Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

USING FAST-TIME SIMULATION TO ASSESS WEATHER FORECAST ACCURACY REQUIREMENTS FOR AIR TRAFFIC FLOW MANAGEMENT

Alexander Klein

AvMet Applications, Inc.
1800 Alexander Bell Drive, Suite 130
Reston, VA, 20191, USA

ABSTRACT

We present the concept and initial results of using fast-time air traffic simulation modeling to assess requirements for convective weather forecast accuracy from a Traffic Flow Management (TFM) standpoint. For strategic TFM applications with longer lead time (2-8 hours), such requirements can be relaxed compared with tactical (0-1 hours) weather avoidance applications. A gradual increase in forecast error does not always cause commensurate gradual changes in operational costs (e.g. delays) associated with a given TFM action (for example, a ground delay program or strategic reroute) which was implemented based on that forecast. But, while such modest inaccuracies in forecast may not initially require any adjustments to TFM actions, at some point when the discrepancy between forecast and actual weather exceeds a certain threshold, it may prompt a different weather avoidance strategy to be initiated. This, in turn, may cause a significant increase in operational costs. This paper demonstrates how such thresholds, indicative of forecast accuracy requirements for TFM, can be determined using parametric forecast accuracy changes in a series of simulations.

1 INTRODUCTION

1.1 Background: Traffic Flow Management and Weather Forecast Uncertainty

Adverse weather, such as low cloud ceilings, wind, winter weather, and especially convective weather (thunderstorms), is the cause behind the majority of delays and, more broadly, excess operating costs (over no-weather baseline) in the U.S. National Airspace System (NAS) today. Strategic Traffic Flow Management (TFM) uses Traffic Management Initiatives (TMIs), including Ground Delay Programs (GDPs), Airspace Flow Programs (AFPs), coordinated reroutes and Miles-in-Trail restrictions, which are key instruments for traffic managers to control demand and avoid unsustainable congestion on weather-impacted days. An important point is that the TMIs are typically applied to aircraft still being on the ground prior to departure, because the cost of ground delays is much lower than the cost of airborne delays. Final adjustments are then made tactically by air traffic control when aircraft may be still on the ground or already airborne.

A major challenge for TFM is the uncertainty in weather forecasts and traffic demand prediction. If the impact of adverse weather is over-forecast, TMIs may turn out to be unnecessarily restrictive. Conversely, if weather impact is under-forecast, too many flights may be allowed into en-route or terminal airspace, exceeding the airspace or airport capacity degraded due to worse-than-forecast weather, which results in excess tactical rerouting, airborne holding and, ultimately, flight diversions.

Improving the accuracy of weather forecasts is therefore a major goal for the aviation community. At the same time, there is a notion that a longer-range forecast, e.g. 4-8 hours ahead, is inherently less accurate than a short-range forecast (e.g. 1-2 hours ahead). In the foreseeable future, strategic TMIs are not likely to become a perfect TFM tool which, once set, doesn't require any further adjustments.

There is also a notion that relatively minor differences in actual weather vs. forecast are unlikely to alter a particular TMI's stance. This is explained, to a large extent, by the step-wise rather than linear nature of weather impact. For example, isolated cumulus clouds in an airspace sector may still allow for mostly smooth flow of traffic (with occasional tactical deviations); as cloud growth increases, the airspace might remain mostly permeable for a while; but at some point, thunderstorm-impacted portions of the airspace will become unavailable, causing a significant degradation of sector capacity and necessitating reroutes and/or delays.

These thresholds or "inflection points" in weather impact are the determining factor in understanding forecast requirements for TFM. It may not be necessary (however desirable) for the forecast to very accurately predict all changes in weather conditions: the minimum requirement is that it should capture the onset or cessation of the inflection points in weather conditions as they relate to TFM.

1.2 Prior Work on Weather Forecast Accuracy Requirements Assessment

Ample literature exists on post-event assessment of the accuracy of convective forecasts, from both meteorological forecast validation standpoint (using metrics such as Probability of Detection (POD)) and weather impact standpoint (using derived metrics such as forecast vs. actual convective weather coverage in an airspace sector). However, little work has been done on the quantitative assessment of maximum allowable errors – from TFM perspective – in predicted thunderstorm location and timing as a function of forecast lead time, as well as airspace type (en-route vs. terminal airspace vs. airport vicinity). Discussions on this subject were initiated over 15 years ago (see, for instance, Weather Forecasting Accuracy Workshop Report 2003) but specific requirements have been formulated more recently (Sauders 2011). These requirements were based on the analysis of TFM concept of operations vis-à-vis potential convective weather impacts on air traffic in en-route and terminal airspace. At that time, it was not possible to conduct numerical evaluation of these requirements through simulations due to lack of large-scale, weather-aware fast-time simulation models. Lately, however, such capabilities have become available (see next sub-section). In 2016, AvMet Applications, Inc. (AvMet) was awarded a contract by the FAA Aviation Weather Division to evaluate forecast accuracy requirements for TFM through simulation modeling.

1.3 DART: A Weather-Aware Fast-Time Simulation Model

DART, a superfast-time weather-aware simulation model developed at AvMet, is capable of processing an entire day of air traffic on a continent-wide scale such as the U.S. National Airspace System (NAS), in the order of 50,000 flights, in high detail including airport traffic management, departure/arrival sequencing and spacing, terminal and en-route airspace, convective and airport surface weather and forecasts, a variety of TMIs, as well as tactical reroutes, delays and delay propagation for individual flights, in less than two minutes on a PC (Robinson and Klein 2013). Its operational outcome metrics, such as arrival delays, cancellations and diversions, are typically within about 10% of actual metrics statistically recorded for the same historical period as measured by Route Mean Square Error (RMSE).

2 RESEARCH APPROACH

2.1 Methodology Using Fast-Time Simulation

The research methodology for identifying forecast requirements for TFM can be formulated as follows.

- Using a weather-aware fast-time simulation model, develop a series of weather-impacted air traffic scenarios in different TFM sub-domains (airport arrival flow management; terminal airspace; en-route airspace) under different adverse weather conditions (convective and non-convective).
- Given the actual weather and assuming a perfect forecast, simulate a particular TMI (a GDP or a set of strategic reroutes, for example) aimed at managing traffic demand when airport or airspace capacity is reduced due to weather. Record the simulated operational outcomes (delays, excess miles flown, etc.) for this perfect-forecast baseline.

- Next, begin to reduce the forecast accuracy, either toward an under-forecast (i.e., weather turns out to be worse than predicted) or over-forecast (weather turns out to be better than predicted) in a series of simulation runs. For instance, shift the forecast location of thunderstorms in the airspace being studied by 10, 20, 50, 100, … miles from actual weather location in a particular direction.
- In each such case, simulate the TMI – for instance, a GDP or airport arrival flow redirection to a different arrival fix (the first point of the approach to the airport after completing the en-route phase of the flight), or a reroute – based on the *forecast* rather than actual weather. When the forecast weather does not differ too much from actual, the operational outcomes of a simulation run are likely to be close to baseline. However, when the forecast "moves beyond an inflection point" vs. actual weather, we expect to see a significant leap in delays or excess miles flown.
    - More specifically, in case of an under-forecast we expect to see excessive airborne delays and some diversions (that is, the TMI would be too "loose"), while in case of an under-forecast we expect to see unnecessarily high ground delays and airspace under-utilization (the TMI would be too restrictive).
- As the discrepancy in predicted vs. actual weather, for example in the location of non-permeable thunderstorms or in the timing of thunderstorm onset, is varied parametrically, instances when we observe a jump in operational outcome costs will indicate the minimum forecast accuracy requirements for a given type of TMI.

In this paper we will limit the analysis to convective forecast accuracy and will present the results of an initial exploration of two convective-weather-impacted air traffic scenarios: one associated with airport arrival flow management and the other with en-route weather avoidance via a series of strategic reroutes.

2.2 Modeling the Effects of Weather Forecast Inaccuracy

For convective weather in the U.S., we used the National Convective Weather Diagnostic (NCWD), a radar reflectivity- and lightning detection-based product (NCWD Overview 1993).

We limited the actual weather to only cover the area being studied, such as a terminal area or a portion of the NAS en-route airspace; all other weather was removed from the simulation. Then, forecast weather was generated from actual by perturbing the timing and location, as well as movement, of actual areas of convection. Here, we utilized AvMet's *AlterWx* tool which processes native NCWD 4-Km weather grid and allows the user, for instance, to shift the actual weather, rotate it or change its intensity to create a specific forecast. Deteriorating forecast accuracy will be represented by growing discrepancies between thus-generated forecast and actual areas of convection.

In this initial series of experiments we shifted the actual weather to produce the forecasts as illustrated in Fig. 1. Each original grid cell was shifted along a randomized vector in the same direction.

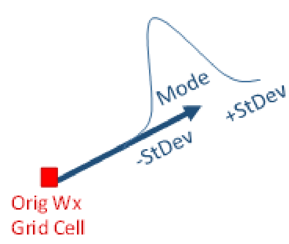

Figure 1: Shifting a weather cell (reporting point) along a randomized vector with a Gaussian distribution. "Wx" is a common aviation abbreviation for "Weather".

Fig. 2 shows an example of actual weather and a 2-hour forecast generated by *AlterWx* using such a randomized vector. Green hexagons indicate areas of weather considered permeable for most flights; red hexagons are non-permeable (delay or deviation around weather is advised); yellow hexagons are "semipermeable" ("proceed with caution"). Permeability thresholds were introduced by Sheth et al (2007), and applied in fast-time simulation modeling (Klein et al 2009).

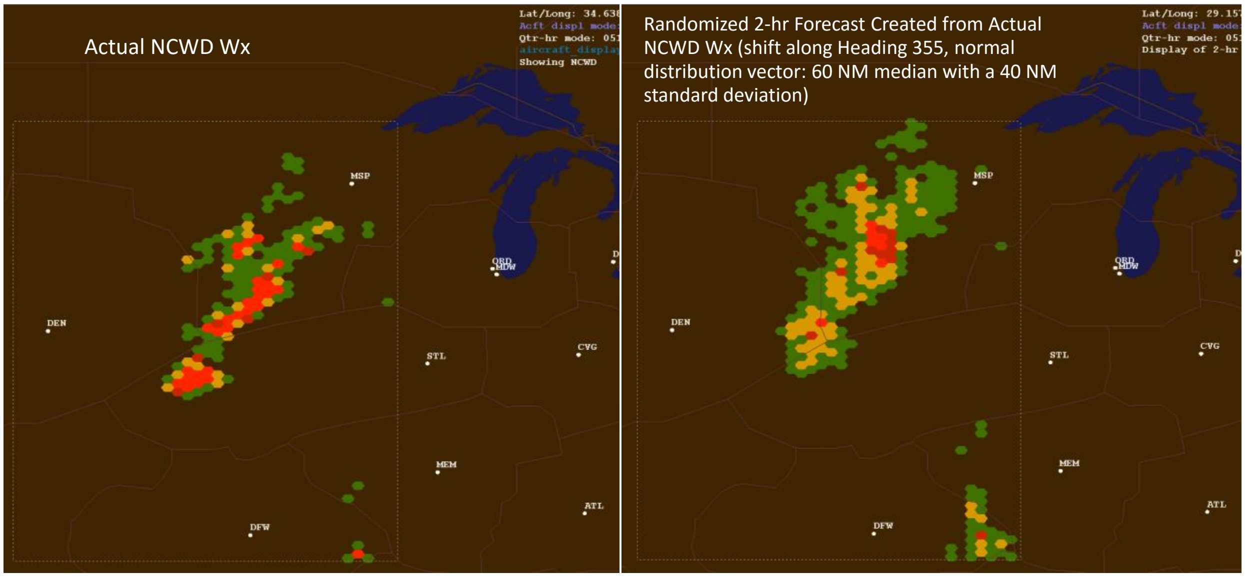

Figure 2: Actual convective weather (left) and a randomized 2-hr forecast generated by *AlterWx* tool (right). The forecast has shifted and is "fuzzier" because of the randomized shift vector.

By varying the difference between the forecast and actual weather and observing the model behavior and its outputs, we sought to identify inflection points (or, at the very least, significant-enough differences) in simulated operational response (e.g., delays, costs, fix or airway blockage, etc.) and trace them back to the forecast accuracy as the input parameter. This would allow us to compare model results vis-à-vis the forecast accuracy requirements.

3 SIMULATION EXPERIMENTS

3.1 Terminal Airspace: Arrival Flow Management

3.1.1 Scenario Description

For this scenario, we considered arrival flow management via the airport's arrival fixes. The TMIs in this case were represented by applying Minutes-in-Trail (MINIT) flight spacing restriction in DART so as to delay flights or, if feasible, by redirecting traffic from a weather-impacted fix to non-impacted fixes. Given the typical range of arrival flow management, simulated TMIs were based on a 1-hr or 2-hr forecast.

The timing and extent of MINIT activation reflects the forecast inaccuracy. Fig. 3A-3C show examples of accurate and inaccurate TMI timing. The dots in these figures indicate the airport's arrival fixes (green if not impacted by weather, red if impacted). A typical four-corner-post arrival fix configuration for a major US airport is shown. Overall traffic flow into the terminal area is managed through TFM; from there, air traffic control (ATC) finalizes arrival flight sequencing and spacing using speed control, radar vectoring and, where necessary, airborne holding. If one or more arrival fixes are impacted by weather, this puts significant strain on ATC's ability to accommodate traffic demand while keeping airborne congestion and delays to a minimum, which is why TFM's role in regulating traffic demand in the presence of non-permeable convective weather is so important.

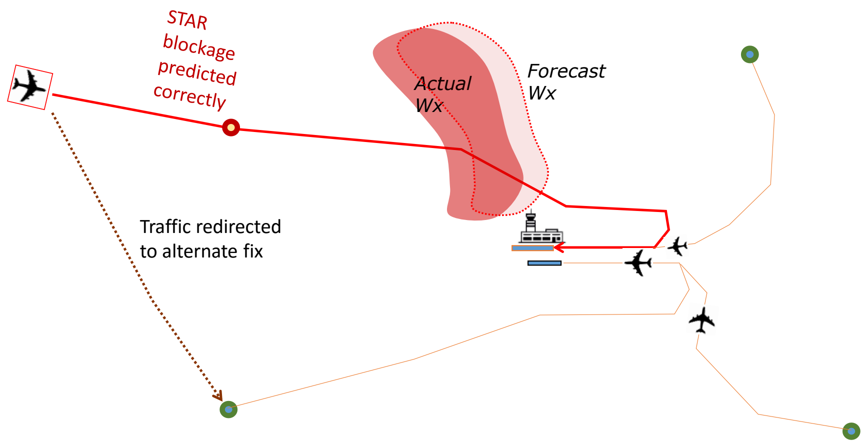

Figure 3A: A largely accurate forecast and the TMI (ground delay or redirection of traffic) reflecting it. "STAR" is the Standard Terminal Arrival Route (arrival procedure) from arrival fix to airport.

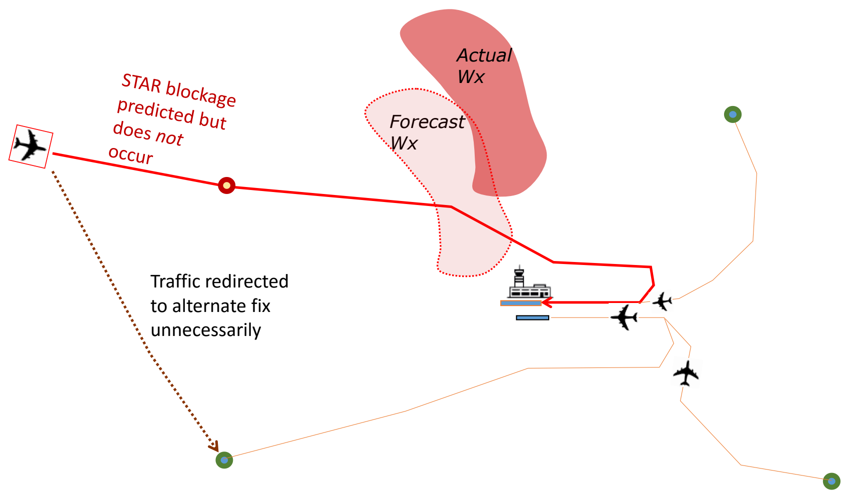

Figure 3B: Inaccurate forecast resulting in a lack of a TMI and excessive holding.

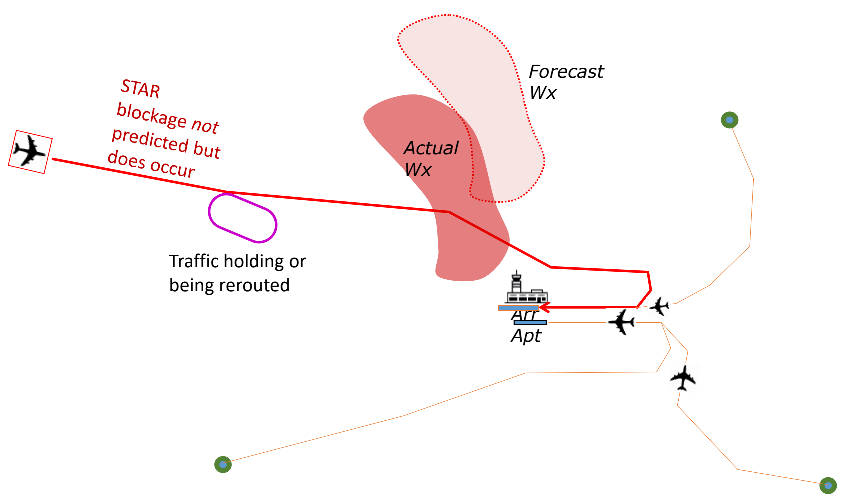

Figure 3C: Inaccurate forecast resulting in an unnecessary TMI.

DART allows the user to specify terminal airspace "wedges" that encapsulate arrival and departure fixes or groups of fixes so as to measure the convective impact on them, as illustrated in Fig. 4. Non-permeable weather (red in Fig. 4) may partially or fully block specific terminal area wedges and the arrival fixes they contain. Accordingly, if a fix becomes closed for any 15-min or longer period, the simulation model may attempt to redirect arrival flows to other fixes, or if that is not feasible, delay the affected flights.

If a MINIT TMI with perfect timing is introduced, the resulting delays and other costs (cancellations, diversions, excess miles) are largely unavoidable. If, however, the TMI timing is incorrect – too early or too late, representing varying accuracy of weather forecast – the model will respond with increased ground delays (in case of over-forecast) or airborne delays and holding (in case of under-forecast).

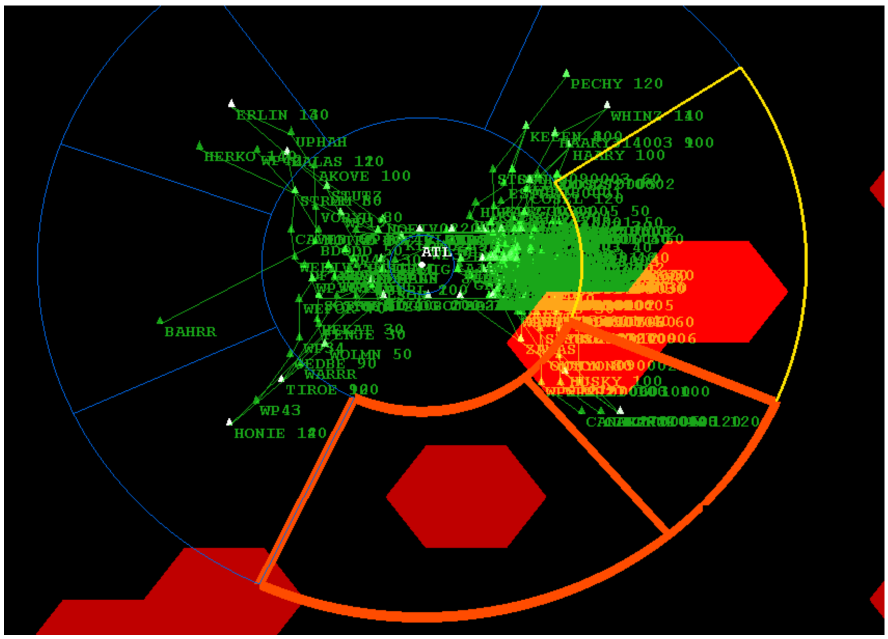

Figure 4: Atlanta Hartsfield (ATL) airport terminal airspace wedges (red if fully blocked) and STARs (green polylines). Non-permeable weather is indicated by red hexagons.

3.1.2 Sample Simulation Results

Figures 5A-5C show DART simulation results for three weather-impacted days at ATL.

To measure operational outcomes in simulations, we used Equivalent Ground Delay (EGD), a weighted sum of ground delays, airborne delays, cancellations and diversions, expressed in minutes of ground delay, with weights determined by relative costs of these factors. Thus, a minute of ground delay is 1 EGD unit, a minute of airborne delay is C_1 EGD, a cancellation is C_2 EGD units, and a diversion is C_3 EGD units. The values of the weighting coefficients C_1, C_2 and C_3 can be obtained by comparing industry-average costs for the four EGD factors (delays, etc.); here, it will suffice to note that C_1 is > 1 while C_2 and C_3 are > 100.

For each simulated day, we only retained weather within approx. 200-NM box around ATL airport so as to eliminate weather effects not relevant to the terminal airspace being studied. Actual (historical) weather was shifted in 17 different directions (20, 40, …, 340 degrees) by the following distances: 2, 4, 6, 8, 10, 12, 15, 20, 30 and 40 NM, to create a 2-hr convective forecast. Each distance was randomized using +/- 25% of it as the standard deviation for a Gaussian distribution. Each bar in Fig. 5A-5C represents the average EGD increase across all weather shift directions for a given shift distance.

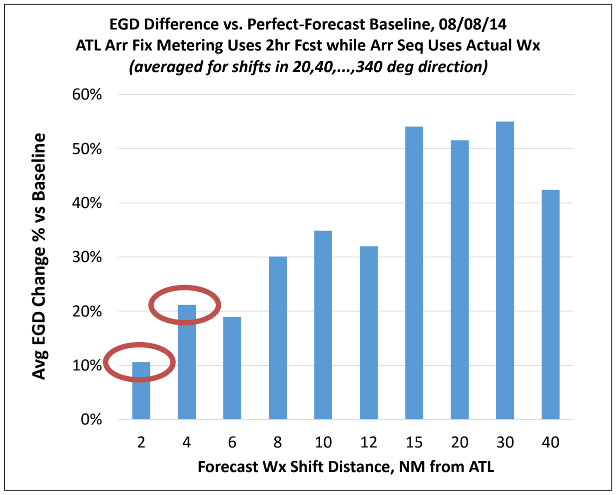

Figure 5A: EGD change percent for ATL terminal convective forecast error, very heavy impact day.

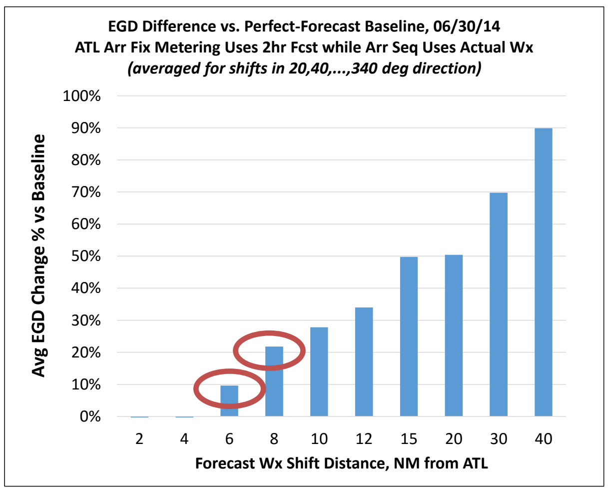

Figure 5B: EGD change percent for ATL terminal convective forecast error, heavy impact day.

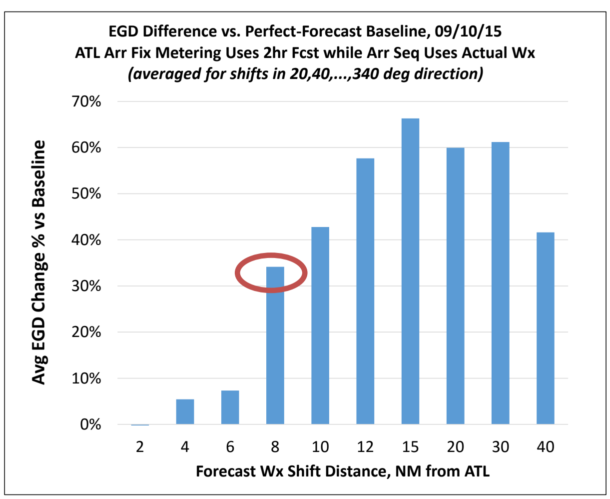

Figure 5C: EGD change percent for ATL terminal convective forecast error, moderate impact day.

The forecast error distances, when we observe a noticeable increase in EGD, are highlighted with red ovals in Figs. 5A-5C above. As can be seen from these initial experiments, the thunderstorm location error threshold varies between 2 and 8 NM, which is of the similar magnitude as the Midterm NAS Terminal Forecast Requirements (Sauders et al 2011). Smaller forecast errors appear to cause larger EGD changes when weather impact is worse, which is to be expected. Individual runs, for example all forecast weather location shift directions at 6 NM randomized distance, may exhibit significant variability – which is understandable due to the nature of convective weather. If it covers a specific sensitive area in terminal airspace (e.g. a busy arrival fix), the result may be a much higher delay than otherwise.

An alternative interpretation of the charts in Figs. 5A-5C may be that, if the thunderstorm location error is around 2-4 NM, a 10-20% EGD increase may be expected; if the location error is within 4-8 NM, a 20-30% increase in EGD may occur; and for location errors of 8-15 NM, a more significant EGD increase, up to 30-50%, may occur in some situations. The extent of the forecast error impact during a particular terminal convection event depends on the magnitude of the error itself, the overall severity of the weather, and the nature of the forecast error ("direction of the weather location displacement").

3.2 En-Route Weather Impact Scenario

3.2.1 Playbook Reroutes

Here, we used DART simulation of the so-called playbook reroutes, with custom-built reroutes that fit the chosen weather patterns. Playbook reroutes can be triggered via the embedded rules using what we call Weather Impacted Areas (WIAs). A WIA is a user-defined polygon in which the simulation model can measure weather permeability based on actual or forecast weather (Fig. 6).

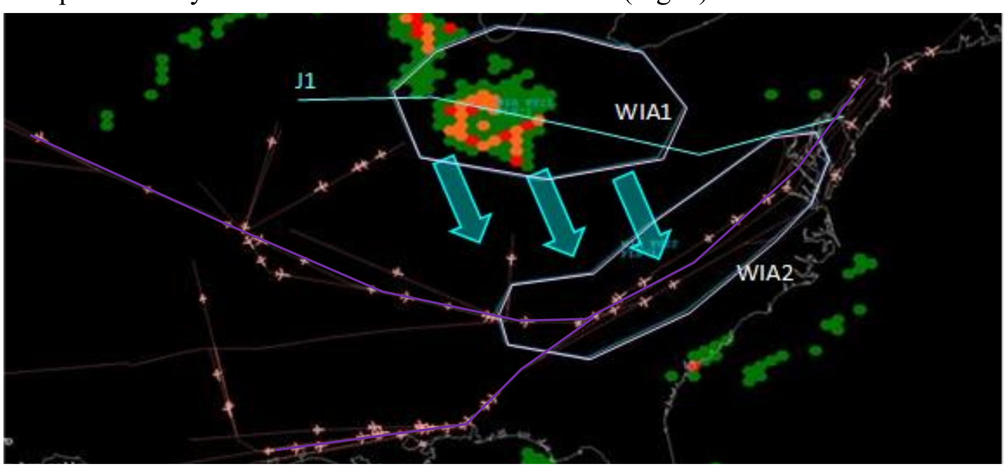

Figure 6: WIAs triggering the use of a strategic playbook reroute called VUZ (indicated by dark purple lines). The eastern portion of the US is shown. Traffic heading from central US toward the north-east is "pushed" further south, over waypoint VUZ at the intersection of two reroute branches.

Essentially, these rules allow the user to tell the model, "If there is significant weather coverage in WIA1 placed over a major route J1, but little weather in WIA2, then reroute traffic originally planning to use J1 to proceed through WIA2". The size and number of WIAs can vary. An example of a rule understood by DART:

```
[...REROUTE HEADER AND TIME TO/FROM...]
WIA1 GPF2 BELOW 70 AND WIA2 GPF2 ABOVE 90
[...REROUTE DEFINED HERE...]
```

In this example, if the permeability of WIA1 according to 2-hour convective forecast (GPF2) is less than 70% (i.e. permeability loss > 30%) but the permeability of WIA2 is > 90% (permeability loss < 10%, i.e.

minor weather impact), then playbook reroute VUZ will be triggered. TMI issue time, start and end times, and the forecast horizon used can be varied, accounting for different lead-time forecast accuracy requirements.

In addition to being able to enforce historical or model-triggered playbook reroutes, DART can simulate tactical weather avoidance; it will apply an economical combination of reroutes and ground delays for pre-departure route adjustments and, once the flight takes off, airborne reroutes or, if that is not sufficient, airborne holding. It will also seek to minimize deviations from originally filed flight plans where feasible.

The concept of airspace unit (sector or user-defined polygon) permeability quantifies the estimated loss of capacity. For example, to comply with aircraft separation requirements and controller workload limitations, an airspace sector might be capable of accepting up to 20 aircraft at any given time (unimpeded capacity = 20). If convective weather blocks part of the sector's airspace, its available capacity might be reduced to, say, 15 aircraft (a 25% loss of permeability). For more information on weather translation, the definition of permeable and non-permeable weather cells, and the sector permeability computations, see Sheth et al (2007), Klein et al (2009), and Cook et al (2011). Impact of convective weather on sector capacity is discussed, for instance, by Song et al (2009).

If actual weather is located at, or close to, the forecast weather specified in a playbook reroute, causing the WIA permeability reduction to trigger the right playbook reroute at the right time, the length and flight time of affected flights will be longer than nominal (more-direct) route, but it would be the "unavoidable" baseline because flights do need to deviate around hazardous weather (Fig. 7A).

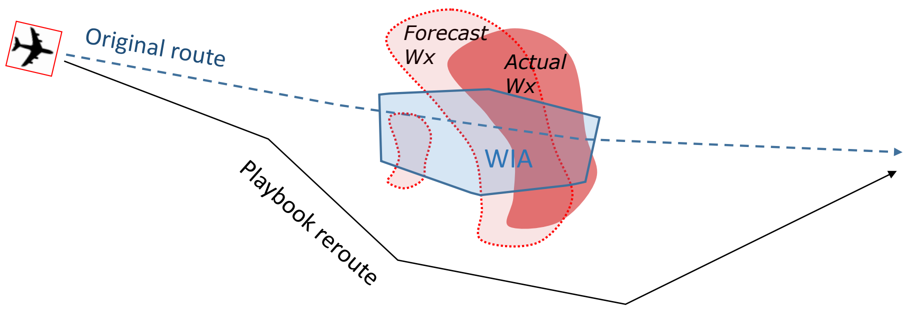

Figure 7A: Actual weather is located close to forecast: correctly applied TMI.

If, on the other hand, actual weather was displaced vs. its forecast location, it could make the TMI redundant or in fact counterproductive (Fig. 7B and 7C, respectively).

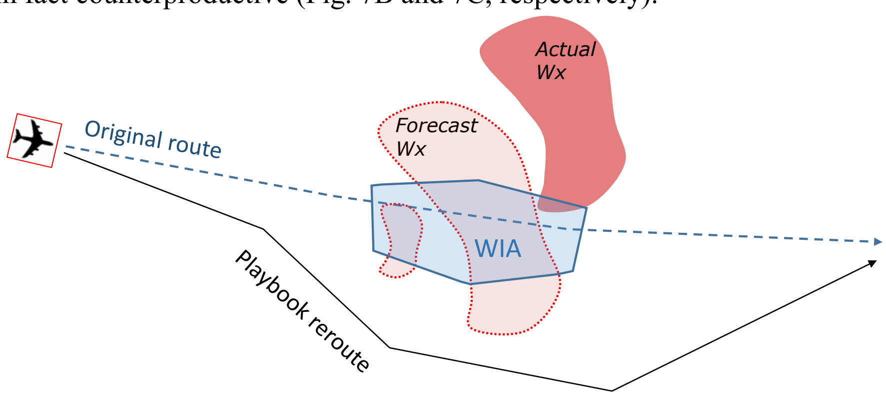

Figure 7B: Actual weather is located away from the forecast location but outside of the playbook reroute: unnecessary TMI.

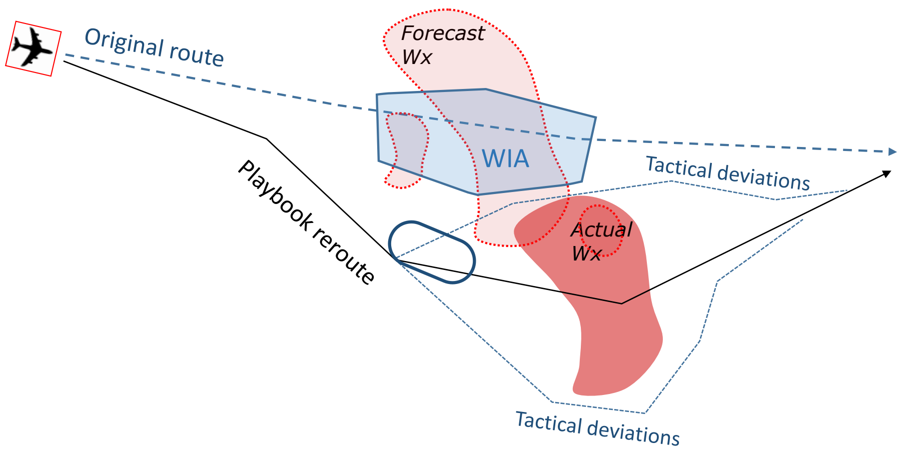

Figure 7C: Actual weather is located away from the forecast location and is blocking the playbook reroute: a counterproductive TMI.

3.2.2 Sample Simulation Results for NAS-Wide Reroutes

A weather-impacted day during a summer convective season had significant weather in Kansas City ATC Center (ZKC) and neighboring airspace, with a north-south line of storms moving east. Accordingly, we selected a "weather box" eliminating any other weather, as illustrated in Fig. 8. We also retained those historical playbook reroutes that were activated on that day to avoid the convective weather, but extended their duration to accommodate all the forecast weather variants being generated. The playbook reroutes are shown in Fig. 8 as light brown polylines.

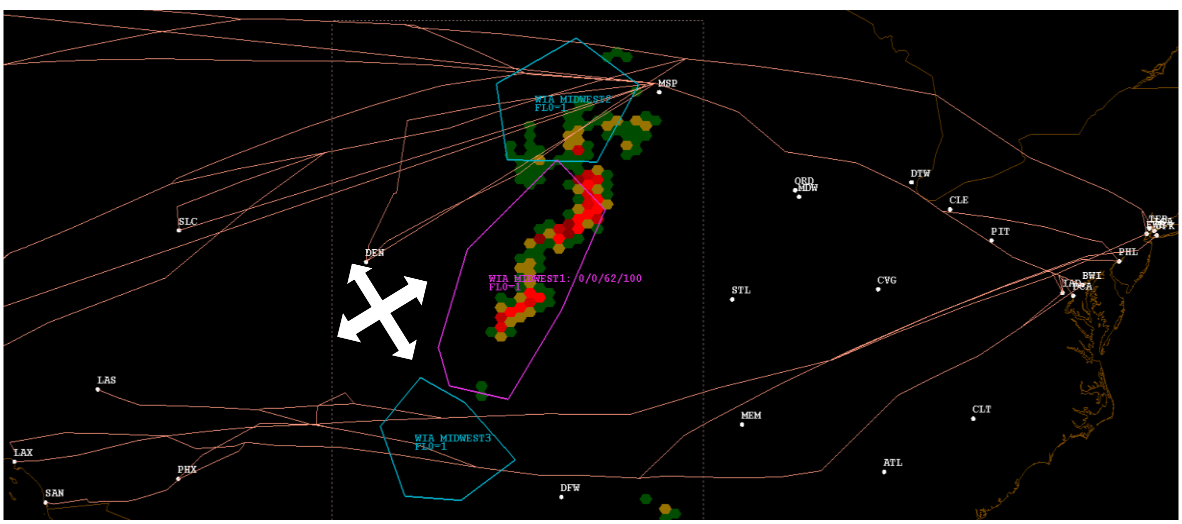

Figure 8: Reroutes around convective weather in mid-section of the US. White arrows indicate the directions in which forecast weather was shifted.

In our experiments we shifted actual NCWD native grid cells along a randomized vector in a set of directions to simulate forecast location errors. Four different directions of the forecast weather shift (along 045, 135, 225 and 315 degree headings) were used, with each individual NCWD grid cell shifted by a randomized vector with D=10, 20, …, 100 NM mean and 25%D standard deviation (see Fig. 1). Each shift angle H and mean distance D represented one DART simulation run, for a total of 4x10 = 40 simulation runs.

We also designed three WIAs, WIA_MIDWEST1, 2 and 3, depicted in Fig. 8, to capture the area most impacted by weather (WIA_MIDWEST1, in the middle) and the two smaller areas north and south of it where traffic could be rerouted if there was little weather in them (WIA_MIDWEST2 and 3, north and south of WIA_MIDWEST1, respectively). As mentioned earlier, DART simulation model allows the user to create WIA rules triggering a reroute depending on WIA permeability.

The simulation results for the 40 simulation runs in this scenario are shown in Fig. 9.

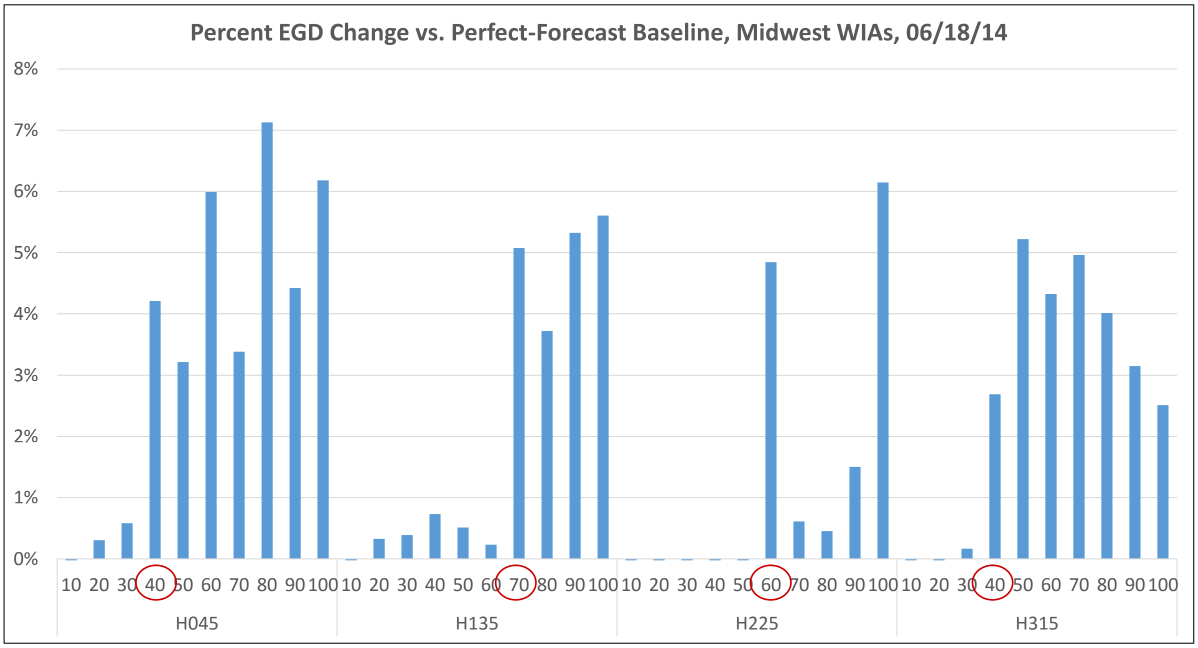

Figure 9: EGD percent change vs. perfect forecast, different 2-hr forecast location errors, en-route airspace, large-scale reroutes. "H045" indicates a shift in 45-deg direction, etc.

These results suggest that, in a strategic TFM scenario with longer lead times and NAS-wide reroutes, a convective weather location forecast error of approx. 40-60 NM would begin to cause noticeable increases in EGD. These en-route scenarios simulate the entire NAS, which is why smaller EGD percent increases, as compared to terminal airspace scenarios (which used a single airport's traffic only), are to be expected.

4 CONCLUDING REMARKS

We have applied a weather-aware fast-time simulation model to assess convective forecast accuracy requirements for TMIs. By varying the forecast accuracy, from perfect to gradually less accurate, and using parametrically driven series of simulations to create a variety of weather location errors, we were able to gauge air traffic system responses. The TMIs simulated by the model were triggered by the forecast weather timing, location and/or perceived severity; the efficacy of these TMIs depended on the forecast error, making them effective (for perfect forecast), too loose (for an under-forecast) or too restrictive (for over-forecast). That in turn changed the balance of ground and airborne delays in the simulation; and with it, altered the overall operational cost. Inflection points in the cost function indicate suggested requirements for convective forecast errors. Our initial results demonstrate that, at a more tactical level (airport arrivals), 2-8 NM errors are significant, while for strategic TMIs seeking to reduce impact of large-scale adverse weather (such as convective fronts or widespread air-mass convection), relatively large location forecast errors, several tens of nautical miles, may be tolerable when a 2-8 hour look-ahead time is considered.

ACKNOWLEDGEMENT

This research was funded under the FAA SE2025 Research & Mission Analysis Contract #DTFAWA-16-D-00002, TORP 5061, with ATAC Corporation (Santa Clara, CA) being the prime contractor.

REFERENCES

Cook, L., B. Wood, and A. Klein, 2011, "Weather Translation Modeling to Support TFM: Report of Methods for Interpreting Gridded Probabilistic Forecast Products". NASA Contract NNA11AA18C Report, Nov 15, 2011.

Klein, A., L. Cook, and B. Wood, 2009, "Directional Demand, Capacity and Queuing Delay in En-Route Airspace". *9th American Institute of Aeronautics and Astronautics (AIAA) Guidance, Aviation Technology, Integration and Operations (ATIO) Conference*, Hilton Head, SC.

National Convective Weather Diagnostic / Nowcast (NCWD/NCWF) Overview, 1993. https://aviationweather.gov/products/ncwf/info/products/ncwf/info.php.

Robinson, M., and Klein, A., 2013, "Evaluating Air Traffic Management (ATM) Benefits of Enhanced Aviation Weather Decision Support Using a Weather-Aware National Airspace Simulator," *16th Conference on Aviation, Range, and Aerospace Meteorology* (ARAM), Austin, TX, 2013, American Meteorological Society (AMS), Paper No. 10.4.

Sauders, S., F. Bayne, R. Showalter, J. Tauss, L. Leonard, and J. May, 2011, "Determining FAA Mid-Term Aviation Weather Requirements for TFM – the Transition to NextGen", *91st American Meteorological Society Annual Meeting*, Seattle, WA, 2011.

Sheth, K., B. Sridhar, and D. Mulfinger, 2007, "Application of Probabilistic Convective Weather Forecasts for Flight Routing Decisions". *7th American Institute of Aeronautics and Astronautics (AIAA) Guidance, Aviation Technology, Integration and Operations (ATIO) Conference*, Belfast, UK.

Song, L., D. Greenbaum, and C. Wanke, 2009, "The Impact of Severe Weather on Sector Capacity", In: *Proceedings of the 8th US-Europe ATM Seminar, Napa, CA, USA.*

Weather Forecasting Accuracy for FAA Traffic Flow Management: A Workshop Report. National Academies Press, Washington, DC, 2003.

AUTHOR BIOGRAPHIES

ALEXANDER KLEIN is Senior Vice-President, R&D at AvMet Applications, Inc. in Reston, VA. He holds a Ph.D. in theoretical mechanics from Moscow State University, Moscow, Russia. Dr. Klein has almost 30 years' experience in developing airport and airspace simulation models for aviation industry in the US and around the world. His research interests include fast-time simulation, National Airspace System performance analysis and metrics, analysis of weather impact on air traffic ways to mitigate it, as well as airline operations. His e-mail address is aklein@avmet.com.