Proceedings of the 11th Conference on Aviation, Range and Aerospace Meteorology, Hyannis, MA 2004

# 2.3 QUANTIFYING DELAY REDUCTION BENEFITS FOR AVIATION CONVECTIVE WEATHER DECISION SUPPORT SYSTEMS*

James E Evans, Shawn Allan, Mike Robinson
MIT Lincoln Laboratory

## 1. INTRODUCTION

In this paper, we summarize contemporary approaches to quantifying convective weather delay reduction benefits. We outline a program to develop a significantly improved capability that can be used to assess benefits of specific systems. This program may potentially accomplish weather impact normalization for studies of National Airspace System (NAS) performance in handling convective weather.

Benefits quantification and NAS performance assessment have become very important topics for the aviation weather community. In an era of significant federal government and airline budget austerity for civil aviation investments, it is essential to quantitatively demonstrate delay reduction benefits of improved weather decision support systems. Major FAA initiatives stress the importance of quantitative system performance metrics that are related to aviation weather. For example, the new FAA Air Traffic Organization (ATO) and the FAA Flight Plan 2004-08 both have quantitative performance metrics that are closely related to reducing convective weather delays. The Flight Plan metrics include: "Improving the percentage of all flights arriving within 15 minutes of schedule at the 35 OEP airports by 7%, as measured from the FY2000-02 baseline, through FY08," and "Maintaining average en route travel times among the eight major metropolitan areas." The ATO metrics include the percentage of on time gate arrivals and the fraction of departures that are delayed greater than 40 minutes. However, these metrics currently do not account for the differences in convective weather severity and changes in the NAS. The dramatic increase in convective season delays in 2004 (Figure 1) due to a combination of severe weather, increases in overall demand, and specific airport issues has demonstrated that one needs to consider these other factors.

Approaches to delay reduction quantification that were viewed as successful and valid several years ago are no longer considered to be adequate by either by the FAA investment analysis branch or by the Office of Management and Budget (OMB).

The paper proceeds as follows. We first discuss at some length the mechanisms by which convective weather delay occurs in the NAS and highlight challenges in delay reduction assessment. We consider this to be very important since one needs to understand how the system operates if one is to design an effective, accurate performance assessment system.

We then consider benefits quantification based on feedback from experienced users of a system. Feedback on "average" benefits from a system at the end of a test period was used to generate delay reduction estimates for the Integrated Terminal Weather System (ITWS) and the Weather and Radar Processor (WARP). This end-of-season interview approach was not viable in highly congested en route airspace. Hence, a new approach was developed for Corridor Integrated Weather System (CIWS) benefits assessment that uses real time observations of product usage during convective weather events coupled with in depth analysis of specific cases.

Next, we discuss the problems that arise when one attempts to quantify delay reduction benefits by comparing flight delays before and after the Integrated Terminal Weather System (ITWS) system was deployed at Atlanta Hartsfield International Airport (ATL). This seemingly simple approach has proven very difficult in practice because the convective weather events in the different time periods are virtually never identical[1] and because other aspects of the NAS may also have changed (e.g., user demand, fleet mix, and

*This work was sponsored by the Federal Aviation Administration under Air Force Contract No. F19628-00-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

[1] In a later section, we show by combinatorial arguments that achieving perfectly identical storm impacts on separate events is extremely unlikely.

other systems that impact convective weather delays). It has become clear that one needs a quantitative model for the NAS that would permit adjustment of measured delay data to account at least for the differences in convective weather and changes in user demand (i.e., flight scheduling).

The paper concludes with recommendations for measuring near term benefits of various classes of convective weather decision support systems.

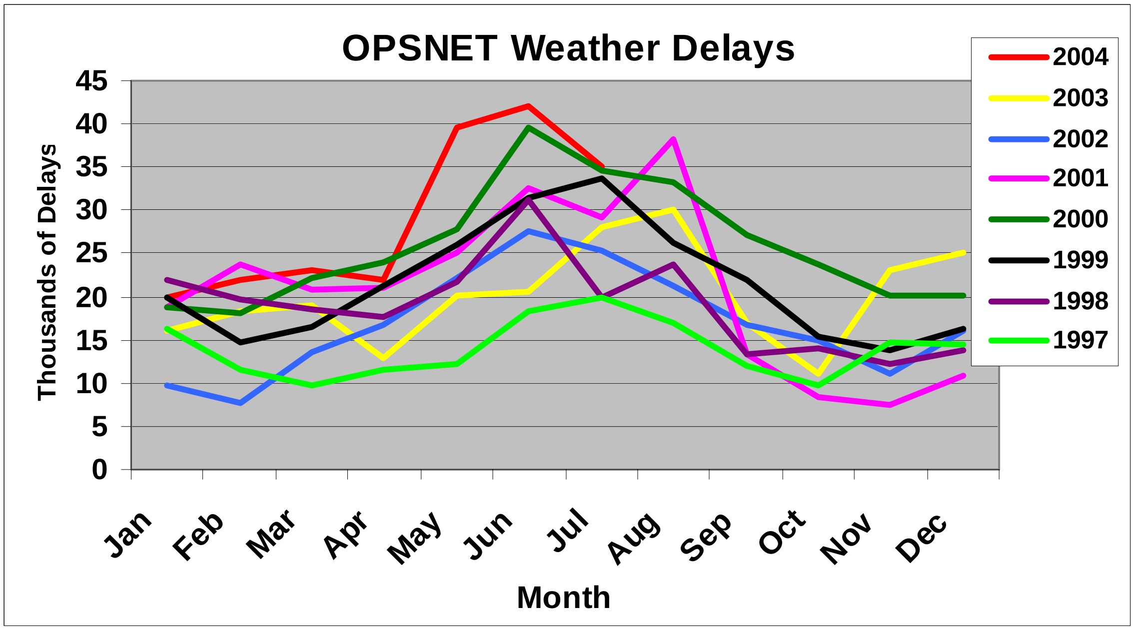


*Figure 1. OPSNET delays by month.*

## 2. CAUSES OF CONVECTIVE WEATHER DELAYS

It is important to have an in-depth understanding of the cause of convective weather delays if one is to design benefits quantification systems and develop effective NAS performance metrics. In this section, we will briefly review insight developed from our analysis of delays over the past decade. Historically, one finds three major theories of convective weather delay causality in the literature:

- blockage of routes between terminal areas by en route weather (the "Spring 2 K" effort, Post et al., 2002 and Callahan et al., 2001),
- thunderstorms near or over airports (Bond, 1997)[2],.
- reduced airport capacity associated with low ceiling-and-visibility conditions at the airport during thunderstorms.

However, it is now clear from the detailed studies of convective weather delays and traffic handling at Dallas, Memphis, New York and the northeast quadrant of the U.S. over the past decade that <u>none</u> of the above theories <u>alone</u> provides an adequate explanation.

### *A. Convective Weather Impacts within the Terminal Radar Approach Control (TRACON)*

The typical structure of traffic flows between major terminals[3] and the surrounding airspace is shown in Figure 2. Planes enter (blue dashed arrows) the TRACON at corner fixes and depart

[2] It should be noted that Bond suggests that such delays cannot be reduced by a change in ATC technology because aircraft cannot fly safely in thunderstorms.

[3] This structure applies at nearly all of the major terminals that encounter convective weather delays with the notable exception of the New York terminal area, which has a much greater number of arrival and departure transition fixes.

(green arrows) through fixes on the sides of TRACON. Overall TRACON width is typically 100 nmi. Within the TRACON, there are a variety of routes from the corner fixes to the runways. The transitions between TRACON and en route are relatively inflexible due to facility differences and the structure of the en route sectors (purple lines). Within the TRACON, there is a great deal of flexibility to vary routes between the airport and the arrival and departure fixes.

Convective weather impacts on the airport (e.g., within 5 nmi of the airport) cause delays because they will often reduce the capacity of the runways. However, convective storms do not close all of the runways very often. R. Ferris of Lincoln Laboratory conducted a study of terminal operations during convective weather events at Orlando (MCO) and Dallas-Ft. Worth (DFW) airports. In only one of 20 storm day cases examined did arrival and departure traffic stop for 15 minutes or more when thunderstorms were over the runways or within 5 nmi of an airport. Storms were within 5 nmi of the airport for 886 minutes over 10 days at MCO and 992 minutes over 10 days at DFW.

Rhoda et al. (1999) and (2002) found that a very high percentage of arriving aircraft are more likely to fly through high-reflectivity storms (typically VIP level 3 or higher) if they are within 10 nmi of the airport. The same aircraft likely would not fly through such storms if the storms were located at one of the arrival fixes shown in Figure 2. Experience at the ITWS demonstration sites (MCO, DFW, MEM, and NY) over the past decade shows that arriving planes will frequently penetrate disorganized convective storms (e.g., air mass) when landing, but generally will not penetrate vigorous squall lines moving across the airport. Moreover, air mass storms typically only block one of the runways at a major airport (e.g., MCO, DFW, ORD, or ATL) so that arrival and departure operations can continue on the other runways. Since vigorous squall lines typically are over the airport for a relatively short period (e.g., less than 30 minutes[4]), the overall time that the runways are not used is a relatively small fraction of the time that the squall line is disrupting terminal operations. Thus, convective weather at or very near the airport may reduce the traffic flow but typically will not halt the flow of arrivals and departures completely.

### *B. Reduced Capacity at Major Airports Due to Instrument Meteorological Conditions (IMC)*

If the IMC capacity of an airport is the main constraint, then convective weather forecasts could do little to reduce the delays. Experience at the ITWS demonstration sites shows that the capacity reductions associated with convective weather within 100 nmi of the airport are generally much greater than the airport capacity reduction associated with IMC conditions. A study of high-delay thunderstorm events at Newark Liberty International Airport (EWR) revealed that half of the events had only visual meteorological conditions (VMC) conditions throughout the event, with the duration of VMC time exceeding IMC time duration for another 18% of the events (Allan et al., 2001). On the other hand, the highest EWR delay events also had a higher fraction of IMC conditions. It is our observation that IMC conditions rarely are the main constraint during a convective event in or near a TRACON. However, IMC at the end of a convective weather impact event is an important factor in the magnitude of queue delays (discussed below) that occur. Such conditions can significantly extend the delay recovery period if a ground delay (GDP) or ground stop (GS) program was put into effect to reduce the arrival demand at the airport.

### *C. Storm Impacts on the TRACON-En Route Arrival and Departure Fixes*

Planes at the TRACON-En Route arrival and departure fixes (Figure 2) are typically flying at altitudes of less than 20 kft, so they cannot fly over convective storms. In addition, at these fixes pilots typically will not penetrate storms with an equivalent reflectivity of VIP level three or higher (Rhoda et al., 1999 and 2001). It is difficult to dynamically change the location of these transition fixes during a convective weather event because both the TRACON and ARTCC have designed their internal route structures and procedures based on the fix locations. The result is that storms at the transition fixes will stop the flow of travel through that fix. When an arrival fix is closed, the arrivals that would normally use that fix are either routed to another arrival fix (e.g., from the northeast to the northwest fix in Figure 2) and/or departures are halted while the departure fix is used to handle arrivals.

[4] A 10 nmi wide squall line moving at 30 mph would traverse the 4 nmi diameter circle that typically covers an airport's runways in about 30 minutes.

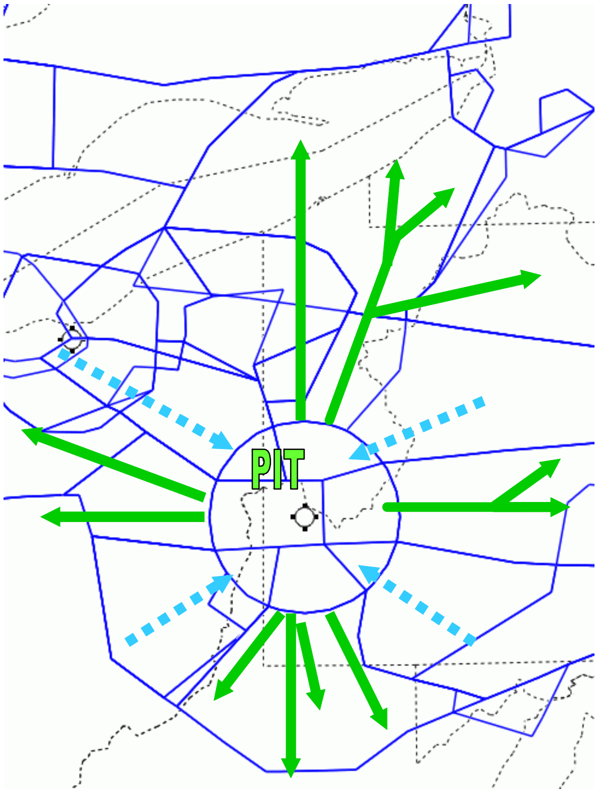


*Figure 2. Typical terminal area arrival and departure route structure.*

Rerouting to another arrival fix, although preferred procedurally, causes delays from two mechanisms.

- Planes are now flying a longer distance than would have been the case had the convective weather not been present (as illustrated in Figure 3).

- There may be queue delays at the other arrival fix if the arrival demand at that fix exceeds the fix capacity[5].

We now discuss how these delay generation mechanisms can be modeled quantitatively.

*1." Linear" delay model*

In Figure 3, the red line is the longer path flown by a plane when an arrival fix for the TRACON is closed. The green line shows the path flown if the plane is proactively routed to another arrival fix. Both paths represent longer distances than the aircraft would have flown had the convective weather not been present. However, the green path (resulting from proactive routing) is shorter than the red path. The extra distance flown along the red path typically corresponds to 20 minutes of additional flight time. With proactive routing, the extra distance flown on the green path depends on how far "upstream" a path change is initiated. The delay mechanism and possible delay reduction approach illustrated apply to many other convective weather delay situations:

- Flying around storms within the TRACON,

- Flying around storms in en route airspace, and

- Major reroutes in en route airspace to avoid a region of convective weather (as opposed to a cell). (See Brennan et al., 2004 for a discussion of rerouting for transcontinental flights.)

In all of these cases, the delay can be written as a product:

Delay ≈ (number of aircraft impacted) x (extra distance per aircraft) x (number of such events) (1)

≈ (demand) x (time duration of an event) x (extra distance flown per aircraft per event) x (number of events per year)

Convective weather delay in this respect is <u>linear</u> in the demand and the extra distance flown. Since the extra distance flown is related to the storm spatial extent, the delay in equation 1 would typically be linear in the storm spatial extent as well.

*2. Non-linear queue model*

Queue delays have a very different, <u>nonlinear</u> dependence on demand and the time duration of events[6]. Conceptually, the capacity associated with an Air Traffic Control (ATC) facility (e.g., an airport or an en route sector) is reduced by

[5] Several examples of such queues at other arrival fixes are discussed in section 8.

[6] The queue model we are discussing here is the classical deterministic queue model that is a standard model in traffic analysis (See, e.g., Newell, [1982] or Daganzo, [1997]). With this model, one can operate quite close to capacity without incurring delays, which is quite different from the random process queue models (e.g., a Poisson process). With random process queues, delay rises rapidly as the demand approaches the capacity.

convective activity from the fair-weather capacity ($C_V$) to a lower convective weather capacity ($C_W$) for a time duration, T. Typically, $C_V$ is greater than the demand, D, but $D > C_W$.

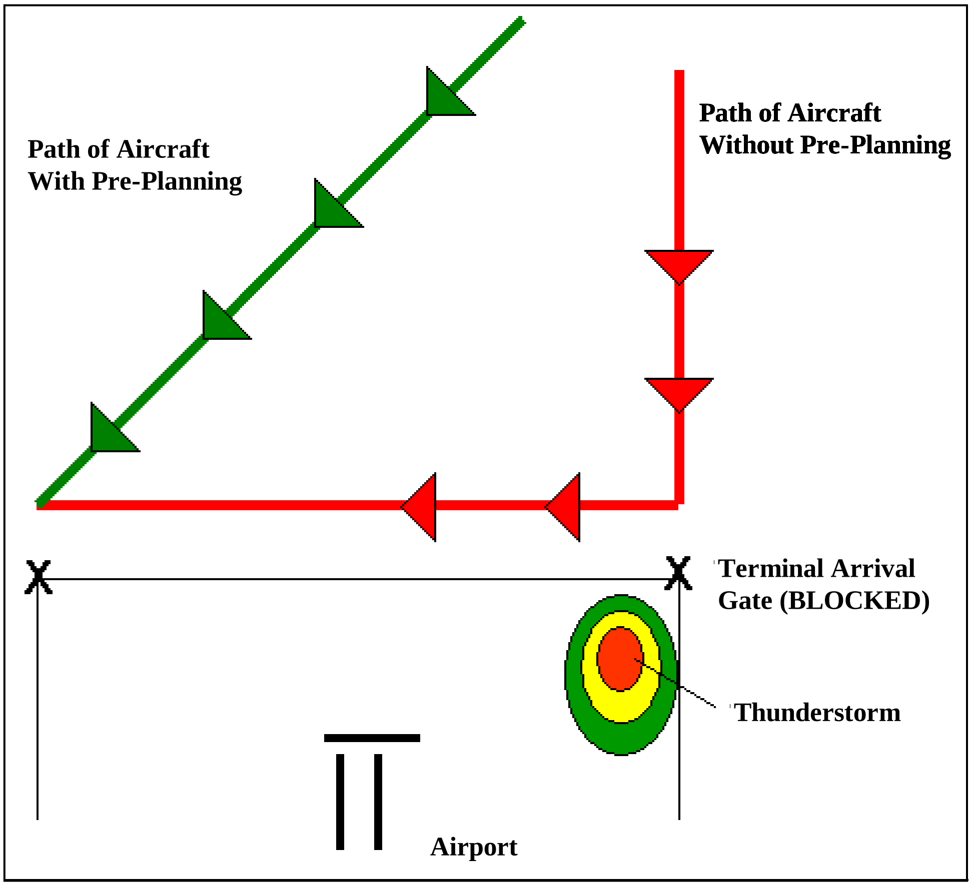


*Figure 3. Example of the "longer path flown" mechanism by which delays occur and how the delay might be reduced through the use of improved aviation weather products.*

During the low capacity event, the number of aircraft that must wait in a queue (either in the air or on the ground) at time t is given by $(D - C_W) \times t$ for $t < T$. Since the waiting time in a queue for a plane joining the queue at time t is simply the length of the queue at that time divided by the capacity (e.g., $C_w$), the waiting time also goes up proportional to t. It is straightforward[7] to show that the accumulated delay for all the aircraft involved in the queue is:

$$\Sigma \text{ (delay to } k^{th} \text{ aircraft)} = 0.5\, T^2\, (D-C_W)\, (C_V-C_W)/(C_V-D) \qquad (2)$$

Where the summation is over k

[7] Validation of the model using actual delay data from Atlanta is discussed in Appendix D of Evans, et. al. (1999)

Here we see that there is a very nonlinear relationship between demand, fair weather capacity, convective weather capacity, and time duration.

To illustrate, a key difference between the very busy airports and lower volume terminals, in terms of sensitivity to convective weather, is that busy TRACONs with major air carrier hubs (such as Atlanta, Dallas, or Chicago) have such high volumes that congestion queues can arise at the terminal fixes in fair weather. At such TRACONs, very large airborne queues (e.g., aircraft in holding patterns) will rapidly develop when a single arrival fix is closed by convective weather[8]. If two or more arrival fixes are closed, ATC at major airports will often use one or more departure fixes to handle

[8] Several examples of airborne holding at ATL due to terminal convective weather are discussed in section 8.

the arrival stream. This loss of departure capacity then results in very long queues of departures on the ground.

Queues are also associated two of the three theories of convective weather delays discussed earlier [closure (or near closure) of an airport due to storms on the runways and delays due to low IMC conditions at the airport associated with convective weather] will nearly always involve queues. Finally, the use of ground delay programs (GDPs) and/or ground stops (GS) to manage terminal convective weather impacts results in queues where the aircraft are held on the ground at the origin airport.

**D. En Route Loss of Capacity due to Convective Weather**

When convective weather shuts down key en route sectors in congested airspace, airborne traffic is typically given priority because of fuel constraints and the disruption that holding patterns cause in congested areas. As a result, departures on the ground end up heavily restricted and long departure delays frequently occur.

The delays that arise when convective weather causes an en route loss of capacity are particularly difficult to model since a given en route sector may be used to handle flights from many different city pairs and there may be multiple flow constraints for flights between a city pair.

As shown in Figure 4, one must explicitly consider the NAS as a complicated network in seeking to understand the delays. Convective storms can block multiple en route sectors. Rerouting aircraft through weather-free sectors can result in overloads and queues in these other sectors, including major problems for departures from the terminals inside the other sectors. This "NAS as a network" concept is nicely illustrated by a major FAA operational initiative for the summer of 2004: the "growth without gridlock" approach to coping with the loss of en route capacity due to convective weather. This initiative is described (McCartney, 2004) as follows: "If the waiting time for takeoff hits 90 minutes at a U.S. airport, the FAA slows down departures from other airports so that the clogged airports can launch more jets. In addition, express lanes are set up for the delayed flights….Storms in one part of the country might delay your flight even though it's sunny where you are, where you are going and even in-between."

There is no convenient quantitative closed-form expression analogous to equation (2) for the delays that arise when there are multiple queues as in Figure 4. What one can say is that the delays are surely a very nonlinear function of both the time duration of the event, the various city pair demands, and the various sector capacities. Since the system is typically operating in a very nonlinear mode, it is not easily possible to decompose the delay that occurs into terminal and en route contributions because:

Overall delay with both terminal and en route convective weather impacts ≠ Delay with only terminal impacts + delay with only en route impacts

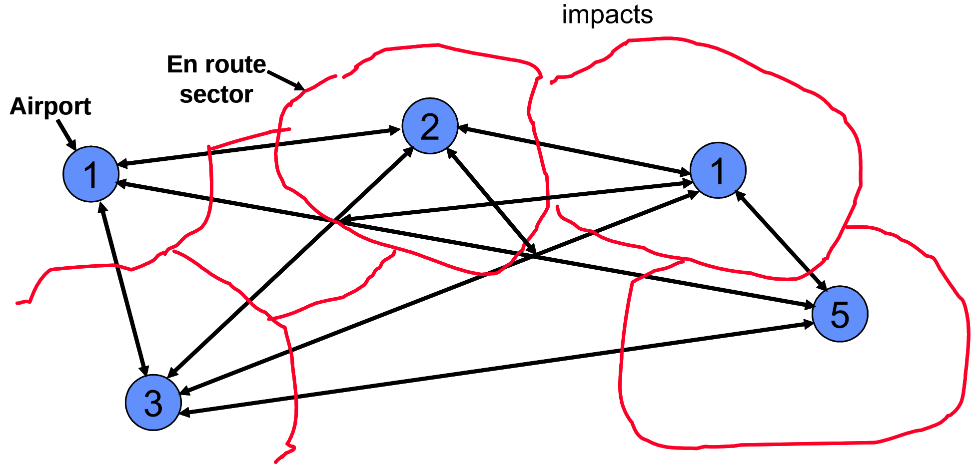


*Figure 4. The National Air System (NAS) as a network.*

The dynamics of the NAS network under time-varying, unpredictable perturbations due to convective weather are quite complicated and not well understood conceptually by many of the people who must manage it in real time. Hence, the decision-making process for adapting to convective weather impacts is itself an important factor in delay causality. Figure 5 illustrates the overall operational decision-making process for managing convective weather impacts. The Operational Decision Loop must be executed in a time period commensurate with a) the time scale over which the weather changes and b) the ability to accurately forecast the weather impact. If this cannot be achieved, then the plans that are executed will not be an appropriate solution for the weather situation.

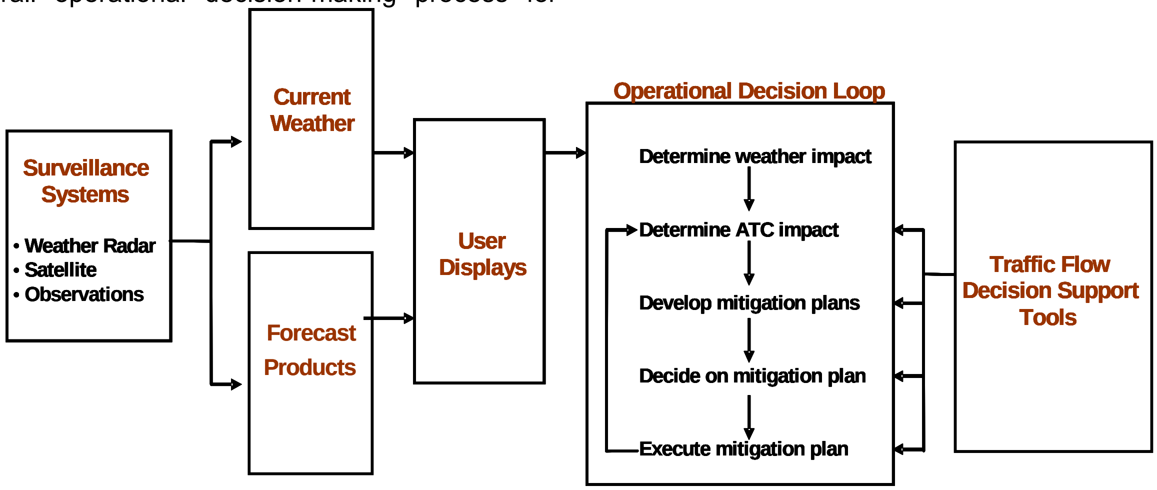


*Figure 5. Overall convective weather impact mitigation process.*

### E. Delay Ripple Effect

When an aircraft is delayed on one leg of a flight (e.g., due to adverse weather), it experiences delay on the next leg (and subsequent legs) flown that day. In cases where the subsequent leg(s) are not weather impacted, the delay may not be attributed to terminal weather even though the initial cause of the delay was weather impact on one leg of flight. The delay that occurs on subsequent legs is not a block-time delay [block-time delay = (actual arrival time – actual departure time) – (planned arrival time – planned departure time)]. Rather, typically it would be a gate departure delay.

A study by American Airlines and Oak Ridge National Laboratory (Beatty et al., 1999) looked at the impact on airline operating resources (specifically aircraft and crews) that results from a first-leg delay. They examined the actual impact on the American Airlines operations schedule as a function of both time and amount of delay. They found that as the delay on the initial leg increases, the number of flights affected increases as well. The downstream impact is a very nonlinear function of the initial delay and the amount of delay. The researchers developed a "delay multiplier" table that characterizes the degree of delay propagation as a function of the time of day the delay occurs and magnitude of the initial delay encountered. The "delay multiplier" ranges from as high as four (for delays > 2 hours early in the day) to a very small number for evening delays. For mid afternoon (e.g., 2 pm) the delay multiplier for a one hour delay is approximately 1.8 (i.e., the downstream impact is about 80% of the convective weather induced initial delay)[9]. Delayed flight operations also impact passengers, cargo, and gate space. However, impacts of delays on those other resources and the extent to which those impacts would further increase the delay could not be quantified. As a result, Beatty et al., suggest that their results are very conservative.

[9] It is interesting to note that a delay multiplier of 1.8 was also determined by Boswell and Evans (1997) from analysis of flights through LaGuardia International airport.

To summarize Section 2:

- Terminal related delays are caused by convective weather impacts on the airport and on the airspace well away from the airport (especially the boundary between terminal and en route airspace).

- Principal delay generating mechanisms are (1) aircraft flying longer paths than desired due to the need to avoid convective weather and (2) the queue delays that arise when demand is greater than effective capacity.

- Queue delays have a very nonlinear dependence on key factors such as the effective capacity of an ATC facility when impacted by convective weather, the demand, and the duration of the capacity impact. In contrast, the "longer path flown" delays tend to be linear in demand, weather event duration, and spatial extent of the adverse weather.

- Convective weather impacts on highly congested en route airspace lead to very complicated, poorly understood multiple queues in the "NAS network". Since terminal and en route impacts contribute to most of these total convective weather impacts, the highly nonlinear nature of the resulting delay means that one cannot easily decompose the total delay into additive terminal and en route contributions.

- The planning/mitigation plan execution process to cope with NAS network problems is a major factor in the delays that occur. Since effective execution of this process requires multi-hour convective forecasts of an accuracy that is currently unachievable, the convective weather-related delays that occur represent a very complicated combination of the actual weather characteristics, the convective weather forecasts, and the decision-making process.

- The "delay ripple effect" arises when an aircraft and/or flight crew are delayed on one leg of a flight (e.g., due to adverse weather), which results in delays on the next leg (and subsequent legs) flown by that aircraft and/or crew on that day. The ensuing downstream delay is often comparable to the initial delay.

## 3. HOW OFTEN DO CONVECTIVE WEATHER IMPACTS OCCUR?

To properly assess the impact of weather on aviation operations, one must consider the climatological frequency of the convective weather impact at specific locations on specific days and times. Climatological estimates of terminal impacts are classically inferred from the long term records of National Weather Service (NWS) station observations (METARS), which report when thunder was heard at the observing station (i.e., a "TS" observation). The accepted criteria for such reports (FMH, 1988) are:

"A thunderstorm is observed at a station when either

a: thunder is heard,

or

b: overhead lightning or hail is observed, and the local noise level is such as might prevent hearing thunder."

The accepted range for auditory detection of thunder is about 5-7 miles. Automated surface observation systems use automated lightning detection and reporting systems (ALDARs) that can provide cloud-to-ground lightning reports out to 10 nmi from the airport. ALDAR reports a thunderstorm at the airport only if the lightning is within a 5-mile radius of the airport (FAA, 1999). Hence, the effective detection range for TS observations from either source is approximately 5 nmi.

We have shown that substantial delays due to convective weather can occur for storms within the TRACON and near the ARTCC-TRACON interfaces. The "transitional en route" airspace, in which arriving planes descend and are vectored to the arrival fixes and departures climb to cruise altitude, is at least another 50 nmi beyond the TRACON boundary. Since the TRACON itself typically has radius of 40-50 nmi, the region of concern for assessing "terminal" convective weather impacts is inside a circle of radius 100 nmi around the airport. Typical thunderstorms are only a few miles in diameter. Clearly situations may arise where the airport and its immediate surrounding area experience fair weather while there is operationally significant convective weather within the TRACON and/or the en route transition area. In these cases, weather may impact the TRACON and transitional area operations even though there is no thunderstorm reported.

Bieringer et al., (1999) used radar data from Dallas (DFW) and Orlando (MCO) to develop a radar-based storm-day climatology for the two sites. A radar-based storm day is defined as a day during which convective weather was detected by radar within 10 nmi of the airport. Comparing the station observation data with the radar-based data, they found that radar data indicated 65% more storm days at DFW and 73% more storm days at MCO than were reported by station observations at the respective airports. The ratio of radar-observed convective weather impact days to station observation thunderstorm days was even greater for a 50-nmi radius circle (i.e., terminal region) centered at the airport. There were 2.48 times as many storm days within the DFW TRACON and 2.01 times as many storms days in the MCO TRACON than were observed at each of the airports. Bieringer et al., attribute the difference in the ratio between the two sites to differences in weather patterns. DFW has more line storms which may enter the edge of the TRACON, while MCO has more air mass type thunderstorms. Air mass storms typically are evenly distributed throughout the TRACON, increasing the likelihood that a thunderstorm will be observed at the airport when air mass storms occur.

These results are a statement about the relative number of days in which thunderstorms occur at some time during the day in various spatial regions (e.g., within 5 nmi or 10 nmi or 50 nmi of an airport) and are <u>not</u> a statement about the relative amount of time in which storm impacts occur on a given day. For example, a 10 nmi wide squall line passing through the TRACON at a speed of 25 knots results in a single thunderstorm observation for the day both at the airport and within the terminal area. However, the duration of the thunderstorm impact on the TRACON is 4.4 hours. The duration of impact on the airport region is only 0.8 hours. The duration of the terminal impact is 5.5 times greater than the airport impact duration.

For air mass thunderstorms, the time ratio of airport impacts to terminal impacts is roughly proportional to the ratio of the area of the storms causing the impact to the area being impacted [e.g., $\pi$ * (region radius + typical storm radius)$^2$]. Assuming an air mass storm radius of 4 nmi, the ratio of the time frequency of storm impacts somewhere in a TRACON to storm impacts at the airport would be about 30:1.

R. Ferris of Lincoln Laboratory conducted a study of the relationship between the times METAR reported a thunderstorm at an airport and times during which convective weather in or near the terminal area was clearly disrupting operations (e.g., causing diversions, significant flight deviations around storms, holding patterns, and/or reduced departure rates). Using ATL METAR data from six such days in 2003, Ferris found:

- The median length of time per day with METAR TS reports was 1.38 hours
- The median time per day during which convective weather delay caused traffic disruptions was 8.1 hours.
- The day-to-day variation in the ratio of thunderstorm delay impact time to METAR TS observation time was 2.7 to 28.2 with a mean value of 10.

For large regions of en route airspace (e.g., the Great Lakes corridor), there is no simple formula relating the thunderstorm day climatology of the various station observations within the region to the climatology of convective storm impacts for the region as a whole. One can, however, identify the frequency of convective weather over the region using radar data and/or lightning strike data. A study of the relative frequency of thunderstorms in the various en route centers in the northeast quadrant of the U.S. has been reported by Robinson et al., (2004). For the 6-month period from April to September, there are typically about 100 days with convective weather impacts in the en route centers between Chicago and Washington and about 60 days in the Boston en route center.

## 4. APPROACHES TO CONVECTIVE DELAY REDUCTION BENEFITS ASSESSMENT

There are two basic approaches to determining the achieved delay reduction benefits. “Direct” measurement compares the delays in a baseline time period when a weather decision support system was not in use to delays in a subsequent time period during which the system was in use. Alternatively, a “Decision/Modeling” approach employs interviews and/or direct observations of applied traffic management decisions to determine the parameters of models that are then used to estimate the delay reduction benefits. The basic assumption is that the weather product is useful only to the extent that it changes user decisions. Thus, one can analyze the various decisions that the users indicated were improved as a result of having access to the weather decision support system under study.

Both of these approaches have been attempted for various systems. The pros and cons of the two approaches are shown in Table 1. Our experience has been that the "direct" method is very hard to carry out in practice even though it appears quite straightforward. In the next section, we discuss two contemporary examples of the "decision/modeling" approach and how one might attempt to execute the "direct" approach.

**Table 1. Pros and Cons of Delay Reduction Determination Methodologies**

| | **"Direct" Method** | **"Decision/Modeling" Method** |
|---|---|---|
| Synopsis | Direct comparison of delay before and after a system is introduced | FAA operational user interviews and questionnaires + delay modeling |
| Good features | Actual delay reflects actual cost incurred<br><br>Easy to explain to recipients of a report | Factors which account for delay reduction are clearly understood<br><br>Extrapolation to changed circumstances (e.g., operations increases, schedule changes, weather time and duration) is relatively straightforward<br><br>Only feasible way to assess potential improvement in system products |
| Problems | Requires very sophisticated knowledge of delay causality to compensate for differences between the "baseline" and "system test" time periods. Factors that must be quantitatively considered are:<br><br>- Weather (severity, time of day, duration)<br>- Weather in other locations<br>- Traffic changes<br>- Airline operations and scheduling<br>- Air traffic procedures<br>- Traffic flow management changes<br><br>Not clear which elements of the system account for the delay reduction | May be difficult to validate the approach in some cases<br><br>Need to make sure that factors considered are independent or that common elements are identified and the impact addressed<br><br>(e.g., one must make sure one is not counting a factor several times by giving it different names) |

## 5. DECISION/MODELING BASED ON POST USAGE INTERVIEWS

### *A. Integrated Terminal Weather System (ITWS) Initial Demonstration System Studies*

The first major study of this type we are aware of was carried out in the context of an assessment of the delay reduction provided by the ITWS (Evans and Ducot, 1994). The first of these studies, which formed the basis for the subsequent ITWS benefits studies, was conducted by L. Stevenson of the Volpe Transportation Center and D. Rhoda of Lincoln Laboratory after the ITWS operational demonstrations in 1994. ITWS operational users were interviewed at the end of the demonstration period to determine:

- Operational decisions that had been improved by use of ITWS products above and beyond the baseline terminal weather information systems at the airports (TDWR and ASR-9),
- The number of aircraft (or time duration) over which the improvement was achieved on a "typical" day with thunderstorm impacts, and
- The benefit (e.g., minutes of reduced delay time) experienced by the individual aircraft.

Based on the interview results, the "longer path flown" model (Figure 3) was used to quantify the delay savings associated with the various ITWS products above and beyond the baseline weather information systems at the airports. This model can be written as

Delay savings = (decision factor coefficient) x (TDWR adjustment) x (WARP adjustment) x (airport type factor) x (operations per day) x (days with thunderstorms) x (convective storm type factor) (3)

Where:

Decision factor coefficient is derived from the typical savings for that decision x number of times that decision might be made on a thunderstorm day,

TDWR and WARP adjustments account for partial sharing of benefits with other programs,

Airport type factor recognizes the difference between double-hub (e.g., Dallas or MEM), single-hub (e.g., IAD), and non-hub airports

Convective storm type factor accounts for relative frequency of squall lines versus air mass storms.

These coefficients were derived from interviews of ITWS users at Memphis, Orlando, and Dallas. A typical "raw" user feedback was that a certain decision (e.g., identifying that an arrival fix would close in 20 minutes) might occur several times on a day and that about 20 minutes would be saved for some number of aircraft. In effect, the user identified the savings per aircraft, the number of aircraft involved, and the number of times the benefit might be achieved on a day with thunderstorms present. The estimates by the operational users at a terminal facility were generally fairly consistent. This is not surprising because many of the operational factors (e.g., the distance between arrival fixes, where planes are held in en route airspace, the time to fly from one arrival fix to another arrival fix, and the average number of aircraft that arrive over a fix per hour) are relatively constant.

The resulting benefits decision factors also seemed fairly consistent between different terminal areas. For example, the "raw" decision factor coefficient for anticipating that an arrival fix would close (the situation illustrated in Figure 3) was 0.00001 at Memphis and 0.000011 at Orlando. The largest "raw" decision factor coefficient was 0.000024, associated with recognizing that one runway would remain open (based on Memphis feedback).

These results were then extrapolated to the other ITWS locations based on the frequency of thunderstorm impacts, the number of operations at the various airports, and type of airport. It is interesting that in this initial ITWS study, the greatest delay reduction benefits, in terms of improved ATC decision making, actually arose from the traffic management units (TMUs) at the ARTCCs containing the TRACONs (Evans and Ducot, 1994). Key high benefits decisions for ITWS (in order of delay reduction obtained) were:

- Anticipation of the closing and reopening of arrival and departure fixes,
- Anticipation of convective weather impacts on runways and runway configurations,

- Optimization of traffic patterns within the TRACON,
- Optimization of airline operations, and
- Higher effective arrival capacity during thunderstorms.

It was subsequently realized that the number of times a given decision would be made per year depended on the climatology of the region in which the decision was being made. In particular, one could not use the climatology of the airport itself to estimate the climatology of the arrival fixes. This led to the Bieringer et al., (1999) study discussed above. That study indicated that one would need to augment equation (3) to include a decision-dependent climatology adjustment factor and also provided numerical values for the high benefits ITWS decisions identified in the Stevenson-Rhoda study.

### B. NY ITWS benefits –An Example of Terminals where Queues Dominate Benefits

An assessment of the operational benefits of the NY ITWS was carried out using the same approach as the first ITWS benefits study (Allan et al., 2001). The NY ITWS study relied heavily on the use of queuing models in determining benefits and on the use of case studies. Although it was recognized in the initial ITWS studies that queues could be a factor in delay causality, queues were not a frequent feature of ATC operations in convective weather at Memphis and Orlando[10]. In New York, situations where the demand exceeded the effective airport/terminal capacity were quite common during adverse weather. Therefore, demand and capacity as a function of time had to be analyzed very carefully to obtain realistic benefits estimates. For example, increasing departure rates during a Severe Weather Avoidance Plan (SWAP) was determined to be the highest convective weather benefit at New York. The benefits of this were estimated from the change in queue delays as a function of departure rate, using the average actual departure rates observed during a SWAP.

Identified benefits at New York that could be modeled using an extension to equation (3) included:

- Determining more efficient (shorter distance) departure routings,
- Landing aircraft before a weather event,
- Not rerouting arrivals or departures to an alternative fix if the storm would miss the desired fix, and
- Recognizing that the airport would remain partially open in a storm event.

The NY TRACON and surrounding en route airspace have much less capability for holding aircraft aloft than do Memphis or Orlando. As a result, the operational responses to convective weather and benefits of various decisions are quite different. For example, if New York ATC is concerned that arrival capacity may be lost in the near future they may impose a ground delay program (GDP) or ground stop (GS) for departures to New York because there is very little airspace capacity for holding planes. Such actions invariably lead to significant queue delays. Therefore, the ability to recognize that a fix and/or airport may remain open is very important.

The projected ITWS delay reduction benefits at NY based on the Memphis/Orlando/Dallas “linear” model were approximately $ 30 M per year. However, after considering the detailed dynamics of convective weather operations at NY, it became apparent that the convective weather delay benefits per year were approximately 3 times greater than had been projected based on the model of equation (3). This very major difference in the estimated benefits at New York emphasizes the need to carefully understand the detailed operations when carrying out a quantitative benefits analysis.

When attempting to compare delay reduction between different time periods, it is important to recognize that system benefits at the New York airports arise largely from changes in queues. In particular, equation (2) shows that the queue delay is a very nonlinear function of demand, capacity, and time duration. Hence, one must be prepared to carry out very detailed analysis of these factors if one is to assess from delay statistics whether delay reduction is or is not being achieved in practice.

### C. Atlanta ITWS Benefits Study

An analysis of the ITWS benefits at Atlanta was carried out by Allan (2004) following the approach used for New York. It was found that Atlanta was intermediate between New York and

[10] Although Dallas provided some benefits estimates for the initial study, there was far more operational usage of the ITWS at Memphis and Orlando when the Rhoda and Stevenson study was carried out.

Memphis/Orlando in the type of modeling that was appropriate. For example, arrivals could be held in en route airspace when there were weather impacts on the runways and/or arrival fixes. However, the immediate holding space would rapidly be filled such that aircraft also had to be held at the boundaries between the Atlanta ARTCC and surrounding ARTCCs. These holding patterns in en route airspace would themselves become a significant impediment to en route traffic flow. As a result, ground delay programs, ground stops, and miles-in-trail (MIT) restrictions on aircraft departing Atlanta were often required to manage the en route congestion problems that arose from terminal area convective weather events at Atlanta.

As a consequence of these terminal-en route dynamic interactions during convective weather, Atlanta required combinations of models discussed above and the inclusion of additional benefits categories that had not been considered previously. For example, maintaining some level of airport operations when a thunderstorm partially impacted the airport was a major benefit that could be assessed using the queue models. However, the benefit of recognizing that an arrival fix would remain open needed to be modeled by a combination of "longer distance flown" and "avoiding a queue at the alternative fix."

Figures 6 and 7 summarize the annual benefits identified at Atlanta. A number of these are quite different in relative magnitude from those discussed above, which had been based on the Memphis/Orlando models and the NY ITWS benefits results. We conclude from the experience at Atlanta and New York that the "mega terminals" such as Atlanta, Chicago, New York and Dallas probably each have to be considered as unique situations and should not have benefits estimated by extrapolation either from less busy airports or from other "mega terminals'.

## 6. DECISION/MODELING BASED ON DIRECT OBSERVATIONS OF PRODUCT USAGE

### *A. Corridor Integrated Weather System (CIWS) 2003 "Benefits Blitz" Observations*

The CIWS operational benefits studies (Robinson et al., 2004) have broken new ground in the methodology for assessing convective weather delay reduction benefits. The CIWS benefits of greatest interest were associated with en route decision making in the most highly congested airspace in the NAS. To assess CIWS benefits, one must come to grips with the NAS as a network (Figure 4) in its full complexity.

Figure 8 summarizes the approach that was unsuccessfully used in 2002 to estimate CIWS benefits. The methodology was clearly derived from the relatively successful ITWS benefits studies discussed previously. This approach did provide useful information on the frequency of various benefits and identified key factors in achieving delay reduction, as well as providing quantitative estimates of the delay reduction for terminal operations. However, in response to questionnaires during post-season interviews, the en route users of CIWS could not provide typical estimates for the key parameters (e.g., number of aircraft impacted by a given ATC decision, the average delay savings for each of the aircraft, flow rates through sectors with and without CIWS, etc) either for a given day or as a seasonal average. Rather, they suggested we obtain the benefits by having observers at the facilities during periods of operationally significant convective weather.

The 2003 data collection design (Figure 9) used Lincoln Laboratory observers at a number of ATC facilities during "benefits blitz" time periods when significant convective weather was expected. These intensive observation periods were treated as a sampling of the population of significant convective weather events at a given facility.[11] Based on both the observations of users utilizing CIWS displays and user statements of ATC decisions made using the CIWS products, very detailed statistics were generated on the

[11] It was assumed that the storm situations "sampled" during the "benefits blitz" events were typical of the population of convective storm events that occurred in the CIWS domain over the summer. During more than one very significant storm event, traffic managers at ZDC informed the CIWS observer that operations for the event were not typical. They suggested that the observer might gain more representative results if the observer visited during smaller-scale events. During some of the larger storm events at ZDC, impacts were so significant that either (a) the storms completely shut down impacted sections of their airspace, leaving little "wiggle room" for dynamic adjustments using the CIWS products or (b) strong storms in neighboring ARTCCs shut off the flows into ZDC. This either reduced traffic so severely that meaningful traffic decisions via CIWS were unnecessary (e.g., demand was so low that alternative reroute and/or miles-in-trail decisions carried little meaning and/or offered little in terms of delay reduction) or rendered potential CIWS-derived opportunities moot since no traffic could reach the "benefit zone."

number of times per significant convective-weather day a given beneficial ATC decision was made using CIWS products. Given this information, straightforward computations (similar to those discussed above) could be used to estimate an annual frequency of those decisions, if one had statistics for the frequency of significant convective weather in a given facility. Once one derived an estimate of the average benefit per beneficial decision per ATC facility, one could then multiply it by the annual frequency of that decision at the ATC facility to arrive at an average annual benefit for each ATC facility. Summing the individual facility benefits over all facilities would result in the annual benefit for a given decision.

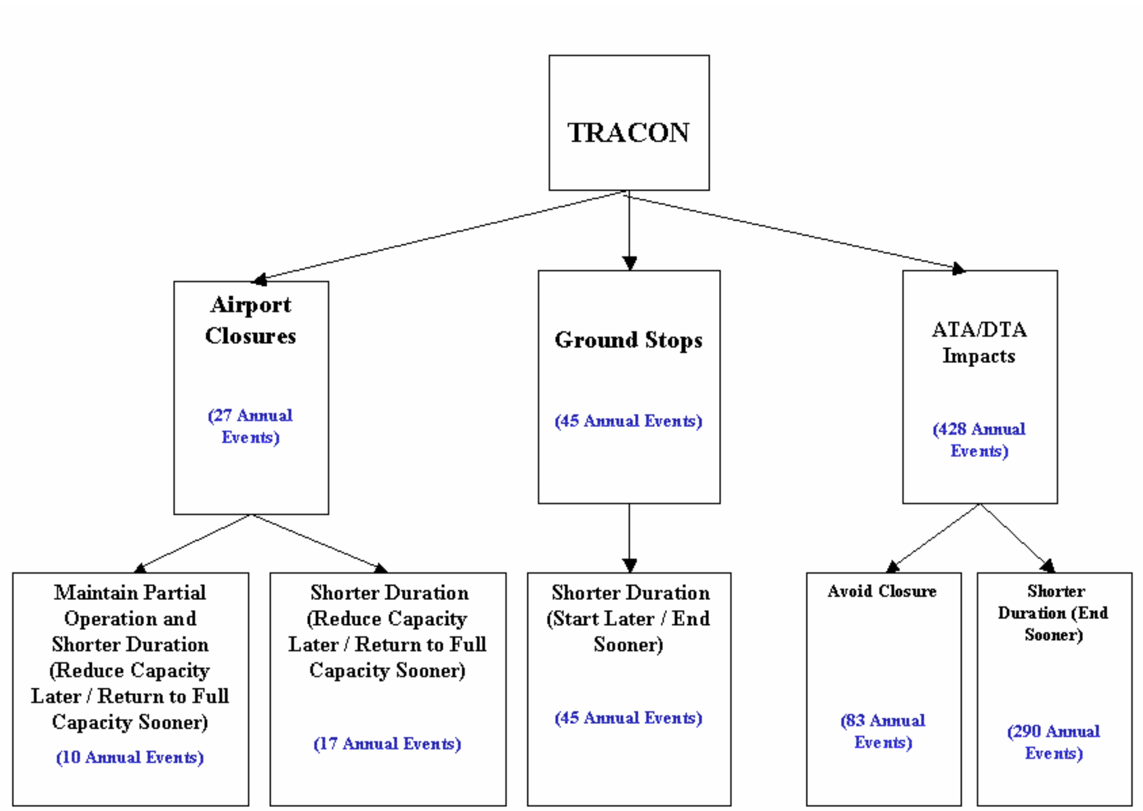


*Figure 6. Identified benefits associated with decision making using ITWS by Atlanta TRACON. ATA is arrival transition area, DTA is departure transition area*

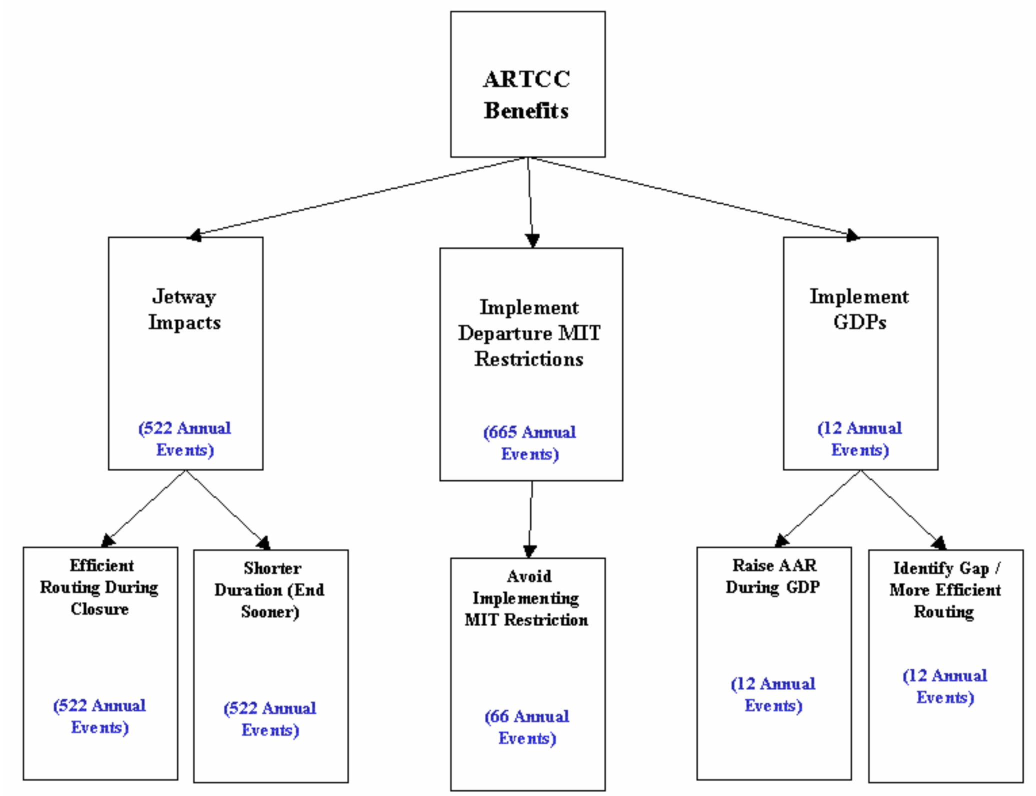


*Figure 7  Atlanta ITWS delay reduction benefits associated with Atlanta en route center decision making. MIT is mile-in-trail spacing; GDP is ground delay program; AAR is airport arrival rate.*

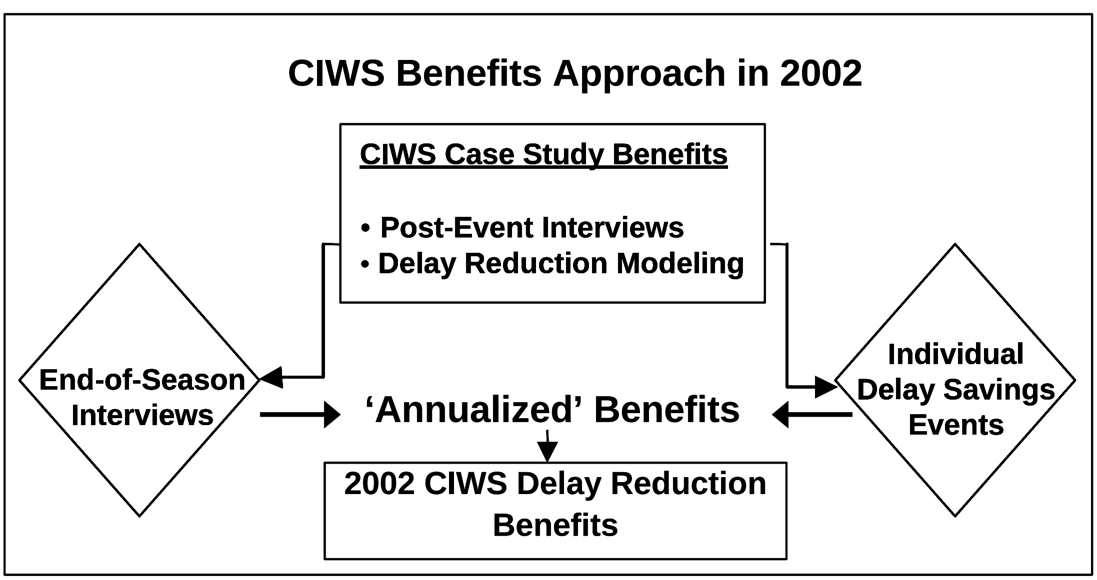


*Figure 8.  Approach used to assess CIWS delay reduction benefits in 2002.*

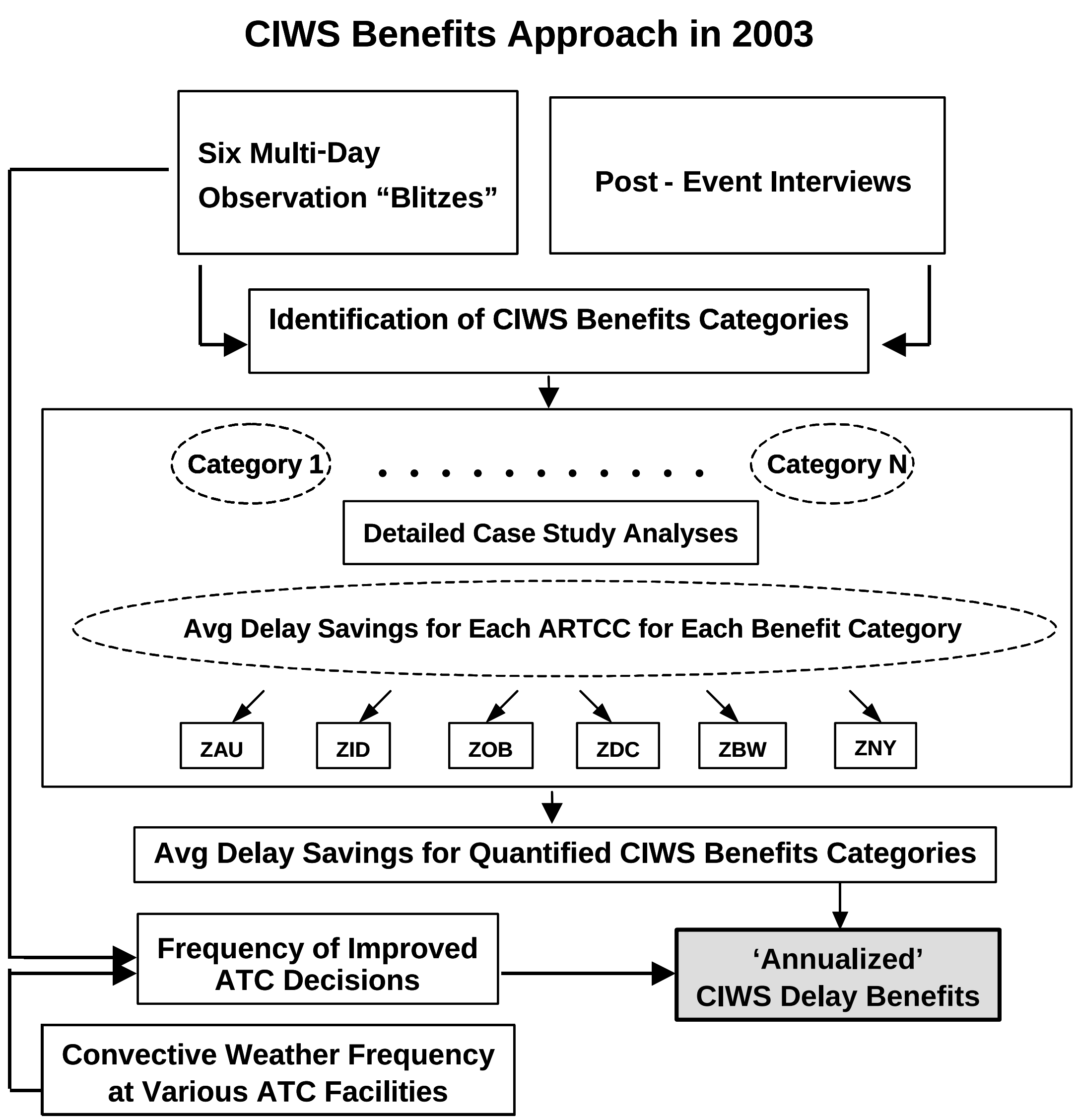


*Figure 9. Approach taken in 2003 to estimate the CIWS annual delay reduction benefits*

Some aspects of the above methodology are relatively straightforward. Table 2 shows the projected frequency of various CIWS benefits decisions by facility based on:

- The frequency with which the CIWS products were used beneficially to make a given decision on the days that an observer was present, and
- The frequency of convective weather events in that facility.

In Table 2, the column labeled "ARTCC Total" does not include ATCSCC contributions. This was done to prevent inflation of benefits resulting from assigning events to more than one facility. In practice, observed usage benefits (from which these roll-ups are based) were only assigned to the ARTCCs using CIWS to initiate traffic decisions, even if coordination with other facilities was needed or if the benefit event occurred along facility boundaries. Exceptions (total benefits occurrences in bold), where ATCSCC benefits occurrences were added to the final totals, include the categories "Interfacility coordination", "Reduced workload", and "Situational awareness". These specific benefits could not be easily separated by facility and may in fact have proved of more importance at ATCSCC compared to elsewhere in terms of enacting efficient delay mitigation schemes. Otherwise, the differences between the various facilities in Table 2 reflect differences in product usage and the frequency of convective weather.

**Table 2. Annual Frequency of Various CIWS Benefits**

| | Benefit Category | ATC Facility | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ZAU | ZOB | ZID | ZDC | ZNY | ZBW | C90 | ATCSCC | ARTCC Total* |
| 1 | Keeping routes open longer and/or reopening closed routes earlier | 75 | 173 | 109 | 357 | 104 | 103 | 14 | 53 | 953 |
| 2 | Closing routes proactively | 0 | 10 | 15 | 89 | 17 | 26 | 0 | 9 | 157 |
| 3 | Proactive, efficient reroutes | 30 | 102 | 36 | 178 | 70 | 9 | 14 | 27 | 439 |
| 4 | Shorter/fewer ground stops | 15 | 10 | 36 | 104 | 0 | 17 | 14 | 0 | 196 |
| 5 | Ground Stop avoided | 0 | 10 | 0 | 0 | 17 | 0 | 42 | 0 | 69 |
| 6 | Reduced MIT restriction | 0 | 20 | 0 | 15 | 17 | 0 | 0 | 0 | 52 |
| 7 | Traffic directed through gaps in weather | 0 | 61 | 22 | 74 | 0 | 9 | 28 | 0 | 194 |
| 8 | Better management of weather impacts on terminal ATAs | 134 | 183 | 109 | 59 | 35 | 26 | 210 | 18 | 756 |
| 9 | Optimization of runway usage; enhanced runway planning | 0 | 10 | 7 | 0 | 0 | 9 | 84 | 0 | 110 |
| 10 | Improved use of GDPs | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 0 | 14 |
| 11 | Greater departures during SWAP | 60 | 20 | 15 | 45 | 35 | 17 | 84 | 0 | 276 |
| 12 | Directing pathfinders | 30 | 20 | 29 | 149 | 52 | 17 | 14 | 27 | 311 |
| 13 | Interfacility coordination | 75 | 214 | 211 | 371 | 35 | 120 | 112 | 151 | **1289** |
| 14 | Improved safety | 0 | 20 | 29 | 59 | 35 | 26 | 14 | 0 | 183 |
| 15 | Reduced workload | 30 | 102 | 95 | 282 | 52 | 77 | 84 | 27 | **749** |
| 16 | Situational awareness | 104 | 458 | 597 | 1054 | 139 | 583 | 210 | 169 | **3314** |
| | | | | | | | | | | |
| | # Convective Weather Days (based on 2002 and 2003) | 89.5 | 91.5 | 102 | 104 | 69.5 | 60 | 42 | 80 | |

* Total occurrences of various CIWS benefits categories do <u>not</u> include ATCSCC contributions, except where the total is shown in bold font.

The challenge was in determining the average benefit associated with a given decision in each facility. This was accomplished using a random sampling approach. Based on facility operations for 22 days, plus feedback from operational users on product usage on other days, a large number of instances of each type of benefit decision were identified for each ATC facility. A subset of these was selected for detailed analysis using a random number generator. Each of these particular instances was then assessed in detail to determine the benefits for a given situation. In many cases, the benefit consisted of increased capacity (e.g., a route was kept open or reopened earlier). In such cases, the benefit is essentially the delay that would have occurred had that additional capacity not been available. Estimating this additional delay involves determining an appropriate approach for handling the aircraft that used the route that was kept open. If the route in question was closed, alternative traffic management decisions include holding aircraft at departure airports and/or using an alternative route from the origin to the destination. In short, one is dealing with a subset of the overall NAS network management problem illustrated in Figure 4[12].

[12] Appendices B and C in Robinson et al., (2004) discuss the results of this analysis for the various individual cases. These case analyses definitively highlight why management of the NAS network in real time with uncertain information on the future en route and terminal weather-related capacity reductions is a very difficult task.

It was found that a large fraction of the individual cases analyzed required use of a tailored queue model and that the benefits varied greatly from case to case as shown in Table 3. This enormous dynamic range in the benefits arises because of vastly varying applications of the queue model in situations where the event durations differ greatly and/or the denominator terms in equation (2) are very small.

**Table 3. CIWS Benefit: Route Open Longer and/or Reopen Closed Route Earlier**
***Individual Case Study Delay Savings Results***

| | | DELAY SAVED (hr) | | | SAVINGS | | | |
|---|---|---|---|---|---|---|---|---|
| **ARTCC** | **Date** | **Primary** | **Downstream** | **Total** | **Operations (DM-1)✦ (DM-2)** | **Passenger (p)** | **Passenger (d)** | **TOTAL (DM-1) (DM-2)** |
| ZAU | 30 Apr | 13.5 | 10.8 | 24.3 | $50,540<br>$33,465 | $29,349 | $23,479 | $103,368<br>$ 86,293 |
| ZAU | 26 Jun | 106.0 | 84.8 | 190.8 | $325,581<br>$191,512 | $230,444 | $184,355 | $740,380<br>$606,311 |
| ZID | 10 Jun | 1.4 | 1.1 | 2.5 | $5,428<br>$3,689 | $3,044 | $2,391 | $10,863<br>$ 9,124 |
| ZID | 10 Jul | 70.4 | 56.3 | 126.7 | $200,312<br>$111,302 | $153,050 | $122,396 | $475,758<br>$386,748 |
| ZID | 23 Jul | 4.8 | 3.8 | 8.6 | $18,656<br>$12,648 | $10,435 | $8,261 | $37,352<br>$31,344 |
| ZOB | 08 May | 5.0 | 4.0 | 9.0 | $14,229<br>$7,905 | $10,870 | $8,696 | $33,795<br>$27,471 |
| ZOB | 06 Jul | 22.0 | 17.6 | 39.6 | $62,924<br>$35,098 | $47,828 | $38,262 | $149,014<br>$121,188 |
| ZOB | 03 Aug | 236.3 | 189.0 | 425.3 | $672,399<br>$373,590 | $513,716 | $410,886 | $1,597,001<br>$1,298,192 |
| ZDC | 22 Jul | 10.3 | 8.2 | 18.5 | $29,248<br>$16,284 | $22,392 | $17,827 | $69,467<br>$56,503 |
| ZDC | 23 Jul | 3.6 | 2.9 | 6.5 | $14,071<br>$ 9,486 | $7,826 | $6,305 | $28,202<br>$23,617 |
| ZDC | 03 Sep | 36.2 | 29.0 | 65.2 | $103,081<br>$57,232 | $78,699 | $63,046 | $244,826<br>$198,977 |
| ZBW | 11 Jun | 4.0 | 3.2 | 7.2 | $11,383<br>$6,324 | $8,696 | $6,957 | $27,036<br>$21,977 |
| ZBW | 05 Aug | 8.9 | 7.1 | 16.0 | $27,404<br>$16,179 | $19,349 | $15,435 | $62,188<br>$50,963 |
| ZNY | 12 Jun | 1.5 | 1.2 | 2.7 | $5,850<br>$3,953 | $3,261 | $2,609 | $11,720<br>$9,823 |
| ZNY | 05 Aug | 49.0 | 39.2 | 88.2 | $139,444<br>$77,469 | $106,526 | $85,221 | $331,191<br>$269,216 |

✦ DM-1 and DM-2 are different down stream delay direct operating cost models described in Robinson et al., (2004) The results of the CIWS operational benefits study are summarized in Figures 10 and 11. In addition to the "bottom line" benefits assessment shown in Figure 10, we were also able to obtain quantitative data as to which products were of greatest use and which ATC facilities might warrant investigation and/or additional training if their usage of the products seemed much lower than the usage of other facilities.

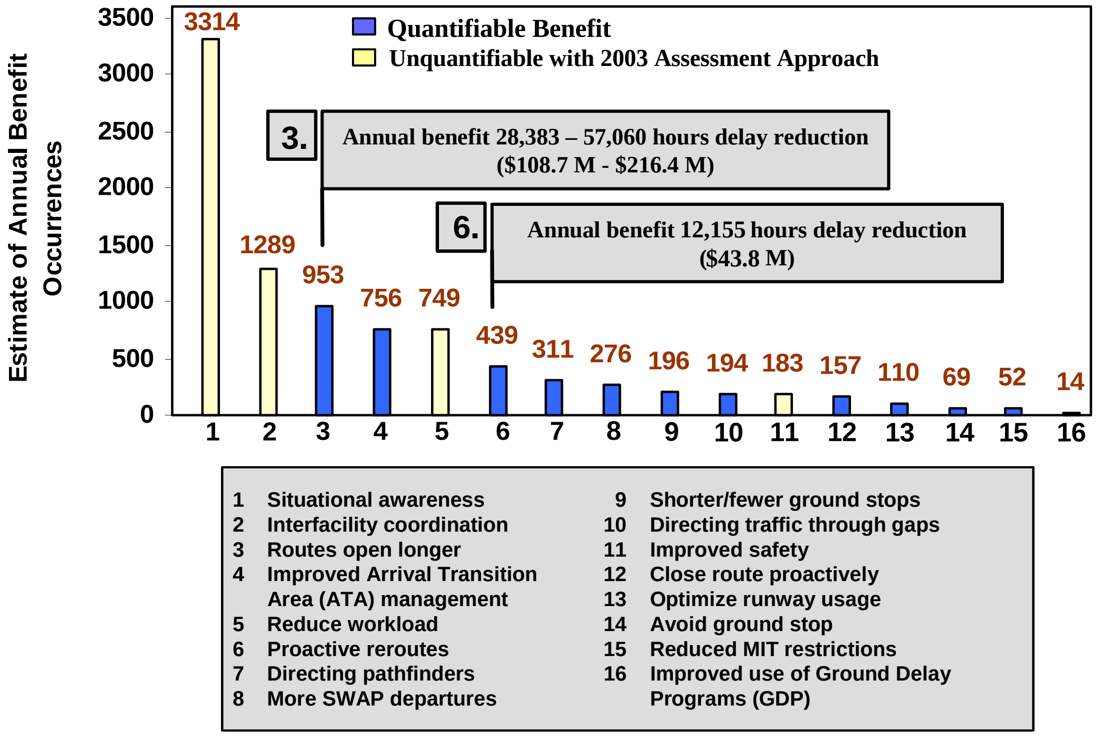


*Figure 10 Estimated annual occurrences of identified CIWS benefits categories. Yellow bars denote unquantifiable benefits. Blue bars denote quantifiable benefits. Annual delay savings estimates associated with the two main categories examined in the initial benefits analyses ("Route kept open longer/reopened closed route earlier" and "Proactive, effective rerouting") are shown.*

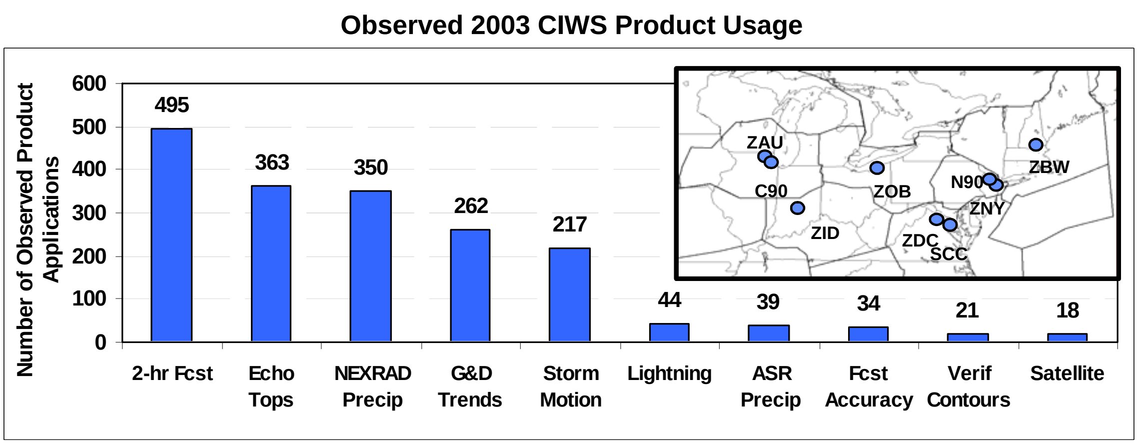


*Figure 11 The number of observed applications of individual CIWS weather products at FAA facilities visited (inset) during 2003 convective weather events.*

## 7. COMBINING USER FEEDBACK WITH COMPARISONS OF TRAFFIC MANAGEMENT DURING CONVECTIVE WEATHER EVENTS

### *A. Near-term Validation of the CIWS Benefits Study*

One of the concerns raised about the CIWS benefits approach discussed above was the reliance on the ATC user judgment regarding improved traffic management decisions during blitz observation periods compared with decisions in the past during similar situations. Hence, a study is being conducted to determine if improved ATC decision making using CIWS can be observed by comparing the management of traffic during similar types of convective weather before and after CIWS was introduced. For example, one of the frequently cited high-benefit situations occurred when

- The storm reflectivity product on ETMS and WARP indicated that routes were blocked by severe weather, but
- The combination of the CIWS precipitation and echo tops products showed that aircraft could easily fly over or around much of the weather.

Work is underway to obtain NEXRAD data and flight track archives for the CIWS test domain for the time period before CIWS was in use. These data would allow comparisons of en route airspace usage (as measured from aircraft flight tracks) before and after CIWS for cases that exemplify the above weather situation.

### *B. WARP Benefits Study*

Such an approach was used in an unpublished study of the Weather and Radar Processor (WARP) delay reduction benefits by MCR, International. Interviews were conducted with air traffic controllers at the Houston and Indianapolis ARTCCs to determine how the availability of NEXRAD mosaics on the controller screens aided in the handling of aircraft. The controllers indicated that they used the improved weather depiction to do a better job of directing aircraft to appropriate gaps in storm systems and proactively rerouted aircraft in certain situations.

MIT Lincoln Laboratory personnel, using archives of ETMS data and a 1 nmi NEXRAD base reflectivity mosaic,[13] compared the handling of aircraft in the Houston ARTCC airspace before and after the WARP composite reflectivity mosaic was displayed on the controller screens. Unfortunately, there was only a single pre-WARP storm case, but there were 12 time intervals from 9 different days of recorded data after WARP products were provided to the controllers. In the post-WARP data sets, it was clear that controllers were directing aircraft through holes in storm systems and rerouting them to an appropriate path well in advance of storm encounters (i.e., as suggested by the green line in Figure 3). However, the single pre-WARP data case did not possess storm geometries required to allow assessment of how effectively controllers could direct planes around weather without WARP. Hence, the comparison of traffic handling before and after WARP was deployed was inconclusive, albeit the basic approach seems sound.

## 8. COMPARISON OF DELAY STATISTICS BEFORE AND AFTER A SYSTEM IS DEPLOYED

### *A. General:*

In the introduction to this paper, we noted that major FAA performance metrics are couched in terms of reduction in delays. Hence, there is a very strong emphasis within the FAA on demonstrating that convective weather delay reduction is being achieved by analysis of actual delay data.

Simply comparing ATC delays before and after a convective weather delay reducing system has been introduced is equivalent to asserting that *correlation* (if it existed) between the introduction of a system and a change in the overall delays, *is evidence of causality* that could be associated with the system under test.

Delays are clearly impacted by many factors including

- Weather differences (both convective and non convective) in the test region,
- Demand (e.g., as exemplified by the FAA/DOT high level discussions with

---

[13] The data provided was by a commercial firm, Flight Explorer. The WARP reflectivity product on the controller displays is a NEXRAD layered composite reflectivity.

airlines regarding scheduling at Chicago O'Hare airport in 2003-04),

- Fleet mix (which impacts use of runways at airports and en route sectors[14]),
- Policies on management of congestion caused by convective weather (e.g., Spring 2000, "growth without gridlock"),
- Airline scheduling and operations procedure changes (e.g., changing the scheduled block time for a flight between two cities and/or deciding when a flight would be cancelled),
- Introduction of other systems (e.g., CCFP, traffic flow management), and
- Weather outside the "test region" (e.g., en route convection, low C/V and/or winds and/or convective weather at airports).

Thus, one of the major challenges in comparing delay statistics is accounting for or minimizing the impact of the factors noted above.

Another important issue in delay analysis is choice of delay statistics. Historically, the principal source of information was the FAA's Air Traffic Operations Network (OPSNET) database. A reportable delay recorded in OPSNET is defined in FAA Order 7210.55B as, "Delays to Instrument Flight Rules (IFR) traffic of 15 minutes or more, experienced by individual flights, which result from the ATC system detaining an aircraft at the gate, short of the runway, on the runway, on a taxiway, and/or in a holding configuration anywhere en route shall be reported." Such statistics include delays due to weather conditions at airports and en route, FAA and non-FAA equipment malfunctions, the volume of traffic at an airport, reduction to runway capacity, and other factors. Flight delays of less than 15 minutes are not reported in OPSNET but may be recorded by Air Traffic facilities.

Non-reportable delays are delays caused by pilot initiated en route deviations around adverse weather (as opposed to reportable delay for weather conditions at an airport); delay caused by mechanical or other aircraft operator/company problems; and delay for taxi time controlled by non-FAA entities (e.g., company/airport ramp towers.)

[14] One of the very significant changes in the past 8 years has been the replacement of turbo prop aircraft by regional jets. The regional jets fly at much the same altitudes as larger jet aircraft and require similar runway lengths.

International delays are caused by initiatives imposed by facilities outside the United States. International delays are recorded in the OPSNET database and are not separated.

Delays are broadly categorized as terminal or en route delays. Terminal delays are incurred as a result of conditions at the departure or arrival airport and are charged to the appropriate airport. En route delays occur when aircraft incur airborne delays of 15 minutes or more as a result of an initiative imposed by a facility to manage traffic. En route delays are recorded by the facility where the delay occurred and charged to the facility that imposed the restriction.

OPSNET data has some good features:

- The data base extends backward for many years, and
- It contains causality associated with the delays (in particular, which delays are attributed to weather), and
- There is information on the category of the delay (e.g., arrival and en route, departure, traffic management system).

However, there are also major deficiencies with OPSNET (Lamon, 2004):

- Delays are only reported if they were at least 15 minutes in a facility. Hence, a flight that was delayed 10 minutes in each of a number of facilities would not be reported. Gate delays are not recorded.
- Reporting methods are subjective and differ widely by facility (e.g., major airports such as EWR and ORD report a much higher fraction of delays that occur than do many less busy airports).
- Delays can be inaccurate due to human effort.
- Delays due to pilot/airline initiated routing around convective weather are excluded. (This would include reroutes developed under collaborative decision making.)

Given these problems, the OPSNET database is not currently regarded as useful for highly detailed quantitative benefits assessments.

Rather, the recent trend has been to use the Aviation System Performance Metrics (ASPM) database. The ASPM data base combines data from the FAA en route system (Host) on aircraft positions, flight plans, OAG schedules, air carrier ASQP and (for some of the major carriers[15]) there is Out/Off/On/In (OOOI) data consisting of:

- Actual gate departure time ("Out")
- Actual flight takeoff time ("Off")
- Actual flight landing time ("On"), and
- The actual gate arrival time ("In)

Hence, it is possible to make much more detailed objective, quantitative studies of where a flight delay occurred than is the case with OPSNET. On the other hand, there is limited causality associated with delays. Additionally, for flights that do not have OOOI data, the estimated takeoff and landing times are based on the Host computer estimates. The accuracy of the Host computer estimates of "off" and "on" times varies widely (due to differences in coverage of the airports by the en route surveillance radars).

The ASPM web site (www.apo.data.faa.gov) provides summary statistics for user selectable filter parameters that include arrival and destination airport, time of day, phase of flight, and type of delay. For example, one can get the fraction of arrival or departure flights delayed 15 minutes or more and various statistics regarding the delay (e.g., mean, median, 90 percentile) associated with delayed flights. In addition, one can access the individual flight records and compute delays with a different set of criteria than are used in the summary statistics.

### *B. Use of ASPM Delay Statistics to Confirm ITWS Benefits at Atlanta*

Some of the difficulties and issues associated with use of ASPM statistics to assess whether an operational system is providing delay reduction benefits have come to light in the course of an ongoing study to confirm the Atlanta ITWS benefits discussed above through the analysis of ASPM statistics.

ITWS was in full operational use in 2003. The 2001 period prior to September 11$^{th}$ was selected as the pre ITWS time period for studies conducted by Lincoln Laboratory and independently by the FAA. The work reported here is based on the Lincoln study alone. This was a time/resource limited study such that one could not hope to develop a comprehensive weather severity index to adjust the measured delays for all the factors discussed above.

One of the major challenges that one faces in assessing delays at Atlanta is that arrival delays can be impacted by both en route weather and weather at the origin airport. Similarly, departure delays can be impacted by both en route weather and weather at the destination airport. Atlanta has approximately 2620 operations a day. Of those, 7 % of the traffic from Atlanta is destined for Florida, which has an abundance of weather impacts in en route airspace and at the terminals, 24% are to the northeast (New York, Washington, Boston, Pittsburgh, and Philadelphia), and 13% are in the north-central region (e.g., Chicago, Detroit, and Minneapolis). The northeast and north-central regions are in the CIWS domain. Given that a) many of the airports in those regions are notorious for delays and b) the CIWS domain frequently has major en route congestion problems in convective weather[16] (Robinson, et. al., 2004), the normalization problem would be very difficult unless one focused on initially analyzing delays that did not involve assessing and compensating for en route and/or terminal weather in the CIWS domain.

To reduce the complexity of the analysis, we focused on traffic handling in convective weather near the ITWS domain. It was assumed that the majority of the reduction in airborne arrival delays attributable to the use of ITWS would occur in close proximity to the airport (within a 50-60 nmi radius). In particular, it was hoped that the change in average flight times from 100 nmi to touchdown on days of terminal convective activity between 2001 and 2003 could be equated to delay reduction. Possible differences in the severity of

[15] Two major carriers report OOOI data in near real time via the ARINC data link; OOOI times for four other carriers are reported once per month as a part of the reports to the Department of Transportation Bureau of Traffic Statistics (BTS).

[16] Robinson, et. al. (2004) show that the Washington and Indianapolis enroute centers that handle the Atlanta traffic into the CIWS domain experienced enroute convective weather over 100 days between April and September in 2003.

the convective weather events would be reduced by using many days from the time periods before and after ITWS was introduced. To maintain compatibility with FAA internal studies (Citrenbaum, personal communication) average flight times as deduced from the ASPM summary statistics would be compared between 22 METAR-identified thunderstorm days from 2003 and the average flight times for 22 METAR thunderstorm days from 2001 (Table 4).

An important factor to consider in design of the data analysis and interpretation is the magnitude of the expected effect, i.e., the change in flight times that one would expect based on the user feedback/modeling analysis discussed above. The analysis discussed above in section 5 C estimated the annual ITWS convective weather delay reduction benefits at Atlanta to be:

- Taxi delay savings: 1613 hours per year
- Airborne arrival delay savings associated with airport storm impacts: 578 hours
- Airborne arrival delay savings not associated with airport impacts: 843 hours.

Our studies of the frequency and duration of convective weather impacts (section 3 of this paper), show:

- Average number of days per year with a thunderstorm observed at the Atlanta airport: 50
- Number of hours per year with a thunderstorm observation at the airport: 81
- Number of days per year with convective weather impacts within 50 nmi of airport: 100

Combining this data with the Atlanta operations (1310 arrivals per day, approximately 70 arrivals per hour between 6:00 am and 11:00 pm local time), we find:

- Expected delay savings per aircraft due to better handling of airport storm impacts is 0.6 minutes for all aircraft on a TS day,
- Expected delay savings per aircraft due to better handling of airport storm impacts is 5.9 minutes for each aircraft scheduled to land during the time period where a thunderstorm was observed at the airport[17], and
- Expected airborne arrival delay savings for non-airport related delay reductions within the ITWS coverage is 0.4 minutes per aircraft on days with convective weather impacts in or near the terminal area.

Thus, we see that the magnitude of the flight time change that one is looking for is on the order of a minute per aircraft on days with a thunderstorm observation at Atlanta airport.

The ASPM summary statistic flight times from 100 nmi to touchdown were determined for each of thunderstorm days shown in Table 4 and for non-thunderstorm days shown in Table 5. It is important to stress that the non-thunderstorm days were identified using a combination of Atlanta METARS and weather radar data from the RTVS validation site[18] (Mahoney, et. al. 2002). There are many days where the METAR does not report a thunderstorm but where convective activity is impacting terminal operations. Hence it is very important to use radar data in determining that there truly are no convective weather impacts on airport operations.

In Table 6, we show the results of comparing the mean and median ASPM summary statistics flight times for the various days in the two years. The difference between average flight times on thunderstorm days and average flight times on non-thunderstorm days was about 5 minutes, which for 2003 could be viewed as corresponding to the average delay per aircraft due to thunderstorms with ITWS in operation. Since the expected delay reduction on such days was about one minute, this suggests that the ITWS delay reduction benefit for airborne arrival delay corresponds to about 16% of the before-ITWS airborne arrival delay.

[17] The time interval during which arriving planes at Atlanta would be impacted by a convective storm impact on runways will in general be much greater than the time interval that the storm is impacting the airport. This is because the queue (i.e., aircraft in holding patterns) that builds up during an airport storm impact lasts well after the storm impact ends.

[18] http://www-ad.fsl.noaa.gov/fvb/rtvs/conv/

**Table 4: ATL Thunderstorm Days (June-August) for ITWS ASPM based delay analysis**

| 2001 | | 2003 | |
|---|---|---|---|
| 6/1/2001 | 7/4/2001 | 6/3/2003 | 7/11/2003 |
| 6/3/2001 | 7/5/2001 | 6/4/2003 | 7/14/2003 |
| 6/4/2001 | 7/9/2001 | 6/11/2003 | 7/15/2003 |
| 6/6/2001 | 7/24/2001 | 6/13/2003 | 7/22/2003 |
| 6/8/2001 | 7/25/2001 | 6/14/2003 | 7/30/2003 |
| 6/14/2001 | 7/28/2001 | 6/16/2003 | 7/31/2003 |
| 6/21/2001 | 7/29/2001 | 6/17/2003 | 8/3/2003 |
| 6/22/2001 | 8/4/2001 | 6/18/2003 | 8/16/2003 |
| 6/25/2001 | 8/13/2001 | 6/19/2003 | 8/19/2003 |
| 6/30/2001 | 8/18/2001 | 7/3/2003 | 8/28/2003 |
| 7/3/2001 | 8/31/2001 | 7/10/2003 | 8/30/2003 |

**Table 5: ATL non-Thunderstorm Days (June-August) for ITWS ASPM based delay analysis. These were determined based on examining both METAR reports and weather radar data.**

| 2001 | | 2003 | |
|---|---|---|---|
| 7/14/2001 | 7/222001 | 6/5/2003 | 6/23/2003 |
| 7/15/2001 | 8/15/2001 | 6/9/2003 | 6/24/2003 |
| 7/16/2001 | 8/21/2001 | 6/21/2003 | 6/25/2003 |
| 7/17/2001 | 8/22/2001 | 6/22/2003 | |
| 8/23/2001 | | | |

**Table 6: Results of ATL ASPM 100 nmi-to-touchdown flight time analysis (minutes).**

| | 2001 | 2003 | Delta |
|---|---|---|---|
| Avg Flt time, Non-Tstorm | 28.2 | 27.6 | -0.6 |
| Average Median Flt time, non-tstorm | 26.7 | 26.1 | -0.6 |
| Avg Flt time, Tstorm | 31.8 | 30.8 | -1.0 |
| Std Dev of daily mean, Non-Tstorm | 1.0 | 1.2 | - 0.2 |
| Std dev of daily median, Non-Tstorm | 0.6 | 0.6 | 0.0 |

The comparison between thunderstorm day average flight times shows a decrease of approximately one minute from 2001 to 2003, which is consistent with the predicted arrival delay reduction of one minute. However, we do not regard this comparison as adequate proof because there are a number of disquieting elements of the ASPM summary statistics:

- Large variability in the daily mean flight times for thunderstorm days. The standard deviation of the daily mean flight times is about two minutes. If we treat the expected delay reduction of one minute as the signal, the "signal-to-noise" in this case for daily means is about $10 \log_{10} (1/4) = -6$ dB. Hence, we have significant concerns about the statistical significance given that there clearly was major variability between different days.

- Large variations in flight times for a number of aircraft even on fair weather days. For example, the median flight time from 40 nmi to the ATL airport was typically on the order of 18 minutes. However, over 10% of the aircraft had flight times that were at least 5 minutes greater than the median and some flights took over 50 minutes to fly from 40 nmi to the airport, and

- Very large 40 nmi to airport flight times (in excess of 70 minutes) were observed for a number of aircraft on thunderstorm days.

Detailed investigation showed two major problems with using the ASPM summary statistics flight times from 100 nmi to touchdown for benefits analyses:

1. *Algorithm used to determine 100 nmi crossing time*

   It has recently been learned that the ASPM-determined time at which a plane comes within 100 miles of the airport is not measured from the time that the plane is within 100 nmi true range from the airport (that is, the plane goes inside a circle of radius 100 nmi centered on the airport).

   Rather, the ASPM 100 nmi to touchdown flight time statistic is based on the time that the plane crosses a line which is perpendicular to the great circle path from the origin airport to the destination at a distance of 100 nmi along the great circle path to the airport to touchdown. This distinction is quite important in convective weather since a plane may be vectored off the great circle path and hence cross the perpendicular line at a point that is much further than 100 nmi true range from the airport.

2. *Inability to capture the region in which airborne delays are incurred due to terminal weather*

   One important issue is the spatial extent of the ATC impact from weather in or near the terminal area. Lamond (2002) shows a number of examples of aircraft that are held at distances much greater than 100 nmi from an airport when there is convective weather at the airport. Apropos the Atlanta ITWS studies, one of Lamond's examples includes a Delta flight into Atlanta that is holding at about 100 nmi such that the range of the aircraft from the airport crosses the 100 nmi distance boundary several times; this may be an explanation for the wide variations seen in flight times from 100 nmi to touchdown that is apparent in the ASPM summary data statistics.

   Figure 12 shows flight tracks and weather for one of the 2003 "thunderstorm days" in Table 3. Arrivals into Atlanta are black plane icons with associated tracks; white plane icons are departures from Atlanta. The Atlanta terminal area is the white circle shown in the center. NEXRAD radar reflectivity is shown in green, yellow and red. A squall line is nearing the Atlanta airport. Note that holding patterns are outside the Atlanta ARTCC (light purple line) in Alabama and Tennessee. Clearly, arrivals are being held in holding patterns at distances greater than 100 nmi from the airport due to loss of terminal/airport capacity by convective weather. Such holding patterns also occur on days where convective weather occurs in the terminal area, but there is not a TS METAR at the airport (Figure 13). Note that the holding pattern to the northwest has arisen because the northwest arrival fix into the terminal area is largely blocked by convective weather.

   Additionally, it has been learned that holding at the arrival fixes occurs virtually every weekday at Atlanta near 8 am and in the afternoon due to congestion at the arrival fixes. Figure 14 shows an example of holding patterns outside 100 nmi on a day with very little convective weather.

Hopefully, if one could utilize the individual flight data records which can be retrieved from ASPM (or ETMS) one could achieve much improved accuracy estimates (e.g., accuracies on the order of a minute) of parameters such as flight time from 100 nmi to touchdown.

The problem of holding beyond 100 nmi is not as easily addressed. What one is seeking to do in the flight time data analysis is to capture the region in which airborne arrival delays due to terminal weather are taken. One option for improving the capture of relevant flights would be to consider greater range from the airport in analysis. The problem with that approach is that the analysis might be impacted by weather outside the region of the ITWS. Another option would be to exclude days with holding patterns outside 100 nmi in the analysis. The problem with that approach is that such days undoubtedly are very high delay days that are important to keep in the analysis data set. We suggest using a threshold for distance from the airport for the long distance holding events if it can be determined from radar data analysis that in doing so, one will not be including additional regions of convective weather that are also causing delays.

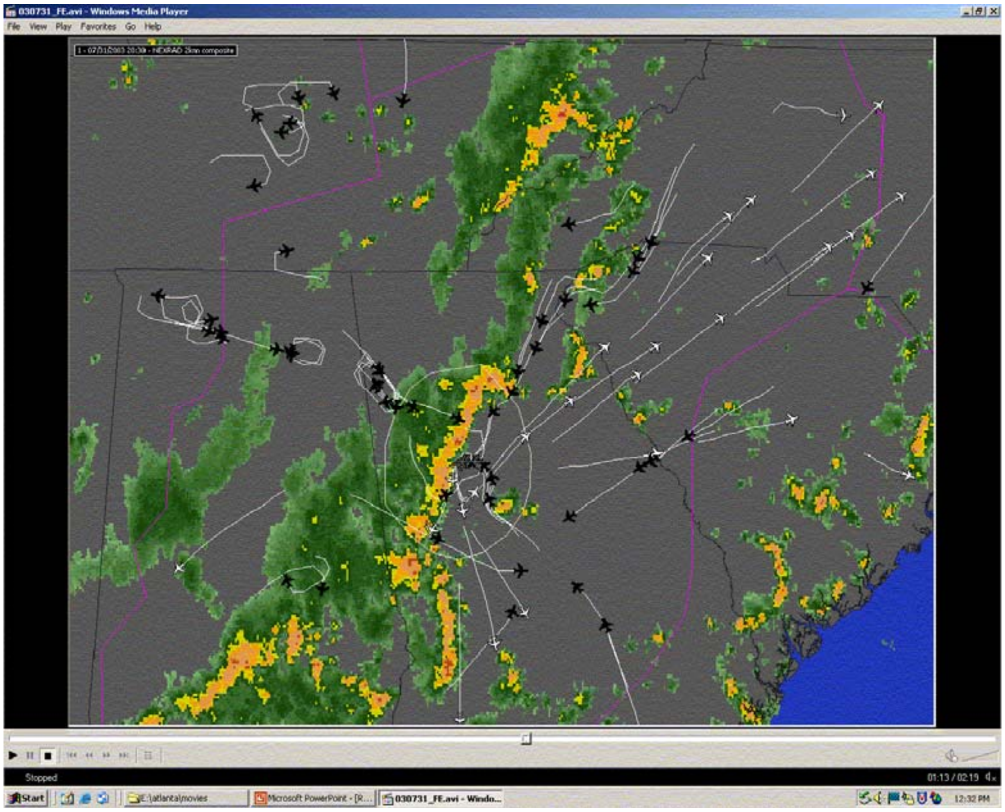

*Figure 12. Aircraft tracks and weather at Atlanta at 2030Z on 31 July 2003 (one of the Atlanta "thunderstorm day" analysis cases in Table 4).*

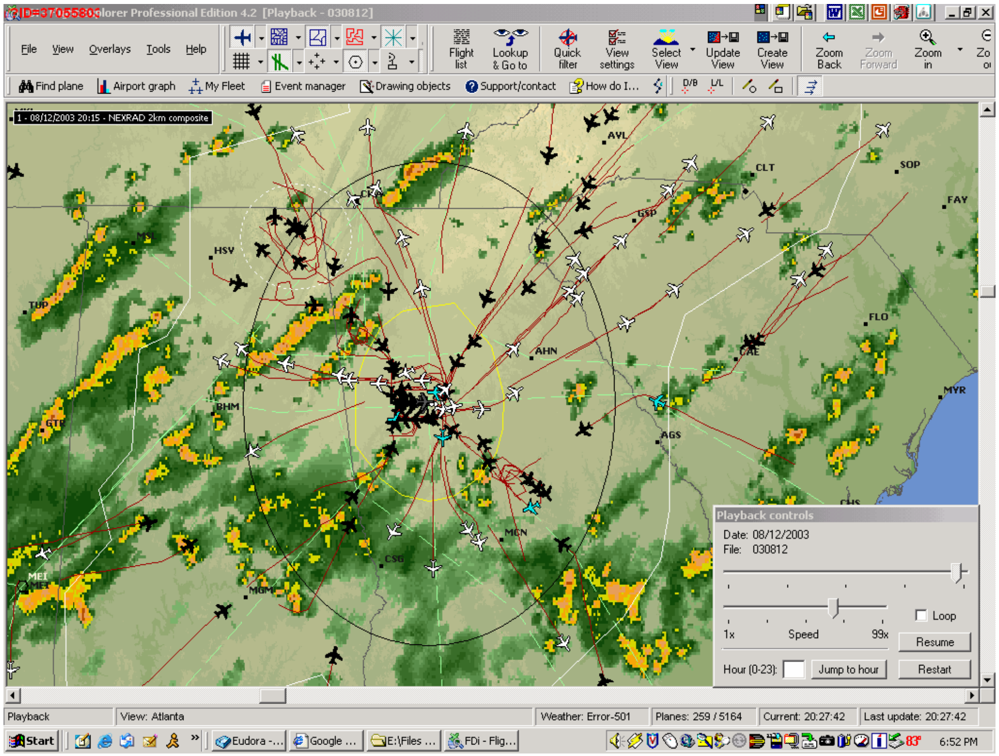

*Figure 13. Aircraft tracks and weather at Atlanta on 12 August 2003 .*

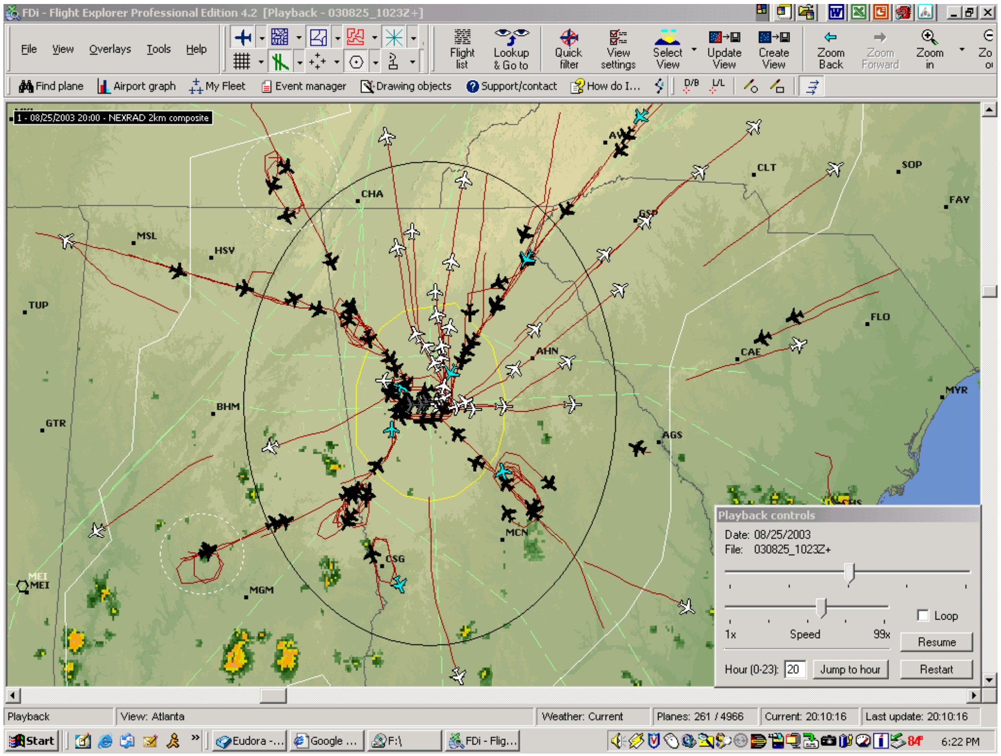


*Figure 14. Aircraft tracks and weather at Atlanta on 25 August 2003.*

One might imagine that the use of 22 thunderstorm days per year would roughly normalize for the differences in weather between 2001 and 2003. However, that appears to be an erroneous assumption. Although 2001 and 2003 had a similar number of "thunderstorm days" between April and August (54 in 2001 and 55 in 2003), the amount of precipitation that occurred (which might be viewed as a surrogate for the severity and duration of convective events) was quite different as shown in Figure 15. For example, during the period June through August, the inches of precipitation at the Atlanta airport in 2003 (16.2 inches) was 60 % greater than the same time period in 2001(10.5 inches).

A far more germane metric is the amount of time that various key points inside and immediately outside the Atlanta terminal area were impacted by convective weather and the extent to which one or more of the arrival fixes into the Atlanta terminal area were impacted by convective weather. We conducted a study of the time duration for weather impacts in the Atlanta terminal area using the NOAA Real-Time Verification System (RTVS) NEXRAD based validation data (Mahoney, et. al, 2002). Our study found that the number of time periods with continuous convective weather at the Atlanta terminal area was <u>significantly</u> higher in 2003.

If one makes the assumption that storms in the terminal area will reduce the effective capacity of the airport, leading to queues as manifested in airborne holding patterns, it follows from equation (2) that one might compare severity of the convective weather delay by comparing the respective sum of the squared durations of terminal weather events. That metric suggests that the 22 thunderstorm days in 2003 used for the Atlanta ITWS study were **47% more severe** in terms of delays than the 22 thunderstorm days in 2001. The RTVS validation data is only provided every 2 hours and that the spatial resolution is poor. Hence, a much more detailed assessment using high time/space resolution storm data is really needed as discussed below. Nevertheless, the results here should raise a warning flag for any convective weather delay comparisons that do not consider in detail the nature of the convective weather events as well as the number of convective events.

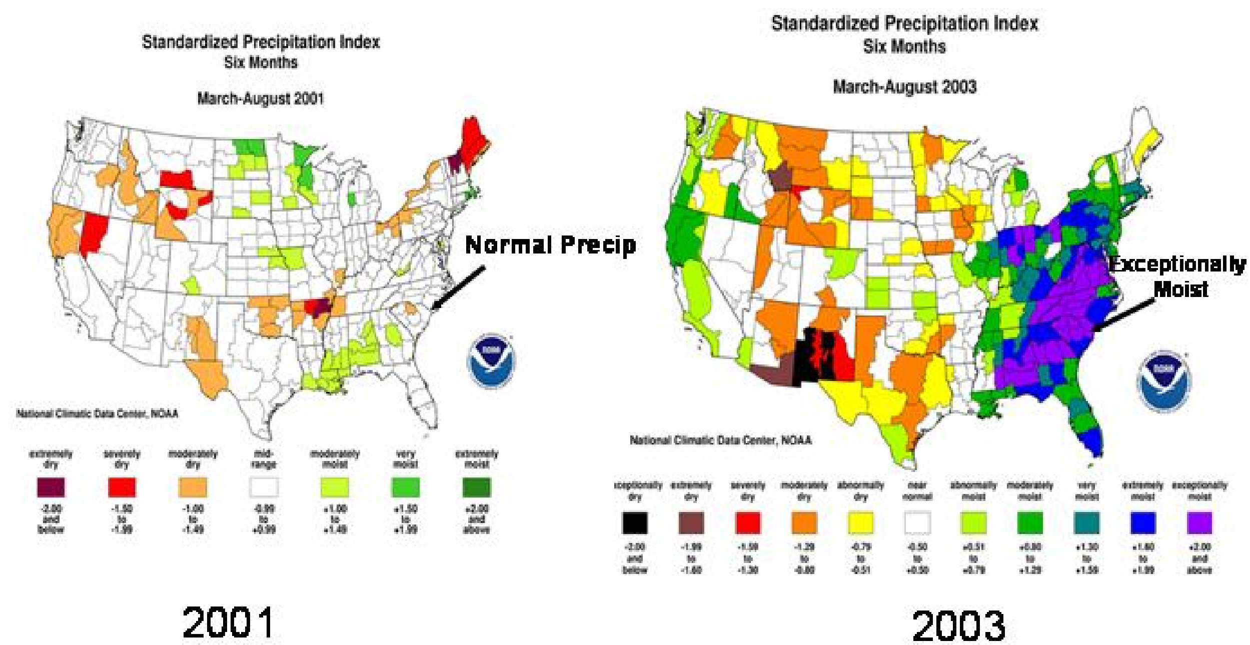


*Figure 15. NOAA National Climatic Data Center (NCDC) Standardized Precipitation Index during the six-month period from March through August for 2001 and 2003.*

As a result of our experience above with ASPM flight time analysis and our analysis of other (unpublished) studies of delay statistics based analyses of ITWS performance at Atlanta, we have concluded that one needs to understand in depth:

- The operational ATC procedures for handling convective weather in the test facilities,
- The limitations (and especially systematic errors) of the ASPM data set,
- The variability in convective weather delay impact between different events, and
- How other systems [e.g., the Center-TRACON Automation System (CTAS)] that were introduced at the same time as a convective weather decision support system would impact the delay statistics.

### *C. What can one do to make delay statistics comparisons a viable measurement option?*

As noted in the introduction to this section, there are many factors which can impact convective weather delays and thus complicate assessing benefits from measured delay data. In assessing systems that have a limited domain (e.g., the ITWS, WARP at an ARTCC, and portions of the CIWS domain), we believe that one should look at a subset of the overall convective delay problem that is appropriate to the system under test.

After discussing that option, we will turn to the question of how one might understand whether improvement has been achieved in the NAS as a whole.

#### 1. Focus attention on specific situations in which one believes that there should be measurable delays

For systems such as ITWS that provide products over a limited domain, it should be possible to systematically define "similar" weather situations before and after system introduction. The word "similar" is quite significant here since a common misconception is that it is easy to find identical convective storm events. It can be shown by combinatorial arguments that the likelihood of identical convective weather events is very unlikely[19]. However, one can seek to find

[19] A detailed discussion of this possibility will be published elsewhere. However, the gist of the argument

predetermined similarities in *degree* of convective impact identified by event duration, *specific* event location, comparable demand/capacity profiles (i.e., time of day, day of week, specific routes/airports involved), and unique city pairs involved.

For example, the distribution of weather in a terminal area can be coarsely characterized by impacts on the airport, the four arrival fixes and the four departure fixes. Even though two events may not have impacts on exactly the same fixes, the likelihood that convective weather will impact a <u>similar</u> number of fixes on different events is much higher than the likelihood that it will be the <u>same</u> fixes. Since the sources of the major traffic flows into a terminal generally have concentrated spatial orientations (e.g., the flow to Atlanta from the northeast is much greater than the flow from the southwest), there may need to be some adjustments to account for the traffic loading differences.

Similarly, one can look at departure delay reduction from terminals if one can find a case where weather is "local" in the sense that few if any of the routes to the destination cities are significantly impeded in en route or terminal airspace (e.g., consider only flights that meet this criterion). In en route domains, one can envision making a similar decomposition and simplification of convective weather impacts. For example, the CIWS benefits analysis (Robinson, et. al., (2004) highlighted the relatively small number of high altitude jet routes that handle east bound arrivals into the New York/Philadelphia terminal complexes. Here again, one can seek to compare delays for cases where the number of arrival and departure routes that are impacted are similar.

We would emphasize, however, that such simplifications should be developed as an <u>outgrowth</u> of <u>in depth</u> understanding of the ATC/convective weather dynamics for the region under consideration. Such understanding should be developed by:

- Discussions with knowledgeable ATC personnel from the facilities of interest, and
- Confirmation of the utility of simplified capacity models by examination of movie loops of traffic plus convective weather

as opposed to "data mining" only from low content databases on delays and causality (e.g., using only ASPM summary statistics and METAR data sets).

**2. Develop a weather impact index that can be used to normalize actual delays**

One of the "holy grails" of convective weather performance analysis would be a convective weather severity index that would enable one to compare delays on a normalized basis (e.g., one compares normalized delay = actual delay/weather severity index). Such an index would be extremely useful for assessing the performance of individual systems that cover large areas (e.g., CIWS) when there are large, complicated weather events. It would also be useful for normalizing NAS delays for FAA performance metrics that relate to on-time performance.

Metrics for the severity of convective weather have been developed by MITRE CAASD (Callahan, et. al. (2001) and the Free Flight Program (Post, et. al., 2002). The common denominator in both of these indices (Figure 16) is that the index is the space/time sum of the product of a weather severity factor (radar reflectivity or lightning) at a location and the fair weather traffic density at that location. Post, et. al, (2002) show an example of how national delays trend up with the Free Flight convective index. Both of these indices are appealing in that they consider weather and demand simultaneously.

However, on close inspection, it is clear that there are significant problems with these indices as a tool for detailed quantitative analysis. One problem is that the indices are not sensitive to the manner in which a route is closed. For example, if a squall line is aligned along a single major route, the index would give roughly the same value as would occur if the squall line were perpendicular to a number of major routes that travel in a given

---

is that one can compactly characterize the ATC impact of a convective weather by assessing which jet routes and fixes have been impacted. Terminals alone would typically have at least 14 such regions which corresponds to 2**14 possible combinations for a given time. Each of these over 16,000 combinations typically would evolve into one of a comparable number of other possibilities roughly every half hour. When one looks at assigning about 100-200 storm events per year to such a large number of combinations, it is clear that the likelihood of two assignments agreeing perfectly is vanishingly small.

direction (e.g., east-west in the Midwest or southwest to northeast along the east coast). However, the disruptive effect is clearly much greater when a number of major routes are blocked than when a single route is blocked.

A major factor in the New York and Atlanta convective weather delays, as well as delays in the CIWS domain, is that queues are a key issue in significant delay situations (e.g., > 40 minutes). The approximate linear dependency of the metrics shown in Figure 16 on weather coverage and duration is clearly at variance with equation (2)[20].

We would argue that instead, one needs to have a weather severity index that explicitly considers:

- Capacity of terminal and en route airspace with the actual weather locations (including consideration of the convective weather features that have been demonstrated to be critical for route usage such as storm reflectivity, echo tops, lightning, storm type, and growth/decay), and
- The degree to which rerouting (which leads to a longer path flown delay) could be used as opposed to ground holding (use of queues at the departure airport) and/or holding in airspace to address the loss of capacity due to convective weather

Developing such an index will require some substantive research and development since:

- Adequately validated models for terminal and en route capacity with convective weather do not yet exist,
- Some very preliminary work has been done on approaches for optimizing the use of rerouting and/or holding, but there needs to be many more detailed studies with comparison to actual events, and
- Although there has been some useful work on how to model "down stream delays," there is some question as to whether those should be included in a severity index. One option for handling this might be to consider indices with and without "down stream" effects considered.

We believe that the development of such an index would benefit the operational use of convective weather decision support systems as well as improve the ability to assess the convective delay reduction benefits of individual systems.

## 9. SUMMARY

In this paper, we have tried to give a reasonably comprehensive view of contemporary practices and new directions in the assessment of delay reduction benefits for aviation convective weather decision support systems. The importance of this topic has increased dramatically in the past few years as a result of the national focus on performance based organizations as well as the increased FAA emphasis on quantitative metrics.

Based on our experience with conducting and analyzing convective weather benefits analyses over the past decade, it has become clear to us that there is much misunderstanding about the causes of convective weather delays that has caused major complications and difficulties in subsequent benefits analyses. In the first few sections, we have argued that the bulk of the convective delays result from the very complicated dynamic behavior of a highly nonlinear NAS network in which both terminal and en route capacity losses due to convective weather lead to delay. Hence the delay that results in many cases cannot be neatly decomposed into terminal and en route additive components.

We have described explicit models for convective delays that can be used in many cases to estimate the benefits of particular systems. Hazards in attempting to assess the frequency and duration of terminal delay convective weather events by use of surface observations alone were detailed.

[20] For example, both indices would view a situation in which convective weather shut down all of the arrival fixes into an airport as being roughly twice as severe as a convective weather situation in which half the arrival fixes were shut down. Clearly, the delays that result from a total shutdown is much worse than twice the delays associated with a loss of 50% of the capacity. Similarly, the two indices suggest a 2 hour convective weather impact at an airport is twice that of a one hour impact whereas equation (2) would suggest that a 2-hour impact causes a delay impact that is 4 times the delay due to a one-hour storm impact.

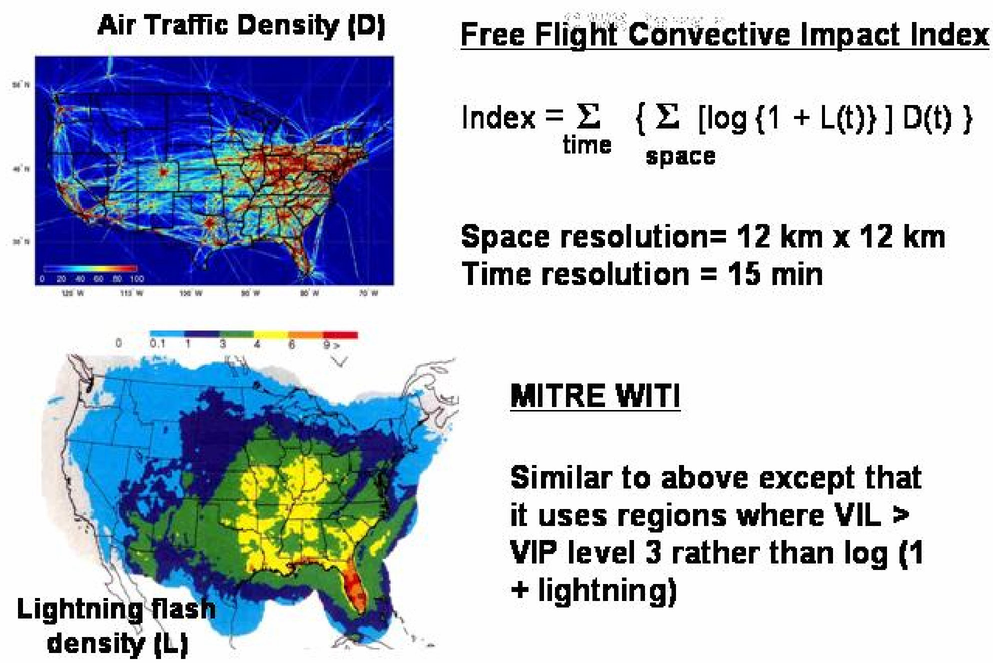


*Figure 16. Convective weather impact indices developed for the Free Flight Program (Post, et. al, 2002) and by MITRE CAASD (Callahan, et. al, 2001)*

One of the major challenges in convective delay assessments is how to utilize the insights of the operational users while still providing quantitative results that are replicable and objective. The initial efforts largely utilized the models described herein applied to post convective season interviews with operational users. The "in situ" observations of usage obtained together with highly detailed case analyses that were used in the CIWS program initial benefits analysis appear to be a significant improvement over the earlier post convective season interviews, albeit they were much more expensive to execute.

We believe that confirmation of user feedback results by comparing air traffic management before and after a system is introduced is another powerful means for validating the results from user feedback. This is accomplished by comparing flight tracks and weather for similar convective weather situations before and after a new capability is introduced. The online NEXRAD archive at the National Climate Data Center plus (limited) availability of ETMS flight track data makes this approach much more plausible than was the case several years ago. This approach appeared useful in a preliminary study for the WARP program and is being attempted in a follow on study to the CIWS benefits study discussed above.

We have recommended that aviation delays be compared for convective weather cases that have been carefully chosen to minimize the impact of factors that are not part of the system under test. A promising approach is to limit the test region to focus on the system under test and first consider only convective weather cases that have "local' effects. We suggest that the design of methodology for such analyses include careful consideration of:

- The ATC (and airline) user feedback on how they are using the system under test and the dynamics of convective weather causality in the test region, and
- Case studies of air traffic management in appropriate convective weather situations before and after the system under evaluation was tested

Moreover, it will be very important to appropriately utilize the ASPM database. In particular, we have noted problems with the accuracy of take off and landing times for a significant fraction of the ASPM data sets as well as the problems that can arise in using the ASPM summary statistics for metrics such as flight time from 100 nmi range to touchdown due to the geometrical algorithms used in ASPM to compute parameters such as the "100 nmi range." One must also be careful to understand the critical difference between the region containing convective weather versus the region in which delay is taken as was exemplified by holding patterns more than 100 nmi from the Atlanta airport due to terminal convective weather.

For cases where the system under test covers a significant fraction of the NAS and studies of the overall NAS performance in handling convective weather, it will be necessary to develop a fairly comprehensive convective weather impact index. We have described the difficulties in appropriately addressing high convective delay situations with the published convective weather impact indices and outline the key elements of an improved index that would explicitly handle the terminal and en route queues that are characteristic of high convective delay events.

**Acknowledgments**

It is a pleasure to acknowledge the contributions of R. Ferris and J. Mellon of Lincoln Laboratory in analyzing data from Atlanta to address various salient issues. We also benefited from analyzing unpublished FAA studies of ITWS delay reduction benefits based on OPSNET and ASPM delay statistics.

## 10. REFERENCES

Allan, S., S. Gaddy, and J. Evans, 2001: Delay Causality and Reduction at the New York City Airports Using Terminal Weather Information Systems, Massachusetts Institute of Technology, Lincoln Laboratory, Project Report ATC-291.

Beatty, R., R. Hsu, L. Berry, J. Rome, 1999: Preliminary Evaluation of Flight Delay Propagation Through an Airline Schedule, Air Traffic Quarterly, 7, pp. 259-270.

Bieringer, P., D. Miller, and D. Meyer, 1999: A Refinement of Thunderstorm Climatology for the Terminal Radar Control Airspace, 8th Conference on Aviation, Range, and Aerospace Meteorology, AMS, Dallas, TX.

Bond, L., 1997: Global Positioning Sense II: An Update, Journal of ATC, pages 51-55, Oct.-Dec. 1997.

Boswell, S. and J. Evans, 1997: Analysis of Downstream Impacts of Air Traffic Delay, Massachusetts Institute of Technology, Lincoln Laboratory, Project Report ATC-257.

Callahan, M., J. DeArmon, A. Cooper, J. Goodfriend, D. Moch-Mooney, and G. Solomos, 2001: Assessing NAS performance: Normalizing for the Effects of Weather, 4th USA/Europe Air Traffic Management R&D Symposium, Sante Fe, NM.

Cooper, J. Goodfriend, D. Moch-Mooney, and G. Solomos, 2001: Assessing NAS performance: normalizing for the effects of weather, 4th USA/Europe Air Traffic Management R&D Symposium, Sante Fe, NM.

Daganzo, Carlos 1997: Fundamentals of Transportation and Traffic Operations, Pergamon, Oxford.

Evans, J.E., and E.R. Ducot, 1994: The Integrated Terminal Weather System (ITWS) Massachusetts Institute of Technology, Lincoln Laboratory Journal, Vol. 7, No. 2, p. 449.

Evans, J.E., T.J. Dasey, D.A. Rhoda, R.E. Cole, F.W. Wilson, and E.R. Williams, 1999: "Weather Sensing and Data Fusion to Improve Safety and Reduce Delays at Major West Coast Airports," Massachusetts Institute of Technology, Lincoln Laboratory, Project Report ATC-290.

FAA, 1999: Air Traffic Bulletin Issue 99-3.

Lamon, K. 2004, Beyond OPSNET: NAS Performance Metrics for the 21st Century, NEXTOR Metrics Workshop, January 2004 (available at www.nextor.org/Jan04/Ken_Lamon_04.pdf).

McCartney, S., 2004: FAA Spreads Flight Delays More Widely, Wall Street Journal, 4 August 2004, p. D-1.

Mahoney, J.L., J.K. Hederson, B.G. Brown, J.E. Hart, A. Loughe, C. Fischer and B. Sigren, 2002: The Real-Time Verification System (RTVS) and its Application to Aviation Weather Forecast, 10th Conference on Aviation, Range, and Aerospace Meteorology, 13-16 May, Portland, OR.

Newell, G. 1982: Applications of Queuing Theory, (2nd edition), Chapman Hall, London.

Post, J., J. Bonn, M. Bennett, D. Howell, and D. Knorr, 2002: The use of flight track and convective weather densities for National Airspace System efficiency analysis, IEEE/AIAA Digital Aviation Systems Conference (DASC).

Robinson, M., J. Evans, B. Crowe, D. Klingle-Wilson and S. Allan, 2004: CIWS Operational Benefits 2002-3: Initial Estimates of Convective Weather Delay Reduction, MIT Lincoln Laboratory Project Report ATC-313 (9 April 2004)

Rhoda, D. and M. Pawlak, 1999: An Assessment of Thunderstorm Penetrations and Deviations by Commercial Aircraft in the Terminal Area, Massachusetts Institute of Technology, Lincoln Laboratory, Project Report NASA/A-2

Rhoda, D., E. Kocab and M. Pawlak, 2002: Aircraft encounters with Thunderstorms in Enroute vs. Terminal Airspace Above Memphis, Tennessee, 10th Conference on Aviation, Range, and Aerospace Meteorology, AMS, Portland, OR.

Wood, B., 2004: Quantifying the impact of severe weather on NAS performance, 11th Conference on Aviation, Range, and Aerospace Meteorology, Hyannis, MA.