln Lab, as well as other forecast products, provide daily statistics such as rates of false positives, false negatives, and a skill score, but these vary daily and by storm, and no historical trends have been tracked or studied.

The natural first step in modeling airspace capacity is to see if historical trends in forecast skill can be correlated with capacity. In effect, it seems desirable to compute the conditional distribution of actual weather given a forecast. This approach is further discussed in the next section.

4 A Pixel-Based Model

A first approach to developing a stochastic model of airspace capacity using deterministic forecast data is to determine a distribution for the probability that hazardous weather occurs at pixel x given the forecast value at pixel x. This probabilistic forecast can then be mapped into a distribution for capacity. This section discusses several issues that we encounter with this approach.

To be more precise, let $w_x(t)$ be the observed weather of pixel x at time t. Let $f_x(t, \tau)$ be the τ-minute weather forecast for pixel x at time t. In other words, $f_x(t, \tau)$ is the forecast taken at time $t - \tau$ for time t. We would like to know the conditional distribution $\Pr(w_x(t) = v \mid f_x(t, \tau) = u)$. Note that this distribution would likely be independent of the pixel x, but might depend on the geographical area.

After trying to approximate this distribution using historical weather data from 8 days of stormy weather conditions surrounding ATL, we found that the deterministic forecasts have large errors, making their use for ATFM problematic.

Figure 2 shows a sample of results from this approach. The two figures contain a histogram of the VIL that actually occurred after a VIL of level 3 (VIL in the range [133, 162]) was predicted, for forecast horizons of 30 and 60 minutes. Although both plots look like they have a roughly Gaussian curve, even the 30-minute forecast results in non-hazardous weather more than half of the time. This translates into wasted capacity if the static forecast were followed.

Several conclusions can be reached from these results. First, it is possible that the forecasts get the general weather trends right, but are off in the position of the weather cells. This turns out to be a major shortcoming of the pixel-based approach. A 10km x 10km storm cell forecast, for example, might be displaced by 10km to the west when observed, resulting in no correct pixel predictions. For planning purposes, however, this forecast is quite good because moving an aircraft's planned trajectory 10km east might be acceptable.

Along similar lines, different storm types have different implications for ATFM. A storm consisting of many small spaced out cells of weather activity (called a popcorn storm), might have very low forecast accuracy in pixel-by-pixel comparisons, because it is hard to predict the exact location of a small cell. However, there might always be routes through the non-stormy sections between the popcorn cells, resulting in no loss in capacity.

These observations suggest that a better way to think about weather forecasts is through a route-centric approach. Identifying persistent routes through weather might identify opportunities for increased capacity even in the presence of storms and of inaccurate forecasts. The remainder of this paper outlines how we validate static forecasts and correlate path characteristics with capacity.

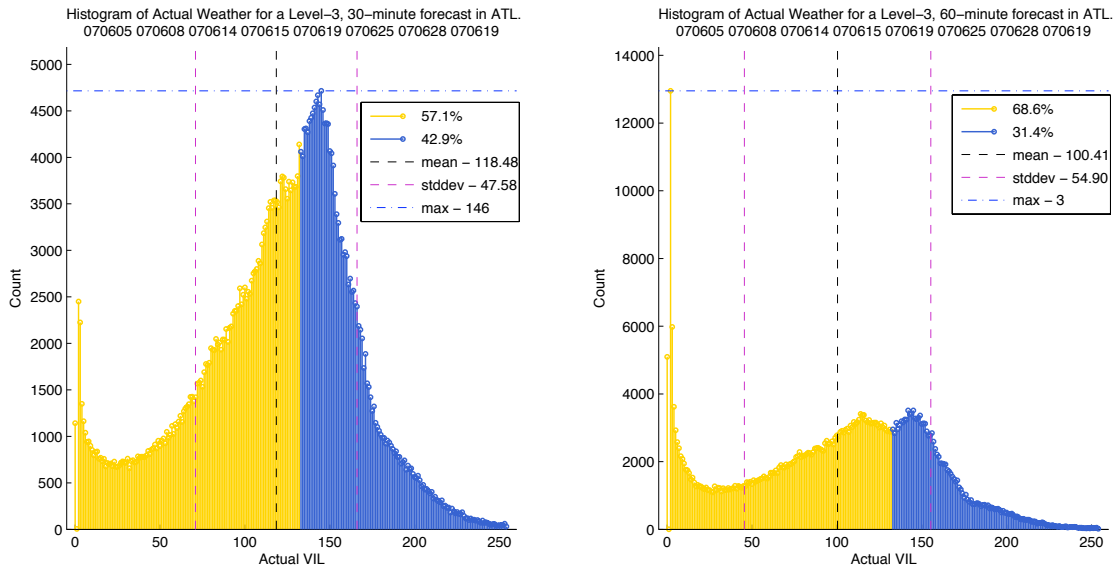

Figure 2: Plot of true VIL when Level 3 VIL (in the range [133, 162]) is forecast. Results for 30-minute and 60-minute forecast are on the left and right, respectively.

5 Route-Based Model of Capacity

The main contribution of this paper is a route-based approach to modeling terminal airspace capacity. In this section we first formalize the terminal area capacity problem. Then we describe our data-driven approach to measuring the robustness of a trajectory through airspace. The high-level steps taken are 1) Compute the theoretical capacity of a region R of airspace using an algorithm for continuous maximum flow developed in [16] and [8]. The resulting minimum cut contains the bottlenecks for flow in R. 2) Use the minimum cut to generate paths through R by solving set of integer programs. 3) For each generated path p, solve another linear program to find a path p' in the neighborhood of p, in the observed airspace. 4) Using data from steps 2 and 3, build a model for the probability that a route will be available by correlating blockage with various features of a path in the forecast space, such as mean distance to weather. 5) Use this model for path robustness to determine a probability distribution for the capacity of R.

5.1 Problem Formulation

Consider the following version of the terminal airspace capacity problem. The input is a terminal area, defined by two concentric circles: an outer circle C_O of radius R, and an inner circle C_I of radius r. The outer circle C_O represents the points at which aircraft first enter the terminal airspace, and R is typically about 40 nautical miles, or 75 km. The inner circle C_I represents the point at which aircraft reach the final arrival waypoint and start their final descent into the airport.

Figure 3 shows a visual model of the terminal area and the input above.

There are two questions that can be answered.

1. If aircraft arrive at C_0 in t minutes, how many will be able to get through to C_I? This is the question of theoretical capacity.

2. If aircraft are routed along trajectories that are clear according to the t-minute weather forecast, what is the probability that that these trajectories will actually be clear in the observed weather?

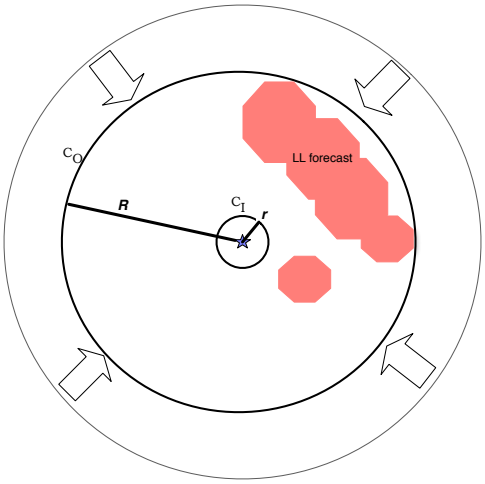

Figure 3: Model of terminal area flow. Flow comes in through the outer circle of radius R, C_O, and into the inner circle C_I. The pink region within the terminal region represents a static weather forecast provided by Lincoln Lab for some time t.

This paper aims to answer the second question. We claim this question is most useful in terms of air traffic control for several reasons. First, it aims to capture trends in how true weather cells differ from predictions. And second, it better takes into account the realities of scheduling aircraft routes. Theoretical capacity might show that N aircraft will be able to enter airspace in the next hour, but that does not indicate the possibility (which is critical for planning), that these aircraft must arrive from the West, for instance. Futhermore, it is possible for a forecast of theoretical capacity to exactly match the true theoretical capacity, yet require that aircraft use trajectories that are very far from original plans.

5.2 Dynamic Weather Grid

Since we assume that aircraft move inwards towards the airport from the outer circle, we can splice together weather data for time t that increases from the outer to inner circle. This effectively creates a dynamic weather grid.

Figure 4 illustrates how this is done, showing a sample aircraft coming in, for reference. We assume aircraft arrive at C_O at time t, after t_0 minutes of look-ahead time. This captures all planning that can occur at time $t - t_0$.

5.3 Synthesizing Paths

Given a t_0-minute look-ahead time, we can model paths through airspace which are likely to stay open. Our approach is a data-driven approach to exploring route robustness. This section outlines the steps taken in creating a dataset of paths through weather-constrained airspace.

Theoretical Capacity

First we compute the theoretical capacity in a dynamic forecast grid. This computation identifies the bottlenecks for flow in the region. To compute the theoretical capacity, we follow the developments on continuous maximum flow extended to the case of airspace in [8], [9], and [16]. This work presents a polynomial-time algorithm for computing the maximum flow through a polygon with holes, from a set of source edges to a set of sink edges. In our case, the polygon represents the terminal airspace, the holes represent weather, and C_0 and C_I are sets of sources and sinks, respectively. The algorithm involves creating

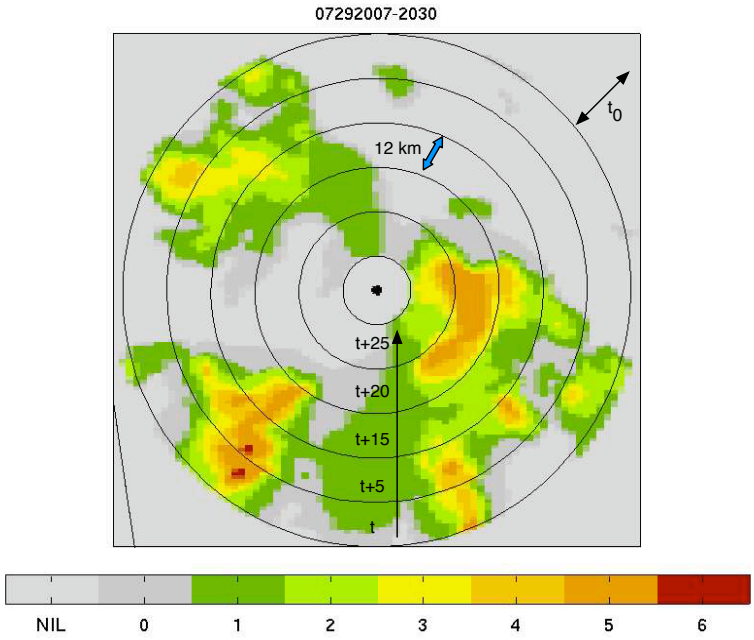

Figure 4: Sample forecast region, created by splicing together 5-min consecutive forecasts. This is for 30-minute look-ahead time on July 29, 2007, where aircraft reach the outer circle at time 21:00. The regions with Level 3 and higher weather become obstacles in the forecast grid.

a discrete *critical graph*, where a shortest path through this graph gives the cost of the minimum cut through the continuous region, and is also equal to the maximum flow.

The minimum cut gives the bottleneck for flow through the airspace, and all trajectories must necessarily fly through this bottleneck region.

Finding paths in the forecast grid

We identify potential aircraft trajectories through the forecast grid by solving a modified shortest path problem as an integer program. The shortest path problem is the following. Given a snapshot of the weather forecast for the terminal region taken at time $t - t_0$, construct a directed graph $G(\mathcal{N}, \mathcal{A})$ such that the set of nodes $\mathcal{N}$ contains all pixels free of weather, and each set of adjacent nodes form an arc $a \in \mathcal{A}$ as long as the arc moves towards the center. At time t, a unit of flow is sent from a set of source nodes $\mathcal{S} = \{s_1, \ldots, s_l\} \subseteq \mathcal{N}$ to a set of sink nodes $\mathcal{T} = \{t_1, t_2, \ldots t_m\} \subseteq \mathcal{N}$, and must pass through a node $K \in \mathcal{N}$. Define $\mathrm{NX}(i, j)$ to be the node $k \in \mathcal{N}$ which constitutes a straight next arc if (i, j) is used. In other words, nodes i, j, k form a straight line in the weather grid, pointing towards the center. The objective is to find the minimum cost flow f such that out of all minimum cost flows, f has the minimum number of turns.

This problem is modeled by the IP below, and can be solved with different sets of sources and sinks and nodes K to generate a large set of candidate paths for a given weather forecast scenario.

$$x_{ij} := \text{flow on arc } (i, j) \in \mathcal{A}$$
$$z_{ij} := 1 \text{ if } (i, j) \in \mathcal{A} \text{ is a turn}, 0 \text{ otherwise.}$$

$$\min \sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij} + \lambda \sum_{(i,j)\in\mathcal{A}} z_{ij}$$

$$\text{s.t.} \quad \sum_{\substack{j\in\mathcal{N}:\\(i,j)\in\mathcal{A}}} x_{ij} - \sum_{\substack{j\in\mathcal{N}:\\(j,i)\in\mathcal{A}}} x_{ji} = b_i \quad \forall i \in \mathcal{N} \tag{1}$$

$$\sum_{\substack{j\in N:\\(K,j)\in A}} x_{Kj} = 1 \tag{2}$$

$$z_{ij} \geq x_{ij} - \sum_{\substack{k\in NX(i,j):\\(j,k)\in A}} x_{jk} \quad \forall (i,j) \in \mathcal{A} \tag{3}$$

$$x \in \{0,1\}^n \tag{4}$$

$$z \in \{0,1\}^n \tag{5}$$

Constraints (1) are the flow balance constraints, with $b_i := -1$ for a supersource $\bar{\mathcal{S}}$, $b_i := +1$ for a supersink $\bar{\mathcal{T}}$, and $b_i := 0$ for all other nodes i in $\mathcal{N}$. Constraint (2) ensures that the path goes through node K. These first two constraints along with the non-negativity constraint (4) finds shortest-paths through the airspace region. However, these constraints alone do not limit the number of turns in the resulting trajectory. This is problematic because aircraft routes must be simple and hence have few turns. Constraints (3) serve to minimize the number of turns in the path without changing the path length. All arcs that follow (i,j) *except* (j,k) for $k = \text{NX}(i,j)$, pay a penalty in the objective function. We set λ to be sufficiently small (less than the maximum length of any path) to ensure that a longer route but with fewer turns is never chosen. Finally, x and z are binary variables because a single path cannot be split up, and the existence of a turn is a binary quality.

We generate a path for many pairs of source and sink sets, each going each segment of the minimum cut.

Validating paths in Observed Weather Grid

Given a set of paths through a region of airspace, we next validate the paths against observed weather. We define a route u as *open* if there exists a corresponding route in the forecast grid which is within B km of u and does not pass through any actual weather.

We synthesize open routes by solving an IP identical to the one defined in the previous section, except over a modified graph $\mathcal{G}'(\mathcal{N}', \mathcal{A}')$, and without the requirement of passing through node K. $\mathcal{G}'$ corresponds to the dynamic grid of *observed* weather in the neighborhood of u. $\mathcal{N}'$ contains all nodes in the truth grid that are within B km of the forecast route u, and $\mathcal{A}'$ contains all pairs of adjacent nodes in $\mathcal{N}'$. The buffer B allows for slight perturbations in the path (on the order of several kilometers), which represents only a slight change from the original planned trajectory of u.

Figure 5 contains a few examples of routes synthesized in the forecast gird, side-by-side with the same routes validated against true weather. Weather scenarios corresponding to 10, 30, and 60 minutes of look-ahead are included to show the difference in pixel accuracy for different forecast horizons. The figure also shows two routes that become blocked, and two that are still open subject to slight changes in the planned trajectory.

5.4 Creating a Dataset of Routes

This section describes details of steps taken to generate a dataset of trajectories through forecast weather with corresponding trajectories through true weather.

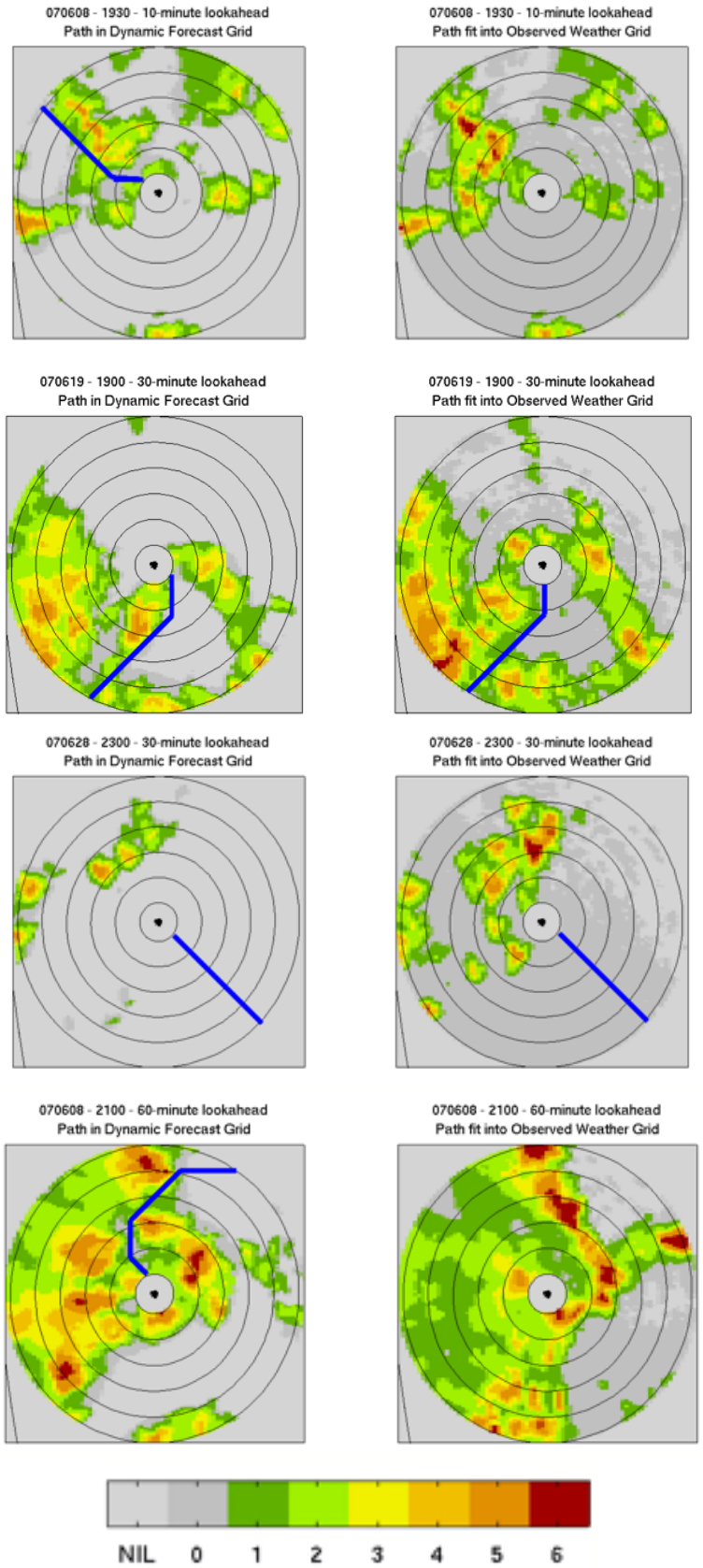

Figure 5: Sample Routes. The left-hand column contains the weather forecast for a given time horizon along with a synthesized route, and the right-hand column contains corresponding observed weather.

Searching for Stormy Weather

We created a dataset containing routes for several weather scenarios during the top 18 worst weather days in June and July 2007, ranked according to weather-related delays. Due to the sheer size of the weather data provided by Lincoln Lab, it was a labor-intensive undertaking to obtain these routes.

Though the Lincoln Lab CWF data as described is just a matrix of integers in the range $[0, 255]$, the archives of this data are kept in a proprietary format, and each month of data takes several hours to extract and decompress, leaving 30 GB of uncompressed binary data. Once the data is extracted, it can be read into MATLAB for processing. To identify the times of day which contain stormy weather, we extracted the forecasts for airspace surrounding ATL and watched visualizations of the entire two month time period to identify the stormiest hours. This resulted in an average of 4 weather scenarios per day, spaced 30 minutes apart, for a total of over 400 trajectories in forecast weather. One dataset was created for each of the look-ahead times of 10, 30, and 60 minutes.

Dataset Details

The following table contains overall rates of route blockage for the datasets.

Look-ahead t_0	dataset size	Viability rate (%)
10	367	88.28
30	385	87.76
60	339	78.17

As the table shows, the vast majority of synthesized routes end up being open in the true weather grid, even for 60-minutes of look-ahead, which had poor pixel-based accuracy. This suggests that subject to minor adjustments, planning at a 10-, 30-, and 60-minute time horizon is quite reasonable. This is very encouraging, and shows that moving away from fixed jet routes gives us the flexibility in the airspace to allow that to happen.

For each path we also compute several features, which will be correlated with route blockage. These 7 features, chosen because intuitively they might be related to route blockage, are listed below:

1 Mean VIL along path
2 Standard Deviation of VIL along path
3 Minimum distance to weather along path
4 Mean distance to weather along path
5 Number of turns taken
6 Theoretical capacity for weather scenario
7 Number of segments in the minimum cut

5.5 Robust Routes

In this section, we correlate path features with blockage and introduce a model for route robustness.

Figures 6 and 7 show histograms of the raw data for 10 and 60-minutes of look-ahead. For each feature f and value v, the green (red) section of the bar shows how many routes having value v ended up being open (blocked) when the route was validated against true weather conditions. In addition, each histogram is overlaid with a plot corresponding to the empirical conditional probability that a route u (taken from the same distribution) is blocked given the feature value. This is computed according to the equation:

$$\begin{aligned}\Pr(\ u \text{ is blocked } | f_i(u) = v) &= \frac{\Pr(f_i(u) = v \ \&\ u \text{ is blocked})}{\Pr(f_i(u) = v)} \\ &= \frac{\#(\text{blocked } \&\ f_i = v)}{\#(f_i = v)}\end{aligned}$$

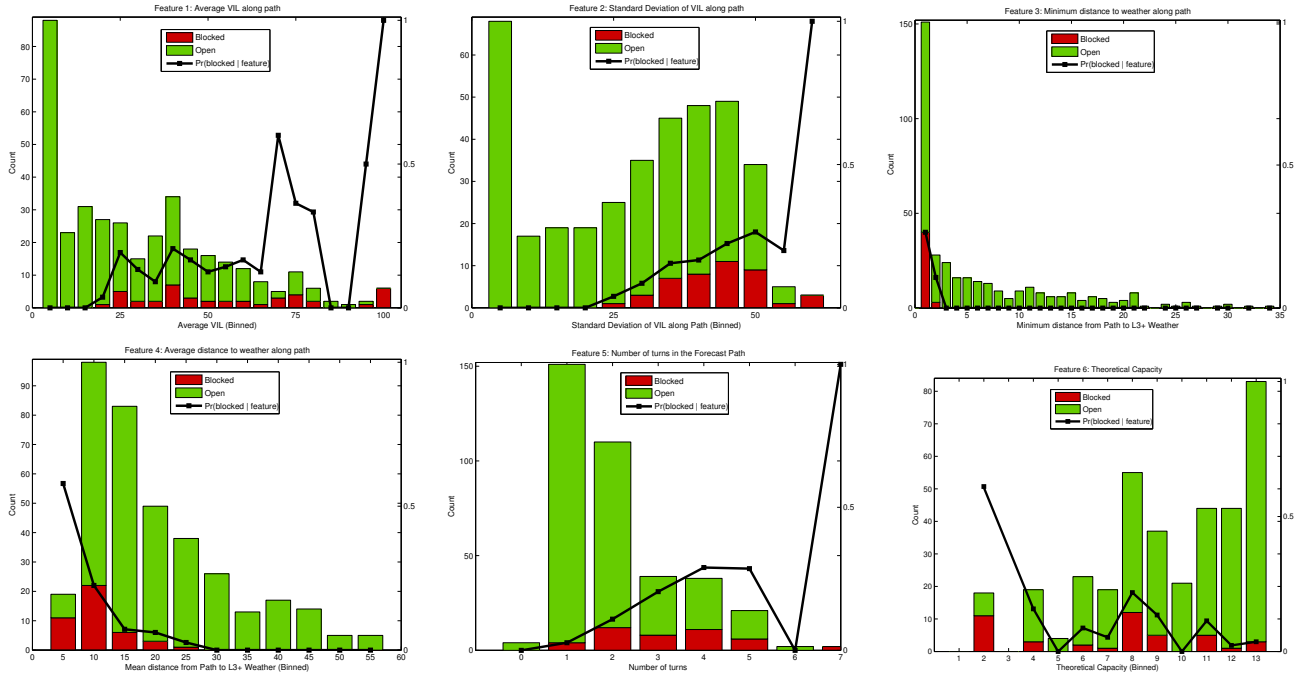

Figure 6: Plots showing how the values of the first 6 features correlate with route blockage, for 10 minutes of look-ahead. The histogram showing the distribution of feature values and blockage is overlaid with the probability that a route is blocked given each feature value. Features 2, 3, and 4 correlate especially well.

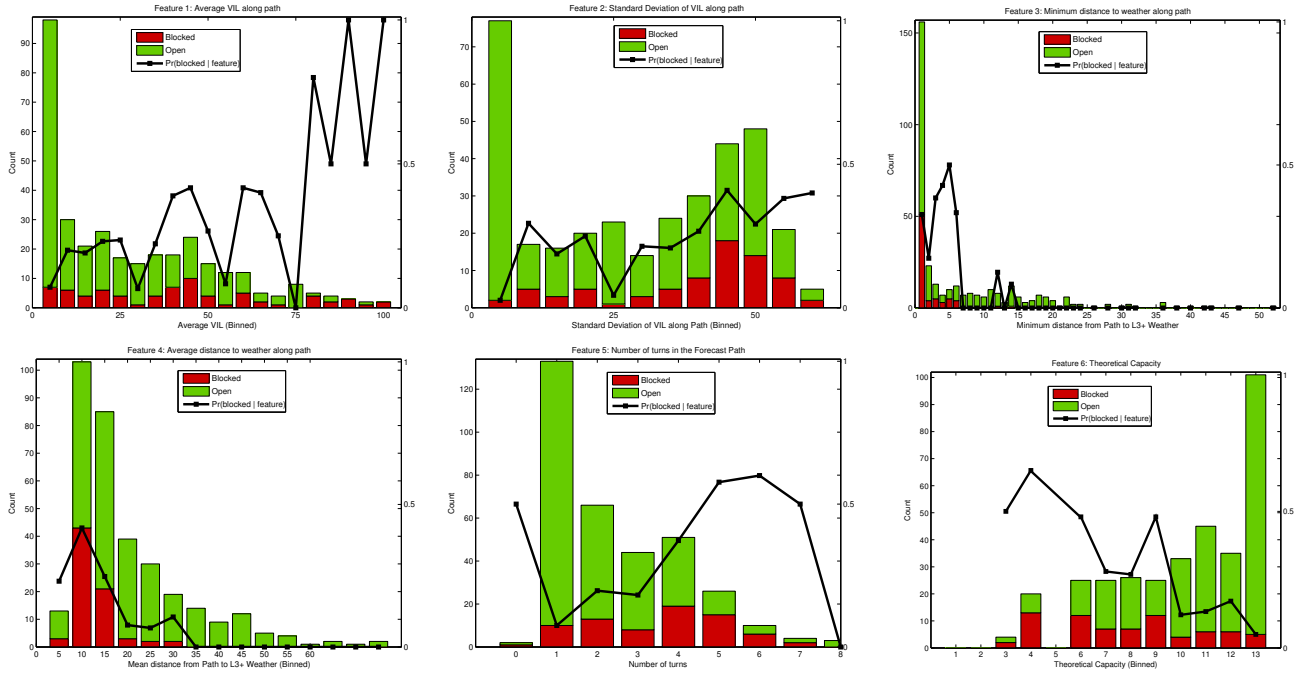

Figure 7: Plots showing how the values of the first 6 features correlate with route blockage, for 30 minutes of look-ahead.

where $f_i(u)$ is the value of Feature i for route u. Features with continuous data (such as average VIL) were binned before the data was plotted.

Looking at the figures, there is correlation between features and blockage, and the correlation is especially strong for the shorter look-ahead time of 10 minutes. This is to be expected, since weather forecasts for a shorter time horizon are known to be more accurate. We see for mean VIL (Feature 1), standard deviation of VIL (Feature 2), and number of turns (Feature 5), that as the value of the feature increases, so does likelihood of blockage. These trends agree with intuition. Indeed, high VIL values along a path indicate the route moves through some weather-affected regions, which are more likely to show up as Level 3 in the true weather. Likewise, higher standard deviation of VIL indicates increased variability in weather conditions along the path, and hence higher likelihood that weather will materialize. And a high number of turns indicate that a route requires lots of maneuvering to avoid weather, so is more likely to be exposed to moving cells. For feature 3 (minimum distance to weather), we see that synthesized routes that are very close to weather end up being blocked more of the time, and that routes that are very far from any forecast weather are safe. The same trend is found with average distance to weather (Feature 4). Finally, theoretical capacity (Feature 6) shows the general trend that increased theoretical capacity correlates with decreased likelihood of being blocked. The plots for Feature 7 were left out, because they showed very little correlation with blockage. This is probably because the number of segments in a minimum cut can occur both if the theoretical capacity is very high (there is 1 large cut segment, for instance), or very low (there is a very small cut segment). The high capacity case would make routes more likely to be viable, since they are less likely to get close to weather cells, while the opposite is true for low capacity.

To get a better understanding of which features best correlate with blockage individually, we compute the Mutual Information between each feature X_i and the blockage label y (+1 for open, -1 for blocked). Mutual information is an information-theoretic measure of the dependence between two random variables X and Y, and measures how much the uncertainty of X is reduced if Y is observed. Note that this measure considers each feature individually and does not capture situations in which two random variables combined correlated very well with y. For discrete random variables X, Y, their mutual information, $I[X;Y]$, can be expressed as

$$I[X;Y] = \sum_{x \in X} \sum_{y \in Y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)}$$

To compute mutual information, it is necessary to have access to the density functions for the corresponding random variables. When the dataset size is much larger than the range size of the joint p.d.f $F_{X,Y}$, we can choose the Maximum Likelihood parameter estimates to the p.d.f.s as good approximations. In other words, use $\hat{P}_X(x) = \frac{\hat{n}_X(x)}{n}$ and $\hat{P}_{X,Y}(x,y) = \frac{\hat{n}_{X,Y}(x,y)}{n}$ to compute the MI. For the case of continuous random variables, the data are discretized by placing points into k equally sized bins, with k set empirically. There are other approaches to approximating MI for continuous distributions, involving setting bin sizes so that the data points are equally distributed between the bins, which is a better approximation to true entropy. But these methods are not used here (see [17, 18]).

Figure 8 shows histograms of mutual information scores between each features X_i and the labels y, for 10 and 30 minutes of look-ahead. Mutual Information is low across the board, meaning that knowledge of the individual features does not reduce uncertainty about the labels by very much. It is also interesting to note that relative mutual information between different features changes with the look-ahead time. For instance, the number of turns (Feature 6) is the best individual indicator of route blockage for higher look-ahead times of 30 minutes, yet for 10-minute look-ahead, features 1 through 4 perform better.

5.6 Classification

We use the dataset to create a model of trajectory robustness. To do so, we train a classifier, based on the features of a candidate route, that classifies the route as open or blocked. The probability that a route will

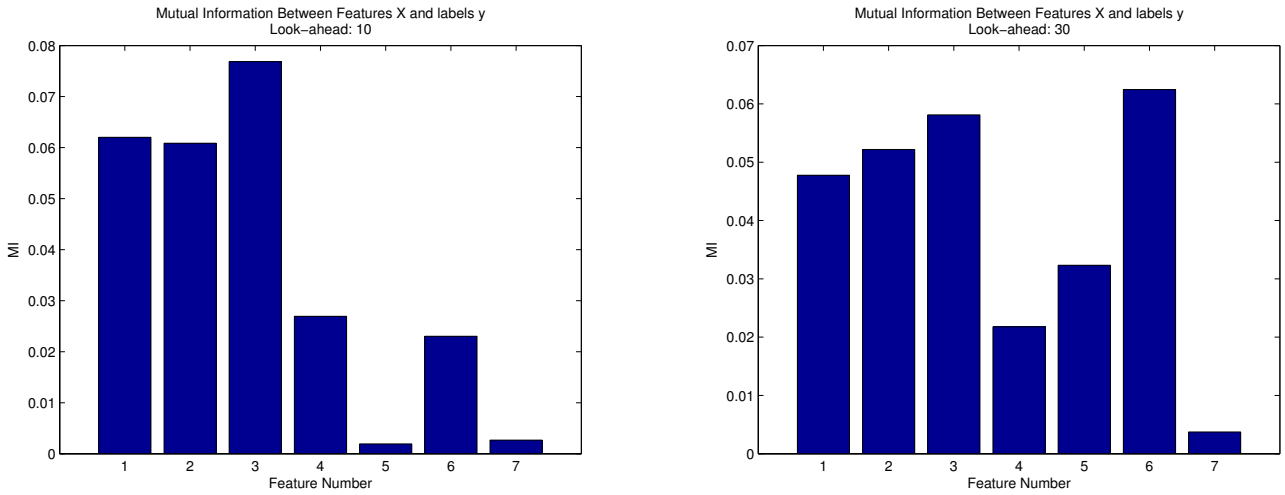

Figure 8: Mutual Information between features and blockage for 10 and 30 minutes of look-ahead.

be blocked given that it is classified as open is defined to be the test rate of false positives of the classifier. Similarly, the probability that a route will be open given that it is classified as blocked is defined to be the test rate of false negatives. This section describes the approach.

We train a Support Vector Machine to try to determine if blockage can be learned based on the route data set. The training set and test set are constructed by partitioning the entire data set by date, with 75% of dates and all accompanying routes placed in the training set and the remaining dates in the test set. By partitioning on weather day rather than on weather scenario, we eliminate any relationship between the training and test data.

First note that the dataset is heavily weighted towards the label +1, meaning that the naive classifier that classifies every route as open performs very well. To get a sense of how much the classifier can learn, we distinguish between two statistics - false positives and false negatives. We still want to be conservative, so want a classifier with very low rate of false positives (the route is classified as open but turns out to be closed) at the cost of a higher rate of false negatives (the route is thought to be closed by turns out to be open). After performing 5-fold cross validation and training the data using multiple kernels (RBF, and polynomial kernels up to degree 3), we found that the classifier performed as well as the naive classifier.

Though this work is ongoing, we can draw two conclusions. First, SVM might not be the most effective classification method, and perhaps some feature selection or other classification model would perform better. Second, perhaps a longer look-ahead time would yield interesting results since in that case the dataset may not be as weighted towards +1 labels.

As a substitute for test error, we can use the overall error rate of the entire dataset for a particular look-ahead time as the probability that a new route v (synthesized by the IP) will be blocked.

5.7 From Routes to Capacity

The route blockage model can be used to create a stochastic model of capacity. We present an initial version of such a model because work here is ongoing. For an airport with m arrival gates (in the case of ATL, $m = 4$), and for a given look-ahead time t_0, we can forecast capacity in the following way. First, synthesize 4 routes through the forecast grid, each sourced from a different quadrant of the outer circle C_O. This can be done using the IP in Section 5.3. Next, for each synthesized route, let the test error of the naive classifier be the probability that the route will be blocked in the true weather grid. Let C be the clear-weather capacity of the airspace. Then the capacity of the the airspace can be forecast as $\frac{Ck}{m}$ with probability Pr(exactly k of the synthesized routes are open), which takes the binomial distribution.

6 Conclusion

A route-based approach has turned out to be fruitful for estimating airspace capacity using actual weather data. By analyzing stormy weather data from the summer of 2007, we were able to show that certain features of a candidate trajectory through airspace correlate with blockage in true weather. We found that VIL levels along the trajectory and in neighboring regions, along with the complexity of the route (number of turns), to be good individual predictors for blockage. Furthermore, we introduced an integer program for synthesizing turn-constrained routes through forecast regions, which turned out to produce routes that tended to be viable in the true weather gird, subject to slight shifts in position. This is promising for short-term route planning under uncertain weather constraints.

This work suggests several directions for the future:

1. Consider departures. This research has only considered arrival traffic, but the same methodology can be applied to departures. This setting can benefit especially well from short-term capacity forecasts and routings because there are always flights on the ground waiting to leave during stormy weather. RAPT has a route-blockage model based on static forecasts for determining when a jet-route is open. However, having more accurate probabilistic estimates as well as adding dynamic changes in routes to the mix might increase departure throughput during stormy weather.

2. Consider echo tops. Storm height should be taken into account when determining if airspace can be flown through. This is because if a storm is at low enough altitude, pilots can often fly over the stormy regions. Lincoln Lab produces echo tops forecasts and truth data that can be incorporated in the capacity models.

3. Use route robustness for ATFM. This would involve enhancing the IP in Section 5.3 to synthesize multiple paths at a time, and to account for trajectories that are too close to forecast weather and to each other. Furthermore, the features that correlate best with blockage can be moved to the objective function and we can minimize the weighted sum of the value of those features, to find paths that are most likely to remain open. This becomes a harder problem, but one worth exploring in more detail.

References

[1] IATA, *Industry Statistics Factsheet*. Retrieved 19 April 2008. `http://www.iata.org/NR/rdonlyres/65A3F2C3-3656-4243-8396-3D77CB31C2FC/0/factsheet_industry_stats_apr20082.pdf`

[2] Bureau of Transportation Statistics, United States Department of Transportation. Retrieved 19 April 2008. `http://www.transtats.bts.gov/`

[3] United States Department of Transportation, *Aviation Delay*. Retreived 19 April 2008. `http://www.dot.gov/PerfPlan2004/mobility_delay.html`

[4] Bureau of Transportation Statistics, U.S. Department of Transportation, *Understanding the Reporting of Causes of Flight Delays and Cancellations*. Retreived 19 April 2008. `http://www.bts.gov/help/aviation/html/understanding.html`

[5] IATA, *Passenger and Freight Forecasts 2007 to 2011*, October 2007, `http://www.iata.org/NR/rdonlyres/E0EEDB73-EA00-494E-9408-2B83AFF33A7D/0/traffic_forecast_2007_2011.pdf`

[6] D. Bertsimas and S. Stock Patterson, *The Traffic Flow Management Rerouting Problem in Air Traffic Control: A Dynamic Network Flow Approach*, Transportation Science, 2000

[7] D. Bertsimas and A. Odoni, *A Critical Survey of Optimization Models for Tactical and Strategic Aspects of Air Traffic Flow Management*, National Aeronautics and Space Administration National Technical Information Service, 1997.

[8] J. S. B. Mitchell, V. Polishchuk, and J. Krozel, *Airspace throughput analysis considering stochastic weather*. AIAA Guidance, Navigation, and Control Conference, Keystone, Colorado, 2006

[9] J. Krozel, J. S. B. Mitchell, V. Polishchuk, and J. Prete. *Capacity Estimation for Airspaces with Convective Weather Constraints*, AIAA Guidance, Navigation and Control Conference and Exhibit, Hilton Head, South Carolina, August 20-23, 2007.

[10] R. DeLaura and S. Allan, *Route Selection Decision Support in Convective Weather: A Case Study*. USA/Europe Air Traffic Management R&D Seminar, Budapest, 23-27 June 2003.

[11] J. Prete, and J. S. B. Mitchell, *Safe Routing of Multiple Aircraft Flows in the Presence of Time-Varying Weather Data*, AIAA Guidance, Navigation, and Control Conference and Exhibit, Providence, Rhode Island, August 16-19, 2004.

[12] J. Krozel, C. Lee, and J. S. B. Mitchell, *Turn-Constrained Route Planning for Avoiding Hazardous Weather*, Air Traffic Control Quarterly, Vol. 14(2), 2006, pp. 159-182.

[13] K. Sheth, B. Sridhar, and D. Mulfinger, *Impact of Probabilistic Weather Forecasts on Flight Routing Decisions*, AIAA Aviation Technology, Integration and Operations Conference, Wichita, KS, September, 2006.

[14] M. Wolfson, B. Forman, K. Calden, W. Dupree, R. Johnson, R. Boldi, C. Wilson, p. Bieringer, E. Mann, and J. Morgan. *Tactical 0-2 Hour Convective Weather Forecasts for FAA*, 11th Conf. on Aviation, Range and Aerospace Meteorology, Hyannis, MA, 2004

[15] S. A. Seseske, M. P. Kay, S. Madine, J. E. Hart, J. L. Mahoney, and B. Brown, *2006: Quality Assessment Report: National Convective Weather Forecast 2 (NCWF-2)*. Submitted to FAA Aviation Weather Technology Transfer (AWTT) Technical Review Panel.

[16] J. Mitchell, *On maximum flows in polyhedral domains*, J. Comput. System Sci., 40, 1990, pp. 88–123.

[17] G. Tourassi, E. Frederick, M. Markey, and C. Floyd. *Application of the mutual information criterion for feature selection in computer-aided diagnosis*, Medical Physics, Volume 28, Issue 12, December 2001, pp. 2394-2402.

[18] R. Battati. *Using mutual information for selecting features in supervisedneural net learning*, Neural Networks, IEEE Transactions on Volume 5, Issue 4, Jul 1994 Page(s):537 - 550.

[19] Chih-Chung Chang and Chih-Jen Lin, *LIBSVM : a library for support vector machines*, 2001. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm