Three Principles of Decision-Making Interactions in Traffic Flow Management Operations[1]

Leonard A. Wojcik
Center for Advanced Aviation System Development
The MITRE Corporation
McLean, Virginia 22102
United States
lwojcik@mitre.org

Abstract

This paper posits three principles that are fundamental to understanding decision-making interactions in aviation traffic flow management (TFM) operations involving airline schedule disruptions caused by weather. The principles have been illustrated with a computer simulation model called Intelligent agent-based Model for Policy Assessment of Collaborative TFM (IMPACT), developed by The MITRE Corporation's Center for Advanced Aviation System Development (CAASD). The principles are as follows:

1. The air traffic management (ATM) authority cannot make the best decisions for the system as a whole without collaboration from the airlines. This is because the ATM authority does not have purview into real airline costs. As a practical matter, only airlines themselves are in a position to know the economic consequences of delays to their flights.

2. Individual airlines, if they make decisions independently, tend to cause excess congestion that is sub-optimal for the airlines themselves. This is true even if the airlines have perfect information about the weather and updates about each other's intent.

3. When future weather information is imperfect, predictability may trade off with expected system efficiency.

An implication of the first two principles, which were originally presented at ATM2000, is that both the ATM authority and the airlines must participate in TFM decision-making to achieve a good result, both from the point of view of individual airlines and of the system as a whole. An implication of the third principle is that the TFM process is, in a sense, self-limiting. If actions are taken to make the system more efficient across many events, then the predictability of the system in any given event may be limited. High predictability and high efficiency (averaged across many schedule disruption events) may be unachievable simultaneously. In a TFM system that is operated to achieve high average efficiency, there can be large fluctuations in the perceived success of TFM decisions made in particular schedule disruption events.

Introduction

The TFM system, as it manifests in a schedule disruption at a major airport, is an example of a complex adaptive system. One aspect of the complexity is that the number of possible decisions among all the decision-makers is extremely large, and the decisions among different decision-makers interact. Another

[1] Copyright 2001 The MITRE Corporation. ALL RIGHTS RESERVED. This work was produced for the U.S. Government under Contract DTFA01-01-C-00001 and is subject to Federal Aviation Administration Acquisition Management System Clause 3.5-13, Rights In Data-General, Alt. III and Alt. IV (Oct. 1996). The contents of this document reflect the views of the author and The MITRE Corporation and do not necessarily reflect the views of the FAA or the DOT. Neither the Federal Aviation Administration nor the Department of Transportation makes any warranty or guarantee, expressed or implied, concerning the content or accuracy of these views.

©2001 The MITRE Corporation. ALL RIGHTS RESERVED.

aspect of complexity is that the goals of the different decision-makers, which include airlines and the ATM authority, may conflict and the information available to the decision-makers typically varies among the decision-makers. The decision-makers, faced with inability to know or control the system in its entirety, adapt to circumstances as the scenario evolves.

MITRE CAASD has developed a set of models to illuminate the complexity of TFM scenarios. The models range from relatively simple models resembling games, to more comprehensive agent-based models that incorporate software agents to represent self-interested, profit-oriented airlines.[2] These models illustrate at various levels of detail how the effects of decision-making evolve in such scenarios.

Can the TFM Authority or Airlines Resolve Schedule Disruptions by Themselves?

In previous work, IMPACT agent-based modeling results have shown that the ATM authority cannot generally resolve schedule disruptions by itself because the ATM authority does not have complete information about airlines' priorities.[3] Even if the ATM authority can accurately predict the weather, it does not, and probably cannot, have enough information about airlines' costs to optimize system performance. Hypothetically, airlines could simply report their costs to the ATM authority for use in optimizing system performance, but other results show that if airlines are deceptive about their costs, the nominal optimal solution can be severely distorted.[4]

If airlines are left to their own volition in schedule disruptions, without action by the ATM authority, another kind of effect occurs. Because the airlines are self-interested, they fail to see the full system benefit of reducing congestion. This effect is an example of the "tragedy of the commons".[5] In economic parlance, the effect can be attributed to "congestion externalities". The effect occurs even when airlines have perfect information about the weather and updates about each other's intent. Thus, when schedules are disrupted, both the ATM authority and airlines have economic roles in managing the set of responses. This is a fundamental result for TFM operations in which the system operates in a first-come-first-served mode.

In TFM operations in the U.S., the Federal Aviation Administration (FAA) and the airlines in the practice of Collaborative Decision-Making (CDM) have recognized these effects. In the 1980s, ground delay programs (GDPs) became part of the FAA's standard repertoire for responding to major weather-related capacity reductions at airports. In the 1990s, the FAA's CDM program facilitated exchange of information between airlines and the FAA and increased decision-making power by airlines to change their flights in a GDP. It also became apparent that airlines by themselves could not resolve the "bow wave" of pent-up demand that typically follows a GDP that ends when the bad weather dissipates. Thus, FAA intervention was needed, usually in the form of an extension of the GDP, beyond the period of reduced airport capacity caused by weather conditions.

Weather Information Quality

The IMPACT model results show that the effects described in the previous section occur even when information about the effect of weather on future airport capacity is perfect. Information about future weather always is imperfect, so what is the effect of this?

Figure 1 shows the results of an IMPACT simulation of an actual weather-induced schedule disruption event that occurred at Washington, D.C. Dulles Airport on August 9, 2000. In this case, the FAA declared a Ground Stop (GS) for

[2] L. A. Wojcik, "Models to Understand Airline and ATM Authority Decision-Making Interactions in Schedule Disruptions: From Simple Games to Agent-Based Models", chapter 37 in *Handbook of Airline Strategy*, ed. G. F. Butler and M. R. Keller. New York: McGraw-Hill, 2001, pp. 549-575.

[3] K. C. Campbell, W. W. Cooper, D. P. Greenbaum, and L. A. Wojcik, "Modeling Distributed Human Decision-Making in Traffic Flow Management Operations", chapter 15 in *Air Transportation Systems Engineering*, ed. G. L. Donohue and A. G. Zellweger. Reston , Virginia: American Institute of Aeronautics and Astronautics (AIAA), 2001, pp. 227-237.

[4] L. A. Wojcik, op cit.

[5] K. C. Campbell, et al., op cit.

©2001 The MITRE Corporation. ALL RIGHTS RESERVED.

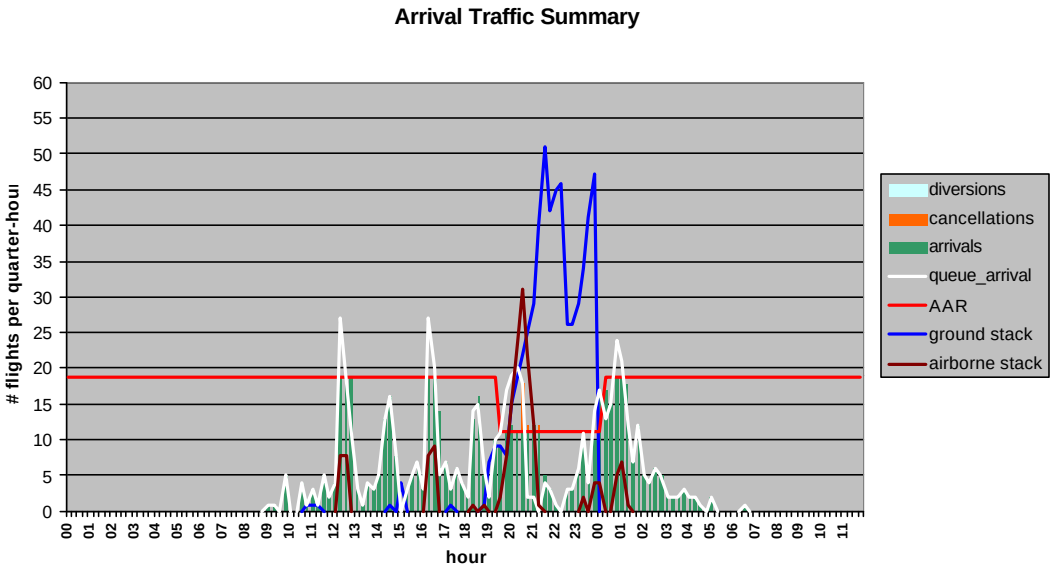

Figure 1. IMPACT Simulation of Ground Stop Scenario at Dulles Airport (IAD)

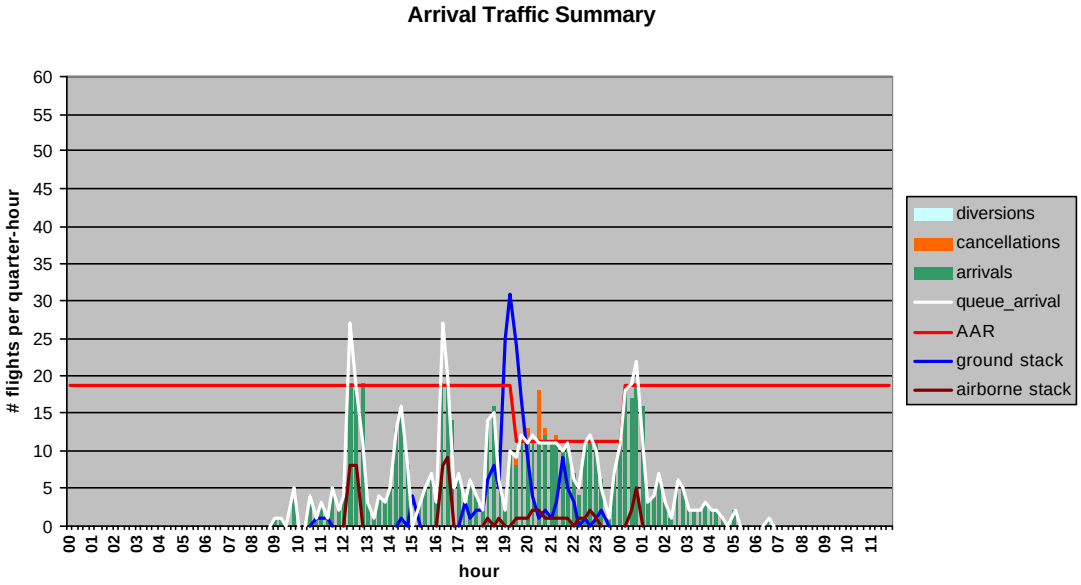

Figure 2. IMPACT Simulation of Same Scenario with an Early GDP

©2001 The MITRE Corporation. ALL RIGHTS RESERVED.

arrivals into Dulles when the weather began at about 19:00 hours, and gradually extended it to include airports more distant from Dulles. In the simulation, the FAA intelligent agent was programmed to behave as the FAA did during the actual event. The dark red line in Figure 1 shows simulated arrival capacity (in number of flights per quarter-hour), and the blue curve shows the ground stack for scheduled Dulles arrivals, which primarily is caused by the GS. In the actual event, the capacity limitation apparently was in airspace near Dulles airport, not the airport itself. In IMPACT we represented this in an approximate way as reduced arrival capacity at the airport. As the simulation evolved, the intelligent agents representing airlines made decisions based on trying to maximize their profit, utilizing perfect weather information. The airlines cancelled some flights, but an airborne stack nevertheless developed during the early portion of the weather event. The overall result (in simulation) was an under-response to the weather during the early stages of the capacity reduction and an over-response during the later stages, and high cost to the airlines. The actual event differed from the simulation in many respects, since the actual system included problems at many other airports and regions of airspace that are not represented in the simulation, plus decision makers took specific actions that are known to be different from the simulation. However, the overall qualitative effect (under-response early in the event, over-response later) was similar, and reports from the field indicate a sense of being "behind the power curve" during the event (not just at Dulles, but over the East Coast in general).[6]

Figure 2 shows the same weather event, except the simulated FAA agent was programmed to declare a GDP match to the capacity reduction well in advance of the bad weather. In this case, the ground stack was smaller and earlier than in Figure 1, and the result was much better for the airlines.

In fact, we have simulated other possible FAA actions during this scenario which show other ways it would have been possible to generate a better result for the airlines than actually occurred (subject to the approximations of the simulation). Does this suggest the FAA made a mistake? Not at all! All these simulations were done after-the-fact, and they do not take into account *the quality of information that was available to the FAA at the time of decision.* Future weather information always is imperfect, and a good after-the-fact analysis must take this into account.

Analyzing Decision-Making with Imperfect Weather Information

There is a well-established methodology for taking into account information quality in decision-making, called decision analysis.[7] Here we apply the decision analysis approach to TFM decision-making for a simple example. The scenario and choices in this initial analysis are much simpler than the real case just described; to date, the tools to facilitate decision analysis of realistic cases have not been developed. We applied IMPACT to generate the distribution of possible outcomes following an initial imperfect forecast, for each of several strategy outcomes.[8] Decision-making should be based upon the entire distribution of possible airport capacity states, not from a single predicted state. Similarly, when decision-making is analyzed after-the-fact, it should be analyzed from the perspective of the distribution of possible airport capacity states at the times of decision, not just from the perspective of the capacity as it actually turned out.

The scenario we considered had a four-hour advance forecast of severe weather. To simplify, the ATM authority was given one of three possible strategy options in a set of simulations. The analysis determined which of the three strategy options generated the best possible result on average, across the set of possible forecast histories and actual capacity states that

[6] Federal Aviation Administration, "ATC Daily Report", Wednesday August 9, 2000.

[7] Raiffa, H, *Decision Analysis*, Addison-Wesley Publishing Company, Reading, Massachusetts, U.S., 1970.

[8] L. A. Wojcik and D. P. Greenbaum, "An Air Traffic Flow Management Decision-Making Model", presented at Transportation Research Board 80th Annual Meeting, Washington, D.C., January 7-11, 2001.

©2001 The MITRE Corporation. ALL RIGHTS RESERVED.

could result from the initial forecast four hours in advance. The three strategy options were:

- Declare GDP immediately after the four-hour advance forecast, assuming the airport arrival capacity reduction matches the forecast,
- Wait two hours after the initial forecast, when the weather forecast should be better than the initial forecast, before declaring the GDP, and
- Do nothing.

If the ATM authority decides early, the capacity forecast information will have poor quality, but few flights will have departed for the affected airport, thus permitting maximum effect of the GDP. If the ATM authority decides later, the forecast will be better, but many flights will already be in the air so the GDP can influence fewer flights.

Using data from an assessment of quality of weather information[9], we generated a set of 100 forecast histories and actual arrival capacities that result following the initial forecast of severe weather. We simulated the three ATM authority strategies against each of the 100 histories. In all three cases, airlines modified their schedules in response to the same imperfect future weather information as the ATM authority. In addition, as a check on sensibility, we ran a case in which neither the ATM authority nor the airlines took any actions to change initial arrival schedules whatsoever, again against the 100 histories. Finally, we ran all four cases when both the ATM authority and the airlines have perfect weather information. With perfect weather information, there is no advantage in waiting for a better forecast, so a GDP immediately after the initial four-hour forecast is expected to be best. Since the no-actions case was independent of information quality, a total of 700 IMPACT simulations were run.

With imperfect future forecast information, the best strategy on average turned out to be the second one, namely to wait two hours before declaring a GDP. With perfect future weather information, a decision immediately after the four-hour forecast generated the best result, as expected. In both the perfect and imperfect forecast cases, no action by both airlines and ATM authority produced the worst result.

Efficiency and Predictability

Since efficient use of airport capacity matches the demand with capacity, we would expect that the best strategy tends to put demand near system capacity. However, in this critical regime where demand and capacity are approximately equal, the system is extremely sensitive to small changes in capacity and demand. In surface transportation systems, behavior in the critical regime has been studied in terms of the individual vehicles and the influence on performance at the system level.[10] These results for surface traffic indicate that efforts to manage the system to make it more efficient are in a sense self-limiting, because they make the system less predictable. Can this kind of effect occur in TFM operations?

Figure 3 shows the distribution of airline costs for the four imperfect-forecast cases, across the 100 histories. (IMPACT estimates delay costs to airlines in TFM scenarios; these costs are based on available published information, but have not been validated by any airlines.) Number of instances in which the average cost per flight to airlines fell into U.S.$500 bins is plotted against the maximum cost in each bin. Figure 4 plots the average cost and standard deviation across all cases with imperfect weather forecasts. Among the imperfect-forecast cases, the one that involved making a decision two hours in advance of the capacity reduction has the lowest average cost across the 100 histories, but its standard deviation is greater than the case where the decision was made four hours in advance. Thus, the decision-making strategy that is most efficient on average is sub-optimal in terms of predictability of total-cost outcome. With perfect information, the average cost is slightly less when the decision was made four hours in advance, as expected.

[9] L. M. Brown, et al., *Model Development Reflecting Uncertainties*, The MITRE Corporation, MTR93W125, January 1994.

[10] K. Nagel, S. Rasmussen, C. L. Barrett, "Network Traffic as a Self-Organized Critical Phenomena [sic]", Santa Fe Institute paper 96-08-056, Santa Fe, February 22, 1996.

©2001 The MITRE Corporation. ALL RIGHTS RESERVED.

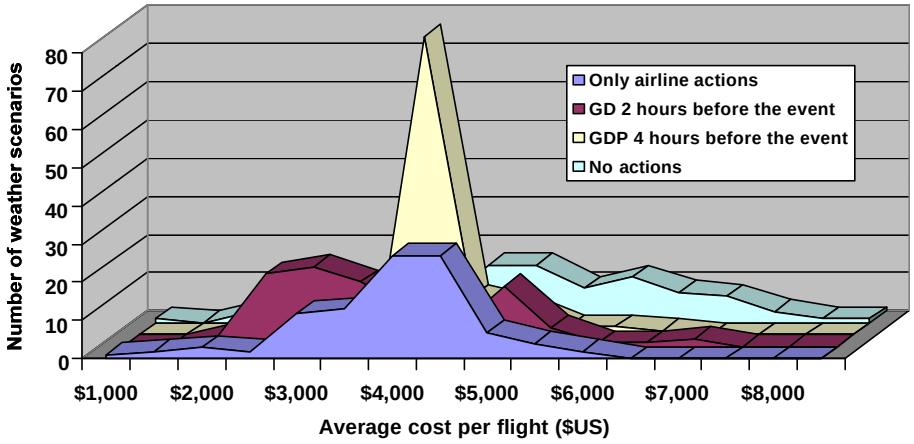

Figure 3. Cost Distribution in IMPACT Simulations for 4 Strategy Cases (100 Weather Scenarios per Case), with Imperfect Weather Forecast Information

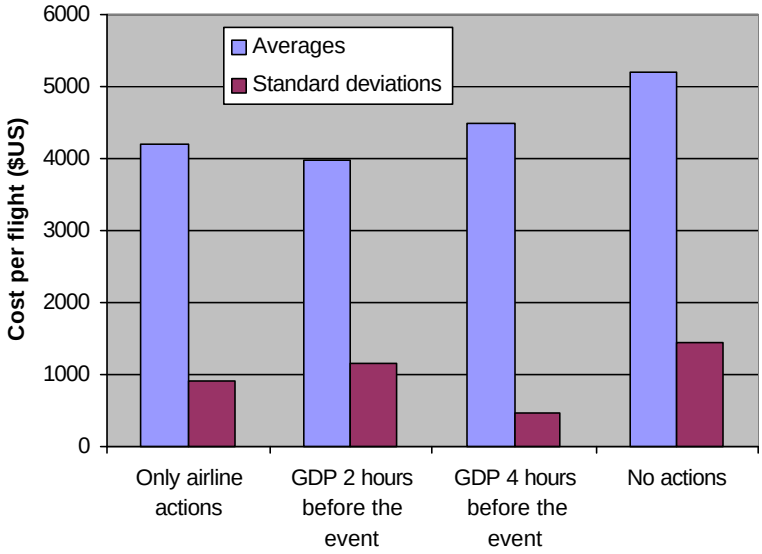

Figure 4. Average Cost per Flight and Standard Deviation Across 100 Weather Histories, with Imperfect Weather Forecast Information

©2001 The MITRE Corporation. ALL RIGHTS RESERVED.

Although we have not tried to show whether this kind of behavior occurs in other instances of imperfect information, the explanation for this behavior is clear. With a severe weather forecast at four hours in advance, forecast performance data indicates that the actual weather on average will not be as severe as the forecast. Thus, an early response to a severe weather forecast tends to clamp the system down to a demand level which is often much less than actual capacity. In this non-critical regime, system performance is relatively insensitive to changes in capacity and demand, so the result has small variance across the possible outcomes. On the other hand, waiting two hours before making a decision matches demand to capacity more accurately and tends to put the system near the critical state where demand and capacity are more nearly equal. Thus, variance is higher. In the other cases, in which the ATM authority does nothing, there is also high variance across relatively high costs because system performance continues to be sensitive to capacity changes when demand exceeds capacity.

In the set of cases where decision makers had perfect weather information, there was no such clear-cut difference in standard deviation between a decision made four hours in advance and a decision made two hours in advance of the event (see Figure 5). As expected with perfect information, a decision made four hours in advance gives the lowest expected cost, although there is very little cost difference for a decision made two hours in advance (see Figure 6).

Attitudes towards Economic Risk

One way of describing the trade-off between expected cost and predictability in TFM operations with imperfect information is through the concept of attitude towards economic risk.[11] If a decision-maker bases decisions on expected cost, the decision-maker is said to be "risk-neutral". If the decision-maker is willing to absorb additional expected cost in order to reduce the probability of a high-cost outcome, the decision-maker is said to be "risk-averse". Finally, if the decision-maker is willing to absorb additional expected cost in order to increase the likelihood of very low-cost outcomes, the decision-maker is said to be "risk-seeking". In the TFM scenario with imperfect information analyzed here, a risk-averse decision-maker may prefer declaring a GDP four hours in advance to waiting two hours. Risk-neutral and risk-seeking decision-makers prefer to wait two hours, since this strategy has lower expected cost and higher variance.

TFM decision-making is analogous to many other kinds of decision-making involving uncertain future information. For example, a trader on the stock market who owns a stock continually faces a decision of whether or not to sell the stock immediately. Imperfect information is available about the future price of the stock. If the trader receives bad news about the stock (e.g., its price starts to decline), the trader may decide to sell the stock to eliminate the risk of possible lower prices in the future. This corresponds (in a very approximate way!) to declaring a GDP early in the TFM scenario. As in the TFM example presented here, selling the stock immediately reduces uncertainty, but in some cases may be worse in terms of expected value than holding the stock. Quantitative risk management is a familiar concept in finance, and the analysis of this paper suggests that it might be useful to TFM decision-making analysis as well.

A key factor in TFM decision-making, however, is that the ATM authority makes decisions on behalf of the entire system. It would seem appropriate to adjust the attitude towards risk to correspond to what is preferred by the system as a whole, if such a thing is possible. The difficulty in reaching a consensus depends on how wide the variation is between the stakeholders on attitude towards risk. It is expected that illustrations generated using an agent-based model like IMPACT could be useful for showing examples of different decision-making strategies and could be helpful as a consensus-building tool.

Conclusions

The decision analysis perspective permits analysis of TFM decision-making that explicitly accounts for the imperfect information available to decision-makers at the times of decision. Without this perspective, and merely analyzing decision-making in terms of actual outcomes, it will be difficult to make progress on improving TFM decision-making.

[11] L. A. Wojcik, op cit., pp. 569-570.

©2001 The MITRE Corporation. ALL RIGHTS RESERVED.

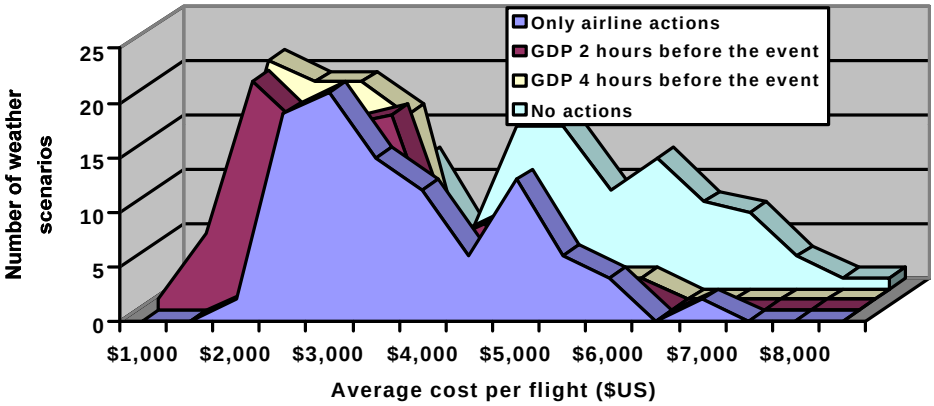

Figure 5. Cost Distribution in IMPACT Simulations for 4 Strategy Cases (100 Weather Scenarios per Case), with Perfect Weather Forecast Information

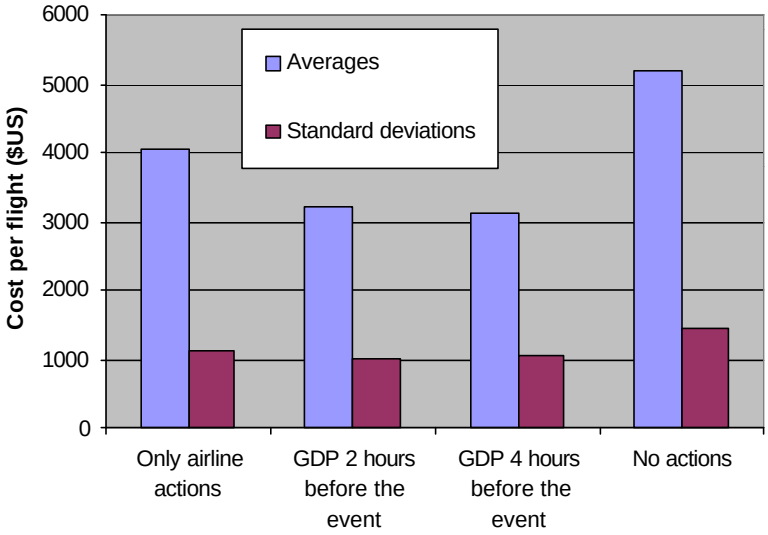

Figure 6. Average Cost Per Flight and Standard Deviation Across 100 Weather Histories, with Perfect Weather Forecast Information

©2001 The MITRE Corporation. ALL RIGHTS RESERVED.

The modeling reported in this paper combines the perspective of decision analysis with agent-based modeling. The results of the modeling show a decision-making regime in which the expected efficiency of the system (as measured by cost to the airlines) trades off against predictability. This trade-off arises when there is a forecast of a severe capacity reduction and the ATM authority has the option to respond conservatively with a ground delay program that severely limits the arrival traffic. Or, the ATM authority can wait until better information about the weather becomes available, but suffer the consequence that more flights are already airborne at the time a GDP is declared. In the single weather scenario simulated in this paper, waiting gives a better result on average, but greater variation in cost outcomes. In other scenarios, it may be better on average to declare a GDP early.

Combining the perspectives of self-interested agent-based modeling with decision analysis, we believe we have a sound basis for understanding and analyzing decision making in TFM operations and possibly other areas of ATM as well. However in the work reported to date, the scenarios are extremely simplified compared to realistic scenarios. Only arrivals at a single airport are considered and the scope of decision-makers' options is very limited. Factors such as propagation of disruptive effects to other system resources[12] [13] are not explicitly modeled. It is a research challenge to determine how to account for the complexities of actual operations in a way that can be validated, while retaining the practicality and clarity of the model. It may be possible to extend the notion of TFM decision-making to explicitly include decision-making with imperfect future information about demand, as well as capacity. Considerable additional work will be necessary to apply the basis described in this paper towards useful tools for assessing operational TFM decision-making strategies in realistic scenarios.

Sponsor Information

The work described in this paper was performed by MITRE CAASD as part of a Mission-Oriented Investigation and Experimentation (MOIE) project sponsored by the FAA.

[12] P. T. R. Wang, F. Wieland, and L. A. Wojcik, "DPAT Flight Delay Modeling and Itinerary Tracking", to be presented at the First American Institute of Aeronautics and Astronautics (AIAA) Aircraft Technology, Integration and Operation Forum, Los Angeles, California, U.S., October 16-18, 2001.

[13] L. Schaefer and D. Millner, "Flight delay propagation analysis with the Detailed Policy Assessment Tool", to be published in *Proceedings of the 2001 IEEE Systems, Man, and Cybernetics Conference.*

©2001 The MITRE Corporation. ALL RIGHTS RESERVED.

Biographical Information

Leonard A. Wojcik is a Project Team Manager with the MITRE Corporation, in the Center for Advanced Aviation System Development (CAASD). He has worked at MITRE for nineteen years. He manages a group that develops and uses state-of-the-art computer models of air traffic to help the U.S. Federal Aviation Administration (FAA) assess system-level improvements. Dr. Wojcik's recent research includes applications of agent-based modeling to air traffic management, web-enabled air traffic management, and system level simulation of Asian air traffic. He is MITRE CAASD's liaison to the Santa Fe Institute, a premier institute for the study of complex adaptive systems. Dr. Wojcik received a B.A. in physics and mathematics from Northwestern University in 1975 and an M.S. in physics from Cornell University in 1979. He received a Ph.D. in engineering and public policy from Carnegie-Mellon University in 1984.

©2001 The MITRE Corporation. ALL RIGHTS RESERVED.