Arithmetic simulation and performance metrics

Jean-Marc Alliot   Geraud Granger   Jean-Marc Pomeret

Abstract

Performance metrics are becoming a strategic issue, and they are getting more and more attention. However, defining such metrics is a difficult problem. In this paper, we show how arithmetic simulations can be used to give performance informations. We also point out that many different metrics can be defined, each of them giving different results regarding efficiency. We conclude that an extreme caution must be applied when interpreting results.

1 Introduction

Performance review is now looked upon as a standard activity in the ATM community [Tea02, WP302]. However, defining performance is not easy. The basic definition usually relies on the ability of a system to handle traffic without generating delays. But such a definition is utterly poor, as it is much more difficult to handle highly dense traffic over the core European area than sparse traffic over deserts in Africa. So, on the one hand, a sector might be generating delays just because its traffic is more difficult to handle. On the other hand it could have a very simple traffic to handle and then should generate no delay at all. Finally, it generates few delays because it is protected by another one. Defining lack of performance simply by generated delays doesn't work. Thus, a new notion steps forward: traffic complexity. But then again, defining complexity is not an easy task either. Many different complexity metrics can be used, each of them being apparently sensible, but all of them giving different results.

Beyond metrics definition, measurement is by itself a difficult process. While some very basic metrics (such as the number of aircraft controlled in one sector for one day, for example) can be easily computed by a statistical analysis of data samples, some others are not. For example, if we want to measure the number of "conflicts to solve" by sector, analyzing radar samples won't help much as, if the controller is doing his job properly, there should be no remaining conflict.

In that case, the use of arithmetic, or fast time, simulations is required. Such simulators can handle one day of traffic in a few minutes. They however, introduce new weak points, as a simulation always relies on some hypothesis that may or may not be correct. In fact, as we will show in this paper, we have mainly the choice between simple reliable metrics, which may not be very meaningful, and more elaborate metrics, probably more meaningful, but not as reliable.

2 Metrics considered

In this paper, we will concentrate on four different metrics: flights controlled per sector, number of conflicts per sector, number of potential conflicts per sector and number of manoeuvres given to aircraft for solving conflict in one sector. The three last values are computed using a fast time simulator, while the first is extracted from flight samples. We also use in the rest of this paper the *sector volume*, which is the sector raw surface divided by the square of the radar separation minimum, multiplied by the number of available flight levels. This is roughly the number of aircraft that could be frozen and packed inside without violating standard separation constraints.

2.1 Number of flights by sector

This metric is certainly the easier to define. Using real flight plans, we just compute the number of flights controlled by sectors. However, to have an idea of the CFMU regulations we also used initial flight plans to compare both results.

2.2 Number of conflicts by sector

A significant part of the controller workload is conflict resolution. Then, we try to count the number of standard separation violations by sector with aircraft following exactly their flight plan routes with no deviations. This is done by fast time simulation, the hypothesis of which being detailed thereafter.

2.3 Number of potential conflict by sector

A large part of controllers workload comes from trajectory monitoring and conflict detection. Different studies show that only one conflict out of three to five detected

and monitored would really result in separation violation. This is directly the consequence of uncertainties affecting aircraft trajectories (wind, unavailable FMS informations, etc), and of human beings inability to handle complex numerical mathematic to compute trajectory predictions. Thus the number of potential conflicts is the number of conflicts detected, given a set of hypothesis regarding aircraft trajectories uncertainties, both in the horizontal and the vertical plane. There again, this value is estimated using fast time simulation.

2.4 Number of manoeuvres by sector

A more subtle indicator is the number and the category of manoeuvres given inside a sector to solve conflicts. Some conflicts involving many aircraft can sometimes be solve by only one manoeuvre, while others may require many manoeuvres. There again fast time simulation is used to estimate this parameter.

3 The OPAS fast time simulator

Why using fast time simulations? There are different reasons that could be summarized as follow:

- there is sometimes no other way to compute some metrics. In real life, the number of conflicts is impossible to compute, as competent controllers are usually doing properly their job, thus suppressing conflicts. The only way to estimate the number of conflicts that would have happened is to simulate traffic without any control actions. The same goes with potential conflict detection, as there is no way to be in the head of the controller.

- some metrics are too difficult to find out. For example, the number of manoeuvres given could be found by a close examination of radar samples, but it would be difficult to find out which manoeuvres are given to solve conflicts and which are given for other purposes.

In the remaining part of this section, the OPAS fast time simulator is presented. This simulator and its different modules have been presented in many different articles [DAB97, DABM97, DAG01b, DAG01a], and we will only concentrate on its main features.

3.1 General principles and conflict detection

The simulator uses a tabulated model for aircraft performances: ground speed, vertical speed, and fuel burn are functions of altitude, aircraft type and flight segment (cruise, climb or descent.) The main dataset for aircraft
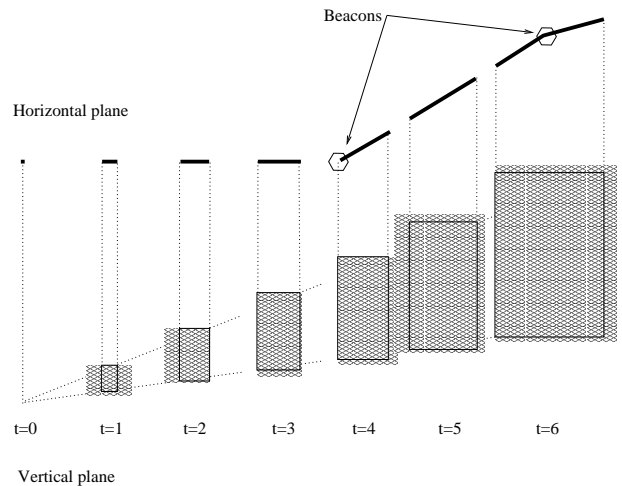

Figure 1: Modeling of speed uncertainties (standard routes).

flight performance is the base of aircraft data (BADA) performance summary tables derived from the total energy model of EUROCONTROL. 69 different aircraft types are described. Synonym aircraft are used to model the rest of the fleet.

Aircraft follow classical routes (from way-point to way-point). The flight model is simple: an aircraft first climbs up to its RFL, then remains leveled till its top of descent, then descends to its destination.

Flight plans are data of the COURAGE system, an archiving system of the operational French CAUTRA Air Traffic Control system.

Aircraft fly with a timestep that can be chosen at the start of the simulation. The timestep is always chosen in order to guarantee that two aircraft face to face flying at 500 kts could not cross without being closer than one radar separation minimum at at least one timestep. For most of our simulation, we use a $15s$ timestep. Separation minima are checked for each aircraft at each timestep.

3.2 Potential Conflict detection

For potential conflict detection, the simulator does a trajectory prediction *each three minutes* by a simulation of a given duration inside the global simulation. This duration is what we call the *anticipation*.

We assume during each of these detection simulations that there is an error about the aircraft future location because of ground and vertical speed prediction uncertainties.

Then, an aircraft is represented by a point at the initial time of the conflict detection window. In the horizontal plane, the point becomes a line segment in the uncertainty direction (the speed direction here, see figure 1). The first

point of the line "flies" at the maximum possible speed, and the last point at the minimum possible speed. These maximal and minimal speeds depend of course on the uncertainty chosen: for 5% uncertainty on ground speed, the first point will fly at a speed of $1.05\,v$ and the last point at $0.95\,v$, if v is the nominal speed of the aircraft.

When changing direction on a waypoint, the heading of the line segment "fastest point" changes as described on figure 1.

To check separation for two aircraft at time t, we compute the distance between the two line segments modeling the aircraft positions and compare it to the separation minima.

In the vertical plane, we use a cylindrical modeling (figure 1). Each aircraft has a mean altitude, a maximal altitude and a minimal altitude. To check if two aircraft are in conflict, the minimal altitude of the higher aircraft is compared to the maximal altitude of the lower aircraft. The radar separation minima used are 6 nautical miles in the horizontal plane and, 1000 ft under FL295 and 2000 ft above (no RVSM) in the vertical plane.

Conflicts detected can be merged: if a conflict is detected a time t_1, and detected again three minutes later, the two conflicts are only considered as one.

3.3 Manoeuvres for conflict resolution

In the horizontal plane, classical manoeuvres given to aircraft are heading deviation. In the simulator, 10, 20 or 30 degrees deviations will be allowed. The deviation starts on a virtual waypoint created on the route (see figure 2). This waypoint is defined by the position of the head of the segment at some time t_0. It ends on a second virtual waypoint, position of the head of the segment at time t_1. An angle criteria is defined to find on which waypoint the modified and initial routes should connect.

A maneuver will be determined by:

- t_0 which defines the first virtual waypoint B_0.

- the deviation angle α.

- t_1 which defines the second virtual waypoint B_1.

In the vertical plane, the aircraft trajectory is divided in 4 periods (figure 3):

- Climbing period. In this period, aircraft can be leveled at a lower than requested flight level to solve a conflict. The aircraft climb is stopped at flight level FL_0 and resumes climbing on a virtual waypoint B_1.

- Cruising period. When aircraft have reached their desired flight level, they may be moved to the nearest lower level to resolve a conflict.
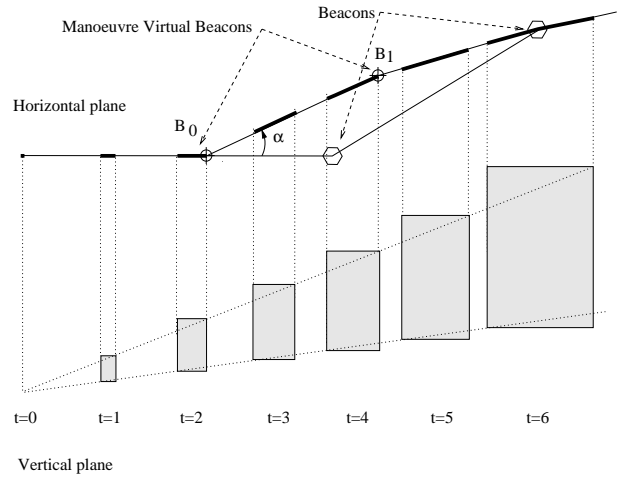

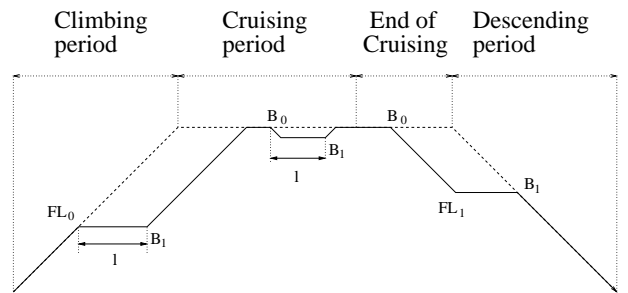Figure 2: Horizontal maneuver modeling.

Figure 3: Vertical maneuver modeling.

- End of Cruising period. When aircraft are about 50 nautical miles away from the top of descent, they may be moved to a lower level to resolve a conflict.

- Descending period. During this period no vertical maneuver is possible.

No maneuver will be simultaneously done in the horizontal and vertical plane.

These manoeuvres were chosen in order to be as close as possible to the controller's behavior.

4 Results

As stated above, we used data (flight plans, airspace, etc) for the French airspace. There are two main reasons:

- we have full access to these data, and have a quite good knowledge of their reliability and their weaknesses

- it would have probably been more interesting to try to compare airspaces of different countries, but results could have been misinterpreted, and would have

	STT	Error	STD	Error	Acft	Error
Aix	30mn	3%	182Nm	2%	3442	2%
Bordeaux	30mn	3%	194Nm	1%	2274	3%
Brest	28mn	2%	197Nm	1%	2431	2%
Paris	21mn	5%	117Nm	3%	3887	4%
Reims	18mn	5%	112Nm	3%	2564	3%

Table 1: Comparison of fast time simulation results and real data for the 5 French ACC

	Flights	Real	Potential	Manoeuvres
1	TC,(P) 1544	SE,(R) 135	DS,(P) 477	SE,(R) 158
2	DS,(P) 1217	DS,(P) 117	LN,(P) 470	XS,(Br) 135
3	LN,(P) 1211	UE,(R) 92	SE,(R) 295	AR,(P) 130
4	TS,(P) 884	ZS,(Br) 91	TC,(P) 192	UE,(R) 130
5	SE,(R) 718	LN,(P) 91	TP,(P) 189	UR,(R) 127

Table 2: Sector ranking (P:Paris, Br:Brest, R:Reims, A:Aix)

	Fl/Vol.	Conf./Vol	Pot./Vol	Man./Vol
1	DS,(P) 6.16	ZS,(Br) 0.68	DS,(P) 2.41	ZS,(Br) 0.90
2	ZS,(Br) 5.01	DS,(P) 0.59	ZS,(Br) 0.92	XS,(Br) 0.75
3	XS,(Br) 3.20	XS,(Br) 0.38	LN,(P) 0.72	XH,(R) 0.38
4	UZ,(P) 2.27	XN,(R) 0.23	XS,(Br) 0.65	SU,(P) 0.31
5	TU,(P) 2.18	XH,(R) 0.21	XN,(R) 0.32	B2,(A) 0.31

Table 3: Sector Ranking

Ranking by number of conflicts : see column 2 of tables 2 and 3.

Ranking by number of potential conflicts : see column 3 of tables 2 and 3.

Ranking by number of manoeuvres : see column 4 of tables 2 and 3.

An excellent example is the TC sector, which is the one having the largest number of controlled flights, but doesn't even appear in the "flight per volume" indicator (rank 10), or in the "number of conflicts to solve" indicator (rank 12). Only one sector (DS) appears in the top five of six classifications (but not in the manoeuvres classificationd), the rest being scattered. Then, after a review of these results, it seems extremely difficult to find the most "loaded" or "efficient" control sectors. Moreover, even the simplest indicator (number of flights going through one sector) should be examined with great care, as it is highly probable that some flights crossing sectors such as TC (arrivals) or DS (departures) are not controlled by these sectors but directly by the approaches of Paris airports. And we are not through yet. . .

4.3 Computing other metrics. . .

We have just presented above some simple, easy to understand metrics. However, other studies have been conducted, and other metrics have been proposed ([DP00, LSBB98, JT91]. . .). So, how deep is the rabbit hole? We decided to compute extensively these metrics, or metrics very similar, to try to find out correlations [Cha01]. Presenting all of them would be a too lengthy task for this paper, and the reader have to read the above references for more precise informations. We are just listing them and giving a very short definition:

Density (dens): counts the number of aircraft with a weighting inversely proportional to distance $\sum_{i,j} \exp^{-\alpha |\overrightarrow{d_{ij}}|}$

Convergence (conv): estimation of convergence of aircraft. The more converging aircraft, the highest this indicator.

been much more subject to controversy. Our goal here was to discuss from a general standpoint the problems linked to performance metrics and fast time simulations, to develop a general methodology, but not to start a discussion about the efficiency of European national air traffic control systems. The present work can be considered as an initial experiment, leaving the field open to more ambitious projects.

4.1 Validation

Using fast time simulations introduces biases. Then, it is mandatory to verify that results fit the reality of traffic before going any further. Then results of fast time simulations must be compared to real radar data statistics. We used IMAGE radar statistics to compare values such as Standard Transit Time (STT), Standard Traveled Distance (STD), and number of aircraft controlled. These were the first results generated, and they are presented in table 1. The error column indicates the difference between radar statistics and fast time simulation results. Differences are insignificant and can be either due to the simulator or to radar data, which are not completely reliable and easy to interpret.

4.2 Computing metrics

We present here the results of simulations for the French sectors. In table 2 and 3, we give the 5 highest ranked sectors regarding the 4×2 indicators defined above.

Ranking by number of flights : see column 1 of tables 2 and 3.

	nb	dens	conv	isp	tsep	desang	desord	confl	cpot
nb	1.00	0.84	0.77	0.58	0.07	0.80	0.75	0.46	0.68
dens	0.84	1.00	0.84	0.48	0.09	0.96	0.87	0.52	0.69
conv	0.77	0.84	1.00	0.55	0.04	0.86	0.85	0.43	0.69
isp	0.58	0.48	0.55	1.00	0.07	0.44	0.57	0.17	0.30
tsep	0.07	0.09	0.04	0.07	1.00	0.08	0.06	0.01	0.02
desang	0.80	0.96	0.86	0.44	0.08	1.00	0.88	0.49	0.67
desord	0.75	0.87	0.85	0.57	0.06	0.88	1.00	0.38	0.61
confl	0.46	0.52	0.43	0.17	0.01	0.49	0.38	1.00	0.53
cpot	0.68	0.69	0.69	0.30	0.02	0.67	0.61	0.53	1.00

Table 4: Correlation matrix

Insensibility (isp): estimation of the sensibility to manoeuvres that can be given by controllers. This indicator is high when giving manoeuvres will result in new conflicts with other aircraft.

Time to separation (tsep): when aircraft are convergent, the controller must monitor them until separation is guaranteed. This indicator was defined by French controllers.

Angular disorder (desang): this indicator is low with highly organized traffic on routes, higher with less organization

Speed disorder (desord): high when aircraft have different speed in the sector

Number of aircraft (nb), number of conflicts (conf) and number of potential conflicts (cpot) were already defined above.

The matrix of correlations is presented in table 4.3. When the figure is close to 1 indicators are correlated. When it is close to 0, they are not. Thus, it is easy to see that the rabbit hole is indeed very deep. Correlation is impossible to discuss, as the table doesn't exhibit any interesting behavior...

5   Conclusion

The main error when discussing complexity or performance metrics would be to believe in a "one-for-all performance indicator for manager" that could summarize all the information in one scalar value, and would enable an easy comparison between different sectors, ACC or Air Traffic Service Providers. Complexity and performance are multidimensional values. Comparison is a difficult problem that must take into account the real meaning of each indicator, and its reliability.

Using fast time simulation is a convenient way to compute many different new metrics, some of them being probably highly interesting. But it is also a way to add a new layer containing a large collection of possible misinterpretations or even errors. For example, counting manoeuvres and giving each of them a cost, is as questionable as any

other method, as it introduces a resolution model which is not necessarily the one used by controllers, and introduces lot of parameters that could be discussed at length.

So the most extreme caution must be used when discussing performance metrics, and specially when using them to compare efficiency of different air traffic control systems.

This might sound obvious, but it is sometimes useful to write it down.

References

[Cha01]   Franck Chatton. Etudes de nouvelles métriques de complexité de la navigation aérienne. Master's thesis, Ecole Nationale de l'Aviation Civile, 2001.

[DAB97]   Nicolas Durand, Jean-Marc Alliot, and Jean-François Bosc. Cats, a complete air traffic simulator. In *Proceedings of DASC97*, 1997.

[DABM97] Nicolas Durand, Jean-Marc Alliot, Jean-François Bosc, and Lionel Maugis. An experimental study of atm capacity. In *Proceedings of Europe-USA conference on Air Traffic Management*, 1997.

[DAG01a] Nicolas Durand, Jean-Marc Alliot, and Geraud Granger. Optimal resolution of en route conflicts. In *Proceedings of ATM2001*, 2001.

[DAG01b] Nicolas Durand, Jean-Marc Alliot, and Geraud Granger. A statistical analysis of the influence of vertical and ground speed errors on conflict probe. In *Proceedings of ATM2001*, 2001.

[DP00]   D. Delahaye and S. Puechmorel. Air traffic complexity: toward intrisic metrics. Technical report, 3rd USA/Europe Air Traffic Management R&D seminar Napoli, June 2000.

[JT91]   M. Janic and V. Tosic. En route sector capacity model. *Transportation Science*, 25(4), 1991.

[LSBB98] L.V. Laudeman, S.G. Shelden, R. Branstrom, and C.L. Brasil. Dynamic density: An air traffic management metric. Technical report, NASA, April 1998.

[Tea02]   Working Together Team. System performance requirements document. Technical report, January 2002.

[WP302] WP30. The atm operational concept document, appendix a. Technical report, Air Traffic Management Operational Concept Panel, March 2002.