FLIGHT MOVEMENT INVENTORY: SAGE-AERO2K

Sophie Michot, Ted Elliff, Eurocontrol Experimental Centre, Brétigny-sur-Orge, France

Gregg G. Fleming, Brian Kim, U.S. Department of Transportation, Volpe Center, Cambridge, MA

Curtis A. Holsclaw, Maryalice Locke, Angel Morales, Federal Aviation Administration, Office of Environment and Energy, Washington, D.C.

Weather and Environment/TRL C

Abstract

A global air traffic emissions database is an essential tool for both policy makers and climate change scientists. Since the last comprehensive aircraft emissions inventories were developed in 1992, an update is necessary. This need is being addressed in the USA through a project entitled SAGE and in Europe through a project entitled AERO2K. Both Europe and the USA have agreed to collaborate on these similar projects. The agreement resulted in the exchange of flight movement data in order to realise an air traffic movement inventory, the essential starting point for estimating global aviation emissions. The objective of the inventory in both projects is to provide 4-D flight trajectories (latitude, longitude, altitude and time) using as much measured data as possible. In AERO2K, to supplement the collection of data, schedule data from the Back Aviation database were also used, complemented by airspace route structure information, where available. Similarly, in SAGE, measured data were augmented by schedule data from the worldwide Official Airlines Guide (OAG), complemented by a vertical and horizontal track dispersion model, which was developed from actual measured 4-D trajectory data. This paper details the aircraft movements data in AERO2K and SAGE.

Keywords: Aviation, Emission, 4-D trajectory, Inventory

1. Introduction

To determine the impact of aviation on global climate change, studies to inventory the air traffic movements in the world are required. In the 90's, projects to produce world flight movement inventories for calculating fuel consumption and emissions were led by NASA, ANCAT and DLR [1, 2, 3]. These inventories were based on ATC and schedule data completed by a Great Circle trajectory between city-pairs. In the ANCAT project, the most significant omission of ATC data was the United States, for which data were unavailable for security reasons. Thus, only timetable data were used for the United States discarding nonscheduled U.S. domestic charters and other flights. To compensate for this problem, fuel usage data were factored up by 10% [2].

Ten years after, two US and European projects respectively named SAGE and AERO2K aim to produce a world inventory using as many real trajectories as possible. A significant improvement is the agreement as part of Action Plan 13 between FAA and Eurocontrol for exchanging flight movement data. These data are four-dimensional trajectories expressed in terms of latitude, longitude, flight level and time. These trajectories are given either by radar tracks or flight plans.

In this paper we examine the origin of the flight movement data for Europe, the USA and the rest of the world and how these data were combined to produce a single inventory. Both projects SAGE and AERO2K adopted different methods, which are described below.

2. SAGE

The Federal Aviation Administration's Office of Environment and Energy (FAA/AEE) with support from the John A. Volpe National Transportation Systems Center (Volpe), the Massachusetts Institute of Technology (MIT) and the Logistics Management Institute (LMI) are developing the **S**ystem for assessing **A**viation's **G**lobal **E**missions (**SAGE**). The development team envisions SAGE as an internationally accepted computer model that can be used for predicting and evaluating the effects of different policy and technology scenarios on aviation-related emissions, costs, aircraft performance, and industry responses. With regard to scope, the model will be capable of analyses on an aircraft, airport, regional, and global level.

The SAGE development effort involves a two-part approach with Version 1 to be completed at the end of 2002 and Version 2 to be completed in the 2006/2007 timeframe. Version 1 provides the basic core capability for assessing commercial aviation's global emissions. Version 2 will be based entirely upon Version 1 while providing increased fidelity in input databases and calculation methodologies; and Version 2 will be expanded to include military and general aviation movements. In addition, an economics assessment capability is currently being considered for inclusion in the Version 2 model. It is also anticipated that interim versions of the model will be produced between the release of Versions 1 and 2.

The basic goal for SAGE is to be a technically sound and internationally accepted computer model used for evaluating technology-related scenarios on aircraft emissions.

Figure 1 shows a simplified overview of the relationships between the main modules in SAGE Version 1.

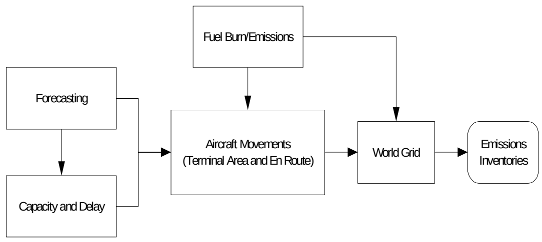

Figure 1. Overview of SAGE Main Modules

In order for SAGE to be considered acceptable, the model must employ technically sound computational algorithms, transparent methodology and processes, and credible databases that are viewed as an acceptable means of estimating global aviation emissions by the international aviation community. International acceptance of the model and the underlying databases and assumptions will greatly contribute to the model's output being accepted and viewed as credible input to the international decision-making processes.

2.1 Components of the Movements Database

To understand the components of the movement's database, one must first understand the scope of SAGE. The model will have a geographical analysis resolution ranging from a single airport, a region, and a global level. In terms of flight analysis, SAGE will be able to analyze a single aircraft in flight (i.e., single city-pair) as well as a fleet of aircraft incorporating multiple worldwide city-pairs. In terms of analysis years, the model will be used to develop baseline emissions inventories initially corresponding to the years 1990, 1992, and 2000. These years generally correspond to the Kyoto Protocol, the Abatement of Nuisances Caused by Air Transport (ANCAT) report, and the SAGE development baseline year, respectively. Baselines will be used to develop derivative inventories from querying operations and will also be used as a basis for comparison (e.g., against forecasted inventories). Baselines will also be developed for 2001, 2002, and future years as data becomes available. Emissions inventories are essentially developed for grid cells. The basic geometric requirement for the world grid involves latitude, longitude, and altitude dimensions of 1 degree by 1 degree by 1 kilometer respectively for each grid cell.

The components of the movement's database within SAGE and their specific usage depend largely on the particular analysis year and geographic region being modeled. Table 1 provides a broad overview of the main functions and geographic coverage of each database component. The Database components can be grouped into three general categories. The first category includes databases that provide schedule, trajectory, and geographic location data. These components include the Enhanced Traffic Management System (ETMS), the Official Airline Guide (OAG), the Central Flow Management Unit (CFMU), and other schedule and trajectory data, as they become available. The second category includes two database components that provide for the computation of aircraft performance. They include the model defined by the Society of Automotive Engineers' (SAE) Aerospace Information Report (AIR) 1845, and the model defined by Eurocontrol's Base of Aircraft Data (BADA). The third group generally consists of data that provides aircraft and engine matching/assignments. These include the Airline Service Quality Performance (ASQP) data, the BACK Aviation Solutions (BACK) world fleet, and an engine distributions database.

Table 1. Summary of Components in SAGE Movements Database

Database	Main Function/Data	Geographic Coverage	Analysis Year			
			1990/ 1992	2000/ 2001	2002	>2003
ETMS	Schedule, Trajectory Data and Dispersion Model	North and Central America and the United Kingdom	–	√	√	√
OAG	Schedule	Worldwide	√	√	√	√
CFMU	Schedule, Trajectory, and Taxi Time Data	Western and Portions of Eastern Europe	–	–	√	√
Other Schedule and Trajectory Data	Schedule, Trajectory Data	Ad Hoc Worldwide	–	–	–	√
Airports	Airport latitude, longitude, and altitude	Worldwide	√	√	√	√
SAE AIR 1845	Terminal Area Performance	N/A	√	√	√	√
BADA	En Route Performance	N/A	√	√	√	√
ASQP	Tail Number and Taxi Time	United States	√	√	√	√
BACK	Equipment Registration	Worldwide	√	√	√	√
Engine Distributions	Engine Assignments	N/A	√	√	√	√

Each database component is discussed in detail below.

ETMS

ETMS is essentially the FAA's electronic record of flight position. ETMS is a combination of flight-identifier encoded radar position reports and a flight's filed flight plan. Each report is time stamped and the two types of records are mixed together and sorted by time stamp. These two reports are called the position report and the flight plan report. ETMS captures every aircraft that flies within North and Central America and the United Kingdom. This coverage accounts for about 50 to 55 percent of worldwide operations. ETMS includes unscheduled, cargo, military, charter, general aviation, and scheduled flights, but military and general aviation are not included in SAGE Version 1. It also captures every flight that files a flight plan, whether that aircraft enters radar-controlled airspace or not.

The radar position reports include flight ID, altitude, and position in digital latitude and longitude. Radar reports are designated with "TZ" or "CZ" headers. The flight plan report contains information such as scheduled departure time, actual departure time, scheduled arrival time, actual arrival time, equipment and origin/destination. Comprehensive data cleaning, parsing and matching programs were developed for the processing of ETMS to ensure data integrity and to streamline the processing of such a voluminous amount of data.

After using a fairly robust cleaning, parsing and matching program, flight plan and radar information with accurate origin/destination locations and times can be expected for about 80 percent of the ETMS flights, which equates to about 25 to 30 percent of global flights.

OAG

The worldwide OAG, a Reed Business Information company and a member of the Reed Elsevier plc Group, provides worldwide lists of scheduled commercial and cargo flights. Since worldwide OAG sells tickets, all scheduled airlines and the majority of scheduled worldwide airlines are represented in the schedules. Although the worldwide OAG includes cargo flights, it does not include unscheduled and charter flights.

Comprehensive data cleaning programs were developed to ensure an accurate picture of scheduled traffic. For example, code-sharing flights are often represented more than once in an OAG schedule, and to avoid double-counting, the code-sharing partners have to be consolidated. Also, airlines in the U.S. regularly cancel overlapping flights at their hubs. After excising these duplicate records and making an allowance for predictable cancellations, the OAG represents the majority of global air traffic.

Some of the key data contained within the OAG include origin and destination airport, the time the flight is scheduled to depart and arrive the origin/destination airport in Greenwich Meridian Time, the equipment type and the flight number.

CFMU

The Eurocontrol's Central Flow Management Unit System provides access to interactive air traffic flow information across Europe. It contains capacity ratings for Eurocontrol sectors and all public European airports. This database is used tactically to grant flight plan authorizations for international flights, and at the same time to reserve and manage the capacity slots associated with Eurocontrol sectors and European airports. It is both a database of flight and capacity information and a tactical, daily-use tool.

Other Schedule and Trajectory Data

Additional sources of schedule and trajectory data will be identified and included in future versions of SAGE. It is currently unclear as to the level of detail available from other regions in the world, but these data will be incorporated into SAGE accordingly.

SAE

The SAE AIR 1845 exists in SAGE as a pre-computed static set of aircraft profiles that describe departure and arrival performance for all aircraft/engine combinations covered by the ICAO emissions data bank [4]. The SAE AIR 1845 methodology exists as the performance model in the FAA's Integrated Noise Model (INM) [5] as well as other models. For each aircraft/engine combination, there will be up to seven takeoff weight classes based on stage lengths. Stage length (synonymous with trip distance) is used as a surrogate for takeoff weight. The relationship between stage and trip distance is consistent for all aircraft. In other words, smaller-sized, shorter-range aircraft may only have a single stage, whereas mid-range aircraft like the B737 would typically have four stages. A single landing weight is assumed for each aircraft/engine combination.

BADA

BADA is managed and updated by Eurocontrol [6]. It provides a set of ASCII files containing performance and operating procedure coefficients for 186 different aircraft types. The coefficients include those used to calculate thrust, drag and fuel flow and those used to specify nominal cruise, climb and descent speeds.

BADA uses a total-energy model. That is, it uses the principle of conservation of energy to determine the rate of change in speed and the rate of change in height. Specifically, the total-energy added to the aircraft (energy added due to thrust minus energy removed due to drag) is allocated to the kinetic energy (proportional to the rate of change in speed) and the potential energy (proportional to the rate of change in height).

ASQP

The ASQP database is developed by the USDOT's Bureau of Transportation Statistics

(BTS) and it contains fleet data for approximately the 10 largest carriers. Specifically, carriers are required to provide fleet data if they account for 1 percent or more of the total domestic scheduled service passenger revenues. The data also includes voluntary reporting by airlines. The database includes information such as flight number, depart and arrival airport, date of operation, day of week, aircraft tail number, and taxi out and in time.

BACK

The BACK registration (fleet) database is developed and managed by BACK Aviation Solutions. It contains a comprehensive list of all registered aircraft worldwide, including those in the Former Soviet States. The database includes information: aircraft manufacturer, type, exact model; engine manufacturer, type, exact model; serial, registration, production line numbers; aircraft age; aircraft status (e.g., active/inactive/stored); noise stage or chapter; equipment classification; and total airframe hours and cycles.

Engine Distributions

Distributions of engines were developed for the top 50 airlines (based on operations) from analyzing the BACK fleet database. An additional (51st) airline was created to represent the aggregation of all other (smaller) airlines. The distributions were developed based on counts of different aircraft and engine types in each of the airline categories. For example, Delta Airlines may have B737-200, B727-100, etc. aircraft types. And under the B737-200 category, Delta may be using 50% JT8D-15A, 30% JT8D-9, etc. engine types. These engine distributions allow for proper statistical assignments (i.e., for a years worth of flight schedules) when the data is not available for an exact identification of engine types. This is critical for OAG derived flights since specific engine information is not provided in OAG.

Airports

A worldwide airport locations file was developed for the set of airports represented in the OAG schedule as flight origins and destinations. Approximately 11988 non-US airports and 2200 US airports are included. The database essentially includes airport name, airport code, and the latitude, longitude, and altitude associated with its location. Sources for this database includes various FAA offices, Eurocontrol's CFMU, and others.

2.2 Movements Database Development

The movements database in SAGE is formed by integrating the terminal area takeoff and approach profile data with en route trajectories. These two sets of data essentially constitute the "backbone" of SAGE by providing aircraft movements information to all other components of the model. Section 2.2.1 describes the development of en route trajectories. Section 2.2.2 describes the development of terminal area trajectories. Section 2.2.3 describes the integration of terminal area and en route trajectories.

2.2.1 En Route Data

In Version 1 of SAGE, a combination of a trajectory generator and the 2000 ETMS database are used to develop the en route trajectories. In later versions of SAGE, the en route trajectories will be augmented by data from Eurocontrol's CFMU, as well as data from other sources. The trajectory generator modifies the Great Circle (GC) trajectory, which is essentially the shortest line fit through two points on the earth's surface. The ETMS trajectory data is favored over those derived from the trajectory generator because ETMS data represents actual measured information. However, the methods used by the trajectory generator were developed from a statistical analysis of a large set of ETMS data. Therefore, the results (e.g., fuel burn) from each set of trajectory data are similar, at least on an aggregate level when looking at large samples of flights. The trajectory generator in concert with the 2000 OAG schedule serves as the default source of aircraft trajectory information with the use of ETMS data occurring whenever available.

In order to allow the trajectory generator to more realistically represent actual horizontal flight paths, a dispersion method is incorporated such that routes offset from the GC is assigned randomly to each flight using pseudo-random (i.e., pre-generated) numbers. These dispersion routes were developed from analyzing a large set of ETMS trajectory data for multiple city pairs. Figure 2 shows an illustration of two flights assigned to dispersion routes around the GC.

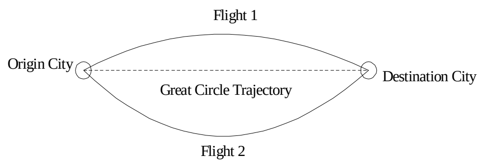

Figure 2. Dispersion Trajectories

The dispersion method employs four (4) dispersion tracks per each of seven (7) stage lengths (trip distances). Table 2 shows these tracks for only one stage length. Each of the 4 tracks has an equal chance of being selected (i.e., probability of 0.25), and each of the track dispersion records consist of 7 offsets from the GC. Offsets are nautical miles perpendicular to the great circle route, and the percentages (e.g., 20%, 30%, 40%, etc.) shown in the table refer to a point along the GC track starting from the origin airport.

Table 2 Trajectory (Lat/Long) Dispersion Distributions Developed from Analyzing ETMS Data

		Perpendicular distance from Great Circle (nm)							
stage	Trajectory No.	20%	30%	40%	50%	60%	70%	80%	Probability
1	1	9.5	12	13.5	13.5	14.5	14	12.5	0.125
1	2	40.5	46	50.5	53	53	52	48	0.125
1	3	-9.5	-12	-13.5	-13.5	-14.5	-14	-12.5	0.125
1	4	-40.5	-46	-50.5	-53	-53	-52	-48	0.125

Similar to the horizontal dispersion, the en route altitudes used with the trajectory generator are also based on a dispersion methodology developed from analyzing a large set of ETMS data. This involves the use of common altitude categories for different trip distances. The current method is to use four categories as shown in Table 3 For flights with a trip distances greater than 200 nm, the altitudes are randomly assigned based on the distributions shown in Table 4 using pseudo-random numbers. Since each distribution point contains a high and low altitude, one of these are chosen randomly on a 50%/50% basis. For flights less than 500 nm where the takeoff and approach profiles cross before they reach the assigned altitudes, a nominal altitude of 20,000 ft are assigned. This was considered more methodically consistent (i.e., more control over altitude assignments) than trying to develop altitudes based on proportioning the takeoff and approach profiles to fit the altitudes. Table 4 shows the altitude distributions derived from ETMS data.

Table 3. Trip-Distance Categories for Developing Cruise Altitudes for use with OAG

Trip Distance	Method
≤ 50 nm	Drop the flight
>50 and <200 nm	Use 15,000 ft
>200 and ≤500 nm	Use distribution in Table 3
>500 nm	Use distribution in Table 3

Table 4. Altitude Dispersion Distributions developed from Analyzing ETMS Data

Altitude Low (ft)	Altitude High (ft)	Probability for flights 200 - 500 nm	Probability for flights above 500 nm
20000	21000	0.118076428	6.24E-04
22000	23000	0.127951911	3.48E-03
24000	25000	0.167453843	1.36E-02
26000	27000	0.190854444	5.78E-02
28000	29000	0.178832117	0.125859
30000	31000	0.136539287	0.260904
32000	33000	5.88E-02	0.240211
34000	35000	0.016960069	0.182321
36000	37000	4.29E-03	9.69E-02
38000	39000	2.15E-04	1.77E-02
40000	41000	N/A	5.35E-04
42000	43000	N/A	8.92E-05

The en route module depends on several databases including OAG, ETMS, and several “look-up” databases developed from statistical correlations. The ASQP and BACK databases need to be used to correlate flight numbers to registration numbers and ultimately, to engine types for computation of fuel burn and emissions.

2.2.2 Terminal Area Data

Although ETMS data contains some near terminal trajectory points (chords), the resolution and inconsistency prevent any detailed modeling of the terminal area using this data. Therefore, the internationally recognized SAE AIR 1845 methodology/data is used to generate the terminal area flight profiles. The input variables which are required to index the appropriate two-dimensional profile data (for takeoff and approach) generated using SAE AIR 1845 include the aircraft ID, the engine type, the trip distance, and the type of operation (i.e., departure or approach). Given these indexing variables, the appropriate profile is identified, and the associated data are extracted. These data include the horizontal distance (either from brake release or runway threshold, depending upon whether the operation is a departure or an approach), the height, the ground speed and the thrust.

The required indexing parameters for the profile database come from a combination of sources, including ETMS and ETMS-like data, augmented by the OAG as appropriate. The ETMS/OAG data are interfaced with the ASQP data and the BACK registration data so that the specific aircraft/engine combination on a particular flight can be identified. If a specific aircraft match cannot be made due to a lack of coverage from the ASQP data, an assignment is made according to the distribution database described in Section 2.1. The trip distance is determined from the ETMS data directly or the OAG as necessary. Determination of takeoff/approach is made using both the ETMS and OAG databases.

Ground movements including taxi times and delays are modeled as idling times for fuel burn/emissions computation purposes. These "movements" (time) are attributed geographically to the airport being modeled. Average taxi times for each airport were developed from a statistical analysis of ASQP data and CFMU data. Airport-specific delays were obtained from a capacity and delay model, LMINET.

2.2.3 Combining En Route and Terminal-Area Data

The terminal-area profile data is fed to a computational module, which integrates the two-dimensional takeoff and approach profile data with the en route trajectories from either ETMS or the trajectory generator. In creating a baseline movements database (e.g., for 2000), ETMS data takes precedence over those from the trajectory generator. That is, data from the trajectory generator is not used if ETMS data is available for the flights in question.

The takeoff profiles are integrated with the cruise segments of flight by modeling the takeoff and cruise modes as shown in Figure 3.The takeoff mode is modeled in fixed altitude steps of 1000 ft, 1500 ft, 2000 ft, 3000 ft, 4000 ft, and in increments of 2000 ft thereafter until the first cruise altitude is reached with a joining chord that connects the last takeoff chord with the first ETMS chord. Similarly, looking backwards from the destination airport, the approach altitudes are incremented as 1000 ft, 1500 ft, 3000 ft, and in increments of 2000 ft thereafter until the last ETMS cruise chord is reached. And a joining chord is used to connect the last ETMS chord with the closest approach chord.

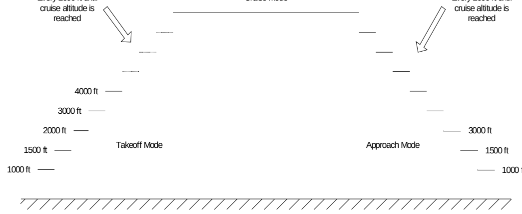

Figure 3. Integration of Terminal Area and En Route Data

These integrated data constitute the aircraft movements database within SAGE.

3. AERO2K

The realization of a world inventory implies the collection of flight data from various sources. A consequence is the existence of variation at the level of the data quality, the data formating and the duplication of information between sources. Standardisation is then required for producing a single inventory. This standardisation is the function of the import tool and the merge tool developed in AERO2K. A prototype tool of importing and merging was first developed in MS Access 2000. The tool was then migrated to Oracle 9i, and the system was modeled using Rational Rose linked to a Change Management Tool Continuus.

3.1 AERO2K Importing Tool

3.1.1 Origin of the data

To compile the global inventory, three main sources were identified: Europe, North America and the rest of the world. Based on 2000 ICAO data, they respectively account for 29, 42 and 29% of total scheduled aircraft departures [7].

European data are obtained from Eurocontrol's Central Flow Management Unit (CFMU). The CFMU is a repository center for flight plans originating in the European Civil Aviation Conference (ECAC) area. CFMU daily flight plan data files are then processed through a simulator called AMOC standing for ATFM (Air Traffic Flow Management) Modelling Capability. AMOC is an integrated ATFM simulation platform developed by the Eurocontrol Experimental Centre, in Brétigny, France. The processing of CFMU flight plans through AMOC results in 4-D trajectories in respect with the route structure and ATFM environment for that day.

American data originate from ETMS and are a combination of radar and flight plan data. Their description has been given in the previous paragraph.

In order to cover as many flights as possible in the world, ATC data were completed with flight schedule data from the Back Aviation database. The coverage of the Back Aviation database is based on the schedule data submitted by airlines according to their forecast for future months. The database contains listings of every scheduled jet and turboprop flight listed by city-pair and airline, e.g. includes departure and arrival times, airplane code, aircraft type and trip frequency.

3.1.2 Organisation of the files

Data generated by AMOC are split into two files named Traffic and Flight. The Traffic table includes data characterising the schedule and the fleet. It reflects the level and geographic distribution of traffic in terms of frequency of flights by origin-destination airport pair. It also contains the daily time distribution of traffic in terms of scheduled departure. It indicates the assignment of a specific aircraft type to each flight. The Flight table includes data characterising the trajectory such as latitude, longitude, altitude, time and speed. The data structure covered by ETMS is discussed in Section 2.

For the rest of the world a complete set of schedules was extracted for the time period selected from Back Aviation database. The fields downloaded were the airport of origin and destination, the published carrier, the equipment type, the Great Circle distance, the flight number, the departure and arrival time, the days of operation and the elapsed time.

3.1.3 Importing format

The structure of the ETMS data being quite similar to the European data: one file listing general information on the airport and aircraft, the other file having information on the flight route, it was decided to keep this structure. Procedures for importing AMOC, ETMS and Back Aviation data were then developed. The common output structure for importing consists of two tables. The first table contains data fields such as the flight callsign, the departure time, the airport of departure, the airport of arrival, the aircraft type, the last event time known and the source of the data. The second table contains flight legs data such as the flight callsign, the flight event time, the latitude, longitude, flight level, the average speed between two consecutive points, the climb/descent rate and the source of the data.

In order to identify each flight as a single flight, a flight unique identifier (FLUID) was created. It consists of the flight callsign concatenated with the airport of departure and the date and time of departure to the closest minute. The process was adapted at each source and includes the conversion of the airport codes to ICAO codes through an airport database, airline codes to three letters codes and date/time to GMT.

The procedure for importing Back Aviation data was a bit more complicated as schedule type of data such as data obtained from Back Aviation, do not have any information on the trajectory between the airport of departure and the airport of destination. Trajectories had to be created. This was done after the merge of AMOC and ETMS data and will be described later.

3.2 AERO2K Merging Tool

The purpose of the merging tool is to select flight information from different sources, after these flights have initially been imported and converted into the AERO2K standard format. Imported flights are loaded into a temporary Flight table and a temporary Flight_Leg table. Two groups of processes are carried out on these data before the real merge. They are a flight leg consistency check and a flight leg assessment. The whole process is first applied to AMOC and ETMS data.

A first check is made such that only consistent flights will be included in the final inventory. Any flights with missing data or not allowing to link the flights to the legs are discarded. A second check consists of updating the legs with identical event times. The rounding to the nearest minute of event-time given by the radar leads to the existence of points with a different position but an identical time. To avoid this problem, the average position of the legs with the same event time is calculated and a single leg is kept. Finally, the tool should allow the production of daily inventory and so the partial selection of flights. Flights starting the day before or finishing the day after should not be totally incorporated in the daily inventory. For all these flights a leg was created at midnight and the position of the point was estimated based on the Great Circle distance between the point before midnight and past midnight. The flight level and speed are calculated as an average between two flight legs, which are the last flight leg for departure day and the first flight leg for arrival day.

The flight leg assessment consists first of updating the average speed, the delta level (i.e. the climb/descent rate) and adding a zone indicator. This last process identifies the position of the leg in the trajectory. Three zones were identified such as departure ‘D’, en route ‘E’ and arrival ‘A’. Within this zone indicator system, each zone was assessed in terms of trajectory consistency (Table 5). For both departure and arrival zones, a “C” for complete was allocated for example to Zone_Indicator equals ‘D’ if the flight had at least one flight leg. Otherwise, a “I” for incomplete was attributed to zone ‘D’. For the en route zone, a flag ‘F’ for fine was allocated if the average speed was inferior or equal to 10,000 knots and the delta level was inferior or equal to 10,000. Otherwise a ‘D’ for dubious was allocated. Therefore a total of eight combinations was obtained. Such an assessment allowed to determine the quality of the flight trajectory. These limits can appear as very high but the experience showed that extremely high values where reached in the case of bad trajectories. Some trajectories did not include flight level or presented inconsistent ground speeds due to successive positions recorded in a short period by different radar centres and rounded to the closest minute.

Table 5. Examples of a zone indicator consistency coding of the trajectory (the zone selected is shown in yellow).

ETMS			AMOC		
D	**E**	**A**	**D**	**E**	**A**
C	F	C	C	F	C
C	F	I	I	F	C
C	F	I	I	F	I
C	F	I	C	D	C
I	F	C	C	F	I
I	F	C	I	F	I
I	F	I	C	D	C
I	F	I	C	D	I
I	F	I	I	D	C
C	D	C	I	F	I
C	D	I	C	D	I

The assessment finished, and the information contained into the temporary flight table and the first and last legs information from the temporary flight leg table is imported into a temporary merge table. This table is scanned such as all flights that have the same FLUID or have the same callsign, departure airport, first and last flight legs within range are identified. Three criteria define the range:

- either the interval of time between flights' departure times is less than 900 minutes
- or the average speed between the first flight leg of the first flight and the last flight leg of the second flight is 0 or within 110 – 600 knots range
- or the average speed between the last flight leg of the first flight and the first flight leg of the second flight is 0 or within 110 - 600 knots range.

Using these criteria, duplicated flights are retrieved and compared in order to identify which flight will provide its flight legs for the final merged flight, and for which one of the flight zones. Comparisons are made on a set of duplicated flights and the first record in the set is replaced with the result of the comparison. The choice of the legs to select is based on the result of the assessment and the number of flight legs for a zone. For example, if the first flight is incomplete for the departure but the flight to compare with is complete then the leg to be selected will be the one from the second flight. If the second flight was also incomplete then the number of legs is counted and the flight with the larger number of legs is kept. This approach assures the best of all trajectory data is retained. Flights for which no duplicated were identified are saved unchanged in the merge table.

As mentioned previously, schedule data from Back Aviation do not have any information on the flight trajectory. For this reason two methods were developed in order to attribute a trajectory to these data. The methods are based on the identification of the city-pair studied among the city-pairs recorded in the merged inventory. In case of the existence of such a city-pair, the aircraft types mentioned in the schedule data and the merged inventory for the city-pair studied are compared. If the city-pair and the aircraft type match, then the flight legs for the schedule data are created using the matched flight legs record. If the aircraft type does not match, a check is done for identifying a matching city-pair with an equivalent aircraft type. The equivalency between aircrafts was determined based on a list of aircraft in BADA [8]. If several aircraft types are available, the aircraft type selected will be the one with a trajectory having the closest departure time.

The second method for creating flight legs is based on the creation of routes between city-pairs and the allocation of altitude and speed defined from template aircraft profile. The National Imagery and Mapping Agency (NIMA)[a] of the United States Department of Defense (DOD) produces a Digital Aeronautical Flight Information file (DAFIF)[b] that gives information on the world route network. A digital map locating world waypoints and route networks were extracted from this file. These data were augmented by a list of city-pairs and an associated airport based on a 40 nm criterion. All this information was then loaded into a tool named CARAT, a Computer Aided Route Allocation Tool developed by Eurocontrol Experimental Centre [9]. For AERO2K, the capacity component was not invoked thus resulting in a route from origin to destination based only on a list of waypoints; capacity and aircraft performances were not considered. The tool delivered a list of shortest paths found between city-pairs. At this stage of the method, latitudes and longitudes of points could be assigned along the path between city-pairs. The second stage consists of determining the aircraft performance in order to attribute a flight level and a speed, to be used for calculating time. For each aircraft type a flight profile pattern can be identified. A flight profile is mainly determined by the aircraft's operational performance and the total flight length. A profile can be subdivided into three sections

Departure profile (climb rate and acceleration)
Cruise (constant flight level and ground speed)
Arrival profile (descent rate and deceleration)

Profiles were determined using a set of ETMS data from which a set of graphs for each aircraft was derived and the trendline assessed visually.
Cruise flight levels are a function of the total flight length and are assumed to be constant between the end of the climb phase and the beginning of the descent phase. So for each aircraft type a graph was created showing the total flight distance on the horizontal axis and the average maximum flight level and average maximum speed on the vertical axis. The average values were calculated by grouping aircraft types into flight distance categories of 50 nm e.g. all flights having total flight distance between 100 nm and 150 nm are grouped into the single category 125 nm (centre point).
For each aircraft type and departure and arrival profile, a number of graphs were created showing the progressive flight distance (up to 500 nm) on the horizontal axis and the average flight level and ground speed on the vertical axis. Separate graphs were created for the different flight ranges (e.g., 0 to 500 nm, 500 to 1000 nm, 1000 plus nm). The assumption was made that for an aircraft type the profile for a short flight might be different than that for a long flight. Progressive distances were grouped into 10 nm categories such as for each category all points having a progressive distance between 10 nm and 20 nm were grouped into the single category 15 nm (centre point).
Knowing the latitude/longitude of two points, the distance was calculated based on the Great Circle function and reporting this distance on the graphs, the flight level and time at each waypoints along the trajectory could be determined. The 4-D trajectory could then be generated.

4 Conclusion

The agreement made by Eurocontrol and the FAA as part of Action Plan 13 is a good example of a US and European mutual beneficial collaboration. The exchange of flight movement data was beneficial to both the SAGE and AERO2K projects. Potential for future progress would be to tend towards a stronger collaboration in order to create a single flight movement inventory. A common strategy for collecting worldwide data has been achieved through the establishment of a questionnaire. The initial idea was to use ICAO regional offices in the world for dispatching the questionnaire to the appropriate ATC authorities. This questionnaire has not yet been disseminated (December 2002) but we think that such a step would contribute greatly to improve the quality of the inventory. Agreements similar to Action Plan 13 between the US and Europe could also be established with large regional air traffic centres such as South Africa, Brazil, Australia, China and Japan for example. An increase in the coverage of measured air traffic data would result in a more accurate estimation of the emissions due to aviation.

References

[1] Baughcum, S.L., Henderson, S.C., Tritz, T.G., Pickett D.C., 1996, Scheduled Civil Aircraft Emission Inventories for 1992: Database Development and Analysis. NASA-CR-4700, National Aeronautics and Space Administration, Langley Research Center, Hampton, VA, USA, 196 pp.
[2] Gardner, R.M. (ed.), 1998, ANCAT/EC2 Global Aircraft Emissions Inventories for 1991/1992 and 2015: Final Report. EUR-18179, ANCAT/EC Working Group, 84 pp + appendices.
[3] Schmitt, A. Brunner, B., 1997, Emissions from aviation and their development over time. In: Final Report on the BMBF Verbundprogramm, Schadstoffe in der Luftfahrt.
[4] SAE A-21 Committee. SAE AIR 1845. Warrandale, PA. Society of Automotive Engineers. 1986.
[5] Olmstead, J.R., Fleming, G.G., et.al., Integrated Noise Model (INM) Version 6.0 Technical Manual, Report No. FAA-AEE-02-01, Washington, D.C., Federal Aviation Administration, 2002.
[6] http://www.eurocontrol.fr/projects/bada/

[7] Costaguta, A., 2001, personal communication. ICAO Chief of Statistics Section.
[8] Nuic, A., 2000, User manual for the base of aircraft data (BADA) Revision 3.3 EEC Note No. 20/00 ERIS IBP Z I2.
[9] Tibichte, A., Dalichampt M., 1997, ATFM modelling capability: AMOC**.** EEC Note No. 28/97, EEC Task E02, EATCHIP Task CFMU

[a] "This product has not been endorsed or otherwise approved by the National Imagery and Mapping Agency, or the United States Department of Defense (10 U.S.C. 425)."
[b] "This product was developed using DAFIF®, a product of the National Imagery and Mapping Agency."

Author Biographies

Sophie Michot, Eurocontrol Experimental Centre
Sophie Michot joined as a consultant, the EEC Environmental Studies Business Area in August 2001. As project leader, she is responsible for producing the world flight movement inventory part of AERO2K project. Her role involves both project management and R&D work.

Ted Elliff, Eurocontrol Experimental Centre
Ted Elliff is head of the EEC business area for Environmental Studies, responsible for a portfolio of R&D projects covering noise, air quality and global emissions from aviation.

Gregg G. Fleming, John A. Volpe National Transportation Systems Center
Gregg G. Fleming, Chief of Volpe's Environmental Measurement and Modeling Division, has been with the Center for 15 years. He is responsible for all aspects of SAGE development.

Brian Kim, John A. Volpe National Transportation Systems Center
Brian Kim of the Air Quality Facility of Volpe's Environmental Measurement and Modeling Division has been with the Center for 4 years. He is responsible for the day-to-day design, development and management of SAGE.

Curtis Holsclaw, Federal Aviation Administration

Maryalice Locke, Federal Aviation Administration

Angel Morales, Federal Aviation Administration

Acknowledgements
The authors of this paper would like to thank the AERO2K flight inventory team members: Nicu Stancioi, Colin MacMullan, Serge Peeters and Sandrine Carlier; and the SAGE development team members: Joseph Ruggiero, Sathya Balasubramanian, Ian Waitz, Joe Lee, J-P Clarke, Peter Kostiuk and Virginia Stouffer. All data providers are also thanked.