Modeling Distributed Human Decision-Making in Traffic Flow Management Operations[1]

Keith C. Campbell
Senior Staff

Wayne W. Cooper, Jr.
Senior Staff

Daniel P. Greenbaum
Senior Staff

Leonard A. Wojcik
Lead Staff

Center for Advanced Aviation System Development (CAASD)
The MITRE Corporation
McLean, VA U.S.

E-mail addresses: keithc@mitre.org, wcooper@mitre.org, dpg@mitre.org, lwojcik@mitre.org

Abstract

This paper describes results from a state-of-the-art computer simulation model of distributed human decision-making in Traffic Flow Management (TFM) operations when weather disrupts airline schedules. The computer model, called Intelligent agent-based Model for Policy Analysis of Collaborative TFM (IMPACT), is believed to be the world's first model to capture the behavioral complexity of human decision-making in TFM operations.

Introduction

TFM is a process in which the economic stakes of the airlines are high, time is precious, the number of possible actions is large, and the interests of decision-makers often conflict. Complexity arises from the way in which prior actions affect later decisions as decision-makers struggle to adapt to a changing environment. IMPACT has been applied to help understand the value of increased information and collaboration between airlines and the air traffic management (ATM) authority in complex, dynamic TFM scenarios involving many airlines. The results show how collaborative decision-making can produce clear economic gains for all airlines and for the system as a whole. IMPACT shows how overall system performance is influenced by airlines whose motivation is to improve their individual performance, and how system behavior depends on the characteristics or "personalities" of the individual airlines. IMPACT also has been used to search for new ways for the whole TFM process to evolve towards better economic performance for airlines and improved service for the flying public.

IMPACT uses agent-based modeling technology, in which both individual airlines and the ATM authority are represented as self-interested, idiosyncratic agents, each with its own volition and ability to make decisions and take actions. Agent-based modeling can represent the evolution of conflict resolution and collaboration by multiple stakeholders in a dynamically changing environment. It also can be used to model possible "gaming" among airlines, as they try to exploit the collaborative system in ways that were not originally intended. Agent-based modeling, as implemented in IMPACT, is believed to be the best approach to date for modeling the complexity of decision-making in TFM operations and for generating and exploring new TFM policies based on information-sharing and collaboration.

TFM Operations and Implications for Modeling

The execution of the TFM function is complicated by the many schedule uncertainties that arise during the course of any day, especially those due to weather. Airline schedules in the U.S., for example, are typically designed for good weather days. When bad weather limits the capacity of one or more airports, U. S. Federal Aviation Administration (FAA) TFM specialists (subsequently referred to simply as "FAA") must institute TFM procedures to delay takeoffs of flights to the destination airports with reduced capacity. The delays given to flights before they take off are called “ground delays”.

The economic consequences of delay for an airline vary considerably among the airline’s flights, especially when the complex economics of hub operations is considered. One flight, for example, may have many connecting passengers to international flights whereas another flight may have relatively few such passengers. Therefore, when the FAA assigns ground delays to particular flights, the FAA is implicitly making significant economic decisions for the airlines. From the standpoint of the FAA's conduct of the ATM system, this poses a serious problem, because the FAA does not have purview into real airline costs. Only the airlines themselves are in position to know the economic consequences of delays to their flights. This has led to the concept and development of Collaborative Decision-Making (CDM) in the execution of TFM. Under CDM, the FAA controls the arrival rate at a reduced-capacity airport, but FAA collaborates with airline flight planners and dispatchers to determine *which* of the airlines' scheduled arrivals should be given priority.

The full potential of CDM on the economic performance of the ATM system is unknown and transcends the arrival substitution question described above. Perhaps most interesting is the question of how overall system performance can be enhanced if driven by decisions made to enhance individual airline performance, and relatedly, how these individual decisions would change with more information sharing among the players, who may be highly competitive with one another.

Over the years, various attempts have been made to model the complex dynamics of the ATM system, including stochastic optimization models and discrete-event dynamic simulations.[2] There has been a previous attempt to simulate multiple-agent decision-making in TFM operations based on a statistical multiple regression model; however, this attempt used a very simplified representation of airline schedules and results were reported only for scenarios in which there were just two airlines.[3] By comparison, ground traffic behavior in congested cities has been simulated at various levels of detail and geographical area, including a detailed model of vehicle behavior for an entire metropolitan area based on a cellular automaton (CA) approach.[4] We saw a need to look beyond these previous approaches to better represent the behavioral complexity of the system in order to address the questions posed above.

The approach presented here builds on pioneering complex-systems research of the Santa Fe Institute (SFI). SFI is the foremost research center for the science of complexity, a field that studies how aggregate phenomena emerge from underlying patterns. Complexity science has been applied in many disciplines, including physics, sociology, anthropology, biology, and economics, and researchers at SFI study the factors that such problems have in common. In the vernacular of complexity science practiced at SFI, the U.S. operational TFM system, including both FAA elements and airline elements, is a complex adaptive system (CAS). It is complex both in the combinatorial sense that there are many possible decisions, and in the behavioral sense that decision-making is distributed among many players who may have very different goals. The TFM system is adaptive, because the airlines and the FAA constantly receive feedback from the system and have the opportunity to change what they do.

A powerful means of studying a CAS like the TFM system in the context of its operations is through the use of agent-based modeling (ABM). The ABM approach typically is most successful when the simulated agents are kept as simple as possible within the constraints of the situation to be simulated.[5] This facilitates the possibility of *emergent* behavior, which is aggregate behavior by the system as a whole that would be extremely difficult to predict from the attributes of the individual agents, which are individually relatively simple to understand.

[2] For example, see Richetta, O. and Odoni, O. (1993), "Solving Optimally the Static Ground Holding Problem in Air Traffic Control", *Transportation Science*, 27(3), 228-238; Lindsay, K. S., Boyd, E. A., and Burlingame, R. (1994), "Traffic Flow Management Modeling with the Time Assignment Model", *Air Traffic Control Quarterly* 1(3), 255-276; and DeArmon, J. S. and Lacher, A. R. (1996), "Aggregate Flow Directives as a Ground Delay Strategy: Concept Analysis Using Discrete Event Simulation", *Air Traffic Control Quarterly* 4(4), 307-323.

[3] Klein, G. L. and Anton, P. S. (1999), "A Simulation Study of Adaptation in Traffic Management Decision Making under Free Scheduling Flight Operations", *Air Traffic Control Quarterly* 7(2), 77-108.

[4] Rickert, M. and Nagel, K. (1997), "Experiences with a Simplified Microsimulation for the Dallas/Fort-Worth Area", *International Journal of Modern Physics C* 8(3), 483-503.

[5] Epstein, J. M. and Axtell, R. (1996), *Growing Artificial Societies: Social Science from the Bottom Up*, Cambridge, MA: MIT Press.

IMPACT, the agent-based computer simulation model of TFM interactions developed by MITRE CAASD, represents individual airlines and the FAA as self-interested, idiosyncratic software agents, each with its own volition and ability to make decisions to take actions. Agents either have economic or policy motivations and have information about their environment and other agents. Once a scenario is populated with agents and a randomly generated system event is introduced, agents are permitted to act according to the rules of the system, and the scenario evolves spontaneously. This is described in more detail below.

Baseline Schedule Disruption Scenarios Modeled by IMPACT

Three kinds of baseline scenarios have been modeled by IMPACT to represent past and present-day TFM operations in the U. S.: these are called No-action, Ground Delay Program (GDP), and Collaborative Decision-Making (CDM). The three kinds of baseline scenarios are simplified representations of approaches that have been used in actual operations to respond to schedule disruptions caused by capacity-reducing weather at airports. The three baseline scenario types function as starting points from which to create new scenarios needed for IMPACT analyses of various kinds.

In all three scenario types, weather conditions are expected to cause a capacity reduction at a single airport for a portion of a single day, and airlines and the FAA respond (within the rules defined by the scenario type) to the resulting schedule disruption as the day's events unfold. To simplify the simulation, only arrivals to the affected airport are modeled. In a typical scenario, on the order of a thousand flights from the Official Airline Guide (OAG) are simulated. About ten to fifteen airlines are represented as decision-making agents, and the FAA is represented as a single agent. In these baseline scenarios, weather information is assumed to be perfect.

No-action is the simplest of the three scenario types. In No-action scenarios, the airlines simply send their scheduled flights to the affected airport with no change in departure times. Similarly, the FAA takes no action to respond in advance to the weather. Thus, in No-action, agents make no decisions, which may result in a large amount of airborne holding. No-action is an exaggerated representation of operations before the early 1980s, when responses to anticipated weather problems were not as well orchestrated as they became after the institution of ground delay programs into standard TFM practice. No-action also represents what could happen in the present system without any response to anticipated or unanticipated future weather conditions.

The GDP scenario represents decision-making after ground delay programs were introduced into standard TFM practice in the early 1980s. In a GDP scenario, the FAA responds to a future anticipated capacity reduction at the affected airport by declaring a ground delay program. The airlines respond by canceling, substituting, exchanging and delaying flights within their individual schedules. Without collaborative decision-making in TFM operations, information exchange among the participants in ground delay programs is very limited, and the model represents this (in an exaggerated way) as no information available to agents about when other airlines' flights will approach the airport. In this paper, two types of GDP scenarios are described, one in which the airlines take no actions following the GDP, and another in which they take actions based on static information about other airlines' flights.

In the CDM scenario, the FAA also responds to a future anticipated capacity reduction by declaring a ground delay program, but has improved information that allows the FAA to compress the arrival schedule by eliminating cancelled flights from the program. The model can be configured to represent different policies with respect to compression and other parameters of collaborative decision-making. Airlines can cancel, substitute, exchange or delay their own flights, but unlike the GDP scenarios discussed above, they have information about each other's intended arrival schedule. This feature represents information sharing of the current CDM program in the U.S. In this program, airlines and the FAA exchange flight intent and other information using a tool called the Flight Schedule Monitor (FSM).[6] Agents modeled in IMPACT make decisions in light of what other airlines as well as the FAA have already decided, just as real airlines involved in the CDM program make decisions in light of information about what other airlines and the FAA have already decided, using FSM. Thus, a comparison of a CDM scenario against GDP scenarios under otherwise

[6]FSM removes airline-specific information, so that users are given expected flight arrival times without airline identifiers and flight numbers.

identical conditions can be used to estimate the value of this kind of information sharing.

To represent actual operational practice in CDM scenarios, the CDM simulation software was developed to include various features that are used in real CDM operations. For example, simulated CDM scenarios include a variable window of time (e.g., 20 minutes) in which airlines are permitted to send flights in advance of the departure time needed to arrive at the capacity-limited airport at the assigned arrival time. This feature gives simulated airlines greater flexibility than they would have without such a window.

Airline and FAA Agents in IMPACT

Airline agents in IMPACT make decisions based on the expected cost to themselves of alternative actions. Modeled costs are based on available industry data, and include direct costs like fuel and crew costs, as well as other, less tangible indirect costs from expected future lost revenue caused by passenger ill will.[7] High incremental costs are assumed for cancelled flights, and very high costs are assumed for diverted flights. These costs are estimates and have not been validated with the airlines.

When an airline agent makes a decision, the model computes the expected cost to the airline of a limited set of possible alternative decisions. In keeping with the spirit of ABM, the agents do not attempt to "optimize" their decisions (such an optimization would be impossible in any case, because they do not know what other agents will do), but they make incremental changes to attempt to improve their situation. Agent attributes can be changed to represent different airline strategies, and airline agents can learn over many simulated events to improve their performance. In the results presented here, the airline agents make decisions by evaluating the approximate cost impact of a limited number of alternative possible decisions. However, the structure of IMPACT also facilitates experimentation with a variety of approaches to airline decision-making, including those based upon powerful heuristic optimization algorithms such as simulated annealing and genetic algorithms. Experimentation with these algorithms is a possible research direction for the future.

[7]Irrgang, M. (1997), "Airline Irregular Operations", in *Handbook of Airline Economics*, D. Jenkins, ed., New York, NY: McGraw-Hill.

The FAA agent in IMPACT makes decisions based on such policy motivations as keeping system demand within the capacity of the relevant resources, and promoting system efficiency and equity between airlines. In scenarios described here, the FAA agent monitors the capacity/demand balance at the weather-influenced airport, and responds within the rules of the scenario being modeled to keep demand within system capacity. In many simulations, we have assumed that the FAA schedules the ground delay program to end as soon as the expected capacity reduction is over, and does not attempt to account for the "bow wave" of pent-up demand that typically follows a restriction on demand. The bow wave is observed in real aviation operations. As with the airline agents, the model facilitates experimentation with a variety of approaches to FAA decision-making.

Sample Results for Baseline Scenarios

Figure 1 shows an arrival traffic summary for an IMPACT simulation of a No-action scenario, for a sample capacity-reducing weather event. Figures 2 and 3 show arrival traffic summaries in two types of GDP scenarios, for the same weather event. In Figure 2, the FAA declares a GDP whose arrival rate exactly matches the storm, and airlines take no further actions to modify their original schedules, which results in a large bow wave following the GDP. Figure 3 shows the outcome when the same GDP is declared by the FAA, but airlines respond based on a static assumption about other airlines' flights. Each airline modifies its schedule following declaration of the GDP, but, in the absence of information about what other airlines are doing, simply assumes the portion of the bow wave created by other airlines does not change. Finally, Figure 4 shows the arrival traffic summary for a CDM scenario with the same weather event and ground delay program. The same original arrival schedule applies across all four scenarios. Table 1 compares key parameters for the four scenarios, where delay and cost values are averaged across all flights for all airlines during the entire day. The total number of flights scheduled to arrive during the day is 1117.

Figures 1 through 4 show the capacity profile (denoted as "SumOfaar") for the weather-influenced airport as a function of time of day. Capacities are in number of flights per quarter hour. The capacity profile is identical for all three scenario types. For the GDP and CDM scenarios, the ground delay program declared by the FAA agent exactly matches the actual capacity profile.

In these simulations, agents have perfect knowledge of the future weather.

Figures 1 through 4 show the numbers of diversions, cancellations and arrivals at the airport as the day unfolds. Diversions and cancellations are shown at the original scheduled times or arrival. Arrivals are plotted at actual time of arrival; note that the number of arrivals per quarter hour never exceeds the capacity of the airport. The curve labeled "queue arrivals" shows the number of flights per quarter-hour interval entering the arrival queue. The curve labeled "ground stack" shows the number of flights held on the ground as a function of time. The "airborne stack" curve shows the number of flights in airborne holding as a function of time. (In real operations, not all flights would have to be held around the airport; they could be held en route as well.)

Comparison of Figures 1 and 2 shows that the GDP eliminates the large airborne stack in the No-action scenario during the storm but creates an even larger stack after the storm. This large bow wave after the GDP occurs because of heavy arrival demand after the storm. In the GDP scenario with airline response to static information (Figure 3) and in the CDM scenario (Figure 4), the airlines reduce the size of the bow wave after the GDP by taking actions to cancel flights, but the airlines do not completely eliminate the bow wave. The sizes of the bow waves in Figures 3 and 4 are similar.

It is possible that more powerful algorithms for airline behavior in IMPACT may permit airlines to reduce the bow wave further than shown in Figures 3 and 4, but there are limits on what airlines can accomplish as long as they are behaving independently of one another. As Figures 1 and 2 show, the original storm and the GDP produce periods of rapidly increasing airborne holding. For a particular airline's flight originally scheduled during this period of increasing airborne holding, there is incentive for the airline to have the flight arrive relatively early, rather than later, since earlier arrival would cause the flight to arrive when the queue is smaller. This applies to *all* flights arriving during the rising portion of the airborne queue, so there is an incentive for airlines to either send flights as soon as possible or cancel them. If all flights were owned by a single airline, that airline would be able to delay its flights on the ground to prevent a large airborne stack. But, with decision-making distributed among a number of airlines, the incentive to send flights early drives the system to maintain a bow wave, even though it is costly to every airline. This situation is analogous to "the tragedy of the commons", in which agents acting in their own self-interest do not necessarily promote the interest of the system as a whole.[8]

Table 1 shows that cost is highest in the GDP-only scenario. This occurs because, without airline action, the GDP pushes the delayed flights into a time interval when originally scheduled demand is already high. This creates a circumstance where flights ground hold for a significant time period, then airborne hold, then divert. These large delays followed by eventual diversion create very high costs for the airline agents.

In the CDM scenario, agents are better able to adapt to the bow wave than in either of the GDP scenarios, so total delays are less. Airline agents in the CDM scenario are able to reduce average cost per flight relative to both the GDP scenarios, as shown in Table 1. A detailed examination of behavior of the agents in the CDM scenario shows that agents use information (such as that available from FSM) to make decisions in response to other agents' decisions to modify their flight schedules. Information sharing permits agents to adapt to the excess demand in the bow wave following the ground delay program. Thus, information sharing facilitates the ability of agents to adapt, not only to the environment, but also to each other. Through such adaptation, agents are able to improve their economic performance, and the economic performance of the system as a whole improves.

Analyses with IMPACT

Analyses have been performed with IMPACT to address a variety of issues related to TFM operations, including the influence of airline characteristics or "personalities" on system economic performance, and the effects of possible changes in rules for collaborative decision-making. Other analyses have focused on decision-making by the ATM authority, for example to look at the effect of changing its planning horizon when weather forecasts are noisy. IMPACT provides a useful platform for investigating decision-making with such noisy information, because it can be run repeatedly to show the range of possible outcomes in scenarios with stochastic elements. Work has also begun to integrate IMPACT with decision analysis to show the effects of changes in decision-making criteria used by the ATM authority.

[8] Hardin, G. (1968), "The Tragedy of the Commons", *Science* 162, 1243-1248.

In one set of experiments, IMPACT was used as an open-ended search tool, to find ways to achieve better aggregate economic performance. Although this application of IMPACT is in a very preliminary stage, the approach could be used to suggest future enhancements to the CDM process to include more extensive forms of collaboration. The open-ended search was accomplished by configuring IMPACT to simulate a weather-induced schedule disruption in which the entire set of airlines behaves as if it were a single composite airline (the composite-airline scenario). In the composite-airline scenario, airlines effectively behave in ways to improve overall system performance, rather than necessarily improve their own performance. Although this kind of behavior is not realistic in current operations, the composite-airline scenario shows how economic improvement might be possible if mechanisms could be produced to ensure that overall system gains are distributed fairly among all participating airlines. These mechanisms could include, but are not restricted to, slot trading and purchasing among airlines within the ground delay program.[9]

Although the preliminary results suggest that additional improvement might be possible through more extensive forms of collaboration, these benefits may be reduced when imperfect weather predictions are introduced. Also, airline flight banking, which is not included in the present results, may influence the outcome. Finally, experiments need to be performed with different cost optimization approaches, such as those based on powerful heuristics like simulated annealing and genetic algorithms, to attempt to improve the performance of the airline agents.

Conclusions

The agent-based IMPACT model has been used to model information sharing and collaboration in weather-induced schedule disruptions. IMPACT is uniquely suited to understanding how overall system performance (and the public good) can be improved in a system where self-interested individual agents each attempt to improve the outcome for themselves. IMPACT permits experimentation and analysis in ways that would be infeasible to try with the real system or with conventional simulation methods. An example of this is in the area of decision making when there is uncertainty about future weather. IMPACT can be applied to show the range of possible outcomes that could result when an uncertain weather prediction is made, and IMPACT can show how the decision-making strategies of individual airlines and the FAA play out in such scenarios. IMPACT also can be used to perform experiments in which airlines or groups of airlines attempt to “game” the system to take advantage of information sharing and collaborative opportunities. These experiments may perform a valuable function in developing confidence among airlines and government regulators in further information sharing and collaboration possibilities for the future. There are also questions regarding the influence of banking operations, and expected or unexpected events in other parts of the system, such as reduced capacity at other airports. Another key area for investigation with IMPACT is how to help the FAA to continue to improve the quality of its decision-making in TFM operations, including guidance on how to evaluate past decisions in light of the uncertainties that existed at the time the decisions were made. By shedding light on these questions, IMPACT can help us understand what is the best path for TFM system evolution towards better economic performance through improved information sharing and collaboration.

Sponsor information

The IMPACT model was developed by MITRE CAASD as part of a Mission-Oriented Investigation and Experimentation (MOIE) project sponsored by the FAA.

Biographical Notes

Keith C. Campbell graduated from the University of Virginia in 1989, and worked in airline economics at GRA in Jenkintown, PA before earning a Ph.D. in Systems Engineering at the University of Pennsylvania. Dr. Campbell joined the Operations Research group of US Airways in 1996, where he developed next-generation revenue management models. Dr. Campbell has been at MITRE CAASD since 1998, focusing on airline economics and collaborative traffic flow management. He is currently working with the FAA's National Command Center as it implements its new strategic traffic flow management planning process.

Wayne W. Cooper received a B.S. in Systems Engineering (with High Honors) from University of Florida in 1981 and a M.S. in Operations Research (with Honors) from Georgia Tech in 1982. He worked on U.S. Army Manpower Planning optimization-based models from 1983-89, as a consultant with Sabre Decision Technologies from 1989-1995, and on electricity

[9] Adams, M., Kolitz, S., Milner, J., and Odoni, O. (1996), “Evolutionary Concepts for Decentralized Air Traffic Flow Management”, *Air Traffic Control Quarterly* 4(4), 281-306.

generation projection models at the U.S. Department of Energy (1995-96). Then he joined MITRE CAASD, where he has worked on historical data analyses and simulations of FAA traffic flow management, and on decision planning tools for airport departures.

Daniel P. Greenbaum has been with MITRE CAASD since June 1998, primarily developing agent based models and air traffic simulations. He was born in Suffern, New York in 1967 and grew up in New City, New York, a suburb of New York City. He received a B.A. in American history from the University of Pennsylvania in 1989, a J.D. from Cornell Law School in 1992, and an M.E. in Systems Engineering from the University of Virginia in 1998. From 1992 to 1996 he worked as a labor law attorney in Washington, D.C.

Leonard A. Wojcik has worked at MITRE for 17 years. He has managed a variety of modeling and simulation projects and serves as MITRE CAASD's liaison to the Santa Fe Institute. Dr. Wojcik was formerly Director of Research at the Flight Safety Foundation, and also has worked at the Office of Technology Assessment of the U. S. Congress. Dr. Wojcik received his Ph.D. in Engineering and Public Policy from Carnegie-Mellon University in 1984. He received his M.S. in Physics from Cornell in 1979 and his B.A. in Physics and Mathematics from Northwestern University in 1975.

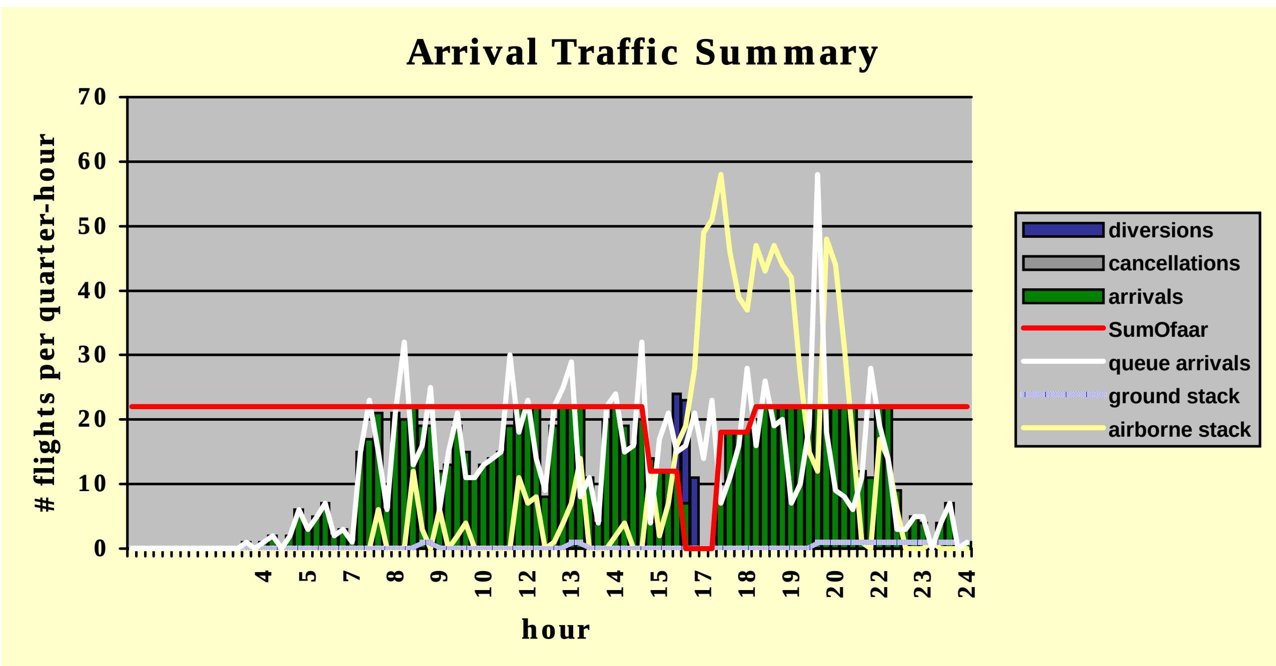

Figure 1
IMPACT output for No-action scenario

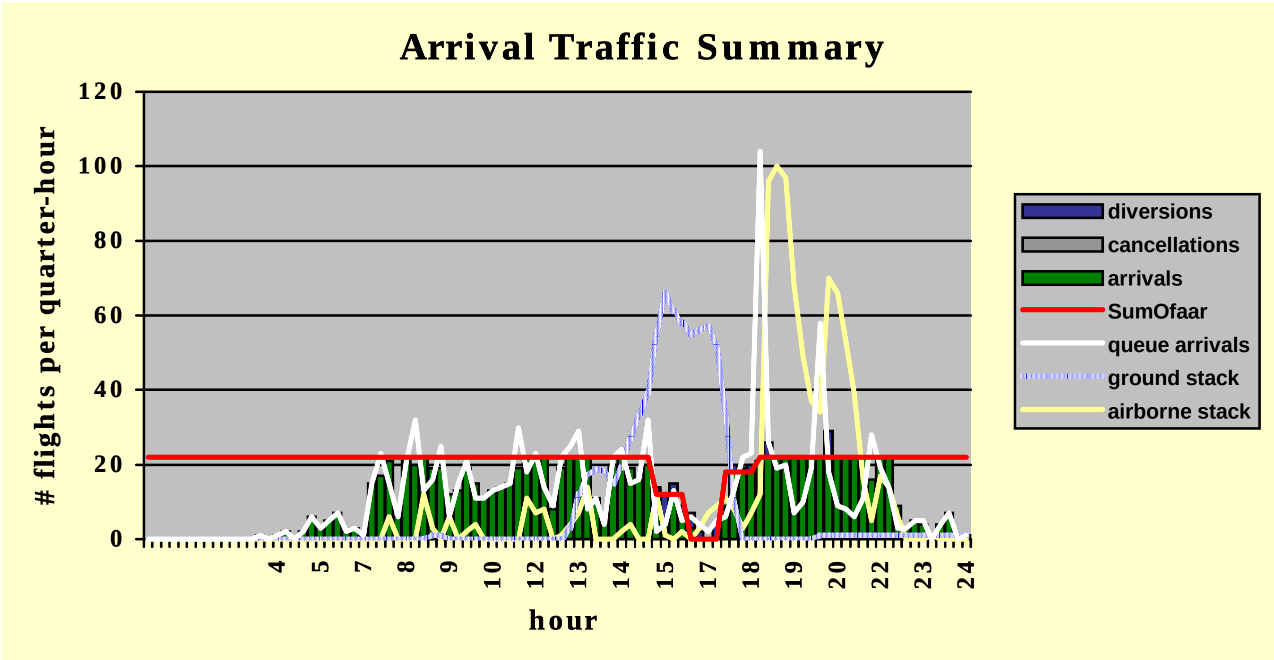

Figure 2
IMPACT output for GDP-only scenario

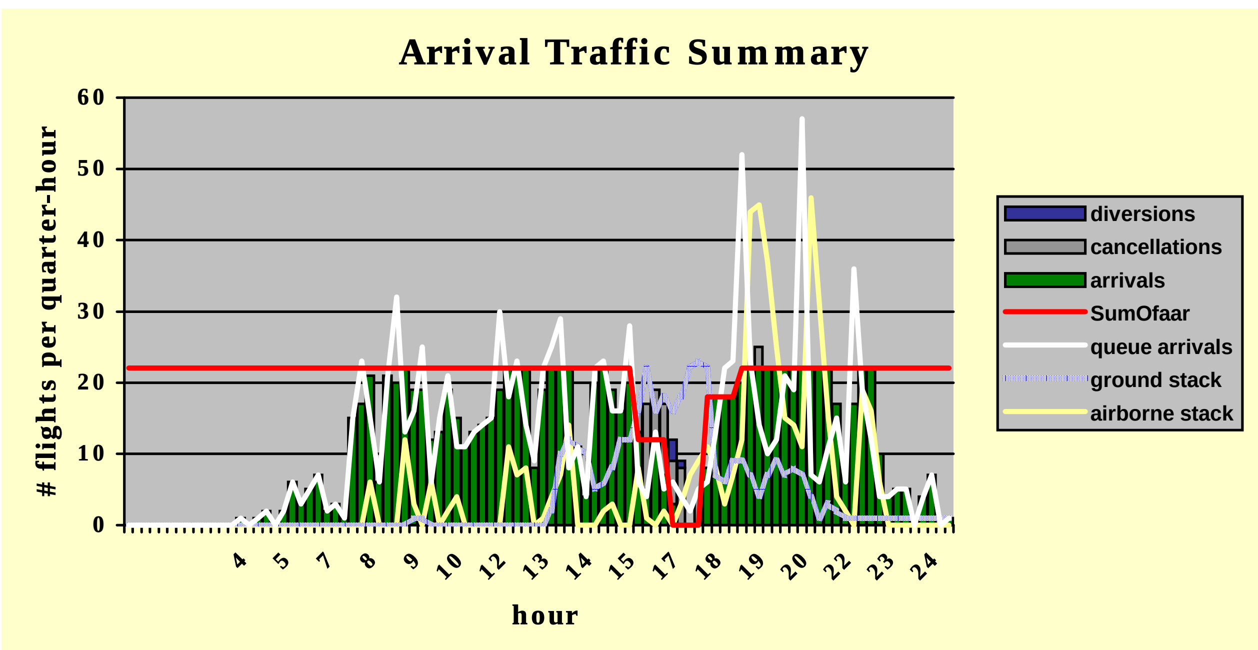

Figure 3
IMPACT output for GDP scenario with airline response to static information

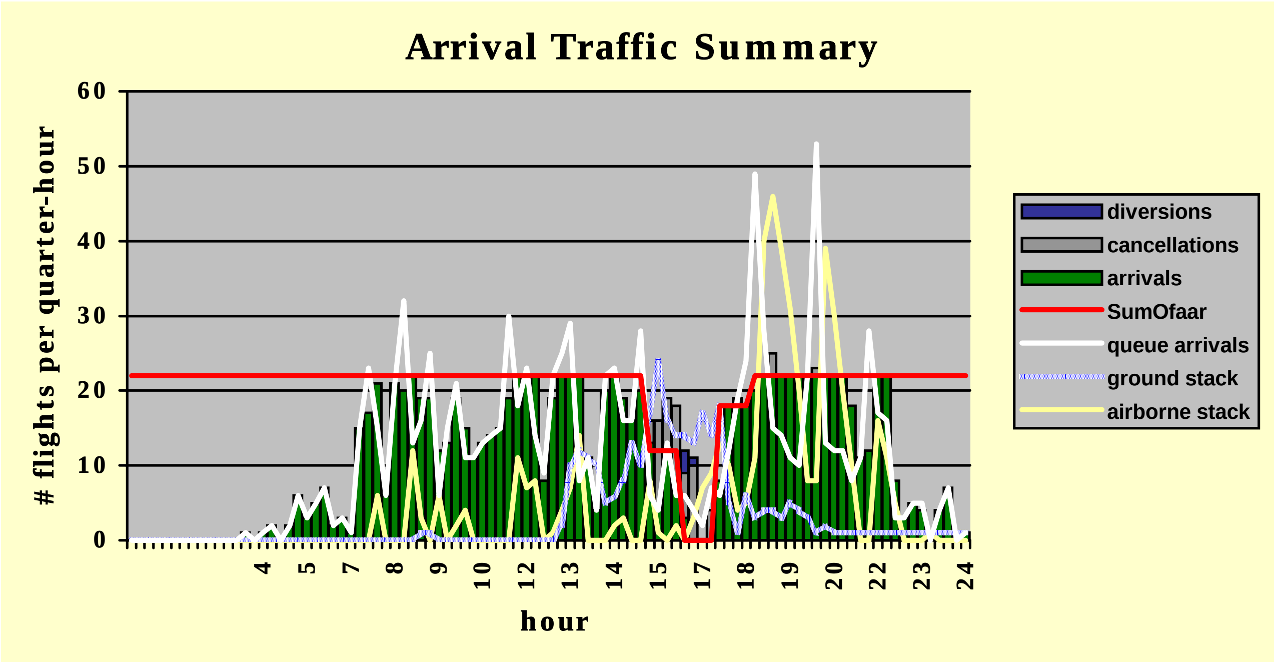

Figure 4
IMPACT output for CDM scenario

	Average airborne holding time (minutes per flight)	Average ground holding time (minutes per flight)	Number of flights cancelled	Number of flights diverted	Average cost per flight (thousands of U.S. dollars)
No-action scenario	12.5	0	0	39	4.5
GDP-only scenario	12.6	9.4	0	42	6.1
GDP scenario with airline response to static information	6.4	5.0	54	4	2.8
CDM scenario	6.4	4.1	57	4	2.7

Table 1
Comparison of the four baseline scenarios