AIAA Guidance, Navigation, and Control Conf., San Francisco, CA, Aug., 2005.

Ground Delay Programs to Address Weather within En Route Flow Constrained Areas

Ray Jakobovits. Ph.D.* Jimmy Krozel, Ph.D.† and Steve Penny‡
Metron Aviation, Inc., Herndon, VA, 20170

This paper investigates an operational concept to control the take off time and route selection of a set of aircraft constrained by en route weather. The concept is based on current airport Ground Delay Program (GDP) procedures. However, instead of controlling the amount of traffic arriving into an airport, we extend the GDP to control the number of aircraft flying into and around en route Flow Constrained Areas (FCAs), which are used to manage the airspace impacted by weather. Because of this relationship with traditional GDPs, we refer to the problem as an FCA-based GDP. We develop new algorithms to assess the FCA-based GDP for an operational scenario. A routing and scheduling algorithm includes ground-delay, route selection, and airborne holding as decision variables for departing and en route flights, and is also designed to align with a Collaborative Decision Making (CDM) philosophy. A dynamic FCA capacity-estimation algorithm uses weather forecast information to produce time-varying entry and exit points as well as maximum flow rates through FCAs. Integration of these algorithms enables assessment of the value of improved weather forecast accuracy, and provides insights into the nature of robust traffic management initiatives.

Nomenclature

AAR = *Airport Acceptance Rate*
ADL = *Aggregate Demand List*
AOC = *Airline Operational Control*
ARTCC = *Air Route Traffic Control Center*
ATCSCC = *Air Traffic Control System Command Center*
CCFP = *Collaborative Convective Forecast Product*
CDM = *Collaborative Decision Making*
CDR = *Coded Departure Route*
CRCT = *Collaborative Routing Coordination Tool*
CRRAT = *Collaborative Routing Resource Allocation Tool*
CTA = *Controlled Time of Arrival*
FCA = *Flow Constrained Area*
FCFS = *First Come First Served*
FEA = *Flow Evaluation Area*
FSFS = *First Scheduled First Served*
GDP = *Ground Delay Program*
NAS = *National Airspace System*
NWP = *Numerical Weather Prediction*
RTA = *Required Time of Arrival*
SCS = *Slot Credit Substitution*
SFO = *San Francisco Intern. Airport*
SGHP = *Stochastic Ground Holding Policy*
SPT = *Strategic Planning Teleconference*
TFM = *Traffic Flow Management*
TFMP = *TFM Problem*
TFMRP = *TFM Rerouting Problem*
TOAD = *Time Ordered Accrued Delay*
Wx = *Weather*

I. Introduction

Traffic Flow Management (TFM) initiatives, including Ground Delay Programs (GDPs), are currently used to resolve imbalances between demand and capacity during severe weather events in the National Airspace System (NAS). However, a greater emphasis has been placed on strategically planning for the terminal conditions (Airport Acceptance Rates (AARs) and matching these AARs via GDPs) in comparison to strategically planning/controlling en route conditions. This paper investigates the planning of takeoff times and route selection to better meet the capacity constraints of en route Flow Constrained Areas (FCAs) during NAS-scale weather events.

Today's airport-based GDPs use an algorithm that determines when to release flights from the ground in order to meet constrained AARs. Air carriers are assigned arrival slots based on their scheduled arrival times, thereby

* Sr. Analyst, 131 Elden Street, Suite 200.
† Chief Scientist, Research and Development, 131 Elden Street, Suite 200, AIAA Associate Fellow.
‡ Analyst, 131 Elden Street, Suite 200.

rationing airport capacity by assigning a ground delay to each flight. Through the FAA's GDP enhancement program, the scope of the GDP algorithm has been modified over the years to allow air carriers the ability to reallocate assigned arrival slots within their operations and perform one-for-one trades with competing airlines, while preserving both equity and efficiency of the overall ground-delay allocation. Effective use of GDPs, implemented through collaboration between the FAA's Air Traffic Control System Command Center (ATCSCC) and the airlines, will continue to be a vital component of the overall management of capacity [1, 2, 3, 4]. Recent research in new GDP algorithms has extended in the direction of fix-based GDPs[5], where the flow over each metering fix of an airport is planned for balanced demand and capacity (e.g., to plan for a reduced capacity due to weather constraints over each arrival metering fix of the airport), and for special purpose GDPs. For instance, fog at San Francisco International Airport (SFO) creates a problem where a GDP must be designed to coincide with the fog burn off; a Stochastic Ground Holding Policy (SGHP)[6] is an effective solution to such a problem.

When capacity is constrained in the en route airspace (e.g. **Figure 1**), a simple application of the GDP algorithm does not provide an adequate solution to the en route resource allocation problem. One reason is that, unlike the airport arrival flow problem, many aircraft may be able to avoid the capacity constraints altogether, and, therefore, should not be assigned ground delays. While flights cannot easily change their arrival airports, carriers might prefer to reroute around an FCA rather than wait on the ground for capacity to be restored within the FCA. Another reason is that while airport capacity levels are reasonably well understood, there is considerable uncertainty in airspace capacity due to weather constraints. This means that plans have to adapt as the perceived situation changes, and plans should, in fact, incorporate the ability to reroute airborne flights into the planning mechanism. It may be better to allow more flights to approach the FCA than the current capacity estimate permits, in order to allow for the possibility that the forecast will improve before the flights must be rerouted (given a contingency plan, of course, if the FCA capacity is overestimated). This introduces the idea of producing a robust plan that represents the best solution taking into account the weather forecast variability.

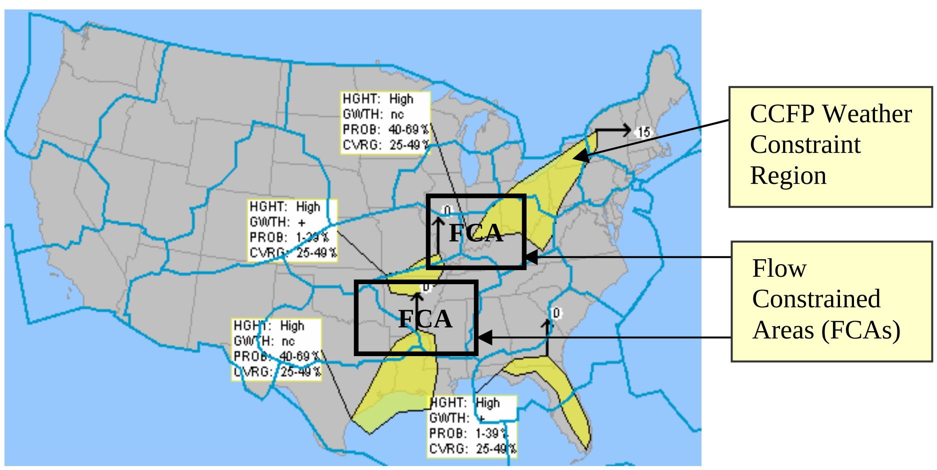

Figure 1: CCFP weather constraints indicate the need for FCAs to constrain the demand in the bottleneck region.

Operations in the NAS are continuous, and ideally, TFM planning should be a continuous process as well. In practice, however, TFM initiatives are planned 1 to 6 hours in advance, and then updated as needed to account for changes in uncertain weather forecasts, updated flight status data, or changes in predicted demand, including flight cancellations. This discrete, iterative planning paradigm is illustrated in **Figure 2**. Each plan is produced based on the latest information available. Plans are updated on a regular schedule, with the time between updates defined as the "planning interval." Each update covers a time window that extends well beyond the planning interval, but only a portion of that plan may be executed before the plan is replaced by an updated version. Under this paradigm, it is important that any control algorithm have the following characteristics:

i. The ability to accept updated status reports on all resources and replan all resources from their current positions forward in time,
ii. The computation time required to produce a plan must be considerably less than the planning interval in order to obtain and put in place any plan revisions,
iii. If robust plans are to be produced, then the algorithm must be able to consider multiple future scenarios before producing a recommended solution, and
iv. The algorithm should be constructed in such a way that air carrier inputs are both encouraged and rewarded.

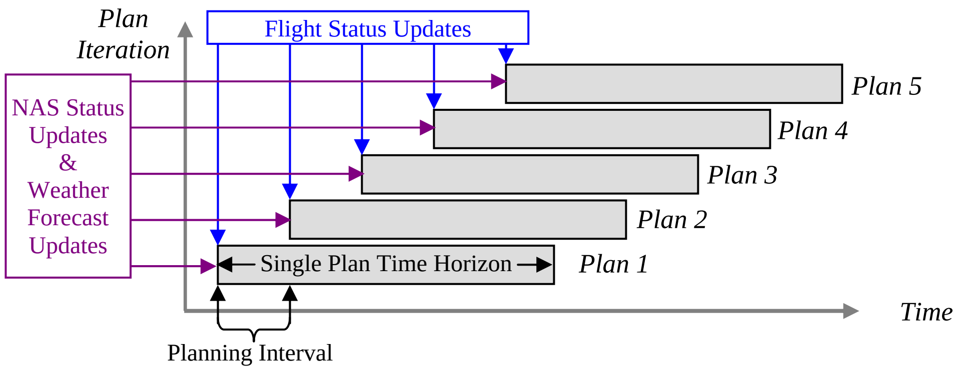

Figure 2: Overview of Planning and Information Flows.

The algorithm described in this paper is designed to address the en route capacity problem. While we refer to the general subject area as "FCA-based GDPs", the decision variables include not just ground delays, but also selection of alternative routes prior to departure, airborne rerouting, and airborne holding, thereby enabling a much richer set of control strategies than are permitted using the standard GDP approach. Specific weather-related questions that the research is designed to address include:

- Given the inherent uncertainty in airspace capacity prediction, what does a robust plan look like?
- If weather-forecasting capabilities improve in the future, what is the payoff in terms of system delays, throughput, and other related measures of system performance compared with current forecasting capabilities?

The approach taken to solving this problem aligns with the Collaborative Decision Making (CDM) philosophy. The fundamentals of CDM are:

i. Create a common view of the problem that is shared between the FAA and airline users,
ii. Create the opportunity and incentive for users to mitigate problems through their own actions and notify the FAA of their intentions,
iii. Give the users flexibility to satisfy their own priorities within the context of FAA-initiated TFM constraints, and
iv. Allow the users to participate in the determination of TFM policies and procedures, for example, through participation in a Strategic Planning Teleconference (SPT), which is currently held in the NAS every 2 hours.

CDM has initiated a shift away from a central authority paradigm by acknowledging that the airlines should play a substantial role in TFM. Resource allocation algorithms that comply with the CDM philosophy are typically a blend of techniques from operations research, optimization, game theory, distributed control, and related disciplines.

The remainder of this paper presents a brief overview of related literature, our approach to aircraft scheduling and routing, examples, and conclusions.

II. Related Literature

Several models have been proposed for the routing and scheduling of aircraft in order to minimize congestion costs, with solutions generally falling into two categories: (1) classical optimization approaches, and (2) heuristics.

In the classical optimization category, the most complete formulation of the problem[7] is given as the Traffic Flow Management Problem (TFMP) when aircraft routes are predetermined, and as the Traffic Flow Management Rerouting Problem (TFMRP) when rerouting of aircraft is allowed. The TFMP problem is formulated as a 0-1 integer programming problem. The objective is to minimize total weighted delay, and the constraints take into account departure and arrival capacities of the airports, sector capacities, sector connectivities, and airport connectivities. Their traffic network consists of traffic control sectors. Sector-to-sector transit times are inputs to the model, and the decision variables determine when each flight enters each sector along its designated route. Their approach is more strategic than tactical. Time is divided into discrete intervals, and all sector entries and exits effectively occur at an interval boundary. Each flight has a single flight path, which can be augmented with ground or en route holding delays, subject to sector capacity restrictions. The number of variables in the problem grows rapidly as the time interval length is reduced, so there is clearly a trade-off between scenario scope and duration, modeling resolution, and computational feasibility.

The TFMP is equivalent in computational complexity to the "job shop scheduling problem"[7], and therefore, is NP-hard[8]. This means that we cannot expect to derive a polynomial runtime algorithm to solve it. However, a linear programming relaxation formulation[7] almost always returned integer values, which provided some confidence that realistically sized problems might be solvable. Their formulation of the TFMP problem extends to the TFMRP problem, but with a potentially significant increase in the number of decision variables. Thus, when addressing the TFMRP problem, the authors employ a hybrid, multi-step approach[9] combining optimization and heuristic methods.

More tactical but related problem formulations have appeared in connection with surface traffic management. Smeltink et al[10] present an integer optimization model for minimizing total airport taxi time of arriving and departing aircraft. Given a set of desired taxi paths, one for each aircraft, the model produces an arrival sequence at each node of an airport network that preserves separation distance constraints among aircraft. Taxi speeds are allowed to vary within defined limits, and holding en route is allowed at designated holding points subject to capacity constraints. A similar formulation by Visser and Roling[11] allows multiple alternate paths to be considered for each aircraft, while holding the taxi speed constant. Both problems are again equivalent to the job shop scheduling problem, and, therefore, do not scale well with the number of flights and/or the size of the network. Each approach uses a "rolling horizon" to obtain a single-airport solution in a reasonable amount of computation time. An optimal solution is obtained for a subset of the flights, and then time is moved forward, bringing new flights into the planning time window.

One of the most common heuristic approaches to solving the generic job-shop scheduling problem involves the use of "dispatching rules." The terms "dispatching rule", "scheduling rule", "sequencing rule" or even "heuristic" are often used synonymously. The general idea is to sort the jobs (flights) according to some criterion and then sequentially optimize each job in priority order. Typical sorting rules are Shortest Processing Time (SPT), First-Come-First-Served (FCFS), or desired completion time. Extensive simulated studies have shown that SPT is the best choice for optimizing the mean value of basic metrics such as total waiting time and system utilization, but can lead to excessively long waiting times for some jobs.

Among the heuristic approaches employing dispatching rules for aircraft routing and scheduling is the recent development of the Collaborative Routing Coordination Tool (CRCT)[12,13] and Collaborative Routing Resource Allocation Tool (CRRAT)[14,15]. These efforts are designed to be consistent with the CDM philosophy.

Based on the assumption that an accurate 2-hour weather prediction is available, CRCT automatically sets up the FCA based on the weather forecast, identifies those aircraft that have flight plans passing through the FCA, allows the user to define candidate routes to alleviate the problem, and assess the merit of the proposed solution on sector loadings. CRRAT implements a resource allocation algorithm that provides each air carrier the ability to trade off ground delay required to stagger a flight's arrival time to the FCA with added costs of flying around the FCA, taking into account the desire of competitors. Users can submit for each flight caught in the allocation situation a primary (lowest cost, first-choice) flight plan along with a list of alternate flight plans. For each alternate flight plan, they submit a delay threshold which implicitly tells the rationing algorithm how many minutes of relative ground delay they would be willing to absorb before switching to that alternate flight plan. The algorithm dynamically selects for each flight the flight plan that the air carrier would most prefer, given the availability of resources (**Figure 3**).

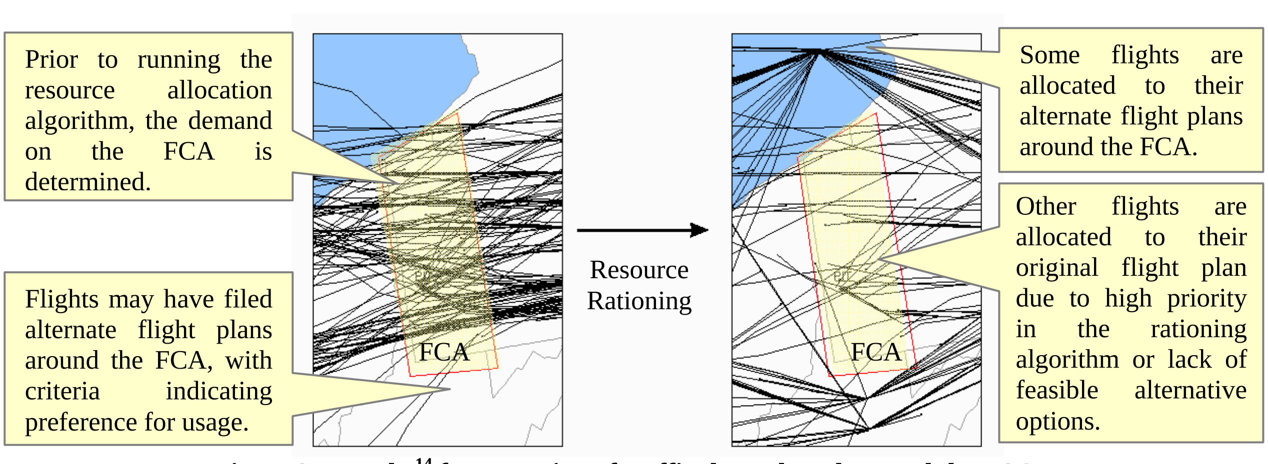

Figure 3: Results[14] for rerouting of traffic through and around the FCA.

CRRAT is priority-based, which means that the flights are batch processed in a specified order; the earlier a flight is processed, the better the chance it receives the resources it wants (flight plan and departure time). The exact priority number (e.g., whether processed 12th or 25th) is determined as the algorithm runs. This is a natural consequence of extending routing and delay options to users. Each time a flight is processed, the algorithm examines the remaining resources and the list of user preferences. An optimum flight plan-delay combination is assigned, given the resource constraints that have already been allocated. The assignment affects the availability of resources for the remaining (unprocessed) flights, which, in turn, may affect the order in which the remaining flights are processed. Airborne and international flights are neither delayed nor rerouted by the algorithm. (As in a GDP, they are treated as *exempt*.) Time-ordered processing (FCFS, First Scheduled - First Served (FSFS), and Time-Ordered Accrued Delay (TOAD) rationing algorithms) is required for system efficiency. That is, by scheduling resources in more or less chronological order, severe outlier delays are avoided.

Finally, market-based strategies[16] may be used to allocate resources; these require sufficient infrastructure to allow for real-time negotiation of resources and FAA oversight of transactions. In a market-based scenario, there is an immediate allocation of the slots (e.g. an auction). From that point on, buying, selling and trading of slots may take place among NAS users, with transactions acknowledged by the FAA. As an alternative to the market-based approach, an initial 'fair allocation' such as a FCFS or FSFS would be required to set the initial slot allocations among users. The re-allocation of slots then could be accomplished using algorithms similar to Slot Credit Substitution (SCS)[17].

III. Operational Concept

An FCA-based GDP is a specific instance of allocating a general en route resource. An FCA-based GDP consists of two stages: (1) defining the impacted airspace (FCA), thereby defining the resource, and (2) allocating the flight demand to this resource (GDP). Implicit in our definition of an FCA-based GDP, we address flights that are already airborne and expecting to use the FCA. Airborne flights have priority over flights on the ground prior to departure.

The process of running an FCA-based GDP is as follows. The ATCSCC evaluates the current weather or traffic forecasts by setting up a Flow Evaluation Area (FEA), and if it is deemed necessary, the FEA is converted to an FCA. A Ground Stop (GS) is immediately enacted for all flights filed through the FCA. Routes that avoid weather within the FCA are generated algorithmically, with the start of the routes defining the entry points to the FCA. An initial allocation assigns airborne flights to slots that correspond with the entry points based on estimates of their arrival times. The size of the slots is based on uncertainty in the capacity of the routes generated for the FCA. FCA slot allocations define Required Times of Arrival (RTAs) to the FCA. Once slots are allocated for airborne flights, slots are then assigned for the flights on the ground. The flights on the ground are released from the GS and assigned new departure times to meet their FCA RTA. With automation, the length of the GS would probably be no more than a few minutes; however, it is desirable to make sure no new flights depart while the allocation is performed. While monitoring flights, if a flight is not expected to meet the RTA, it will either be reassigned to a later slot that is not already assigned, or denied access to the FCA (being forced to fly around it). If the delinquent aircraft desires to use an already assigned slot, it must negotiate with the owner of that slot. These negotiation decisions must be made quickly, and will likely be performed by professionals at the Airline Operational Control (AOC) level, not on the flight deck.

An FCA-based GDP cannot be run independently of other system resource allocations. Allocation of slots in an FCA-based GDP would be back-propagated from the future use of resources by the same flights (fixes in a fix-based GDP, or the runway in a standard airport-based GDP). The same follows for cases with multiple FCA-based GDPs – any FCA-based GDP sharing flights with other FCA-based GDPs would require coordinated RTAs. We assume any flight that does not wish to utilize constrained airspace files a normal flight plan.

IV. Operational Scenario

Based on the flown, filed, and scheduled routes used by historical traffic, we determine a network within which flights may travel. We then introduce weather constraints and create an FCA over the network, defining reduced capacities for routes traversing the weather and FCA. We report the effect of weather uncertainty on system performance (measured by throughput and delay). We also report the effects of limiting capacity due to weather, and how the offload routes are impacted by the increase in demand. We describe a regional-scope scenario that is at the level of the ATCSCC.

In this operational scenario, a major convective weather system is impacting the east coast (**Figure 4, Figure 5**). Today, traffic managers at the ATCSCC apply reroutes through Playbook Plays in conjunction with GDPs and GSs. An FCA-based GDP combines the rerouting capability attained through Playbook Plays while monitoring airspace capacity and providing greater control over traffic flows than a typical airport-based GDP.

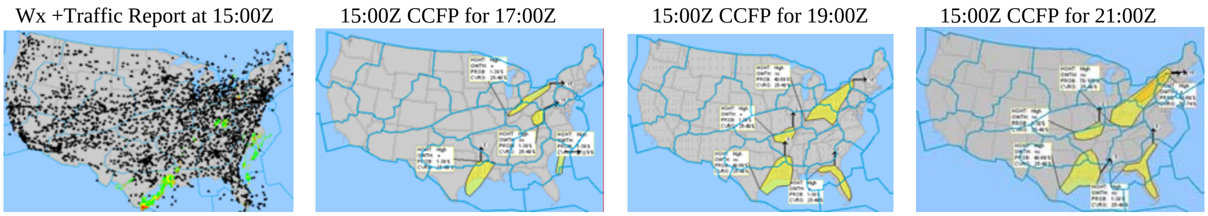

Figure 4: Weather and traffic at 15:00 Z and CCFPs generated at 15:00 for 17:00, 19:00 and 21:00.

Due to the predicted constraint across ZOB and ZID, identified by the CCFP, Playbook routes were selected to move traffic away from these impacted areas while GDPs and GSs were implemented at both EWR and LGA to reduce traffic volume (**Figure 5**). However, many of the flights received excessive delay and airlines were forced to implement a number of cancellations. The circuitous routes defined by the playbook are much longer than the standard routes, imposing increased fuel usage on the airlines and requiring traversal of Canadian airspace.

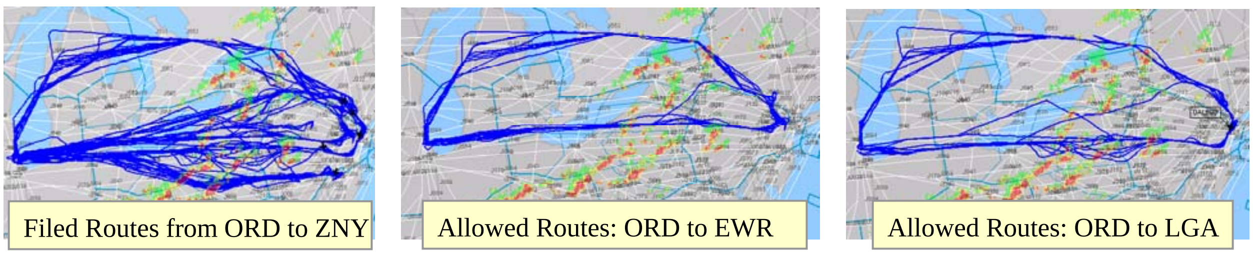

Figure 5: Eastbound flights comparing the filed *vs* allowed routes.

To achieve maximum eastbound flow through ZOB airspace, an FCA-based GDP is used over the impacted area spanning the eastern half of ZOB (**Figure 6**). The approximate points where the standard jet routes intersect the FCA, designated as entry points, are metered with combined en route time-based metering and GDPs. Playbook-like routes are generated for flows around the FCA to manage the excess demand on the airspace regulated by the FCA. Note that the westbound flows can be managed similarly with additional routes that avoid the eastbound routes or with routes generated at alternate altitudes.

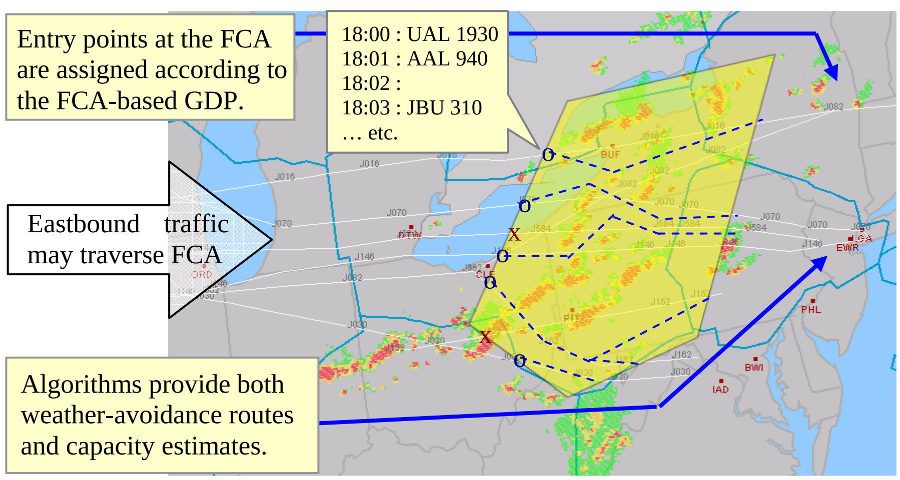

Figure 6: An FCA is created over the impacted airspace.

V. Modeling

A. Problem Statement

The problem can be described as follows. Given:

- A set of scheduled flight start (wheels off) times, S_i for $i = 1,2,...,N_A$ for a total of N_A aircraft in the FCA-based GDP.
- A set of resources r_j for $j = 1,2,...,N_R$, nodes and arcs of an appropriate traffic network, e.g., jet routes or other fixes along a route at which traffic is metered. A typical en route network affected by a FCA is in **Figure 7.**
- A set of possible paths for the i^{th} flight, P_{ik} for $k = 1,2,...,NP_i$
- A set of rules defining inter-operation separation minima and resource holding capacities, and
- A pair of weights ω_o and ω_1 representing the relative costs of gate/ground holding and en route holding

The problem is to find a set of paths, one for each flight, that minimizes the weighted sum of the delays experienced by all flights.

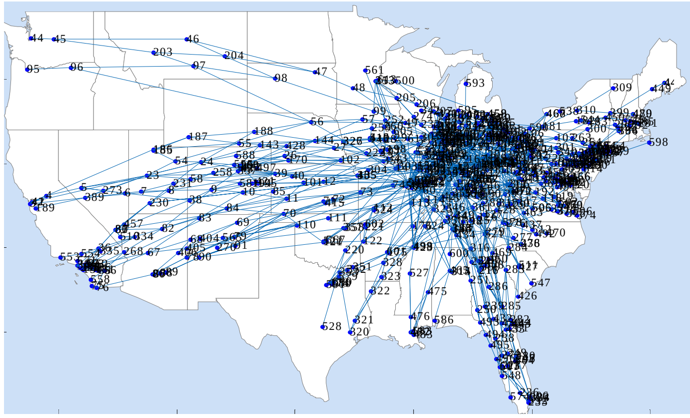

Figure 7: A network generated for an FCA just south of ORD.

B. Analysis

Each path consists of an ordered set of resources (nodes and arcs of the network) and associated transit times:

$$P_{ik} = \left\{ r_{ikj}, \tau_{ikj} \right\}, \ \mathrm{j} = 1, 2, ..., \mathrm{J}_{\mathrm{ik}}.$$

Constants $\{J_{ik}\}$ denote the number of resources on each path *k*. Transit times $\{\tau_{ikj}\}$ may depend on characteristics of the aircraft and the route, and represent times to traverse the set of resources along the path *k*. Once a solution to the TFMRP problem is obtained, there will be an associated transit "appointment" or "slot" for each resource *j* having duration τ_{ikj}. Note that appointments for consecutive resources along a path may not be contiguous in time if the flight is held at a resource prior to being allowed to continue on its path. In particular, we denote holding times at the j^{th} resource by $\left\{ \Delta_{ikj} \right\}$, where Δ_{ik0} denotes holding time prior to the start time S_i, i.e. gate or ground holding. With this notation, we have:

$$\text{Total time to traverse the path } P_{ik} = \Delta_{ik0} + \sum_{j=1}^{J_{ik}} \left(\Delta_{ikj} + \tau_{ikj} \right) \tag{1}$$

And

$$\text{Arrival time at destination using path } P_{ik} = S_i + \Delta_{ik0} + \sum_{j=1}^{J_{ik}} \left(\Delta_{ikj} + \tau_{ikj} \right). \tag{2}$$

Since an alternate route may cover a longer distance than the preferred route, both holding delays and vectoring delays result in additional air carrier cost. In order to identify what constitutes holding and vectoring delay, we first need to compute the (undelayed) flight time for the shortest route* for each flight:

$$FT_{i,\min} = \min_{1 \le k \le NP_i} \sum_{j=1}^{J_{ik}} \tau_{ikj} \quad . \tag{3}$$

The global problem can then be written as:

$$\text{minimize } \sum_{i=1}^{N_A} \left[\omega_0 \Delta_{ik0} + \omega_1 \left(\sum_{j=1}^{J_{ik}} \left(\Delta_{ikj} + \tau_{ikj} \right) - FT_{i,\min} \right) \right] \delta_{ik} \tag{4}$$

where

$$\delta_{ik} = \begin{cases} 1 & \text{if path k is selected for flight i} \\ 0 & \text{otherwise} \end{cases}$$

$$\sum_{k=1}^{NP_i} \delta_{ik} = 1$$

and $\left\{ \Delta_{ikj} \right\}$ are selected based on the availability and holding capacity of each resource.

The objective function eq. (4) represents the goal of minimizing the total delay cost taken over all flights, but, using the iterative approach described below, other objective functions are easily substituted (such as minimizing the maximum delay over all flights).

C. Iterative Algorithm

The approach employs a dispatching rule, solving the problem for each aircraft given all of the previously scheduled events. Flights are sequenced according to one of several priority rules discussed below. When a flight *i* is evaluated, the algorithm considers each route *ik* in turn, computing the minimum total weighted delay if that route is

* Note: Some other metric like fuel optimized route or wind optimized route can also be used.

employed. For each such route, let m = 0, 1, 2, … index the iterations, and define the following quantities that will be computed at each iteration:

α^m_{ikj} = earliest possible arrival time at resource j after the m^{th} iteration

z^m_{ikj} = earliest possible time at which transit can begin at resource j after the m^{th} iteration.

If resource j has no holding capacity, then $\alpha^m_{ikj} = z^m_{ikj}$, but otherwise, it is possible that an aircraft can arrive at resource j and hold for some period of time prior to being allowed to transit. In general, we have:

$$\alpha^m_{ikj} <= z^m_{ikj}. \tag{5}$$

The initial values for α^m_{ikj} and z^m_{ikj} (iteration 0) are obtained by assuming that the flight does not need to be delayed anywhere along its path. In that case, the earliest arrival time at resource j is equal to the start time plus the sum of the transit times for each of the preceding resources along the path:

$$\alpha^0_{ikj} = z^0_{ikj} = S_i + \sum_{j'=1}^{j-1} \tau_{ikj'}. \tag{6}$$

Given the arrival times and transit times for each resource after the m^{th} iteration,

$$\left\{\alpha^m_{ikj}, z^m_{ikj},\ \text{j=1,...,J}_{\text{ik}}\right\}, \tag{7}$$

iteration $(m+1)$ proceeds in two steps, described next.

Step 1. Feasible holding and transit slots. Compute the minimum arrival delay $\delta\alpha^{m+1}_{ikj}$ and the minimum transit delay δz^{m+1}_{ikj} for each resource based on the current schedule for that resource:

$$\left\{\delta\alpha^{m+1}_{ikj}, \delta z^{m+1}_{ikj}\right\} = F\left(z^m_{ikj}\right). \tag{8}$$

where $F(\cdot)$ is a function that is determined computationally. **Figure 8** illustrates how this computation is performed. A possible event schedule is shown for aircraft i at the end of Step 1 of iteration m. A "possible" transit event is shown as the dotted rectangle, beginning at time z^{m-1}_{ikj} and lasting a duration τ_{ikj}. The transit event is termed "possible" because the resource schedule is not actually updated until the iterations have converged and a final routing decision is made for flight i. The time α^{m-1}_{ikj} is the earliest that the flight could arrive, prior to the possible transit event, since the resource's holding capacity (2 flights) is fully utilized prior to that time.

As shown in **Figure 9**, due to upstream effects, it is later determined (see Step 2) that the earliest transit time for resource j is z^m_{ikj}, which is later than the possible transit slot found in the previous iteration. As a result, the first transit interval available after z^m_{ikj} becomes the new "possible" transit event. Due to the holding capacity of the resource that also defines a new earliest arrival time. The definitions of $\delta\alpha^{m+1}_{ikj}$ and δz^{m+1}_{ikj} are then computed as illustrated in **Figure 9**. The illustration in the figure also assumes that α^{m-1}_{ikj} was unchanged during Step 2 of iteration m, so that $\alpha^m_{ikj} = \alpha^{m-1}_{ikj}$.

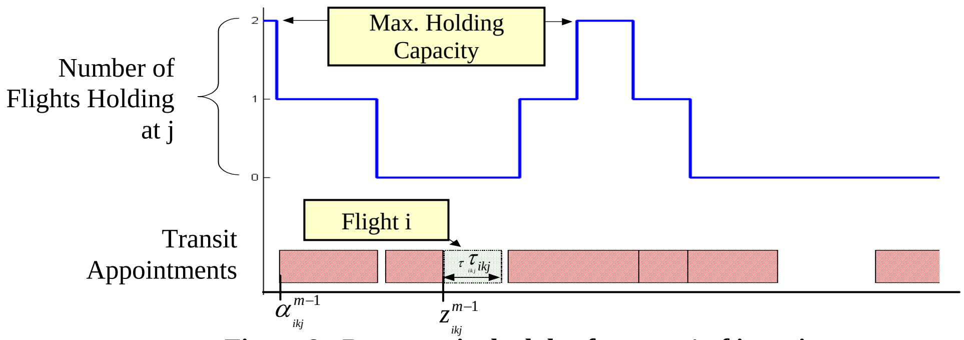

Figure 8: Resource *j* schedule after step 1 of iteration *m*.

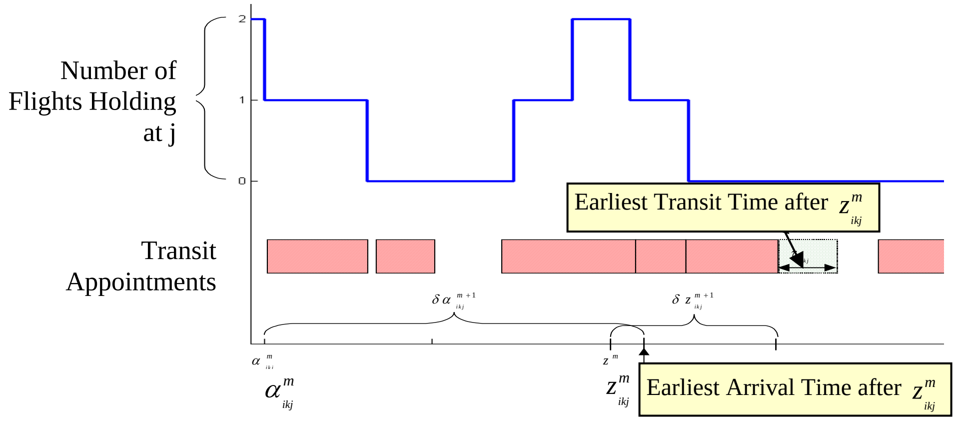

Figure 9: Resource *j* schedule after step 1 of iteration *m+1*.

Step 2. Transit and holding delays across all resources. Compute α^{m+1}_{ikj} and z^{m+1}_{ikj}, for $j = 1,\ldots,J_{ik}$, where:

$$\alpha^{m+1}_{ikj} = \max\left\{\alpha^{m}_{ikj} + \delta\alpha^{m+1}_{ikj}, z^{m+1}_{ik(j-1)} + \tau_{ik(j-1)}\right\} \quad (9)$$

$$z^{m+1}_{ikj} = \max\left\{\alpha^{m+1}_{ikj}, z^{m}_{ikj} + \delta z^{m+1}_{ikj}, \alpha^{m}_{ik(j+1)} - \tau_{ikj}\right\} \quad (10)$$

Equation (9) states that the earliest arrival time is the maximum of the earliest arrival time computed in Step 1 above and the current end of the transit appointment from the preceding resource along the path. Equation (10) states that the earliest transit time for resource *j* is the maximum of (i) the earliest arrival time, (ii) the earliest time transit can begin computed in Step 1 above, and (iii), the earliest arrival time at the next resource along the path minus the time it takes to transit resource *j*. Thus, transit delays propagate forward along the path (from source to destination) using eq. (9), and arrival delays propagate backwards along the path using eq. (10) as the iterations proceed.

Note that α^{m+1}_{ikj} and z^{m+1}_{ikj} are nondecreasing in *m*, so the iterative process is guaranteed to converge, and when both remain constant for an iteration, then the delays $\delta\alpha^{m+1}_{ikj}$ and δz^{m+1}_{ikj} returned from Step 1 are zero.

The iterations proceed until Step 1 returns arrival delays and transit delays equal to zero for all resources $j = 1,\ldots, J_{ik}$. When the iterations have finished, we denote the limiting values of α^{m}_{ikj} and z^{m}_{ikj} by α^{*}_{ikj} and z^{*}_{ikj}. The holding time at each resource is then just the difference between the transit time and the arrival time at that resource:

$$\Delta_{ikj} = z^{*}_{ikj} - \alpha^{*}_{ikj} \quad \text{for } j = 1, \ldots, J_{ik} \quad (11)$$

and the initial (gate/ground) holding time is the difference between the desired departure time and the arrival time at the first resource:

$$\Delta_{ik0} = \alpha^{*}_{ik1} - S_i \quad . \quad (12)$$

Finally, note that in the case that a resource has no holding capacity, the iterations simplify, since we need only worry about δz^{m}_{ikj} In that case, we can replace eq. (9) and eq. (10) in Step 2 with:

$$z^{m+1}_{ikj} = z^{m}_{ikj} + \max_{1\le j'\le J_{ik}}\left\{\delta z^{m+1}_{ikj'}\right\} \quad (13)$$

D. Flight Sequencing

Since the algorithm employs a dispatching rule, the priority order of the flights can have a large impact on the final result. Ordering schemes that are being tested include time order and conflict score. The latter ordering scheme reduces the scheduling priority of those flights that have the potential to cause the most disruption to the schedule.

E. Dual Dijkstra Algorithm

Among the required inputs to the algorithm is a set of alternative routes for each flight. These routes can be provided externally, based on, for example, air carrier flight routing software, Coded Departure Routes (CDRs), or Playbook plays, as illustrated in **Figure 10**. In contrast, the rerouting capability provided by the algorithm allows these routes to be generated and changed at any time and considered whenever the plan is updated. The run time of the algorithm increases linearly in the number of paths specified per flight. Since alternative routes are important during periods of congestion, there is a tradeoff between considering a large number of alternatives and the amount of computation time required to obtain a good solution.

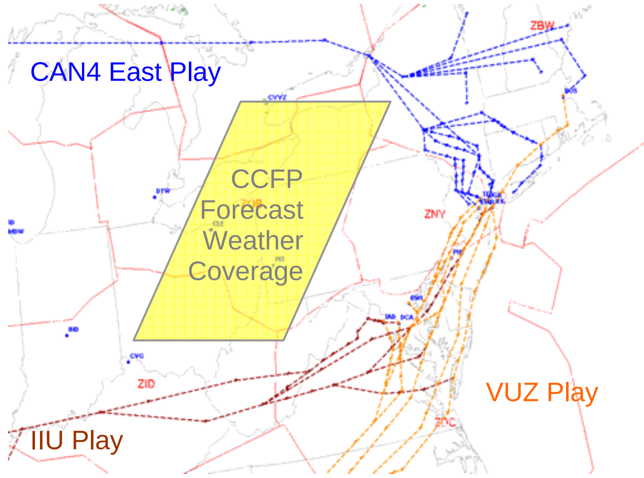

Figure 10: Playbook plays may provide well-defined solutions to avoiding CCFP weather constraints.

For research purposes, routes are generated internal to the algorithm based on the traffic network specification. The standard approach for deriving shortest paths in a network is Dijkstra's algorithm[18]. Given a source and a sink node, Dijkstra's algorithm searches to find the shortest path between the two nodes. A variation of this is a Dual-Dijkstra algorithm[19] that allows the number of alternative routes per flight to be controlled and also produces topographically distinct routes. That algorithm is described below.

When alternative paths are desired, we would like to have efficient paths that use different parts of the network, but still hold the total number of alternative paths considered to some manageable number. Standard algorithms for generating the *k* shortest paths do not consider how topologically different the paths are. The second shortest path, for example, tends to be very close to the same as the shortest path in most networks. The field of robotics has promoted academic interest in finding topographically distinct paths (to avoid collisions among multiple robots, for example). One such method is the Dual Dijkstra algorithm.

Our version of the Dual Dijkstra approach is as follows:

1. Given a source node *S* and a destination node *D*, perform Dijkstra's algorithm twice for every other node *n* in the network. That is, find the shortest path from *S* to *n* and then find the shortest path from *D* to *n*. Of course, combining those two paths yields the shortest path from *S* to *D* that goes through *n*. The number of resulting paths will vary, but will always be less than the number of nodes in the network. (To see this, note that the shortest path from *S* to *D* will be the shortest path that goes through every node along that path.)
2. Eliminate any duplicate paths, and then sort the paths obtained in Step 1 from shortest to longest. The topmost path will be the shortest path from *S* to *D*.
3. Prune the list of paths based on a topological separation rule. The rule we have employed is to specify a "separation parameter" *p* and, starting with the first path on the list, eliminate any path that does not contain at least *p* different nodes from any previously selected path. (We also eliminate any path that uses the same link more than once.) When $p=1$, this procedure returns all of the paths generated in Step 1. When *p* is large enough, only the shortest path will be returned.

For generating en route paths, we have added a distance filter which limits the total length of a path relative to the length of the shortest path. An example of applying this rule for a network is shown in **Figure 11(a)**. By eliminating nodes and arcs of the network corresponding to an FCA and then controlling the amount of additional distance that may be included in any alternate route, the Dual Dijkstra algorithm can generate paths that circumvent the FCA. An FCA added to the example network produces the flight paths shown in **Figure 11(b)**. Thus, we have the ability to schedule routes that specifically fly through the FCA or avoid the FCA, based on the capacity of the FCA and the equity between users.

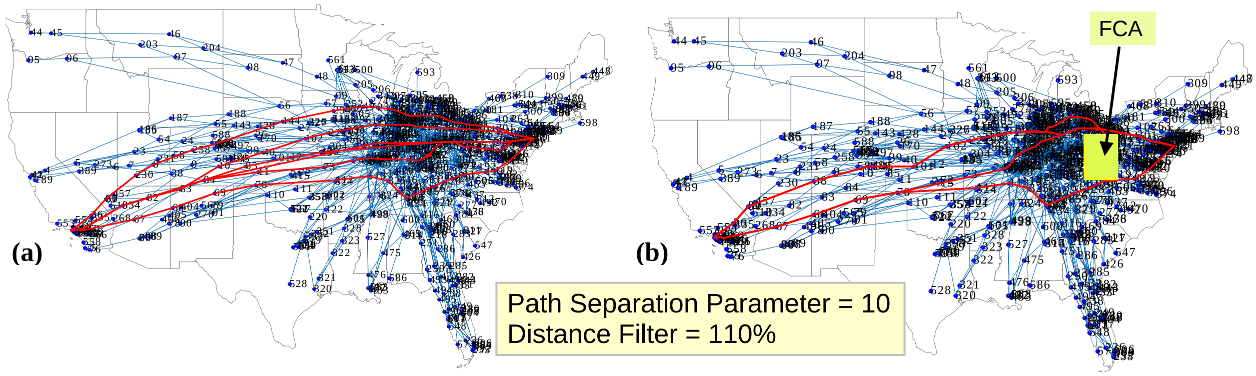

Figure 11: Sample Coast-to-Coast Routing Around and Through Restricted Area generated with Path Separation and Distance Filter.

VI. Methodology

We compare five cases for the operational scenario: (1) good weather, (2) zero capacity FCAs, (3) FCAs with dynamically-varying capacity, (4) FCAs with dynamically-varying weather that is better than planned for, and (5) dynamically varying weather that is worse than planned for - with respect to metrics such as total ground delay, total airborne delay due to holding, and total airborne delay due to rerouting. As a basis for comparison, we first obtain results under the assumption that weather forecasting abilities are perfect. We then apply the planning paradigm of **Figure 2** using weather forecasts that degrade in quality as we look further into the future.

Weather scenarios were modeled by synthesizing weather avoidance routes for varying weather cases within the FCA. The sections of the network that fell within the FCA were removed. Each weather avoidance route was connected to the network via intersection points between the network and FCA boundary. The process is illustrated in **Figure 12**. Each link removed from the original network is replaced by a link connecting a node outside the FCA to the nearest entry or exit point for one of the weather avoidance routes.

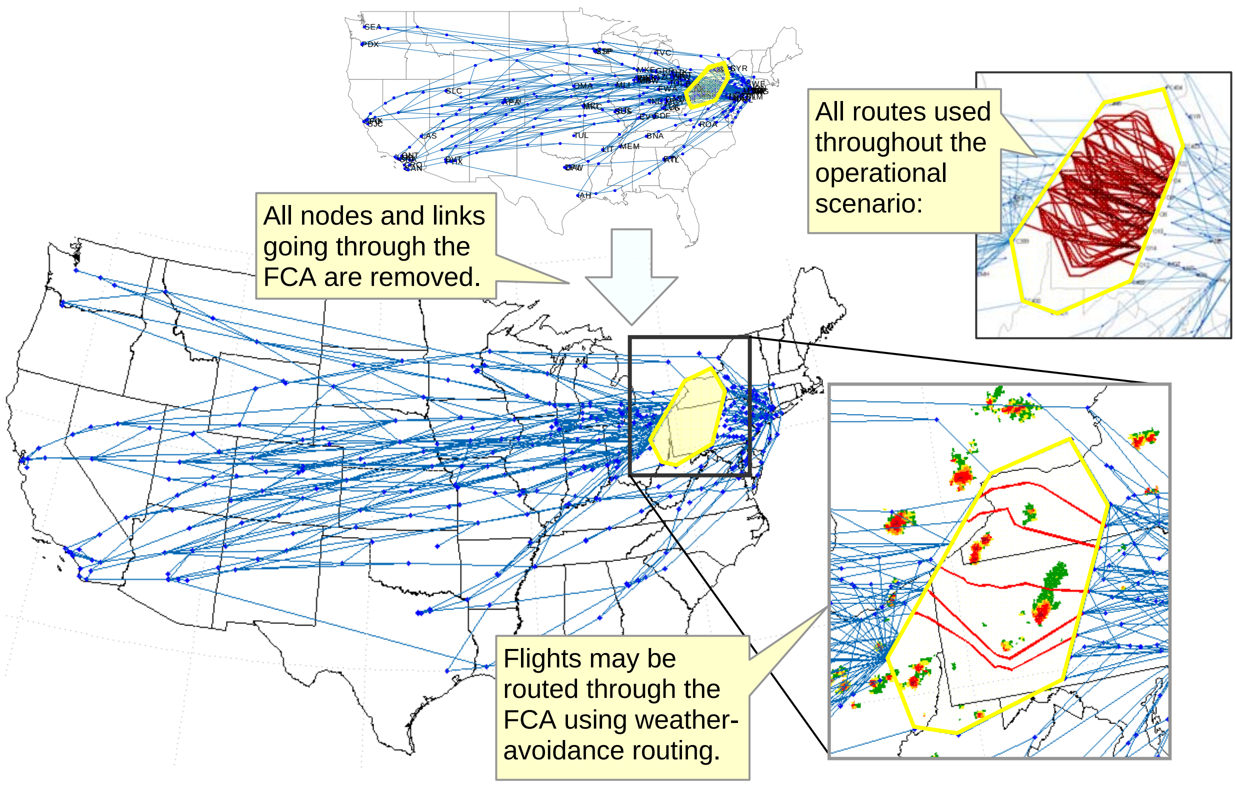

Figure 12: Network generated with nodes and links inside the FCA removed and replaced with Synthesized Weather Avoidance routes.

In order to route aircraft through the FCA, an estimate had to be made for the expected capacity of the FCA given a weather forecast. Applying the flow-based routing technique[20,21] to an FCA not only ensured safe routes across the FCA, but also facilitated an accurate capacity estimate (as the route status, flight distance, and flight duration were known). Flow-based routing was used to ensure safe separation from hazardous weather for flows of aircraft over a specified duration of time. By specifying a direction of flow, a minimum safe aircraft separation (both lateral and longitudinal), and a minimum safe distance for separation from hazardous weather, these routes could be generated for the FCA. Furthermore, these routes could be updated throughout the time span of the FCA as new weather forecasts were generated.

Weather forecasts were varied using increased and decreased severe weather coverage within the FCA. Weather avoidance routes were then synthesized for each of these forecasts. The status of each route, either open or closed, was used to forecast the flow rate through the FCA. The status of the weather avoidance routes within the defined FCA for the operational scenario is shown in **Figure 13** and **Figure 14**. The route is either indicated as open or closed. In all cases, the fraction of time that the routes are open decreases as hazardous weather coverage increases. However, because routes are deconflicted, it is possible that a particular route will be available under a more severe weather case when it is not available in a less severe weather case. This is evident when comparing the Actual Weather case with the Increased Weather case for Route 2.

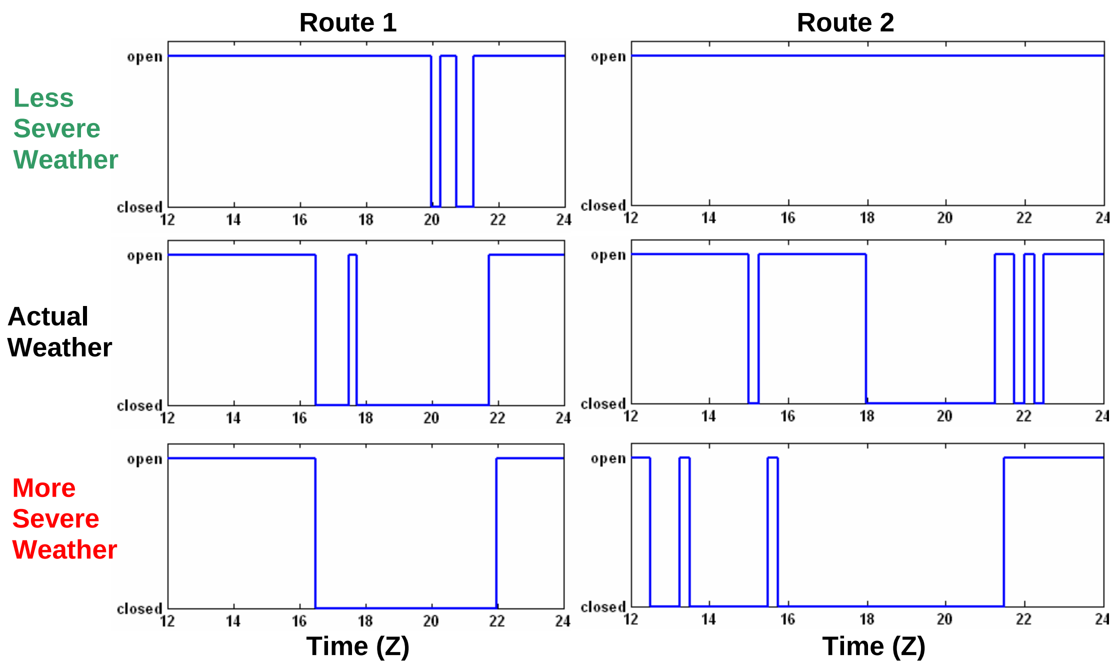

Figure 13: Open/Closed conditions for Routes 1 and 2.

The total flow rate through the FCA is estimated by summing the open/closed conditions of all routes. Results are shown in **Figure 14** for the actual weather and varied weather cases. With less severe weather, the flow rate is near maximum for most of the scenario. Given the actual weather data, there is a period with zero throughput and overall the throughput is lower. With more severe weather, this period of zero throughput is extended by over an hour, while the overall throughput rate is reduced further.

The operational scenario begins at 12:00Z and continues for approximately 8 hours. Under the assumption that we have a perfect weather forecast for the entire 8-hour period, we need only develop a single plan at 12:00Z and use that plan to assign delays and routes to all flights. If, more practically, the forecast is imperfect and changing, then the plan developed at 12:00Z must be periodically revised using updated forecast information. To study the latter situation, we define the "accurate forecast time horizon" to be the length of time looking forward into the future that the weather forecast correctly describes the upcoming weather. Beyond that time horizon, the forecast may indicate that the weather will be more severe or less severe than the actual weather will turn out to be.

A variety of weather scenarios were created by piecing together the open/closed status information from a spectrum of weather forecasts (e.g. actual, increased coverage, decreased coverage). In order to create a forecast with 2-hour accuracy, for example, the actual weather was used for the first 2 hours and spliced together with either an increased or decreased weather case for the remainder of the operational scenario. **Figure 15** illustrates the process.

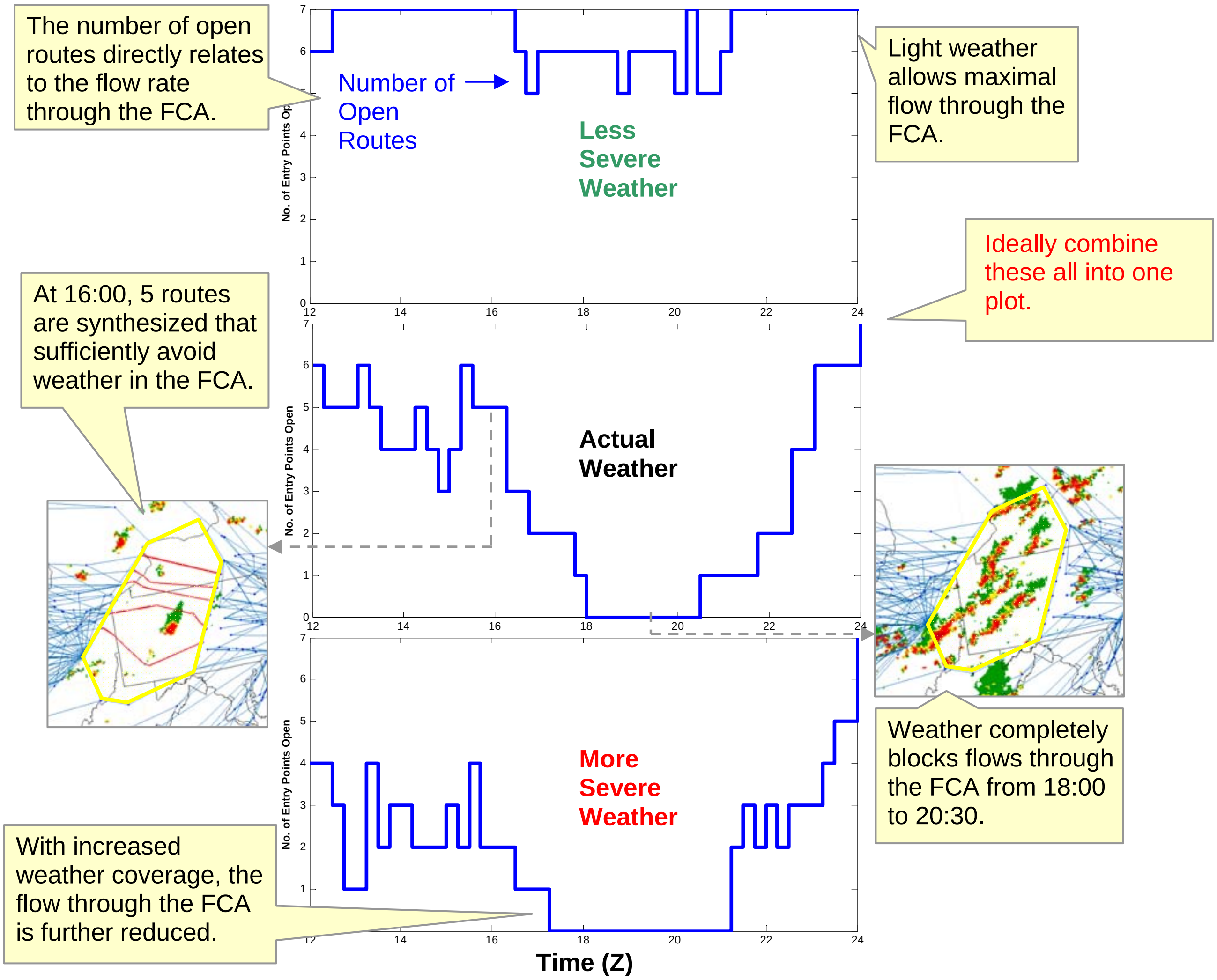

Figure 14: Number of open routes throughout the operational scenario.

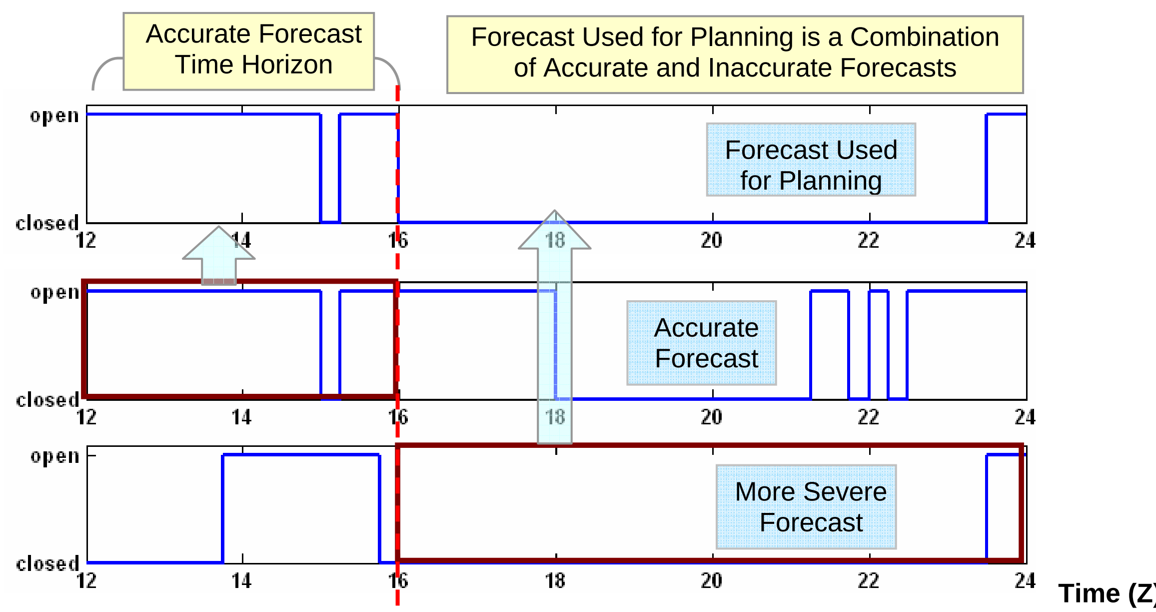

Figure 15: Development of weather forecast scenarios based on route status.

In this example, the first two hours from the accurate forecast are spliced with the remaining 8 hours from the more severe weather forecast. By planning with this new forecast at 12:00Z, the traffic managers would have overestimated the impact of the weather after 14:00Z. Each hour within the operational scenario, the plan would be updated using a similarly constructed forecast. At 13:00Z, the planning forecast would be accurate until 15:00Z, and then overestimate weather severity beyond 15:00Z. In the test cases described below, we varied the accurate forecast horizon from 1 to 6 hours.

VII. Results

Results for the five cases described above are summarized in **Table 1**. All results presented in the table are obtained under the assumption that the associated weather forecasts are perfect. There are a total of 168 flights for which departure times and routes have to be assigned. When the weather is clear, all of the flights depart on time and follow a set of deconflicted routes to their respective destinations. Forcing all flights to fly around the FCA results in a total of 97.8 hours of weighted delay, most of which is due to additional distance flown relative to the clear-weather case. (In computing weighted delay, it is assumed that ground delay is one-half the cost of rerouting delay, i.e. one hour of ground delay is equal to one-half hour of weighted delay.) When weather avoidance routes through the FCA can be used, the total weighted delay ranges from 9.3 to 62 hours, depending upon the severity of the weather. As is evident from the results shown in the table, the actual weather case and the more severe weather case very close in terms of total weighted delay cost.

Table 1: Results for Various Weather Situations.

Perfect forecasting assumed.

Case	Delay (Hours) Ground Delay	Rerouting Delay	Total Weighted Delay
Clear Weather	0.0	0.0	0.0
Avoid FCA	0.4	97.6	97.8
Dynamically Route Through FCA			
Actual Weather	3.9	53.7	55.7
Less Severe Weather	2.6	8.0	9.3
More Severe Weather	3.3	60.4	62.0

As described above, imperfect weather forecasts were created by first selecting an accurate forecast time horizon and then combining the actual weather forecast with an adjusted forecast that was either more severe or less severe than the actual weather. Weighted-delay results for accurate forecast horizons between 1 and 6 hours are shown in **Figure 16**. The curves are based on a planning update interval of 1 hour. At the end of each hour, positions of all flights are updated, a new forecast is received, and the current plan is revised based on changes in the future availability of weather avoidance routes through the FCA. This may result in rerouting some flights, either adding or reducing the total distance flown, and also in adjustments to the assigned ground delays for flights that have not yet departed. The points plotted in the figure represent the total delays realized by all flights at the end of the scenario, after all such hourly plan updates have been implemented.

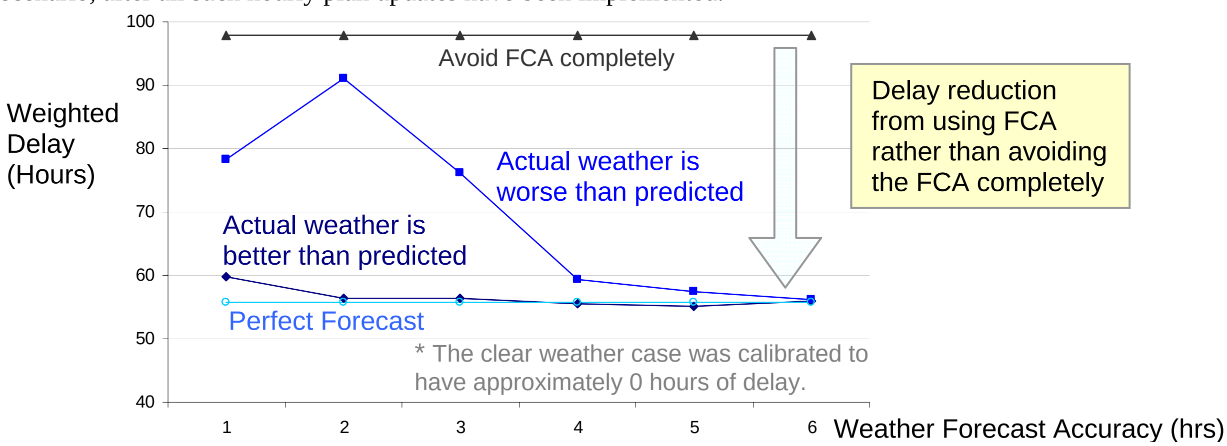

Figure 16: Delay occurring when weather materializes either better or worse than the forecast (weighted delay considers air delay to be twice as costly as ground delay.)

In the opposite case, when the actual weather is worse than predicted, the cost of forecast inaccuracy is much higher unless the accurate forecast horizon is on the order of four hours or more. For shorter forecast horizons, a less severe forecast results in more flights being rerouted as they approach the FCA and learn that the expected path through the FCA will not be available. Also, higher levels of ground delay are imposed on flights because it appears that routes through the FCA will open soon enough to justify holding those flights on the ground rather than sending them on more circuitous (and expensive) paths around the weather. Decisions of the latter type are the principle reason for the apparent anomaly at the 2-hour forecast horizon. In that particular case, many flights are assigned long ground delays, one hour at a time, because it keeps appearing that the weather will clear in another hour, and one hour of ground delay costs less than sending the flights on alternate paths around the FCA that require more than 30 minutes of additional flight time. Because the algorithm employs a look-ahead heuristic approach to scheduling, it is possible that better information can result in poorer decision-making, as is evident in the sample scenario. However, this anomaly does suggest additional enhancements to our model of the planning process which would reduce or eliminate this anomaly, including:

- use of nonlinear costs for ground-holding (ω_0 in eq. (4) above),
- use of a more detailed network, enabling shorter reroutes around the FCA,
- use of a shorter replanning cycle,
- use of stochastic forecasts (e.g. through probabilistic weightings applied to alternate forecasts).

The above list suggests directions for additional research. In addition, incorporation of many of the features of the current research into the CRRAT algorithm[15], followed by human-in-the-loop testing, should lead to improved traffic management decision-support tools for dealing with weather-related resource constraints.

VIII. Conclusion

This paper investigates the planning of takeoff times and route selection to better meet the capacity constraints of en route Flow Constrained Areas (FCAs) during large-scale weather events. A Ground Delay Program (GDP) approach is investigated. FCA-based GDP strategies demonstrate that greater control can be exhibited using flow management on traffic impacted by en route constraints. A dynamic FCA capacity-estimation algorithm uses weather forecast information to produce time-varying entry and exit points to FCAs as well as maximum flow rates for FCAs. Then, a variant of the airport-based GDP algorithm enables traffic to be scheduled for optimal use of the FCA, while allowing for holding or rerouting of airborne flights as the weather constraints in the FCA become known with greater certainty. This method of traffic flow management provides a mechanism for the assessment of the value of improved weather forecast accuracy, and provides insights into the nature of robust traffic management initiatives.

IX. Acknowledgment

This research was funded by NASA Ames Research Center under contract NAS2-02075. We are greatly appreciative to our NASA Technical Monitor, Matt Jardin, Ph.D., for his guidance, which helped focus the effort. Finally, we appreciate the financial support of the sponsor of the research, NASA Ames Research Center and the Virtual Airspace Modeling and Simulation (VAMS) Project Manager, Mr. Harry Swenson.

X. References

[1]Ball, M.O., Chen, C.-Y., Hoffman, R., and Vossen, T., "Collaborative Decision Making in Air Traffic Management: Current and Future Research Directions", in *New Concepts and Methods in Air Traffic Management*, Bianco, L, Dell'Olmo, P, and Odoni, A., Editors, Springer-Verlag, 2001.

[2]Ball, M.O., Hoffman, R.L., Knorr, D., Wetherly, J., and Wambsganss, M., "Assessing the Benefits of Collaborative Decision Making in Air Traffic Management," in A*ir Traffic Systems Engineering*, Donohue, G.L., and Zellweger, A.G., Editors, AIAA Press, Reston, VA, 2001.

[3]Wambsganss, M.C., "Collaborative Decision Making in Air Traffic Management," in *New Concepts and Methods in Air Traffic Management*, Bianco, L, Dell'Olmo, P, and Odoni, A., Editors, Springer-Verlag, 2001.

[4]Wambsganss, M.C., "Collaborative Decision Making through Dynamic Information Transfer," *Air Traffic Control Quarterly*, Vol. 4, pp. 107-123, 1997.

[5]Hoffman, B., Krozel, J., and Jakobovits, R., "Potential Benefits of Fix-Based Ground Delay Programs to Address Weather Constraints," *AIAA Guidance, Navigation, and Control Conf*., Providence, RI, Aug., 2004.

[6]Ball, M.O., R. Hoffman, A. Odoni, and R. Rifkin, "Efficient Solution of a Stochastic Ground Holding Problem," *Operations Research*, Vol. 51, pp. 167-171, 2003.

[7]Bertsimas, D., and Stock Patterson, S., "The Air Traffic Flow Management Problem with Enroute Capacities," *Operations Research*, Vol. 46, No. 3, May-June 1998.

[8]Garey, M. R., and Johnson, D. S., *Computers and Intractability*, Freeman Pub., 1979.

[9] Bertsimas, D., and Stock Patterson, S., "The Traffic Flow Management Rerouting Problem in Air Traffic Control: A Dynamic Network Flow Approach," *Transportation Science*, Vol. 34, No. 3, Aug., 2000.

[10] Smeltink, J. W., Soomer, M. J., de Waal, P. R., and van der Mei, R. D., "Optimisation of Airport Taxi Planning," *Elsevier Science*, June, 2004.

[11] Visser, H. G., and Roling, P. C., "Optimal Airport Surface Traffic Planning using Mixed Integer Linear Programming," *Proc. 3rd Annual Aviation Technology, Integration, and Operations Technical Forum,* Denver, 2003.

[12] Sherry, J.E., Ball, C.G., and Zobell, S.M., "Traffic Flow Management (TFM) Weather Rerouting", *Proc. of the 4th USA/Europe ATM R&D Seminar*, Santa Fe, NM, Dec., 2001.

[13] Rhodes, L.S., Rhodes, L.R., and Beaton, E.K., *CRCT Capabilities Detailed Functional Description*, Tech. Report 00W0000302, MITRE, Center for Advanced Aviation System Development, McLean, VA, March, 2001.

[14] Burke, J.M., *Implementing and Evaluating Alternative Airspace Rationing Methods*, M.S. Thesis, University of Maryland, College Park, 2002.

[15] Hoffman, R., J. Burke, T. Lewis, A. Futer, M. Ball. "Resource Allocation Principles for Airspace Flow Control", *AIAA Guidance, Navigation and Control Conf.*, San Francisco, CA., Aug., 2005.

[16] Neels, K., "Pricing-Based Solutions to the Problem of Weather-Related Airport and Airway System Delay", Air Traffic Control Quarterly, Vol. 10, No. 3, pp. 261-284, 2002.

[17] Howard, K., *ETMS/ATMS System Requirements*, Version 1.2, Volpe Research Center, Report to the Dept. of Transportation, Nov. 8, 2002.

[18] Dijkstra, E.W., "A Note on Two Problems in Connection with Graph Theory," *Numerische Mathematik*, Vol. 1, pp. 269-271, 1959.

[19] Fujita, Y., Nakamura, Y., and Shiller, Z., "Dual Dijkstra Search for Paths with Different Topologies," *Proc. of the IEEE Intern. Conf. on Robotics and Automation*, Sept, 2003.

[20] Krozel, J., Penny, S., Prete, J., and Mitchell, J.S.B., "Comparison of Algorithms for Synthesizing Weather Avoidance Routes in Transition Airspace," *AIAA Guidance, Navigation, and Control Conf.*, Providence, RI, 2004.

[21] Prete, J., and Mitchell, J.S.B., "Safe Routing of Multiple Aircraft Flows in the Presence of Time-Varying Weather Data," *AIAA Guidance, Navigation, and Control Conf.*, Providence, RI, 2004.