An Airspace Planning and Collaborative Decision-Making Model: Part I—Probabilistic Conflicts, Workload, and Equity Considerations

Hanif D. Sherali • Raymond W. Staats • Antonio A. Trani
Grado Department of Industrial and Systems Engineering (0118),
Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
Department of Operational Sciences, Air Force Institute of Technology, Wright Patterson AFB, Ohio 45433
Charles Edward Via, Jr. Department of Civil and Environmental Engineering (0105),
Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
hanifs@vt.edu • raymond.staats@afit.edu • vuela@vt.edu

We present a large-scale, airspace planning and collaborative decision-making model (APCDM) to enhance the management of the U.S. National Airspace System (NAS). Given a set of flights that must be scheduled during some planning horizon, along with alternative surrogate trajectories for each flight as prompted by various airspace restriction scenarios imposed by dynamic severe weather systems or space launch special use airspaces (SUA), we develop a mixed-integer programming model to select a set of flight plans from among these alternatives, subject to flight safety, air traffic control workload, and airline equity constraints. The model includes a three-dimensional probabilistic conflict analysis, the derivation of valid inequalities, the development of air traffic control workload metrics, and the consideration of equity among airline carriers in absorbing costs related to rerouting, delays, and possible cancellations. The resulting APCDM model has potential use for both tactical and strategic applications, such as air traffic control in response to severe weather phenomena or spacecraft launches, FAA policy evaluation (separation standards, workload restrictions, sectorization strategies), Homeland Defense contingency planning, and military air campaign planning. The model can also serve a useful role in augmenting the FAA's *National Playbook* of standardized flight profiles in different disruption-prone regions of the national airspace. The present paper focuses on the theory and model development; Part II of this paper will address model parameter estimations and implementation test results.

1. Introduction

The airline industry is a highly competitive business that has a significant impact on the overall U.S. economy. This was never more evident than during the extraordinary aftermath of the terrorist attacks of September 11, 2001, that resulted in catastrophic revenue losses, subsequently requiring the federal government to provide $15B in emergency economic relief. It is in the nation's continued interest to ensure safe and dependable use of the National Airspace System (NAS), even as increased demand strains air-traffic management capabilities. Furthermore, it is

Transportation Science © 2003 INFORMS
Vol. 37, No. 4, November 2003, pp. 434–456
0041-1655/03/3704/0434
1526-5447 electronic ISSN

essential that appropriate steps are taken to encourage new entrant airlines and to develop services for smaller communities to ensure a healthy and competitive industry in the long term. Accordingly, the Federal Aviation Administration (FAA) is sponsoring an overall 10-year, $11.5 billion effort to increase NAS capacity by 30%. Currently, air traffic density is such that a single severe weather system can cause takeoff and landing delays that cascade throughout the entire NAS. During 2000, the top 55 airports conducted more than 20.8 million such operations, with 425,000 of them subject to delays (see Crawley 2001). Although not directly related to capacity enhancement, the model developed herein contributes indirectly towards this goal via a more effective management of air traffic, particularly under scenarios of disruptions because of the closure of sections of the airspace as precipitated by severe weather or by space launches.

More specifically, this paper presents an airspace planning and collaborative decision-making (APCDM) model, which is a significant extension of a preliminary version (called the airspace planning model—APM) described in Sherali et al. (2002). The principal subroutines of this model are depicted in Figure 1, along with a conceptual overall framework in which this model can be embedded.

Given a set of flights that must be scheduled during some planning horizon over some region of interest, the overall intent is to select a set of flight plans from among alternatives, called *surrogates*, subject to flight safety, air traffic control workload, and airline equity constraints. These surrogate flight plans for each flight are alternatives proposed by the corresponding airline, which might differ in departure/arrival times, altitudes, and trajectories. Each individual flight plan taken by itself is assumed to be feasible to the system constraints and might typically be generated to avoid dynamically moving severe weather systems (see Hansen et al. 2002 for an example). (Some comments on the mix of surrogates from an equity perspective are given in §5.) In this regard, it is also worth pointing out that, as is typically the case during severe weather disruptions, the mix of flight plan surrogates used in this paper do not consider any airport arrival slot ownerships by airlines. However, it should also be noted that the FAA is considering the abolishment of slot ownerships altogether in the future (Keagan 2003).

The first subroutine implemented in the APCDM model is the air occupancy model (AOM) that has been developed by Sherali et al. (2000). This subroutine identifies the three-dimensional, nonconvex sectors traversed by each surrogate flight plan, along with the respective sector occupancy time intervals. Using AOM, we assess the average aircraft-occupancy times within each sector along with the maximum monitoring load. Accordingly, we devise a suitable (nonlinear) workload function to ascribe a cost associated with safely monitoring and directing the composed mix of flight plans within each sector.

The second subroutine also contributes toward sector workload issues while principally addressing safety considerations. This subroutine is the

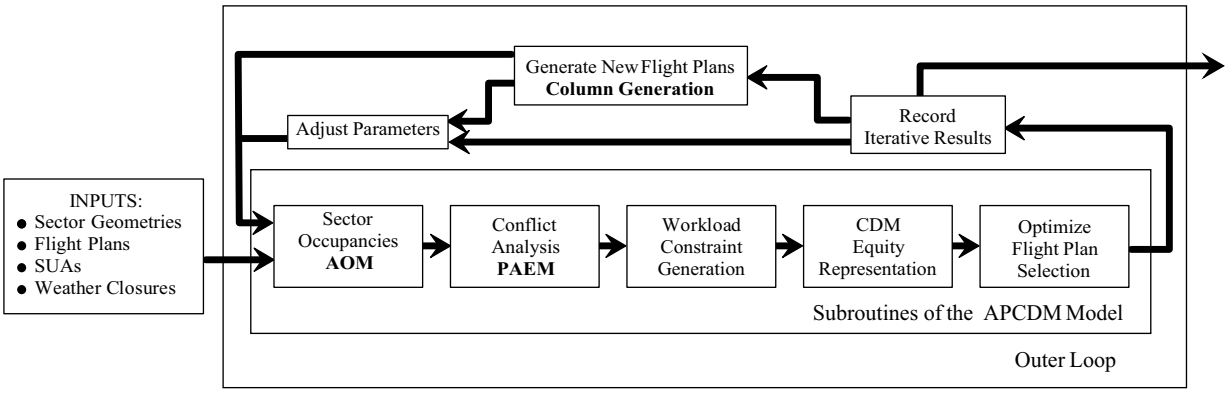

Figure 1 Conceptual Extended Operational Framework for the Proposed Model

Probabilistic Aircraft Encounter Model (PAEM) that examines conflict risk situations in the en route airspace. In contrast with the deterministic conflict analysis conducted in the APM model of Sherali et al. (2002), we consider random errors for each aircraft about the planned or filed trajectory, recognizing that aircraft are subject to pilot, navigation, and wind-induced errors. We develop two alternative representations of maximal displacement errors at each waypoint and use them to describe a set of probabilistic trajectory realizations. By conducting a conflict analysis between pairs of realizations, PAEM identifies duration intervals over which pairs of flight plans conflict with at least some specified threshold probabilities. These stochastic conflict intervals are then used to generate a set of constraints that restrict the conflict resolution workload imposed on each sector to be within its respective handling capacity. In contrast to the related analysis conducted in APM, which was based on a discretization of the time horizon, we employ a continuous time formulation and prove that this representation adds significant flexibility in representing the conflict resolution capacity of the different air traffic control sectors. In addition, we conduct a polyhedral analysis of the developed conflict resolution constraints to derive several classes of valid inequalities that yield provably stronger linear programming relaxations. The combined analyses conducted by AOM and PAEM therefore produces a set of sector workload constraints, as represented by the third box in the set of APCDM subroutines in Figure 1.

The fourth subroutine addresses the issue of equity among airline carriers in absorbing the costs related to rerouting, delays, and cancellations within the context of a *collaborative decision-making* (*CDM*) environment. We introduce a concept of collaboration efficiency for each airline based on the fuel and delay costs incurred in the overall solution prescribed by the model versus the corresponding costs related to individually optimized decisions. This concept is used to develop an equity formulation that assesses the distribution of collaboration efficiencies achieved by the different airlines. Several alternative strategies for representing equity are also discussed.

The foregoing sector workload (including collision safety) and equity-based constraints developed are integrated within a mixed-integer programming formulation that seeks to prescribe a mix of flight plans that minimize fuel, delay, and workload-related costs, along with inefficiency and inequity measures. This constitutes the principal APCDM model discussed in this paper. Further analysis can then be conducted as in the outer loop of Figure 1 to readjust key parameters and to generate new flight plans that are superior (in terms of achieving lower objective function values) to the existing surrogates, based on the model outputs. This model generation and execution can then be reiterated as desired. (Our focus in this paper, however, resides with the basic APCDM model within this framework.) The APCDM model has potential use mainly for strategic purposes such as FAA policy evaluations (e.g., with respect to aircraft separation standards, sector workload restrictions, and air traffic control sectorization strategies), Homeland Defense contingency planning, and military air campaign planning. However, conceivably, it could also be used for tactical applications, such as air traffic control in response to diversions and delays prompted by severe weather phenomena or the cordoning of airspace because of spacecraft launches (see Sherali et al. 2000, 2002). The model can also serve a useful role in augmenting the FAA's *National Severe Weather Playbook* of standardized flight profiles in different disruption-prone regions of the national airspace (FAA 2003).

Several other papers in the literature relate to different aspects of the proposed APCDM model. For example, Reich (1966) examines aircraft-separation standards in the along-track, cross-track, and vertical axes to safeguard against the risk of collisions. Similar to our approach, these separation restrictions are modeled by using a geometric box, called a *proximity shell*, aligned with the so-called *focal* aircraft's trajectory direction, and constructed about its instantaneous position. A penetration of an *intruder* aircraft into this proximity shell is identified as a conflict that would need to be addressed or resolved by an air traffic controller (ATC). Reich also stressed the importance of the geometry of intrusions into the proximity shell of the focal aircraft. Today,

FAA standards require that aircraft maintain a five nautical-mile horizontal separation and a vertical separation of 2,000 feet for aircraft flying above 29,000 feet (and 1,000 feet below this altitude).

Paielli and Erzberger (1997) noted that during flight, the prediction of aircraft trajectories is not an exact science. For example, wind modeling is still quite imprecise, aircraft navigation and control equipment are not error-free, and finally there is the human element—the pilot. They found that errors in the cross-track axis were relatively constant in a 30-minute period, as they are routinely corrected by the pilot or the aircraft's flight-management system. The along-track errors were found to be significantly more variable because they are created primarily by headwinds and tailwinds, which are not well predicted or easily corrected during flight.

Bakker et al. (2000) experimented with both geometric and probabilistic approaches to conflict prediction. In contrast to previous work, they examined trajectory errors in three dimensions simultaneously, rather than assigning independent error distributions for each axis. They concluded, as previous researchers have, that a probabilistic approach is a necessary part of modeling collision prediction. The Center-TRACON Automation System (CTAS) (Erzberger et al. 1997) also contains a conflict prediction module called a *conflict probe*, which additionally computes the flight corrections necessary to bring the probability of a separation violation below a programmed threshold. CTAS is designed to be a real-time tool to aid air traffic managers. Paielli and Erzberger (1999) discuss its testing and validation at two air route traffic control centers (ARTCC). Our model also considers a three-dimensional probabilistic analysis, but unlike Paielli and Erzberger (1997), who were mainly interested in identifying conflicts that occur with a specified threshold probability, we also develop duration intervals over which this conflict risk persists as well as identify the sectors occupied during such conflict events.

Bertsimas and Stock Patterson (1998) develop a model to control the flow of aircraft by either adjusting their release times into the NAS (ground holding times) or their en route speed, with the objective of minimizing the total ground and airborne-delay costs. Their model considers arrival and departure capabilities at airports and simultaneous occupancy capacities of sectors. However, unlike our model, no other sector workload metrics are incorporated. Also, the flight plan trajectories are assumed to be fixed and conflict-free, and no equity considerations are addressed. CDM, in general, is essentially a business practice that advocates decentralized cooperative decision making between the various participants in a common endeavor. CDM takes advantage of many of the "total quality management" precepts (Turner et al. 1993) that were advanced and popularized by W. Edwards Deming in Japan during the 1970s and 1980s, and later in the United States during the early 1990s. The advent of the "free-flight" paradigm further complicates air traffic control in an increasingly crowded airspace (e.g., see Sherali et al. 2000 and the references cited therein). In an attempt to actively involve all the users and service providers in the decision-making process, the FAA is sponsoring a comprehensive CDM effort to improve air traffic management, increase capacity, and reduce costs (see Ball et al. 1998, 2000; Carlson 2000; and Wambsganss 1996 for a discussion on related issues and benefit assessments).

The implementation of the Ground Delay Program Enhancement (GDPE), the first major thrust of CDM (Metron Aviation 2002), is an impressive example of CDM success. The GDPE effort implements three central tenets of CDM: common situational awareness; a distributed planning environment where the roles and responsibilities of the players are agreed on and well understood, limited resources are rationed equitably, and all participants have a voice in the decision process; and a performance measurement and feedback process to facilitate system improvements. In particular, GDPE procedures reduce disincentives for airlines to report flight schedule changes, particularly delays and cancellations, by utilizing a ration-by-schedule (RBS) procedure along with a compression algorithm. Ball et al. (1998, 2000) and Chang et al. (2001) report that this algorithm is a "win-win" situation for the carriers, in that compression reduces delays while, at worst, preserving the status quo for any given carrier. Over a two-year period, the compression algorithm alone has reduced delays by 4.5 million minutes versus the previous ground delay procedures.

A general discussion of the principles of equity in the context of allocating limited resources to activities is given in Luss (1999) (also, see Young 1994). Examining performance factors for each activity to reflect, for example, the weighted shortfall of allocated resources from a specified target or goal, Luss defines an equitable allocation to be one in which no performance function can be feasibly improved without degrading another activity's performance value that is greater than or equal to this one. Accordingly, Luss shows that a lexicographical minimax solution (i.e., a lexicographically smallest vector of performance values that are sorted in nonincreasing order) yields an equitable solution. In particular, in the context of the RBS and compression procedures implemented by the FAA, Vossen and Ball (2001) formalize the allocation of arrival slots to airlines and the exchange of slots between airlines via assignment optimization models having appropriately defined lexicographical minimax objective functions. For example, they show that the RBS procedure essentially lexicographically minimizes the vector of the number of flights having decreasing levels of delay. Vossen and Ball also describe how interairline slot exchanges can be viewed as a bartering process in which the FAA behaves as a mediator broker. It should be noted that while these RBS and compression algorithm-based equity concepts have been implemented and largely accepted by the airlines, the equity constructs proposed in the present paper are new and untested ideas that need to be studied and validated in detail to promote acceptance by the participating players in this process.

The remainder of this paper is organized as follows. In §2, we present the PAEM constructs. Based on this development, we address the formulation of various conflict-resolution and associated workload constraints in §3. We derive enhanced formulations for these restrictions and prove that they progressively provide tighter linear programming relaxations. Next, in §4, we provide a further characterization of airspace sector monitoring workloads and use this characterization to develop a set of workload constraints. The model's equity representation is proposed in §5, where we describe the concepts of collaboration efficiency and collaboration equity. These are subsequently incorporated into the model via suitable functions and penalty terms. The overall mixed-integer APCDM model is then summarized in §6. Finally, in §7, we address the proposed model's potential contributions toward improving NAS operations and conclude by recommending several avenues for future research in this rich area of great contemporary public interest and national importance.

2. Probabilistic Aircraft Encounter Model (PAEM)

In this section, we establish a probabilistic framework for conducting a conflict analysis between some focal aircraft A and an intruder aircraft B. The proposed Probabilistic Aircraft Encounter Model (PAEM) ascertains whether an intruder aircraft can possibly penetrate the protective standard separation shell (considered to be a rectangular box) enveloping some focal aircraft, assuming that both the trajectories of the focal and intruder aircraft are subject to either randomized or wind-induced probabilistic errors. The PAEM module determines the duration for which a penetration into the standard separation shell (called a Level-1 conflict) or a penetration into a shell of half this dimension (called a Level-2 conflict) occurs. In addition, if the intrusion penetrates a tight enveloping box having a half-dimension of 500 feet in the in-trail and cross-track dimensions, and 100 feet in the altitude dimension, then a *fatal conflict* is declared, and such pairs of flight plans are necessarily prohibited from coexisting.

Consider a three-dimensional, Cartesian X-space system having its origin at the center of the earth, and suppose that in this system, the nominal trajectory of any aircraft A is described by a piecewise linear trajectory corresponding to the finite sequence $\{X_A^1, X_A^2, \ldots, X_A^p\}$ of waypoints (where p depends on A), and that the actual path can be defined by a sequence of displaced waypoints given by $\{X_{Ak}^1, X_{Ak}^2, \ldots, X_{Ak}^p\}$ for some realization k, $k = 1, \ldots, n_A$, having an associated probability p_{Ak}, where $\sum_{k=1}^{n_A} p_{Ak} = 1$. Note that because an aircraft must take off and land at the specified airports, we assume that the initial and final positions on the trajectory are

error-free, i.e., $\{X_A^q\} = \{X_{Ak}^q\}$ for $q = 1$ and $q = p$, $\forall k = 1, \ldots, n_A$.

Now, consider any waypoint $q \in \{2, \ldots, p-1\}$, and let us examine the spatial segment $[X_A^{q-1}, X_A^q]$ leading to this waypoint. To describe the assumed probabilistic deviation of the aircraft from the nominal position X_A^q, consider the following Y-space definition with respect to the nominal trajectory segment $[X_A^{q-1}, X_A^q]$. Let the Y_1-axis correspond to the in-trail direction $d_A^q = X_A^q - X_A^{q-1}$; let the Y_3-axis be orthogonal to the Y_1-axis and lie in the plane spanned by d_A^q and the position vector X_A^q (or X_A^{q-1}) emanating from the origin such that it makes an acute angle with the latter position vector; and let the Y_2-axis be orthogonal to the (Y_1, Y_3) plane, with the positive Y_2-axis being oriented along the left wing of the aircraft. (Note that we assume that the Y_1-axis and the position vector X_A^q (or X_A^{q-1}) are not collinear, i.e., the aircraft is not moving vertically with respect to the earth.) Hence, as derived in Sherali et al. (2000), there exists an orthonormal matrix Q_A^q such that the transformation of vectors (with a common origin) from the X-space to the Y-space, and vice versa, occurs according to

$$X = Q_A^q Y, \quad \text{i.e.,} \quad Y = \left[Q_A^q\right]^T X. \tag{1}$$

Accordingly, to describe a general form of the trajectory realizations, consider any waypoint $q \in \{2, \ldots, p-1\}$. Let us define a *maximal displacement region* as a smallest possible bounding region centered at X_A^q that is assumed to contain all the displaced waypoint position realizations X_{Ak}^q, $k = 1, \ldots, n_A$. Figure 2 displays two shapes of such a maximal displacement region that we consider, namely, a rectangular and a cylindrical region (detailed derivations for each of these regions are given in §§2.1 and 2.2, respectively). Furthermore, suppose that there exists some spatial three-dimensional probability density function (PDF) defined over this maximal displacement region that produces the probabilistically displaced waypoint realizations X_{Ak}^q. This region is centered at X_A^q, and the particular (upright) orientation displayed in Figure 2 is assumed to coincide with the transformed Y-space defined by (1), with its horizontal cross-section lying in the (Y_1, Y_2) plane. For the sake of simplicity, let us examine a discretization of this region into n_A suitable subregions, as illustrated in Figure 2, for the rectangular or cylindrical bounding region. Taking ξ_{Ak}^q to be the centroid of the kth such subregion in the Y-space relative to the origin translated to X_A^q, for $k = 1, \ldots, n_A$, we ascribe a probability p_{Ak} to this realization ξ_{Ak}^q by integrating the three-dimensional PDF over the corresponding subregion. Now, assume that each of the waypoints $q = 2, \ldots, p-1$ has an identically distributed waypoint displacement PDF, and that the corresponding displaced waypoints X_{Ak}^q for $q = 2, \ldots, p-1$ jointly produce a kth piecewise linear trajectory realization (see Figure 3). Hence, p_{Ak} is the probability associated with this complete kth trajectory realization, and by the nature of this conceptualization $\{\xi_{Ak}^q\}_{k=1}^{n_A}$ is independent of q, $\forall q = 2, \ldots, p-1$. However, to accommodate different levels of control or navigation accuracy with respect to the different waypoints as discussed below, we scale the maximal displacement region centered at the qth waypoint by a factor β_q, where $0 \le \beta_q \le 1, \forall q = 2, \ldots, p-1$.

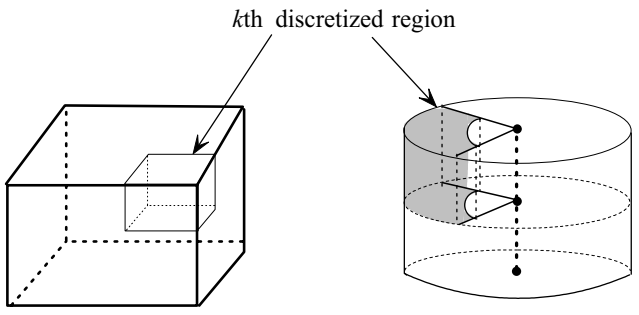

Figure 2 Discretization of Maximal Displacement Regions

Hence, noting that $Q_A^q \xi_{Ak}^q$ denotes the transformation of ξ_{Ak}^q into the original X-variable space based on (1), the displaced waypoint realizations that effectively define the trajectory realizations in the X-space are defined as

$$X_{Ak}^q = X_A^q + \beta_q Q_A^q \xi_{Ak}^q, \quad k = 1, \ldots, n_A, \ q = 2, \ldots, p-1. \tag{2}$$

The limiting displacement region defined in this manner moves with the aircraft (shrinking or expanding in dimension as it traverses the different piecewise linear airway segments depending on the β_q-parameters in (2)) being centered on the nominal

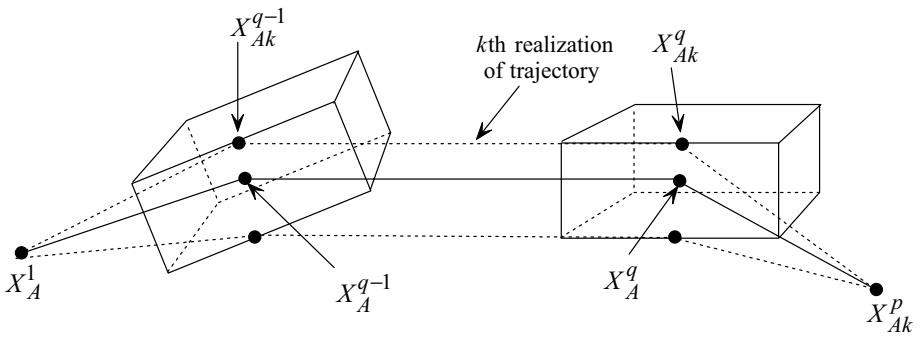

Figure 3 Three-Dimensional Rectangular Dispersion of Trajectories

flight plan trajectory, thus describing a corridor that contains any realized trajectory of the aircraft (see Figure 3). The factor β_q is a function of the air traffic intensity in the airspace surrounding the qth waypoint under consideration. The motivation for this is that in high traffic density areas, pilots might navigate more accurately, and positional errors might be more likely communicated to pilots by air traffic controllers. This, in effect, reduces the width of the corridor. Note that in practice, an aircraft is not controlled along-track to conform with the filed flight plan, except to ensure an adequate maintenance of miles-in-trail separation between strings of aircraft flying on similar airway trajectories. Although cross-track checks with respect to filed flight plans are also not routinely done, air traffic controllers will usually probe the pilot's intent and possibly force a corrective action if the aircraft deviates approximately five nautical miles from a filed flight plan (Downer 2001). Notwithstanding this practice of navigational control, note that the foregoing discussion simply characterizes a set of probabilistic realizations of actual trajectories based on a filed flight plan and maximal assumed deviations at each waypoint.

We now proceed to present two particular cases of the foregoing maximal displacement regions, along with the associated PDFs that describe the discretized trajectories and their corresponding probabilities.

2.1. Rectangular Displacement Region: Random Errors

As a first case, suppose that each displaced waypoint is constrained to be within some rectangular maximal displacement region centered on the respective nominal waypoint as shown in Figure 3. We denote $r_{\max}$ as the maximum in-trail displacement, $c_{\max}$ as the maximum cross-track displacement, and $v_{\max}$ as the maximum altitude displacement.

Suppose further that the displacements from the nominal trajectory in each of the three Y-space axes are mutually independent. Hence, the joint three-dimensional PDF that describes the deviation of each of the waypoints $q = 2, \ldots, p-1$ from their nominal positions is simply the product of these three univariate PDFs. Specifically, let us assume that the displacement in the in-trail direction follows a triangular distribution as shown in Figure 4, with $P[|r| \geq r_{\max}] = 0$. The associated PDF, $f_1(r)$, $-r_{\max} \leq r \leq r_{\max}$, is given by

$$f_1(r) = \frac{1}{r_{\max}} - \frac{|r|}{r_{\max}^2} \quad \text{for} \quad -r_{\max} \leq r \leq r_{\max}. \qquad (3)$$

To determine a set of discrete realizations of the three-dimensional PDF, we begin by dividing the in-trail dimension into n_1 segments according to the discretization $-r_{\max} \equiv r_0 < r_1 < \cdots < r_{n_1} \equiv r_{\max}$. Integrating (3), we compute the probability of the

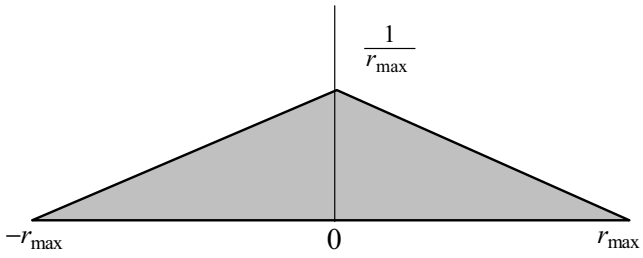

Figure 4 In-Trail Trajectory Error Distance PDF

displacement distance r lying in the interval $r_{a-1} \le r \le r_a$, for $a = 1, \ldots, n_1$, as follows:

$$P[r_{a-1} \le r \le r_a] = \begin{cases} \dfrac{r_a - r_{a-1}}{r_{\max}} - \dfrac{|r_a^2 - r_{a-1}^2|}{2r_{\max}^2} & \text{if } 0 \notin (r_{a-1}, r_a) \\ \dfrac{r_a - r_{a-1}}{r_{\max}} - \dfrac{(r_a^2 + r_{a-1}^2)}{2r_{\max}^2} & \text{if } 0 \in (r_{a-1}, r_a). \end{cases} \quad (4)$$

Next, we define c as the cross-track displacement and let this random variable have some PDF $f_2(c)$ for $-c_{\max} \le c \le c_{\max}$, with $P[|c| \ge c_{\max}] = 0$. Likewise, we define v as the vertical displacement and let this random variable have some PDF $f_3(v)$ for $-v_{\max} \le v \le v_{\max}$, with $P[|v| \ge v_{\max}] = 0$. Paielli and Erzberger (1997) noted that both cross-track and vertical errors are approximately constant with respect to time. Empirical data indicates that vertical errors rarely exceed ±200 feet while cross-track errors lie generally within ±0.5 nm (for aircraft equipped with a flight management system). As errors in the cross-track and vertical dimensions are small relative to in-trail errors, we shall assume for simplicity a uniform error distribution for the region respectively, spanned by $-c_{\max} \le c \le c_{\max}$ and $-v_{\max} \le v \le v_{\max}$, with $c_{\max} =$ 0.5 nm and $v_{\max} = 200$ feet.

For the sake of a discrete approximation, we partition the cross-track and vertical dimensions into n_2 and n_3 segments, respectively, given by the discretizations (assumed equispaced for simplicity) $-c_{\max} \equiv c_0 < c_1 < \cdots < c_{n_2} \equiv c_{\max}$ and $-v_{\max} \equiv v_0 < v_1 < \cdots < v_{n_3} \equiv v_{\max}$. Hence, the rectangular maximal displacement region is partitioned into $n_A = n_1 n_2 n_3$ subhyper rectangles. For any $q \in \{2, \ldots, p-1\}$, the kth such hyperrectangle that is characterized by, say,

$$r_{a_1-1} \le r \le r_{a_1}, \quad c_{a_2-1} \le c \le c_{a_2}, \quad \text{and} \quad v_{a_3-1} \le v \le v_{a_3}, \quad \text{for some } (a_1, a_2, a_3) \in \{1, \ldots, n_1\} \times \{1, \ldots, n_2\} \times \{1, \ldots, n_3\}, \quad (5)$$

the corresponding centroid ξ_{Ak}^q for use in (2) is taken as the probability mass centroid given by (noting that $f_2(\cdot)$ and $f_3(\cdot)$ are uniform distributions),

$$\xi_{Ak}^q = \left[\frac{\int_{r_{a_1-1}}^{r_{a_1}} r f_1(r)\, dr}{\int_{r_{a_1-1}}^{r_{a_1}} f_1(r)\, dr}, \frac{c_{a_2-1} + c_{a_2}}{2}, \frac{v_{a_3-1} + v_{a_3}}{2} \right]^T. \quad (6)$$

The first component in (6) is given by, noting (3) and (4),

$$\frac{(r_{a_1}^2 - r_{a_1-1}^2)/2r_{\max} - |r_{a_1}^3 - r_{a_1-1}^3|/3r_{\max}^2}{(r_{a_1} - r_{a_1-1})/r_{\max} - |r_{a_1}^2 - r_{a_1-1}^2|/2r_{\max}^2} \quad \text{if } 0 \notin (r_{a_1-1}, r_{a_1}),$$

and

$$\frac{(r_{a_1}^2 - r_{a_1-1}^2)/2r_{\max} - (r_{a_1}^3 + r_{a_1-1}^3)/3r_{\max}^2}{(r_{a_1} - r_{a_1-1})/r_{\max} - (r_{a_1}^2 + r_{a_1-1}^2)/2r_{\max}^2} \quad \text{if } 0 \in (r_{a_1-1}, r_{a_1}). \quad (7)$$

The corresponding probability p_{Ak} associated with this kth realization is given by (noting (4))

$$p_{Ak} = \frac{1}{n_2 n_3} P[r_{a_1-1} \le r \le r_{a_1}]. \quad (8)$$

2.2. Cylindrical Displacement Region: Wind-Induced Errors

As a second alternative that is motivated by wind-induced errors, we examine a maximal displacement region that is given by a cylinder, as shown in Figure 2, having a radius of $r_{\max}$ and with its height extending to $h_{\max}$ in each of the positive and negative Y_3-axis directions. For any waypoint $q \in \{2, \ldots, p-1\}$, we generate the possible realizations ξ_{Ak}^q, $k = 1, \ldots, n_A$, for use in (2), along with the associated probabilities p_{Ak}, $k = 1, \ldots, n_A$, as follows.

Consider a wind direction in the X-space as given by the vector w. By (1), this vector translates to $[Q_A^q]^T w$ in the Y-space for waypoint q. Let the projection of this vector onto the (Y_1, Y_2) space (obtained simply by suppressing the third component of $[Q_A^q]^T w$ to zero) be given by $w^q \equiv (w_1^q, w_2^q, 0)^T$. This direction w^q makes an angle θ^q with the Y_1-axis (the direction of the flight path in the Y-space for the current segment), where

$$\theta^q = \text{sign}[w_2^q] \cdot \cos^{-1}\left(\frac{w_1^q}{\|w^q\|} \right)$$

and where

$$\text{sign}[w_2^q] = \begin{cases} 1 & \text{if } w_2^q \ge 0. \\ -1 & \text{if } w_2^q < 0. \end{cases} \quad (9)$$

Note that this conforms with the convention that $\cos^{-1}(\cdot) \in [0, \pi]$ and that positive angles are measured

counterclockwise with respect to the Y_1-axis and negative angles are measured clockwise. We assume that any consequent angular displacement in the Y-space that is induced by this projected wind vector lies in the cone spanned by $\theta^q \pm \Delta$ for some $0 \le \Delta \le \pi$ radians. Accordingly, we define the three-dimensional displacement PDF in the Y-space as described by the independent distributions $f_R^q(r)$ for the radial displacement r, $0 \le r \le r_{\max}$; $f_\Theta^q(\theta)$ for the angular displacement θ measured with respect to the Y_1-axis, $\theta^q - \Delta \le \theta \le \theta^q + \Delta$; and $f_H^q(h)$ for the altitude displacement h, $-h_{\max} \le 0 \le h_{\max}$. Note that for corresponding isomorphic discretized regions $k = 1, \ldots, n_A$ of the type depicted in Figure 2 based on such distributions for the different waypoints $q \in \{2, \ldots, p-1\}$, we have the same associated probability p_{Ak}, which therefore relates to the probability of realizing the corresponding trajectory k. (In a similar manner, we could consider different wind directions for different waypoints, if necessary, so long as the foregoing property holds true. Furthermore, in either of these scenarios, variations in wind intensities at the different waypoints q could also be absorbed within the factor β_q in (2).)

Now, similar to the previous rectangular error distribution case, we assume that the PDF $f_R^q(r)$ for radial displacements is described by a triangular distribution given by

$$f_R^q(r) = 2\left(\frac{1}{r_{\max}} - \frac{r}{r_{\max}^2}\right), \qquad \text{for } 0 \le r \le r_{\max}. \quad (10)$$

In addition, motivated by the foregoing discussion, we let $f_\Theta^q(\theta)$, $\theta^q - \Delta \le \theta \le \theta^q + \Delta$ and $f_H^q(h)$, $-h_{\max} \le h \le h_{\max}$ be described by uniform distributions. For obtaining discrete realizations ξ_{Ak}^q along with the associated probability p_{Ak} for $k = 1, \ldots, n_A$, given any $q \in \{2, \ldots, p-1\}$, we discretize the partial cylindrical region that is described by a positive probability support as follows. We segment the radius r into n_R partitions according to the discretization $0 \equiv r_0 < r_1 < \cdots < r_{n_R} \equiv r_{\max}$. Similarly, we partition θ over the cone spanned by $\theta^q \pm \Delta$ into n_Θ segments as given by the discretization $\theta^q - \Delta \equiv \theta_0 < \theta_1 < \cdots < \theta_{n_\Theta} \equiv \theta^q + \Delta$ and we partition the height h into n_H subintervals discretized by $-h_{\max} \equiv h_0 < h_1 < \cdots < h_{n_H} \equiv h_{\max}$. Again, for simplicity, we assume that the discretized values of θ and h are equispaced. Hence, each discretized subregion is of the type depicted in Figure 2, and there are $n_A = n_R n_\Theta n_H$ such subregions. For the kth discretized subregion that is characterized by, say,

$$r_{a_1-1} \le r \le r_{a_1}, \quad \theta_{a_2-1} \le \theta \le \theta_{a_2}, \quad \text{and} \quad h_{a_3-1} \le h \le h_{a_3}$$

for some

$$(a_1, a_2, a_3) \in \{1, \ldots, n_R\} \times \{1, \ldots, n_\Theta\} \times \{1, \ldots, n_H\}, \quad (11)$$

we compute the corresponding ξ_{Ak}^q for use in (2) based on the probability mass center of this region. Note that if (r_k, θ_k, h_k) denotes this mass center in the (r, θ, h)-space, we have

$$\xi_{Ak}^q = [r_k \cos\theta_k, r_k \sin\theta_k, h_k]^T. \quad (12)$$

Following a derivation similar to (7), we have

$$\begin{aligned} r_k &= \frac{\int_{r_{a_1-1}}^{r_{a_1}} r f_R^q(r)\, dr}{\int_{r_{a_1-1}}^{r_{a_1}} f_R^q(r)\, dr} \\ &= \frac{(r_{a_1}^2 - r_{a_1-1}^2)/2r_{\max} - (r_{a_1}^3 - r_{a_1-1}^3)/3r_{\max}^2}{(r_{a_1} - r_{a_1-1})/r_{\max} - (r_{a_1}^2 - r_{a_1-1}^2)/2r_{\max}^2}, \\ \theta_k &= \frac{\theta_{a_2-1} + \theta_{a_2}}{2}, \quad \text{and} \quad h_k = \frac{h_{a_3-1} + h_{a_3}}{2}. \end{aligned} \quad (13)$$

The corresponding probability p_{Ak} is given by

$$\begin{aligned} p_{Ak} &= \frac{1}{n_\Theta n_H} P(r_{a_1-1} \le r \le r_{a_1}) \\ &= \frac{1}{n_\Theta n_H}\left[\frac{2(r_{a_1} - r_{a_1-1})}{r_{\max}} - \frac{r_{a_1}^2 - r_{a_1-1}^2}{r_{\max}^2}\right]. \end{aligned} \quad (14)$$

2.3. Conflict Analysis

Now, suppose that the focal aircraft A is traversing (part of) its linear trajectory segment $q_1 - 1$ to q_1 during some time interval, say $[0, T]$, and that an intruder aircraft B is traversing (part of) its corresponding linear trajectory segment $q_2 - 1$ to q_2 during the same time interval. Consider any corresponding realizations $k_1 \in \{1, \ldots, n_A\}$ and $k_2 \in \{1, \ldots, n_B\}$ of trajectories for aircraft A and B, respectively. Let the corresponding linear segments of aircraft A and B over

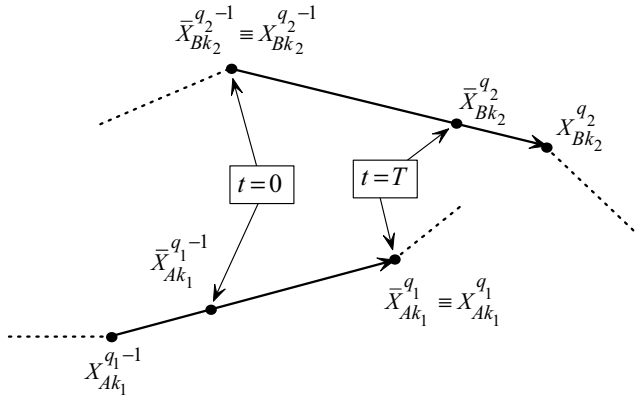

Figure 5 Trajectory Segments for Aircrafts A and B

the time interval $[0, T]$ be described by the following for the respective realizations k_1 and k_2:

$$X_A = \bar{X}_{Ak_1}^{q_1-1} + td_{Ak_1}^{q_1} \quad \text{and} \quad X_B = \bar{X}_{Bk_2}^{q_2-1} + td_{Bk_2}^{q_2}. \tag{15}$$

Observe that as depicted in Figure 5, each endpoint of the time interval $[0, T]$ would necessarily correspond to some breakpoint(s) in the piecewise linear trajectories traversed by aircraft A and B. Assume that the protective shell around the focal aircraft A is rectangular in dimension $2\delta_1 \times 2\delta_2 \times 2\delta_3$ with the aircraft centered in this shell, and with the rectangular shell having its dimensions oriented along the Y-axes according to the trajectory of aircraft A. Then, examining the transformed positions of aircraft A and B in the Y-space as given by $Y_A = [Q_A^{q_1}]^T X_A$ and $Y_B = [Q_A^{q_1}]^T X_B$ according to (1), we will have a penetration of B (taken as a point) into the protective shell of A whenever $-\delta \le Y_B - Y_A \le \delta$, where $\delta = (\delta_1, \delta_2, \delta_3)^T$. From (15), this occurs whenever

$$-\delta \le [Q_A^{q_1}]^T \left(\bar{X}_{Bk_2}^{q_2-1} - \bar{X}_{Ak_1}^{q_1-1}\right) + [Q_A^{q_1}]^T \left(d_{Bk_2}^{q_2} - d_{Ak_1}^{q_1}\right) t \le \delta. \tag{16}$$

Observe that in (16), the analysis is based on assuming the orientation of the protective shell being aligned with the nominal trajectory. A more precise analysis could consider this shell aligned with each particular trajectory realization k_1 of aircraft A. This would require a derivation of (16) based on a transformation that is specific to each trajectory realization, for each piecewise linear segment. While this is theoretically possible, it would be computationally prohibitive, and in our experience yields only a marginal improvement in the model's output accuracy. Also, note that a parallel analysis is conducted with B as the focal aircraft and A as the intruder aircraft, because a penetration of B into A's protective shell does not imply the converse and vice versa.

The system (16) is comprised of six linear inequalities and will either have no feasible solution for any $0 \le t \le T$, or else will yield a solution for some interval $[t_1, t_2] \subseteq [0, T]$. In the latter case, there exists a collision risk of severity level determined by δ during the interval $[t_1, t_2]$, with the corresponding associated probability being $p_{Ak_1} p_{Bk_2}$. Using (16) for all possible combinations $(k_1, k_2) \in \{1, \ldots, n_A\} \times \{1, \ldots, n_B\}$, we can easily glean this type of information for each pair of realizations for the present pair of linear trajectories. Repeating this over all such piecewise linear segment pairs, we can obtain a net partitioning of the time horizon into intervals where a collision risk due to B penetrating A's protective shell exists with some derived associated probability. For example, suppose that the only nonzero conflict probabilities of this type over some duration $[0, T]$ for a particular pair of piecewise linear segments are obtained as depicted over the time intervals illustrated in Figure 6a, where each interval corresponds to a conflict between some pair of realizations k_1 and k_2 having the displayed probability $p_{Ak_1} p_{Bk_2}$. By accumulating the independent probabilities over incremental time segments, we obtain the aggregate conflict probability shown in Figure 6b for each specific subinterval that is created by examining the union of the end-points of the various intervals depicted in Figure 6a.

Based on this analysis, for each sector during any interval of time, we can determine whether any occupying pair of aircraft pose a conflict risk, along with a specification of the corresponding time subintervals when penetration occurs and their associated probabilities. While conducting this conflict risk analysis, we can simultaneously assign these duration intervals of conflict risk (with the associated probabilities) to the various sectors traversed. Note that we ascribe this conflict risk to the sector of the focal aircraft using the sector occupancy information obtained via the model AOM (Sherali et al. 2000) based on its nominal flight path. (While it is possible to determine sector

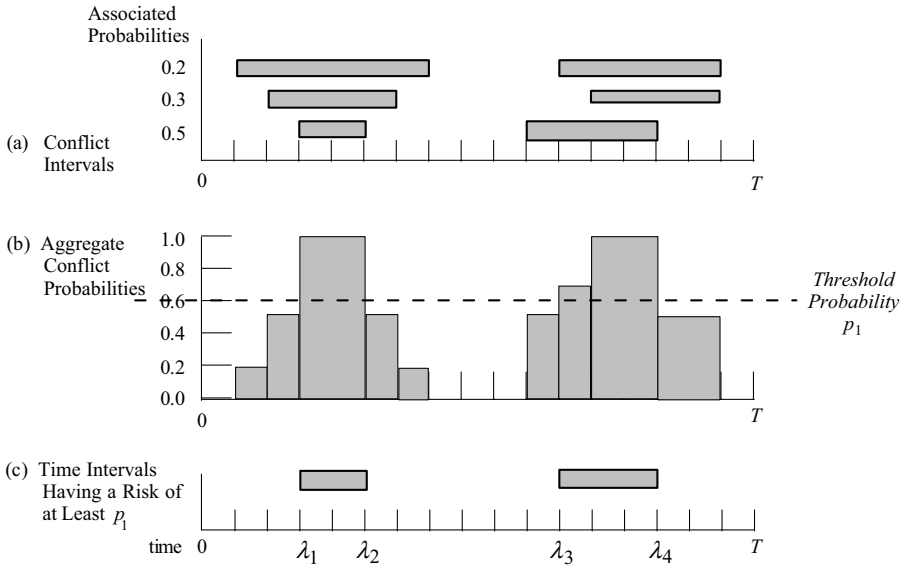

Figure 6 Probabilistic Conflict Aggregation Example

occupancies for each trajectory realization in a likewise fashion, this would impose a prohibitive computational burden.) *Therefore the PAEM produces, for each sector, a set of duration intervals for which pairs of aircraft flight plans have an identified probability of collision risk.*

Let us now define a threshold probability parameter p_1 with respect to the standard separation box of dimension δ, such that we will consider collision risk only if its probability is at least this value. Figure 6c depicts the time intervals during which the conflict risk is at least $p_1 = 0.6$. Observe that PAEM permits, via (16), a simultaneous analysis of conflicts having varying degrees of severity as determined by the dimension δ of the protective shell box. Accordingly, suppose that we consider a conflict based on the standard separation criteria, δ, as having Level-1 severity and consider conflicts based on a box of dimension $\delta' = \delta/2$ having a Level-2 severity. Correspondingly, we specify a pair of threshold probabilities p_1 and p_2, with $p_2 < p_1$, and record conflict intervals that are given by the union of (a) conflict intervals of Level-1 severity that occur with a probability of at least p_1 and (b) conflict intervals of Level-2 severity that occur with a probability of at least p_2. This information generated by the two-tuple (p_1, p_2) will be used to derive our conflict resolution constraints as described in the next section. (Part II of this paper establishes baseline values $(p_1, p_2) = (1/3, 1/6)$ and then studies the empirical behavior of the model to variations in these threshold probabilities.) Additionally, we also consider an inviolable airspace around each aircraft as described by a box having dimension $(500', 500', 100')^T$ and label any penetration of an intruder within this tightly enveloping airspace of a focal aircraft with a probability exceeding some small threshold p_3 as a *fatal conflict*. (Part II of this paper assigns $p_3 = 1/18$ as the baseline value.) Such fatal conflicts will be explicitly precluded from occurring by our model constraints. The other nonfatal conflicts identified will be considered as *resolvable*.

3. Conflict Resolution Workload Formulation

In this section, we begin to construct our model APCDM based on the preprocessing conducted by the modules AOM (Sherali et al. 2000) and PAEM (§2). These modules respectively determine occupancy intervals within each sector for each flight trajectory, and the intervals for each pair of aircraft flight plans where a conflict risk occurs with at least some threshold probability p_1 at Level-1 severity or with probability $p_2 < p_1$ at Level-2 severity.

To begin, let us introduce some preliminary notation. Let $[0, H]$ denote the planning horizon time interval, and suppose that we are given a set of flights $f = 1, \ldots, F$ that are relevant to a certain region of airspace under present consideration, and whose duration overlaps this horizon. For each flight f, let P_f denote the set of alternatives, or surrogates, submitted by the respective airlines that differ in departure and anticipated arrival times, along with time-space trajectories and cruising altitudes, while traversing between the designated origin and destination for flight f. (We assume $|P_f|$ is manageable in size, $\forall f$, and is typically of order 4–6, as used in the computations given in Part II of this paper.) We shall also let plan $p = 0$ designate the event of canceling any given flight f and, accordingly, denote $P_{f0} = P_f \cup \{0\}$. (Note that if the cancellation of a particular flight is not an alternative that is possible (e.g., an airborne flight) or that is tendered by the corresponding airline for CDM consideration, then its associated cost is assigned a prohibitively high value.) Hence, our principal (binary) decision variables are defined as:

$$x_{fp} = \begin{cases} 1 & \text{if plan } p \in P_{f0} \text{ is selected for flight } f, \\ 0 & \text{otherwise,} \end{cases}$$
$$\forall f = 1, \ldots, F, \ p \in P_{f0} \quad (17)$$

where

$$\sum_{p \in P_{f0}} x_{fp} = 1, \qquad \forall f = 1, \ldots, F. \quad (18)$$

We shall also find it convenient to index the entire collection of flight plans contiguously, and we will accordingly use upper-case letters to designate these indices. Hence, for example, in this notation, x_P would represent the x-variable for some flight plan of a certain flight. Typically, we will use this notation whenever we are not particularly interested in identifying the flight plan P with a certain flight. Also, this contiguous indexing will facilitate a referral to ordered pairs of flight plans P and Q with $P < Q$.

Now, to formulate a suitable set of conflict resolution constraints, let us introduce the concept of an overall *conflict graph* $G(N, A)$. This graph has a node set N comprised of all the flight plans P under consideration and a set of (undirected) arcs (or edges) A that record all pairs of resolvable conflicts identified by PAEM, i.e., $A \equiv \{(P, Q)$: $P < Q$, and the flight plans P and Q yield a resolvable conflict over some duration within $[0, H]\}$. Note that as alluded to earlier, denoting FC as the set of flight plan pairs (P, Q), $P < Q$, that pose a fatal conflict risk, we explicitly prohibit these conflicts by imposing the specific restrictions

$$x_P + x_Q \le 1, \qquad \forall (P, Q) \in FC. \quad (19)$$

To facilitate the formulation of our conflict resolution restrictions, we additionally define a product variable $z_{PQ} = x_P x_Q$, $\forall (P, Q) \in A$, that takes on a value of one if and only if $x_P = x_Q = 1$. This relationship will be enforced via the linear restrictions

$$z_{PQ} \ge x_P + x_Q - 1, \qquad z_{PQ} \ge 0, \quad (20)$$

where an objective penalty term $\sum_{(P,Q)\in A} \phi_{PQ} z_{PQ}$, with $\phi_{PQ} > 0$, $\forall (P, Q) \in A$, will ensure in combination with (20) that $z_{PQ} = 1$ if $x_P = x_Q = 1$, and $z_{PQ} = 0$ otherwise. Note that the parameter ϕ_{PQ} would depend on the geometry of the conflict itself, and the intensity of the required conflict resolution actions. (PAEM also generates this conflict geometry information.) Part II of this paper prescribes a nominal value for ϕ_{PQ} based on an assumed duration and intensity of the conflict.

Suppose that while running PAEM we encounter a conflict (with respect to the threshold probabilities p_1 and p_2) between a pair of flight plans P and Q, $P < Q$, over some duration $[t_1, t_2]$, such that this conflict is to be resolved in sector $s \in \{1, \ldots, S\}$. We record this interval along with the conflicting pair (P, Q) in a *conflict Gantt chart* for sector s, denoted CGC_s. Note that if P and Q lie in different sectors over the duration $[t_1, t_2]$, the resolution of this conflict is assigned to the sector that contains the focal aircraft at time t_1. In the event that the focal aircraft lies at the boundary of more than one sector at time t_1, the conflict is arbitrarily assigned to the smallest-indexed sector among these sectors.

In addition, depending on the sector s and the severity and characteristics of the identified conflict, we define a *prep-buffer* duration $b^s_{PQ} > 0$ that serves to represent a preparatory duration required by the air traffic controller in sector s to address the imminent conflict between flight plans P and Q. We augment

the interval $[t_1, t_2]$ represented in the CGC_s by adding the prep-buffer, i.e., by replacing t_1 with $t_1 - b^s_{PQ}$.

Having constructed the CGC_s for each sector s, we now impose a conflict-resolution workload constraint, which requires that the

$$\left\{\begin{array}{l}\text{number of simultaneous (using}\\ \text{augmented durations) conflicts to}\\ \text{be resolved in } s \text{ at any time}\end{array}\right\} \le r_s, \qquad (21)$$

where $r_s \ge 1$ is some designated workload parameter for sector s, $s = 1, \ldots, S$. Using the algorithm prescribed by Sherali and Brown (1994), suppose that we identify the entire collection of maximal overlapping sets M_{sk}, $k = 1, \ldots, K_s$, for CGC_s, where each M_{sk} is comprised of pairs of conflicting flight plans (P, Q), $P < Q$, in CGC_s that overlap at some time and is maximal in the sense that it is not a strict subset of any other set of simultaneously occurring pairs of conflicting flight plans. Figure 7 illustrates a hypothetical CGC_s and its associated maximal overlapping sets M_{sk}, $k = 1, \ldots, K_s$. Observe that, as in the case of the pair (A, B) in Figure 7, disconnected conflict intervals for any pair of flight plans are treated as separate conflicts in this definition.

Accordingly, we can formulate (21) by imposing

$$\sum_{(P,Q)\in M_{sk}} z_{PQ} \le r_s, \qquad \forall k = 1, \ldots, K_s, \ \forall s = 1, \ldots, S. \qquad (22)$$

Observe that (22) is a valid representation of (21) because any set of overlapping conflicts must be a subset of some M_{sk}, $k \in \{1, \ldots, K_s\}$, where the sets M_{sk} themselves contain overlapping conflicts. Let us refer to the collection of constraints (22) for each $s = 1, \ldots, S$ as *M_s-inequalities*.

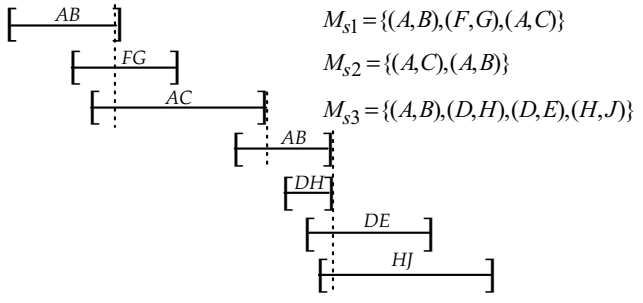

Figure 7 Conflict Gantt Chart CGC_s and Its Associated Maximal Overlapping Sets

In the previous formulation of conflict resolution workload constraints proposed by Sherali et al. (2002) given any sector s, the horizon was discretized into slots $t = 1, \ldots, T_s$, each of duration t_s, and accordingly, a set A_{st} was used to record the pairs of flight plans (P, Q), $P < Q$, which yield a conflict that (partially) overlaps the duration of slot t in CGC_s (without any prep-buffers). Similar to (22), the corresponding conflict resolution restrictions imposed (using the same generalized parameter $r_s \ge 1$) can be stated as

$$\sum_{(P,Q)\in A_{st}} z_{PQ} \le r_s, \qquad \forall t = 1, \ldots, T_s, \ \forall s = 1, \ldots, S. \qquad (23)$$

Let us refer to the collection of discretized constraints in (23) for each s as *D_s-inequalities*, for $s = 1, \ldots, S$. The following result relates the previous discretized formulation (without prep-buffers) to the presently proposed continuous formulation.

PROPOSITION 1. *Consider any sector $s \in \{1, \ldots, S\}$.*

(a) *Suppose that $b^s_{PQ} \equiv 0, \forall (P, Q) \in A$. Then for any value of $t_s > 0$, the satisfaction of the D_s-inequalities implies the satisfaction of the M_s-inequalities.*

(b) *Conversely, suppose that given a value of $t_s > 0$ for each corresponding slot $t = 1, \ldots, T_s$, we compute*

$$\tau_s(t) = \max\{0, t_2(t) - t_1(t)\}, \qquad (24)$$

where

$$t_1(t) = \min\Big\{t'' : [t', t''] \text{ is a conflict interval for some } (P, Q) \in A_{st}\Big\}, \qquad (25)$$

$$t_2(t) = \max\Big\{t' : [t', t''] \text{ is a conflict interval for some (possibly other) } (P, Q) \in A_{st}\Big\}. \qquad (26)$$

Then, $\tau_s(t) \le t_s$, $\forall t = 1, \ldots, T_s$, and moreover, for

$$b^s_{PQ} \ge \tau_s(t), \qquad \forall (P, Q) \in A_{st}, \quad \forall t = 1, \ldots, T_s, \qquad (27)$$

we have that the satisfaction of the corresponding M_s-inequalities implies the satisfaction of the D_s-inequalities.

PROOF. (a) Consider an arbitrary M_{sk} and its corresponding inequality in (22). Because all the conflicts between pairs $(P, Q) \in M_{sk}$ occur simultaneously at some point in time (with $b^s_{PQ} = 0, \forall (P, Q) \in M_{sk}$), there

exists some time slot t for which $A_{st} \supseteq M_{sk}$. Hence, we have that the corresponding D_s-inequality in (23) implies this M_s-inequality in (22).

(b) Conversely, given the hypothesis of part (b), consider any slot $t \in \{1, \ldots, T_s\}$ and the corresponding inequality (23) for the accompanying set A_{st}. If $t_2(t) \le t_1(t)$, then $\tau_s(t) \equiv 0 \le t_s$. Moreover, all the conflicts in A_{st} overlap at some point in time because if not, then some conflict in A_{st} ends strictly before another begins when we would have $\tau_s(t) > 0$. Consequently, there exists an M_{sk}, $k \in \{1, \ldots, K_s\}$, such that $M_{sk} \supseteq A_{st}$, so (22) implies (23).

On the other hand, if $t_2(t) > t_1(t)$, let $t_1(t)$ and $t_2(t)$ respectively correspond to some conflicting pairs of flight plans (G, H) and (U, V). Because the conflict (U, V) begins after (G, H) ends, and because they both overlap time slot t, we have that both these foregoing start and end events occur within time slot t, so $\tau_s(t) \equiv t_2(t) - t_1(t) \le t_s$. Moreover, examining the time $t_1(t)$, for any conflict $(P, Q) \in A_{st}$ that occurs over some duration $[t', t'']$ we have that (i) if $t' < t_1(t)$, then since $t'' \ge t_1(t)$, we get by (24) that

$$t' - \tau_s(t) \le t_1(t) \le t'' \tag{28}$$

and (ii) if $t' \ge t_1(t)$, then since $t' \le t_2(t)$, we again get that (28) holds true because $t' - \tau_s(t) = t' - t_2(t) + t_1(t) \le t_1(t) \le t' < t''$. However, (28) asserts that under (27) all the conflict pairs in A_{st} would overlap at the particular time $t_1(t)$ in CGC_s, and so there exists an M_{sk}, $k \in \{1, \ldots, K_s\}$ for which $M_{sk} \supseteq A_{st}$. Therefore, we again have that the corresponding (22) would imply (23). □

Observe from Proposition 1 that if we use a prep-buffer duration of zero, i.e., we do not augment the conflict intervals, then the D_s-inequalities impose a more stringent set of conflict resolution workload constraints (for any slot duration) than do the M_s-inequalities. In this case, the slot-based restriction constrains not only the simultaneously occurring conflicts, but also recognizes that nonoverlapping conflicts that might occur in relatively quick succession impose a stressful workload on the air traffic controller. Indeed, as Proposition 1 illustrates, as the prep-buffer begins to increase, the condition (21) related to the M_s-inequalities begins to accommodate the consideration of conflicts that occur in relatively quick succession within the workload formulation, and the M_s-inequalities all imply the D_s-inequalities once the prep-buffer becomes sufficiently large. Proposition 1 demonstrates that this occurs at or before a prep-buffer value of t_s, the slot duration for the D_s-inequalities.

Now, for each sector s, let $G_{sk}(N_{sk}, M_{sk})$ be the conflict graph that is constructed for the overlapping set $M_{sk}, k = 1, \ldots, K_s$, where N_{sk} is the set of nodes representing the kth overlapping set of flight plans traversing sector s during the time horizon and M_{sk} is the set of edges connecting the corresponding nodes in N_{sk}, representing simultaneously occurring conflicts between pairs of flight plans.

We now develop a series of progressively tighter representations for the conflict resolution constraints (21). To facilitate comparisons with the previous APM model of Sherali et al. (2002), let us assume for now that $r_s = 1$, $\forall s = 1, \ldots, S$. For this case, Sherali et al. proposed the following set of conflict resolution constraints in the x-space, denoted C_1:

$$C_1 = \left\{ x: \sum_{P \in S_k} x_P \le |S_k| - 1, \ \forall k \in K_{NR}, \ x \text{ binary} \right\}, \tag{29}$$

where, translating the development in Sherali et al. (2002) to the present context of conflict graphs G_{sk}, each S_k is a set of nodes in some conflict graph G_{sk} at which designated pairs of edges in M_{sk} are incident (whence $|S_k|$ equals three or four), and where K_{NR} records the collection of all such nonredundant constraints over the entire set of conflict graphs G_{sk}, $\forall (s, k)$.

Furthermore, Sherali et al. (2002) also proposed an alternative representation essentially embodied by (22) and (20), given by:

$$C_2 = \left\{ (x, z): \begin{array}{ll} \sum_{(P,Q) \in M_{sk}} z_{PQ} \le 1, & \forall (s, k) \\ z_{PQ} \ge x_P + x_Q - 1, & \forall (P, Q) \in A \\ z \ge 0, \ x \text{ binary} & \end{array} \right\}, \tag{30}$$

where recall that A is the arc set of the overall conflict graph. The experiments conducted by Sherali et al. (2002) reveal that C_1 is preferable for sparse conflict

situations, while C_2 is preferable for relatively more dense conflict graphs. More extensive computational tests with C_1 versus C_2 in the context of generalized vertex packing problems (see Sherali and Smith 2003) have revealed a significant empirical advantage of C_2 over C_1. We shall now augment C_2 to derive a provably stronger representation than that given by C_1. The following example motivates this derivation.

EXAMPLE 1. Consider a conflict graph having nodes P, Q, and R, and with edges (P, Q) and (P, R). The corresponding set of constraints in C_2 is given by

$$z_{PQ} + z_{PR} \le 1, \qquad z_{PQ} \ge x_P + x_Q - 1,$$
$$z_{PR} \ge x_P + x_R - 1, \quad z \ge 0, \ x \text{ binary}. \tag{31}$$

If we maximize $\{x_P + x_Q + x_R\}$ subject to the continuous relaxation of (31), we obtain the fractional extreme point solution $z_{PQ} = z_{PR} = 1/2$, $x_P = 1/2$, $x_Q = x_R = 1$. However, observe that the constraint $x_P + x_Q + x_R \le 2$, which would be inherent in the definition of C_1, deletes this fractional solution.

Motivated by this example (as well as by Proposition 2 below) let us first examine which constraints from C_1, at a minimum, need to be added to C_2 to obtain a provably tighter representation than C_1. The answer for this lies in the following set of "nonclique" triplets (P, Q, R):

$$T_{NC} = \left\{ \begin{array}{l} (P, Q, R), \ P < Q < R\text{: for some } (s, k), \\ \text{a subgraph of } G_{sk} \text{ that is induced by the} \\ \text{nodes } P, \ Q, \text{ and } R \text{ contains precisely} \\ \text{two edges, but no such subgraph for} \\ \text{any } (s, k) \text{ contains three edges} \end{array} \right\}. \tag{32}$$

Accordingly, let us define

$$C_3 = \left\{ \begin{array}{ll} (x, z)\text{:} \displaystyle\sum_{(P,Q)\in M_{sk}} z_{PQ} \le 1, & \forall (s, k) \quad (33a) \\ x_P + x_Q + x_R \le 2, & \forall (P, Q, R) \in T_{NC} \quad (33b) \\ z_{PQ} \ge x_P + x_Q - 1, & \forall (P, Q) \in A \quad (33c) \\ z \ge 0, \quad x \text{ binary} & \quad (33d) \end{array} \right\}.$$

Consider the following result, where for any set C_i, the set $\overline{C}_i$ denotes its continuous relaxation, obtained by replacing x binary with $0 \le x \le e$, where e will always denote a (conformable) vector of ones.

PROPOSITION 2. *C_3 yields a tighter representation of the conflict constraints than does C_1 in the sense that the constraints defining $\overline{C}_3$ imply those defining $\overline{C}_1$.*

PROOF. Consider any constraint in $\overline{C}_1$ where $|S_k| = 3$ with, say, $S_k = \{P, Q, R\}$ and $P < Q < R$. If $(P, Q, R) \in T_{NC}$, then the corresponding constraint is directly present in (33b). Otherwise, there exists some (s, k) for which G_{sk} contains a clique on the node set $\{P, Q, R\}$. Accordingly, (33c) contains the constraints $z_{PQ} \ge x_P + x_Q - 1$, $z_{QR} \ge x_Q + x_R - 1$, and $z_{PR} \ge x_P + x_R - 1$. Summing these three constraints and using $z_{PQ} + z_{QR} + z_{PR} \le 1$ from (33a), we get $1 \ge z_{PQ} + z_{QR} + z_{PR} \ge 2(x_P + x_Q + x_R) - 3$, which implies that $x_P + x_Q + x_R \le 2$. Hence again the corresponding constraint in $\overline{C}_1$ is implied in the continuous sense.

Now, consider a constraint in $\overline{C}_1$ for which $|S_k| = 4$ and is given by $x_P + x_Q + x_R + x_S \le 3$ based on edges (P, Q) and (R, S) in M_{sk} for some graph G_{sk}. Again, summing the constraints $z_{PQ} \ge x_P + x_Q - 1$ and $z_{RS} \ge x_R + x_S - 1$ from (33c), and using $z_{PQ} + z_{RS} \le 1$ from (33a) for this (s, k), we obtain that this constraint is implied. □

We can actually derive an even tighter representation than C_3 by replacing the T_{NC}-cuts with higher-dimensional underlying star graph convex hull constraints as explained next. For any (s, k), consider the conflict graph $G_{sk}(N_{sk}, M_{sk})$ and examining the edges in M_{sk} that are incident at any given node $P \in N_{sk}$, define

$$J_{sk}(P) = \{Q \in N_{sk}\text{: } (P \sim Q) \in M_{sk}\}, \tag{34}$$

where $(P \sim Q)$ denotes (P, Q) if $P < Q$, and denotes (Q, P) otherwise, noting our convention for describing any edge in M_{sk}. Likewise, we denote by (PQ) either PQ if $P < Q$, or QP otherwise. Consider the following result.

PROPOSITION 3. *Let $G_{sk}(N_{sk}, M_{sk})$ be the kth conflict graph for sector s, and for any $P \in N_{sk}$, let $J_{sk}(P)$ be given by* (34)*. Then, the following is a valid inequality for the conflict constraints*:

$$\sum_{Q \in J_{sk}(P)} z_{(PQ)} \le x_P. \tag{35}$$

PROOF. If $x_P = 0$, then (35) implies that $z_{(PQ)} = 0$, $\forall Q \in J_{sk}(P)$ (because $z \geq 0$), which is valid because $z_{(PQ)}$ represents the product $x_P x_Q$. If $x_P = 1$, then this is again valid because the conflict constraint asserts that we must then have $\sum_{Q \in J_{sk}(P)} x_Q \leq 1$. □

Based on Proposition 3, consider the following collection of constraints of the type (35), where we have imposed the condition $|J_{sk}(P)| \geq 2$ because, as we shall show in Proposition 4 below, the constraints for the case $|J_{sk}(P)| = 1$ are inconsequential with respect to the projection onto the original x-space.

$$\sum_{Q \in J_{sk}(P)} z_{(PQ)} \leq x_P, \qquad \forall P \in N_{sk}$$

such that

$$|J_{sk}(P)| \geq 2, \ \forall (s,k). \tag{36}$$

Note that if we consider the overall conflict graph $G(N, A)$, and accordingly define, similar to (34),

$$J(P) = \{Q\colon (P \sim Q) \in A\}, \qquad \forall P \in N, \tag{37}$$

then several subsets of $J(P)$ will be used in the different graphs G_{sk} to generate constraints of the type (36) for any $P \in N$. Naturally, if $J_{s_1k_1}(P) \subseteq J_{s_2k_2}(P)$, then the constraint (36) that is based on $J_{s_1k_1}(P)$ may be dropped because it is implied by that for $J_{s_2k_2}(P)$. Hence, let us define

$$I^* = \big\{P \in N\colon \text{at least one constraint of the type (36) is generated}\big\} \tag{38}$$

and for each $P \in I^*$, let $J_r(P)$, with $|J_r(P)| \geq 2$, for $r = 1, \ldots, r_P$, be the collection of sets of the type $J_{sk}(P)$ that yield a nonredundant system of constraints in (36). Accordingly, let us define the conflict constraint set

$$C_4 = \left\{ \begin{array}{lll} (x,z)\colon \displaystyle\sum_{(P,Q)\in M_{sk}} z_{PQ} \leq 1, & \forall (s,k) & \text{(39a)} \\ \displaystyle\sum_{Q \in J_r(P)} z_{(PQ)} \leq x_P, & \forall r = 1, \ldots, r_P & \\ & \forall P \in I^* & \text{(39b)} \\ z_{PQ} \geq x_P + x_Q - 1, & \forall (P,Q) \in A & \text{(39c)} \\ z \geq 0, \quad x \text{ binary} & & \text{(39d)} \end{array} \right\}.$$

Consider the following result, which justifies the omission of constraints (36) corresponding to $|J_{sk}(P)| = 1$ in (39).

PROPOSITION 4. *Let C_4^+ be defined as in C_4 with the additional constraints*

$$z_{(PQ)} \leq x_P, \qquad \forall P \in N_{sk}$$

such that

$$J_{sk}(P) \equiv \{Q\}, \qquad \forall (s,k). \tag{40}$$

Accordingly, define $X = \{x\colon (x,z) \in \overline{C}_4\}$ and $X^+ = \{x\colon (x,z) \in \overline{C}_4{}^+\}$, where $\overline{C}_4$ and $\overline{C}_4{}^+$ respectively, denote the continuous relaxations of C_4 and C_4^+. Then $X = X^+$.

PROOF. It is clear that $X^+ \subseteq X$ because of the additional constraints (40) defining $\overline{C}_4^+$. To show the converse, note that for any $\bar{x} \in X$ we can take $\bar{z}_{PQ} \equiv \max\{0, \bar{x}_P + \bar{x}_Q - 1\}$, $\forall (P,Q) \in A$, such that $(\bar{x}, \bar{z}) \in \overline{C}_4$. However then $\bar{x}_P \geq \max\{0, \bar{x}_P + \bar{x}_Q - 1\} = \bar{z}_{(PQ)}$, $\forall (P,Q) \in A$, so that constraints (40) are satisfied by $(\bar{x}, \bar{z})$. Therefore, we have $\bar{x} \in X^+$ or $X \subseteq X^+$. □

We now exhibit that C_4 provides a tighter representation than C_3.

PROPOSITION 5. *C_4 provides a tighter representation for the conflict constraints than does C_3 in the sense that $\overline{C}_4 \subseteq \overline{C}_3$.*

PROOF. Let $(x,z) \in \overline{C}_4$. To show that $(x,z) \in \overline{C}_3$, it is sufficient to demonstrate that x satisfies any T_{NC}-inequality of the type $x_P + x_Q + x_R \leq 2$ where, without loss of generality, suppose that $\{P,Q,R\} \subseteq N_{sk}$ with $\{Q,R\} \subseteq J_{sk}(P)$ for some (s,k). Hence, there exists a constraint of the type (39b) defining $\overline{C}_4$ for which $\{Q,R\} \subseteq J_r(P)$. From this constraint, using (39c) and $z \geq 0$, we get $x_P \geq \sum_{Q \in J_r(P)} z_{(PQ)} \geq z_{(PQ)} + z_{(PR)} \geq (x_P + x_Q - 1) + (x_P + x_R - 1)$, or $x_P + x_Q + x_R \leq 2$. □

To illustrate that C_4 can provide a strictly tighter representation than C_3, consider a conflict graph having nodes P, Q, R, and W and with edges (P,Q), (P,R), and (P,W). For this graph, the set $\overline{C}_3$ is given by

$$\overline{C}_3 = \left\{ \begin{array}{l} (x,z)\colon z_{PQ} + z_{PR} + z_{PW} \leq 1 \\ \qquad x_P + x_Q + x_R \leq 2, \ x_P + x_Q + x_W \leq 2, \\ \qquad\quad x_P + x_R + x_W \leq 2 \\ \qquad z_{PQ} \geq x_P + x_Q - 1, \ z_{PR} \geq x_P + x_R - 1, \\ \qquad\quad z_{PW} \geq x_P + x_W - 1 \\ \qquad z \geq 0, \ 0 \leq x \leq e \end{array} \right\}. \tag{41}$$

The seven structural inequalities defining $\overline{C}_3$ are linearly independent, and their intersection yields the feasible fractional vertex $z_{PQ} = z_{PR} = z_{PW} = 1/3$ and

$x_P = x_Q = x_R = x_W = 2/3$. However, note that $\overline{C}_4$ has the constraint $z_{PQ} + z_{PR} + z_{PW} \le x_P$ of type (39b) that deletes this fractional solution. Observe that this constraint, along with (39c), implies the valid inequality $x_P \ge z_{PQ} + z_{PR} + z_{PW} \ge (x_P + x_Q - 1) + (x_P + x_R - 1) + (x_P + x_W - 1)$; that is, $2x_P + (x_Q + x_R + x_W) \le 3$, which also deletes the prior solution. In fact, as shown by Sherali and Smith (2003), this is a facet for the underlying star conflict graph, and that in general for star graphs, the constraints of $\overline{C}_4$ characterize the complete convex hull representation for the conflict constraints, actually implying an exponential collection of facets in the original projected x-space representation. Part II of this paper demonstrates that while C_3 improves the performance over C_2 for some instances, it has the tendency to generate far too many constraints that would need to be filtered to be practically effective. On the other hand, the representation C_4 is more compact and significantly enhances the overall computational performance over C_2, resulting in an average savings of about 29% for more challenging problem instances.

In closing this section, let us revisit the case where the number of conflicts to be allowed simultaneously in any sector s is restricted to be no more than some $r_s \ge 1$ in general. The extended form of C_4 then replaces (36), or (39b), by

$$\sum_{Q \in J_{sk}(P)} z_{(PQ)} \le r_s x_P, \qquad \forall P \in N_{sk}$$

such that

$$|J_{sk}(P)| \ge r_s + 1, \ \forall (s,k). \tag{42}$$

Note that we can show that the projection of this representation onto the x-space remains unaltered if we include the constraints of the type (42) that correspond to $|J_{sk}(P)| \le r_s$, even if we tighten (42) in this case to $\sum_{Q \in J_{sk}(P)} z_{(PQ)} \le |J_{sk}(P)| x_P$. Furthermore, as in (39b), we can identify and eliminate redundant constraints within (42). In particular, if for some (s_1, k_1) and (s_2, k_2) (where possibly $s_1 = s_2$), we have for some $P \in I^*$ defined similar to (38) that $J_{s_1 k_1}(P) \subseteq J_{s_2 k_2}(P)$ and $r_{s_1} \ge r_{s_2}$, then the constraint in (42) corresponding to (s_1, k_1) may be dropped for this P. Hence, scanning (42) in this fashion for each $P \in I^*$ for which at least two constraints of type (42) are generated, we can eliminate the implied inequalities from this set.

4. Sector Workload Constraints

In the foregoing section, we imposed certain workload limitations on ATC personnel for each sector by restricting the number of conflicts that might need simultaneous resolution at any point in time, including the consideration of suitable preparatory buffer times. In practice, ATC operators routinely monitor several aircraft that are simultaneously traversing their respective sectors. Naturally, when the occupancy workload (maximum simultaneous occupancy of a number of aircraft) becomes too high, a potentially dangerous or untenable situation can arise. Accordingly, in the present section, we further constrain the workload imposed on each sector in terms of the average number of aircraft being simultaneously monitored, as well as in terms of the variation of the peak workload from this average. This is in contrast to the workload measure used in Sherali et al. (2002) that was based simply on the peak monitoring load.

Toward this end, consider the total collection of flight plans $\bigcup_{f=1,\ldots,F} P_f$. These plans involve trajectories between certain pairs of fixes that might traverse through some sector s under present consideration. Define the *occupancy workload* for any such sector s at any moment in time to be the number of aircraft resident within that sector at that given instant in time. To characterize this type of workload for each sector $s = 1, \ldots, S$, we can examine the occupancy durations of the various flights within s over the time horizon H. The model AOM of Sherali et al. (2000) provides this information by constructing a Gantt chart of flight plan occupancy intervals for each sector.

Hence, define Ω_s as the set of flight plans (f, p) that occupy the sector s during a horizon of length H (minutes), where the total occupancy time interval for flight plan p of flight f in sector s is given by t^s_{fp} (minutes), as generated via the model AOM. The *average occupancy workload* for sector s can then be computed as

$$w_s = \frac{1}{H} \sum_{(f,p) \in \Omega_s} t^s_{fp} x_{fp}, \qquad \forall s = 1, \ldots, S. \tag{43}$$

The average occupancy workload is penalized in the objective function in a linear fashion. The rationale here is that the average occupancy workload

represents the nominal state of monitoring activity. As such, personnel and equipment can be scheduled as a direct function of the expected amount of work to be performed (assuming, of course, that the expected workload is within the sector's capacity). Hence, we include in the (minimization) objective function the term

$$\sum_{s=1}^{S} \gamma_s w_s = \sum_{s=1}^{S} \sum_{(f,p)\in\Omega_s} \frac{\gamma_s t^s_{fp}}{H} x_{fp}, \tag{44}$$

where γ_s is a suitable constant penalty factor.

Furthermore, for each sector $s = 1, \ldots, S$, let $i = 1, \ldots, I_s$ index the collection of maximal overlapping sets C_{si} of flight plans (f, p), where we use the Sherali and Brown (1994) algorithm to determine these sets. Hence,

$$C_{si} = \left\{ (f,p): \begin{array}{l} \text{flight plan } (f,p) \text{ belongs to the } i\text{th} \\ \text{maximal overlapping set for sector } s \end{array} \right\}, \quad \forall i = 1, \ldots, I_s,\ s = 1, \ldots, S. \tag{45}$$

Note that we permit C_{si} to possibly include alternative surrogates corresponding to the same flight. Now, define the variable n_s to represent the peak occupancy workload as given by the maximum number of flight plans that overlap in occupancy within each sector s. Note that n_s is given by the largest number of flight plans selected from any of the maximal overlapping sets C_{si}, $i = 1, \ldots, I_s$, i.e.,

$$n_s \geq \sum_{(f,p)\in C_{si}} x_{fp}, \qquad \forall i = 1, \ldots, I_s,\ \forall s = 1, \ldots, S, \tag{46}$$

because any other overlapping set is a subset of some maximal overlapping set. In the model formulation, the variable n_s is bounded above by some absolute maximum number $\bar{n}_s$ of overlapping flights, as determined by the capacity of sector s.

With respect to workload, an operations tempo that is steady is preferred to one that is erratic. When the ATC workload varies significantly, additional personnel and equipment resources are required that might remain idle during nonpeak periods. To accommodate this feature, we shall assign a penalty in the objective function corresponding to the maximal variability defined as the difference between the peak occupancy workload and the nominal (average) occupancy workload over the horizon. This maximal variability, $n_s - w_s$, is penalized in the objective function using a penalty factor that increases nonlinearly in an appropriate fashion with an increase in this type of workload. The motivation here is that if this variability in a sector increases from one to three, for example, the associated penalty should likely more than triple. To impose this nonlinear penalty while retaining linearity in the model, we employ a piecewise linear convex structure. Let μ_{sn} be the penalty cost assessed when the peak monitoring workload (n_s) in sector s exceeds the average workload (w_s) by n. (Part II of this paper derives appropriate values for μ_{sn}.) We assume that $\mu_{s1} \geq \mu_{s0}$ and $\mu_{sj} \geq 2\mu_{s(j-1)} - \mu_{s(j-2)}$, $\forall j = 2, \ldots, \bar{n}_s$, thereby imparting a convex penalty structure. Accordingly, we represent the maximal variability via

$$(n_s - w_s) = \sum_{n=0}^{\bar{n}_s} n\, y_{sn}, \qquad \forall s = 1, \ldots, S, \tag{47}$$

where the y-variables represent convex combination weights, satisfying

$$\sum_{n=0}^{\bar{n}_s} y_{sn} = 1, \quad y_{sn} \geq 0, \qquad \forall n = 0, \ldots, \bar{n}_s, \quad \text{for each } s = 1, \ldots, S, \tag{48}$$

and where we accommodate into the objective function the related penalty term $\sum_{n=0}^{\bar{n}_s} \mu_{sn} y_{sn}$, for each $s = 1, \ldots, S$. Observe that by the convexity of the penalty function for each $s = 1, \ldots, S$, either a single or at most two adjacent indexed values of y_{sn}, $n = 0, \ldots, \bar{n}_s$, will be positive at optimality.

Remark 1. Note that the quantity $(n_s - w_s)$ is not necessarily integral. Also, it should be evident that the difference between the peak and average workloads is nonnegative, which justifies the lower bound of zero used in the representation (47). Indeed, if $n(t)$ is the number of aircraft overlapping at time t, we have that $w_s = (1/H) \int_0^H n(t)\, dt \leq (n_s/H) \int_0^H dt = n_s$.

5. Equity Considerations in a Collaborative Decision-Making (CDM) Framework

In this section, we discuss the CDM-related concepts that are incorporated within the APCDM model.

(These equity principles are new recommendations that are not a part of current FAA-airline practice.) Let $\alpha = 1, \ldots, \bar{\alpha}$ index the involved airlines, and let A_α be the set of flights f, with respective surrogates $p \in P_f$, that belong to airline α. We begin by defining the total cost associated with a flight plan $p \in P_{f0}$ of flight f to be given by

$$c_{fp} = F_{fp} + D_{fp}, \qquad \forall p \in P_{f0},\ f = 1, \ldots, F, \tag{49}$$

where F_{fp} and D_{fp} are respectively the fuel- and delay-related costs associated with flight plan p of flight f. (Also, recall that c_{f0} is given a prohibitively high value if the cancellation of flight f is not tendered as an option by the corresponding airline.) Observe that there is a potential cost trade-off between the fuel efficiency of the flight trajectory and the timeliness of the flight's arrival at the destination airport. Part II of this paper develops and tests a comprehensive flight plan (including flight cancellation) cost model.

For each of the flights $f = 1, \ldots, F$, in the absence of constraints (e.g., no competition for scarce resources or conflicts with other flights), airlines would select the least-cost flight plan, c_f^*, from among the available surrogates, where

$$c_f^* = \min\{c_{fp}\colon p \in P_f\}, \qquad \forall f = 1, \ldots, F. \tag{50}$$

However, such a selection may not be feasible when subject to collaborative decision considerations. With this motivation, we first define the *airline collaboration cost* function as the normalized ratio of the total weighted cost incurred to the total weighted individually optimized cost, which is given by

$$d_\alpha(x) = \frac{\sum_{f \in A_\alpha} W_f \sum_{p \in P_{f0}} c_{fp} x_{fp}}{\sum_{f \in A_\alpha} W_f c_f^*}, \qquad \forall \alpha = 1, \ldots, \bar{\alpha}, \tag{51}$$

where for each airline α, $W_f \geq 0$ is a relative priority weight attached to flight $f \in A_\alpha$, such that $\sum_{f \in A_\alpha} W_f = 1$.

Designating $D_{\max} > 1$ as the maximum allowable ratio for any airline of its cost pertaining to the set of surrogates selected through the CDM process to its individually optimized set of surrogates, we impose

$$d_\alpha(x) \leq D_{\max}, \qquad \forall \alpha = 1, \ldots, \bar{\alpha}. \tag{52}$$

Notice that the ratio $D_{\max}$ provides a minimal degree of cost efficiency for all the airlines. The computational analysis in Part II of this paper suggests $D_{\max} = 1.2$ as an appropriate value for this parameter. Note also that the imposition (52) might render the problem infeasible; Remark 2 below addresses this aspect. Accordingly, we then define the *airline collaboration efficiency*, $E_\alpha(x)$ as the linear function that achieves:

$$E_\alpha(x) = \begin{cases} 1 & \text{if } d_\alpha(x) = 1 \\ 0 & \text{if } d_\alpha(x) = D_{\max} \end{cases},$$
$$\text{and linear in between,} \quad \forall \alpha = 1, \ldots, \bar{\alpha}. \tag{53}$$

This yields, using (51) within (53),

$$E_\alpha(x) = \frac{D_{\max} \sum_{f \in A_\alpha} W_f c_f^* - \sum_{f \in A_\alpha} \sum_{p \in P_{f0}} W_f c_{fp} x_{fp}}{(D_{\max} - 1) \sum_{f \in A_\alpha} W_f c_f^*}, \qquad \forall \alpha = 1, \ldots, \bar{\alpha}, \tag{54}$$

where we note from (50) and (52) that the collaboration efficiency defined in (54) automatically satisfies $0 \leq E_\alpha(x) \leq 1,\ \forall \alpha$. (However, in lieu of imposing (52) in the model, we will directly impose $E_\alpha(x) \geq 0$, for $\alpha = 1, \ldots, \bar{\alpha}$.) Observe that in certain situations, we might prefer to use a value $D_{\max}^\alpha$ specific to each airline α in (54). For example, we might wish to give some airline α_1 preferential treatment by letting $D_{\max}^{\alpha_1} < D_{\max}^\alpha,\ \forall \alpha \neq \alpha_1$. This might be desirable if airline α_1 is a small or new entrant airline that requires lower operating costs to initially enter a market or to establish a new service that is beneficial to the overall airline industry.

Additionally, let us define the *ω-mean collaboration efficiency* as the weighted sum of the individual airline collaboration efficiencies:

$$\sum_\alpha \omega_\alpha E_\alpha(x), \quad \text{where} \quad \sum_{\alpha=1}^{\bar{\alpha}} \omega_\alpha = 1, \quad \omega_\alpha \geq 0, \qquad \forall \alpha = 1, \ldots, \bar{\alpha}. \tag{55}$$

The relative weights ω_α could be selected based on the proportion of flights operated by each airline (whence $\omega_\alpha = |A_\alpha|/F,\ \forall \alpha$), or based on the proportion of passengers/passenger-miles serviced.

Based on this efficiency measure, let us now define a *collaboration equity* function for each airline as the

deviation of the corresponding airline's collaboration efficiency from the ω-mean collaboration efficiency:

$$E^e_\alpha(x) = E_\alpha(x) - \left(\sum_{\alpha=1}^{\bar{\alpha}} \omega_\alpha E_\alpha(x)\right), \qquad \forall \alpha = 1, \ldots, \bar{\alpha}. \quad (56)$$

Observe that $E^e_\alpha(x)$ can take on positive or negative values, respectively indicating whether airline α has achieved an individual collaboration efficiency that is better or worse than the ω-mean collaboration efficiency. Accordingly, we then denote the *ω-mean absolute collaboration inequity* by a variable x^e as given by,

$$x^e \equiv \sum_{\alpha=1}^{\bar{\alpha}} \omega_\alpha \left| E^e_\alpha(x) \right|, \quad (57)$$

and using suitable nonnegative parameters μ^e, μ^D, ν^e, and $E^e_{\max}$, we incorporate in the (minimization) objective function the respective inefficiency and inequity terms:

$$\mu^D \left[1 - \sum_{\alpha=1}^{\bar{\alpha}} \omega_\alpha E_\alpha(x)\right] + \mu^e x^e, \quad (58)$$

and include in the constraints the additional bounding restrictions

$$x^e \le \nu^e, \quad \omega_\alpha E^e_\alpha(x) \ge -E^e_{\max}, \quad \forall \alpha = 1, \ldots, \bar{\alpha}, \quad \text{and}$$
$$E_\alpha(x) \ge 0, \quad \forall \alpha = 1, \ldots, \bar{\alpha}. \quad (59)$$

Remark 2. Observe that the first objective function term in (58) attempts to increase the ω-mean collaboration efficiency, while the second term works towards reducing the weighted spread of the individual airline efficiencies from this mean value. Hence, this is a mean-variance minimization type of strategy, where the variance factor is being represented by a mean absolute deviation term to retain linearity, and where each airline's efficiency is assumed to be duplicated in this mean-variance representation in proportion to its corresponding weight. When used in concert with a policy that would require the airlines to offer surrogates for each flight that conform with some consistent distribution of flight plan costs, an equity formulation based on this efficiency measure would obviate gaming strategies on the part of airlines to bias the collaborative decision towards individually optimized surrogates. Moreover, at least in the sense that the efficiency measure (54) is normalized by the individual airline's optimized operating costs, we alleviate the concern that a more efficient airline might subsidize a less efficient airline in the model. (Also, see Wambsganss (1996) for a related cost-based equity discussion.) Furthermore, to control the resulting relative efficiency and equity values somewhat more determinably, we impose certain bounds on some related terms in (59). Here, in addition to imposing a hard upper bound ν^e on x^e to accompany the objective term $\mu^e x^e$ in (58), we require each airline's weighted collaboration equity to be no more negative than $-E^e_{\max}$. (One could also explore here an unweighted uniform lower bound $E^e_\alpha(x) \ge -E^e_{\max}$, $\forall \alpha = 1, \ldots, \bar{\alpha}$.) Note that if we choose $\nu^e < \min\{1, \bar{\alpha} E^e_{\max}\}$, given that $E^e_{\max} > 0$, then the upper bound ν^e on x^e is not necessarily implied. In this connection, note that these bounds could make the overall problem infeasible. In this case, within the CDM framework, some subsequent sensitivity analysis runs could be made with suitable relaxed values of these bounds, including a relaxation of the nonnegativity restrictions on the efficiency values, which in effect, would relax (52) (while still using the parameter $D_{\max}$ within (54)). As an alternative, we can soften (59) to some extent by deleting the restriction $x^e \le \nu^e$, declaring $E^e_{\max}$ to be a variable rather than a constant and further augmenting the objective terms in (58) according to

$$\mu^D \left[1 - \sum_{\alpha=1}^{\bar{\alpha}} \omega_\alpha E_\alpha(x)\right] + \mu^e x^e + \mu^e_{\max} E^e_{\max}, \quad (60)$$

where $\mu^e_{\max} \ge 0$ is a suitable parameter. Observe that similar to x^e, the variable $E^e_{\max}$ is another measure of dispersion that essentially reflects the range of weighted efficiency values. Essentially, $E^e_{\max}$ can be viewed as the maximum weighted collaboration inequity. Hence, with (60), in addition to minimizing the total weighted collaboration inequity x^e, the objective function would attempt to also minimize the maximum weighted collaboration inequity.

A couple of examples might help elucidate the foregoing discussion. Suppose that the model permits the following two feasible solutions, S1 and S2, having respective efficiency values as stated below.

S1: $E_1(x) = 1.0,\ E_2(x) = 1.0,\ E_3(x) = 0.8,\ E_4(x) = 0.7,\ E_5(x) = 0.5,\ E_6(x) = 0.5$;

S2: $E_1(x) = 1.0,\ E_2(x) = 0.9,\ E_3(x) = 0.8,\ E_4(x) = 0.7,\ E_5(x) = 0.6,\ E_6(x) = 0.5$.

If we assume equal weights for each airline (i.e., $\omega_\alpha = 1/6,\ \forall\alpha$), the corresponding objective function terms in (58) for S1 are $0.25\mu^D + 0.18\mu^e$, while for S2 they are $0.25\mu^D + 0.15\mu^e$. Observe that both solutions have the same total (weighted) collaboration efficiency. Moreover, they have the same range as well as extremal values. Hence, the modeling construct used in the previous model APM (Sherali et al. 2002), when based on these efficiency values, would not distinguish between the solutions S1 and S2. However, by also incorporating a variance-type term x^e that examines the mean absolute deviations of the efficiencies from their (weighted) mean value, the (partial) objective formulation (58) identifies S2 to be a more equitable solution because it exhibits a more uniform distribution of efficiencies as compared with S1.

As another example (proposed by an anonymous referee), consider the following two solutions having the specified efficiency values, where $\omega_\alpha = 1/10,\ \forall\alpha = 1, \ldots, 10$.

S1: $E_i(x) = 0.6$ for $i = 1, \ldots, 5$, and $E_i(x) = 0.4$ for $i = 6, \ldots, 10$;

S2: $E_1(x) = 1$, $E_2(x) = 0$, and $E_i(x) = 0.5$ for $i = 3, \ldots, 10$.

Note that for both of these solutions, the objective function terms in (58) yield the values $0.5\mu^D + 0.1\mu^e$; hence these two solutions are indistinguishable via (58). However, it might appear that airline $\alpha = 2$ is unfairly disadvantaged in the solution S2. Indeed, the equity model in Sherali et al. (2002) would have strictly preferred S1 in this case because it offers a smaller range in efficiency values. Nonetheless, given that from (59) we have $\omega_2 E_2^e(x) \geq -E_{\max}^e$ (i.e., $-(0.1)(0.5) \geq -E_{\max}^e$), S2 does conform with evidently agreed-upon equity lower bounds. Notwithstanding this, observe that when $E_{\max}^e$ is taken as a variable on the other hand, we obtain $E_{\max}^e = 0.01$ for S1, whereas $E_{\max}^e = 0.05$ for S2, so that the alternative objective representation (60) does indeed prefer S1 over S2.

Remark 3. In the foregoing equity formulations embodied by (58) and (60), observe that we have assumed a linear utility function for each airline based on its achieved efficiency for the sake of computational tractability. This linear relationship might not adequately reflect an airline decision maker's risk attitude (i.e., risk tolerance) with respect to costs incurred via the collaborative process. While we could adopt more general multiattribute utility-theory constructs as expounded by Kirkwood (1997) in the present context, this would inject nonlinearity and nonconvexity into the model, thereby complicating it beyond its present scope. We therefore leave the consideration of such nonlinear utility functions for future investigation.

6. Overall Model Formulation

We summarize below the APCDM model using the C_4 conflict resolution constraints, and introducing the new variables $\nu_\alpha,\ \alpha = 1, \ldots, \bar{\alpha}$, to linearize (57) via $(61m, n)$. Observe that we have also explicitly included an overall system cost as the first term in the objective function (61a) in spite of the fact that the third objective term includes this factor, because the latter does so only in a scaled normalized fashion. Furthermore, we have stated this model using the equity objective terms given in (60) rather than (58) with $E_{\max}^e$ being a variable for the sake of generality. As expounded in Remark 2, the user might wish to use $E_{\max}^e$ as a suitable agreed-upon constant, as well as explore variations in the other bounds included within (61o) (see Remark 2 above).

Minimize

$$\sum_{f=1}^{F} \sum_{p \in P_{f0}} c_{fp} x_{fp} + \sum_{s=1}^{S} \sum_{n=0}^{\bar{n}_s} \mu_{sn} y_{sn} + \mu^D \sum_{\alpha=1}^{\bar{\alpha}} \omega_\alpha \left[1 - E_\alpha(x)\right] + \mu^e x^e + \mu_{\max}^e E_{\max}^e + \sum_{s=1}^{S} \gamma_s w_s + \sum_{(P,Q) \in A} \varphi_{PQ} z_{PQ} \qquad (61a)$$

subject to:

$$\sum_{p \in P_{f0}} x_{fp} = 1, \qquad \forall f = 1, \ldots, F \qquad (61b)$$

$$\sum_{(f,p) \in C_{si}} x_{fp} \leq n_s, \qquad \forall i = 1, \ldots, I_s,\ s = 1, \ldots, S \qquad (61c)$$

$$w_s = \frac{1}{H} \sum_{(f,p) \in \Omega_s} t_{fp}^s x_{fp}, \qquad \forall s = 1, \ldots, S \qquad (61d)$$

$$n_s - w_s = \sum_{n=0}^{\bar{n}_s} n y_{sn}, \qquad \forall s = 1, \ldots, S \tag{61e}$$

$$\sum_{n=0}^{\bar{n}_s} y_{sn} = 1, \qquad \forall s = 1, \ldots, S \tag{61f}$$

$$x_P + x_Q \le 1, \qquad \forall (P, Q) \in FC \tag{61g}$$

$$\sum_{(P,Q) \in M_{sk}} z_{PQ} \le r_s, \quad \forall k = 1, \ldots, K_s,\ s = 1, \ldots, S \tag{61h}$$

$$x_P + x_Q - z_{PQ} \le 1, \qquad \forall (P, Q) \in A \tag{61i}$$

$$\sum_{Q \in J_{sk}(P)} z_{(PQ)} \le r_s x_P, \qquad \forall P \in N_{sk}\colon \left| J_{sk}(P) \right| \ge r_s + 1, \quad \forall k = 1, \ldots, K_s,\ s = 1, \ldots, S \tag{61j}$$

$$E_\alpha(x) = \frac{D_{\max} \sum_{f \in A_\alpha} W_f c_f^* - \sum_{f \in A_\alpha} \sum_{p \in P_{f0}} W_f c_{fp} x_{fp}}{(D_{\max} - 1) \sum_{f \in A_\alpha} W_f c_f^*}, \quad \forall \alpha = 1, \ldots, \bar{\alpha} \tag{61k}$$

$$E_\alpha^e(x) = E_\alpha(x) - \left(\sum_{\alpha=1}^{\bar{\alpha}} \omega_\alpha E_\alpha(x) \right), \quad \forall \alpha = 1, \ldots, \bar{\alpha} \tag{61l}$$

$$\nu_\alpha \ge -E_\alpha^e(x), \qquad \nu_\alpha \ge E_\alpha^e(x), \qquad \forall \alpha = 1, \ldots, \bar{\alpha} \tag{61m}$$

$$x^e = \sum_{\alpha=1}^{\bar{\alpha}} \omega_\alpha \nu_\alpha \tag{61n}$$

$$\begin{aligned} &z_{PQ} \ge 0, \qquad \forall (P, Q) \in A, x_{fp} \text{ binary}, \quad \forall p \in P_{f0},\ \forall f = 1, \ldots, F, \\ &y_{sn} \ge 0, \qquad \forall n = 1, \ldots, \bar{n}_s,\ s = 1, \ldots, S, \\ &E_\alpha(x) \ge 0, \qquad \forall \alpha = 1, \ldots, \bar{\alpha}, \\ &n_s \le \bar{n}_s, \quad \forall s = 1, \ldots, S, \\ &x^e \le \nu^e, \qquad E_{\max}^e \ge -\omega_\alpha E_\alpha^e(x), \qquad \forall \alpha = 1, \ldots, \bar{\alpha}. \end{aligned} \tag{61o}$$

7. Summary and Conclusions

In this paper, we have presented the Airspace Planning and Collaborative Decision-Making Model (APCDM) that significantly enhances a preliminary formulation APM as described in Sherali et al. (2002). Specifically, we have first incorporated two alternative representations of randomized aircraft trajectory errors into the model, in contrast with using deterministic flight paths, recognizing that aircraft are subject to pilot, navigation, and wind-induced errors. By studying various conflict risk thresholds, this construct can also offer new insights that can be used for investigating possible revisions to FAA's aircraft separation standards. Second, we have employed a continuous time consideration of conflict risk incidents to provide a more flexible means to represent sector conflict resolution capacities. Third, we have proposed two new model formulations of conflict resolution constraints via suitable, valid inequalities that serve to tighten the linear programming relaxations, and thereby potentially improve the solvability of the developed large-scale, mixed-integer optimization problem. Finally, we have proposed a new collaboration equity concept that effectively scrutinizes the distribution of individual airline schedule costs relative to the overall cost and takes better account of the dispersion of relative airline cost efficiencies than in the preliminary APM model, which focused mainly on ranges and extremal values of equity functions.

As previously indicated, Part II of this paper deals with a detailed description of implementing the APCDM model, including the development of a comprehensive cost model and a study for prescribing a set of appropriate parameter values for the overall model. This sequel paper also presents results using the Enhanced Traffic Management System (ETMS) database to illustrate the utility of the proposed valid inequalities, and to study the sensitivity of the model with respect to various pertinent model parameters.

The APCDM can be used in practice for various tactical decision-making purposes as well as for strategic planning studies. Potential tactical contexts include decisions pertaining to air traffic control diversions and delays during spacecraft launch operations or during severe weather conditions, military theater operations (such as damage assessment, search and rescue, ground support, and counter-operations), and the generation of alternative flight plans (via the feedback outer loop in Figure 1). Strategic applications include air traffic control (ATC) policy evaluations (e.g., with respect to separation standards, ATC-workload restrictions, and resectorization strategies), Homeland Defense contingency planning, spaceport location planning, military theater air campaign planning, and the construction of

a priori plans to respond to various disruption scenarios (e.g., augmenting the FAA's *National Playbook* 2003). Furthermore, the model can be used to study the incorporation of the Small Aircraft Transportation Systems (SATS) into the NAS. Detailed investigations of these and other applications of the APCDM model will be explored in future research.

Acknowledgments

This research has been supported by the National Science Foundation under Grant 0094462 and by a Federal Aviation Administration grant under the National Center of Excellence for Aviation Operations Research (NEXTOR) program. Thanks are due to Dr. Hojong Baik for his assistance in data acquisition and computations and to an associate editor and two referees for their detailed and constructive comments that have led to an improved version of this paper. The views expressed in this article are those of the authors and do not reflect the official policy or position of the U.S. Air Force, Department of Defense, or the U.S. government.

References

Bakker, G., H. Kremer, H. Blom. 2000. Geometric and probabilistic approaches towards conflict prediction. *Proc. 3rd USA/Europe Air Traffic Management R&D Seminar*, Napoli, Italy (June 13–16).

Ball, M., R. Hoffman, W. Hall, A. Muharremoglu. 1998. Collaborative decision making in air traffic management: A preliminary assessment. NEXTOR Technical Report RR-99-3, University of Maryland, College Park, MD.

——, ——, D. Knorr, J. Wetherly, M. Wambsganss. 2000. Assessing the benefits of collaborative decision making in air traffic management. *Proc. 3rd USA/Europe Air Traffic Management R&D Seminar*, Napoli, Italy (June 13–16).

Bertsimas, D., S. Stock Patterson. 1998. The air traffic flow management problem with en route capacities. *Oper. Res.* **46**(3) 406–422.

Carlson, P. 2000. Exploiting the opportunities of collaborative decision making: A model and efficient solution algorithm for airline use. *Transportation Sci.* **34** 381–393.

Chang, K., K. Howard, R. Oiesen, Shisler. 2001. Enhancements to the FAA ground-delay program under collaborative decision making. *Interfaces* **31** 57–76.

Crawley, J. 2001. U.S. lays out 10-year plan to reduce flight delays. *Reuters News Service* (June 16).

Downer, R. 2001. Washington ARTCC Traffic Management Coordinator. *Personal Communication*.

Erzberger, H., R. Paielli, D. Isaacson, M. Eschow. 1997. Conflict detection and resolution in the presence of prediction error. *Proc. 1st USA/Europe Air Traffic Management R&D Seminar* (June 17–20).

Federal Aviation Administration. 2003. *National Severe Weather Playbook*. http://www.fly.faa.gov/playbook.html.

Hansen, M., A. Nilim, L. El Ghaoui, V. Duong. 2002. Dynamic routing of aircraft under weather uncertainly. *NEXTOR CDM Project Rev.*, University of California, Berkeley, CA (June 10).

Keagan, C., Plenary Speaker. 2003. *NEXTOR Conf. Air Traffic Management and Control*. Falls Church, VA (June 2–3).

Kirkwood, C. 1997. Notes on attitude toward risk taking and the exponential utility function. Working paper, Arizona State University, Tempe, AZ.

Luss, H. 1999. On equitable resource allocation problems: A lexicographical minimax approach. *Oper. Res.* **47**(3) 461–478.

Metron Aviation, Inc. 2002. CDM overview. http://www.metronaviation.com.

Paielli, R., H. Erzberger. 1997. Conflict probability estimation for free flight. *J. Guidance, Control, Dynamics* **20** 588–596.

——, ——. 1999. Conflict probability estimation generalized to non-level flight. *Air Traffic Control Quart.* **7**.

Reich, P. 1966. Separation standards, I. *J. Instit. Navigation* **19** 88–98.

Sherali, H. D., E. Brown. 1994. A quadratic partial assignment and packing model and algorithm for the airline gate assignment problem. M. Pardalos, H. Wolkowicz, eds. *Quadratic Assignment and Related Problems, Vol. 16. DIMACS Series in Discrete Mathematics and Theoretical Computer Science*. American Mathematical Society, 343–364.

——, J. C. Smith. 2003. A polyhedral study of the generalized vertex packing problem. *Math. Programming*. Forthcoming.

——, ——, A. A. Trani. 2002. An airspace planning model for selecting flight plans under workload, safety, and equity considerations. *Transportation Sci.* **36** 378–397.

——, ——, ——, S. Sale. 2000. National airspace sector occupancy and conflict analysis models for evaluating scenarios under the free-flight paradigm. *Transportation Sci.* **34** 321–336.

Turner, W., J. Mize, K. Case, J. Nazemetz. 1993. *Introduction to Industrial and Systems Engineering*, 3rd ed. Prentice Hall, Englewood Cliffs, NJ.

Vossen, T., M. Ball. 2001. Optimization and mediated bartering models for ground delay programs. Working paper, Robert H. Smith School of Business, University of Maryland, College Park, MD.

Wambsganss, M. 1996. Collaborative decision making through dynamic information transfer. *Air Traffic Control Quart.* **4** 107–123.

Young, H. P. 1994. *Equity in Theory and Practice*. Princeton University Press, Princeton, NJ.

Received: May 2003; revision received: July 2003; accepted: July 2003.