EUROPEAN ORGANISATION
FOR THE SAFETY OF AIR NAVIGATION

EUROCONTROL

EUROCONTROL EXPERIMENTAL CENTRE

Flight Delay Propagation

Synthesis of the Study

EEC Note No 18/03

M3·SYSTEMS

REPORT DOCUMENTATION PAGE

Reference EEC Note No.18/03	**Security Classification** Unclassified
Originator **M3 Systems**	**Originator Location** :
Sponsor EEC - NCD (**N**etwork **C**apacity and management **D**emand)	**Sponsor** (Contract Authority) Name/Location EUROCONTROL Agency Rue de la Fusée, 96 B-1130 BRUXELLES Telephone: +32-(0)2-729 90 11

Title :

Flight Delay Propagation
Synthesis of the Study

Authors Willy Vigneau	Date October 2003	Pages xii+33	Figs 21	Tables 4	Annex 2	References -
	Project PFE F EC		**Sponsor Task No.** -		**Period** -	

Distribution Statement :
    (a) Controlled by :   Head of NCD
    (b) Special Limitations (if any) :   None
    (c) Copy to NTIS :   No

Descriptors (keywords) :

Abstract :
This document presents the conclusions of the study of propagation of flight delay.

DOCUMENT CHANGE RECORD

The following table records the complete history of the successive editions of the present document.

EDITION	DATE	REASON FOR CHANGE	SECTIONS PAGES AFFECTED
1	June 2003		
2	3/7/2003	Application of the Eurocontrol common format	
3	16/09/03	Revision according to Eurocontrol comments	all
4	6/10/03	Revision according to Eurocontrol comments	all
5	13/10/03	Review of English language	all

TABLE OF CONTENTS

List of Tables and Figures

Executive Summary

The aim of this study, based on a partnership between Eurocontrol (Performance, Flow Management, Economics and Efficiency Business Area) and the local civil aviation authority (DAC Sud), was to understand the relation between the progression of a given aircraft on its daily schedule and its delay at successive station stops.

Repeated itineraries through the major French airports were identified within a reliable and large data set. Local delays models were derived for each airport. The local delays models were connected iteratively to assess the relation between the progression of an aircraft on its daily schedule and its delay at successive station stops.

The analysis at the airport level shows that :

✓ *Delays on arrival* could be predicted essentially from the delay on departure (80%) and to some extent from the plane load factor (16%).

✓ The airline whose itineraries have been studied allows a very small margin between the scheduled and the real flight and taxiing times, which does not allow absorption of any part of departure delays.

✓ *Delays on departure* could be predicted essentially from the departure plane load factor (70%) when the plane is not significantly delayed on arrival.

✓ On the other hand, there is a strong correlation between the delay on arrival and the next delay on departure.

✓ There is some evidence that the scheduled stop time plays an important role in absorbing the arrival delays. For instance, a 45 minutes scheduled stop time does not allow a delay recovery if the departure plane is 80% loaded. The shorter the station stop time, the greater the sensitivity of the delay on departure to the delay on arrival as shown in the Figure 0-1, here below.
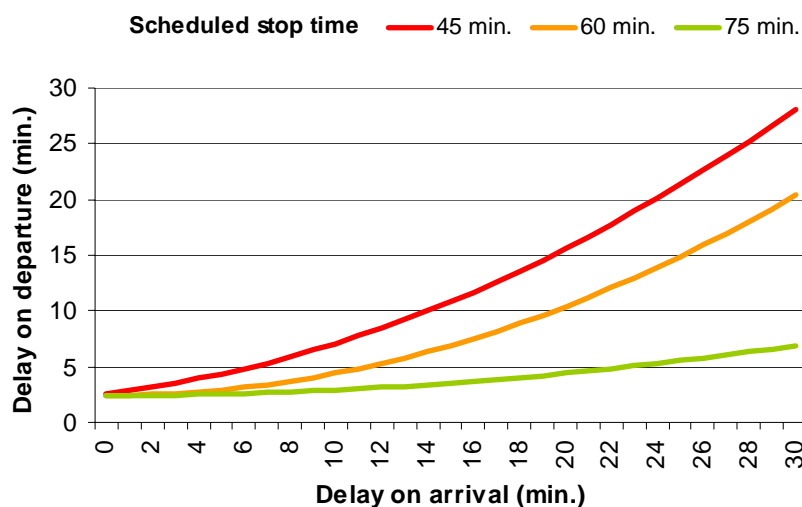
Figure 0-1 : Departure delay sensitivity to the delay on arrival

The average delay through stations is presented in the Figure 0-2, here below. It shows that :

✓ delays present a relative peak in the morning ;

✓ delays are partially recovered on departures from the 3rd and 4th stations ;

✓ in the evening, delays are continuously increasing up to the last station.
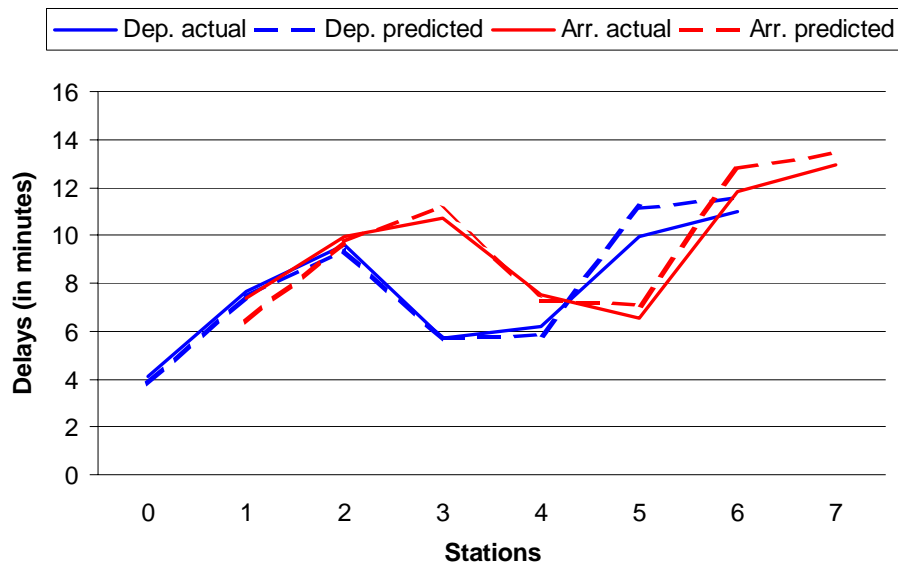
Figure 0-2 : Average delay at station rank

The analysis of the propagation pinpoints that :

✓ Flights experiencing long delays through their daily schedule face the occurrence of an ATFM regulation or an exceptional event, whatever its cause : airlines, airports or passengers.

✓ However, local conditions (ground airport operation elements) and propagation effect play an important role in maintaining a high level of delays : a flight experiencing a disruptive event or an ATFM regulation at a station is very likely to undergo, at the next station, a long delay due to the propagation and local factors.

✓ The propagated delays could also generate extra delays when they do not allow the flight to cope with an ATFM slot at the next station.

As shown in the Figure 0-3, here below :

✓ By the middle of the day (stations 3 and 4), the load factor is low and stop times are long. These favorable conditions allow the airline to minimize local and propagated contributors and the average delay curve reaches a minimum.

✓ Conversely, in the morning and in the evening, the demand for transport is high. Thus, planes are heavily loaded but the airline correlatively reduces stop times in order to increase the turnover. As a result, peaks in delays occur in the morning and in the evening since propagation and local effects sustain event delays or ATFM delays at a high level.
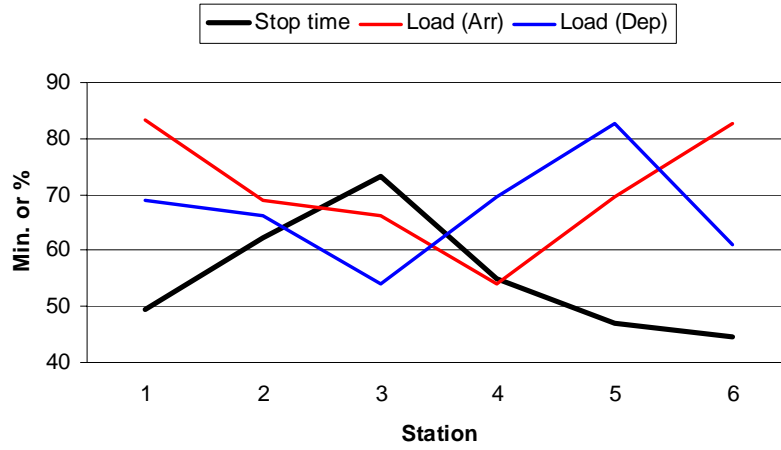
Figure 0-3 : Average stop time and aircraft load per station

	EUROCONTROL Propagation of Flight Delay	Ref: M3S/RP/ATM/DELAY/04 Date: 14/10/03 Issue: 1 Rev: 4	

M3·SYSTEMS

EUROCONTROL

1   CONTEXT AND OBJECTIVES

1.1   Context

In the commercial flight sector, delays are a source of great concern for the actors as they generate disruptions and costs for airlines, airport operators, ground handlers and eventually for passengers.

In order to decrease flight delays, it is important to identify their origins and causes as well as to understand their mechanisms of formation and propagation. This has been done extensively with ATC delays (due to ATFM regulation) through sundry studies. As a result, ATC delays are well known and Air Navigation Service Providers develop continuous efforts to reduce them.

It is current to split total delays into two types: ATC delays and non ATC delays. Non-regulated flights, when delayed, undergo exclusively non ATC causes (e.g. late arrival of the aircraft, plane breakdown, etc.). Non ATC delays causes are qualitatively known as the delay borne by a delayed flight is attributed to a cause listed in a standard IATA codification.

A brief look at non ATC delays causes shows that these delays are either related to causes which seem unpredictable (e.g. plane breakdown, mandatory security, non availability of aircrew, etc.) or associated with reactionary causes.

In a former study, conducted at Toulouse airport, M3 Systems in partnership with CS Communication & Systems studied the formation of departure delays in order to explain them by a set of parameters. This study proved that the non ATC delays were predictable and a predictive model of the overall departure delay was built.

Amongst the results, the study highlighted that, for non-regulated flights, departure delays are influenced by the following elements:

✓ The delay on arrival is responsible for the major part of the overall delay on departure for station stop times less than 60 minutes.

✓ The departure delay is also very sensitive to station stop time as a reduction of station stop time leads to a significant increase in the delay on departure.

✓ The pax load of the aircraft obviously plays a part in the departure delay through check-in, boarding and unboarding operations. To a lesser extent, terminal congestion (the number of passengers in terminal) also plays a role in delays.

1.2 Objectives

The conclusions of the previous study on departure delays explanations demonstrated that non regulated flights face delays notably because of the characteristics and organisation of airlines, airports and ground handling operations. As departure delays have an impact on delays on arrival which in turn generate delays on departure, it could be thought that for a given aircraft, airlines, airports and ground operations organisations, create and propagate delays during its daily schedule.

The present study aims at understanding the propagation of delays for a given aircraft through successive stations stops. In particular, the study aims at :

✓ Establishing a model for the relation between the progression of a given aircraft on its itinerary and its delay at successive station stops.

✓ Assessing the impact of major contributors, regulations and non-traffic dependent causes of delays on the propagation process.

2   METHODOLOGY [1]

The methodology is broken down into three steps.

The **first step** consists in sorting out the available data and identifying a set of repeated itineraries and airports upon which it will be possible to establish local models.

In a ***second step***, a local delay model will be developed for each airport. The model differentiates delays on arrival and departure:

- ✓ the delay on arrival will be described as a function of the delay on departure from the opposite station and a set of local parameters ;

- ✓ the delay on departure will be defined as a function of the delay on arrival and a set of local parameters including the scheduled station stop time, the load of the aircraft and the local capacity.

Local models are built in order to assess the non-ATC delay at each station as a function of the former delay and local parameters. They are represented by the following generic equations:

$$DLD = f_D \text{ (DLA, Local parameters)}$$

$$DLA = f_A \text{ (DLD, Local parameters)}$$

Where DLD and DLA represents respectively the delays on departure and arrival at a given airport.

The ***third step*** consists in devising a general model through which it will be possible to assess the relation between the progression of the aircraft on its itinerary and its delay at successive station stops.

At first, the local models are connected iteratively to each other which brings out a delay model for non ATC delays.

In addition to the delays propagation study, the ***general model*** highlights the impact of major contributors to delay propagation. In particular, the assessment of overall delay shows the impact of regulation and non-traffic dependant causes of delays on the overall itinerary process.

---

[1]) Definitions and detailed methodology are in annex I.

3 DATA [2]

3.1 Sample definition

This model has been developed from a large sample of data which come from sundry reliable sources:

✓ Slots were extracted from the French national archives.

✓ Local ATS on airports provided data for throughput and capacity.

✓ Air France provided flight records data.

Initially, a sample of some 45,000 flight of A319 and A320 records between September 16th and December 13th of 2002, was available. In order to produce accurate and unbiased base on which relevant delays models could be built, this sample has been processed:

- Week-ends, legal holidays, days with industrial actions or any major failure in ATC systems have been discarded.

- Flight data were provided by a single airline. In order to avoid variability due to heterogeneous module sizes, information for an homogenous type of aircraft was kept. These information dealt with 102 specific aircraft (A319 and A320 type).

This initial process resulted in the definition of some 325 repeated itineraries going through seven stations or more, presented in the table below.

Itinerary id.	STATION #								Number of cases	Cumulative frequency
	0	1	2	3	4	5	6	7		
1	MRS	ORY	MRS	CDG	TLS	CDG	TLS	CDG	45	4%
2	ORY	SXB	CDG	MRS	CDG	SXB	ORY	SXB	28	6%
3	ORY	TLS	ORY	SXB	ORY	MPL	ORY	MRS	23	8%
4	MRS	CDG	LYS	ORY	MRS	ORY	MRS	ORY	22	10%
5	ORY	TLS	ORY	TLS	ORY	TLS	ORY	TLS	22	11%
6	TLS	ORY	LYS	CDG	BOD	ORY	TLS	ORY	22	13%
7	NCE	ORY	BIA	ORY	MLH	ORY	BES	ORY	21	15%
8	SXB	ORY	TLS	CDG	TLS	ORY	TLS	ORY	20	17%
9	MRS	ORY	MRS	ORY	TLS	ORY	TLS	ORY	20	18%
10	ORY	SXB	ORY	MPL	ORY	BES	ORY	MLH	19	20%
…	…	…	…	…	…	…	…	…	…	…
36	ORY	TLS	ORY	BES	ORY	TLS	ORY	BOD	10	50%
…	…	…	…	…	…	…	…	…	…	…
325	ORY	MLH	ORY	MRS	ORY	MLH	ORY	BES	1	100%
All									1 225	100%

Initially, 2,228 individual itineraries covering the major French airports, with an average of 6.3 stations per itinerary, were identified. In order to assess propagation at late station stops,

---

[2] Details on data and their processes can be found in annex II.

1225 individual itineraries (composed of 7 station stops or more) have been extracted from the previous itineraries sample. These itineraries are represented on the figure here below :
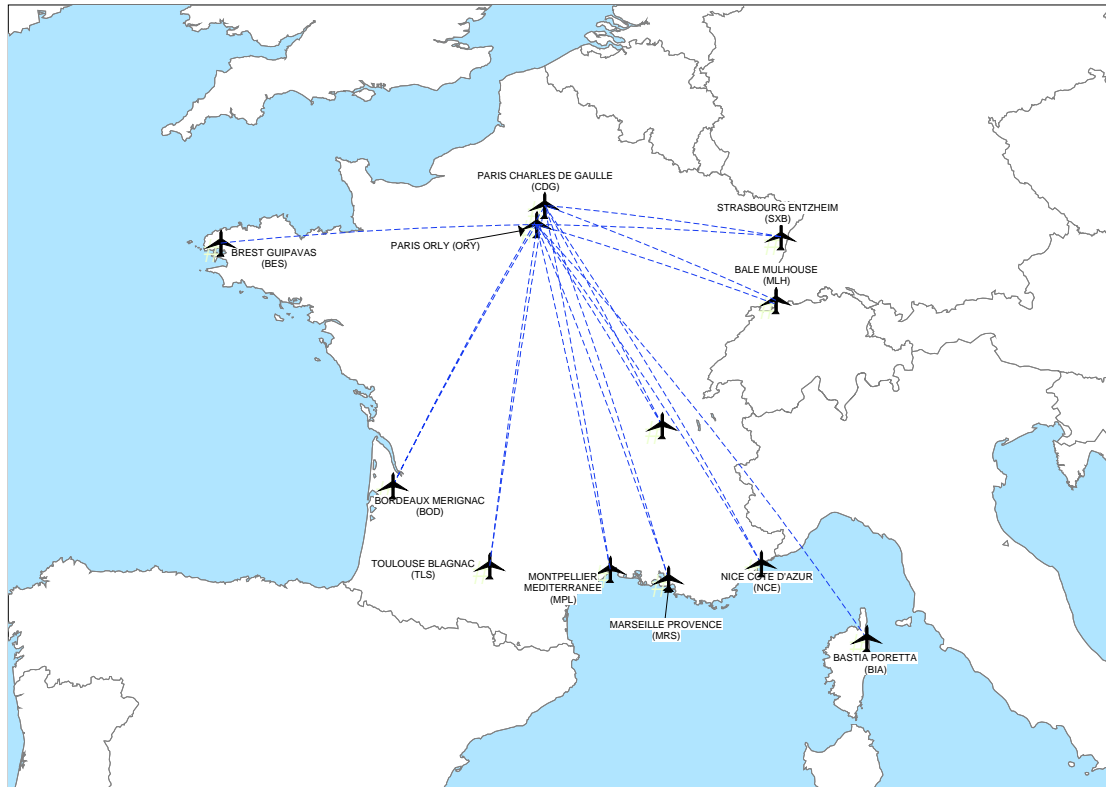
Figure 3-1 : Individual itineraries illustration

For further applications, 9 standard itineraries have been selected as a sub-sample i.e. 223 individual itineraries and more than 1,500 turnarounds.

With regard to the number of the data available (high enough to represent the behaviour of the entire population), the quality of the processed samples and the reliability of the data sources, there is strong confidence that the data statistical processes produce unbiased and reliable results.

3.2 Progression of delays through stations

As a first approximation, the assumption that the flight delay is only a function of the station number has been assessed. In that case, the flight delay can be expressed by the following expression:

$$Delay = f(n) \text{ where n designates the number of the station.}$$

In order to evaluate this assumption, the average delays across stations were calculated from the 1,225 individual itineraries defined before. These results are shown by the diagram below:
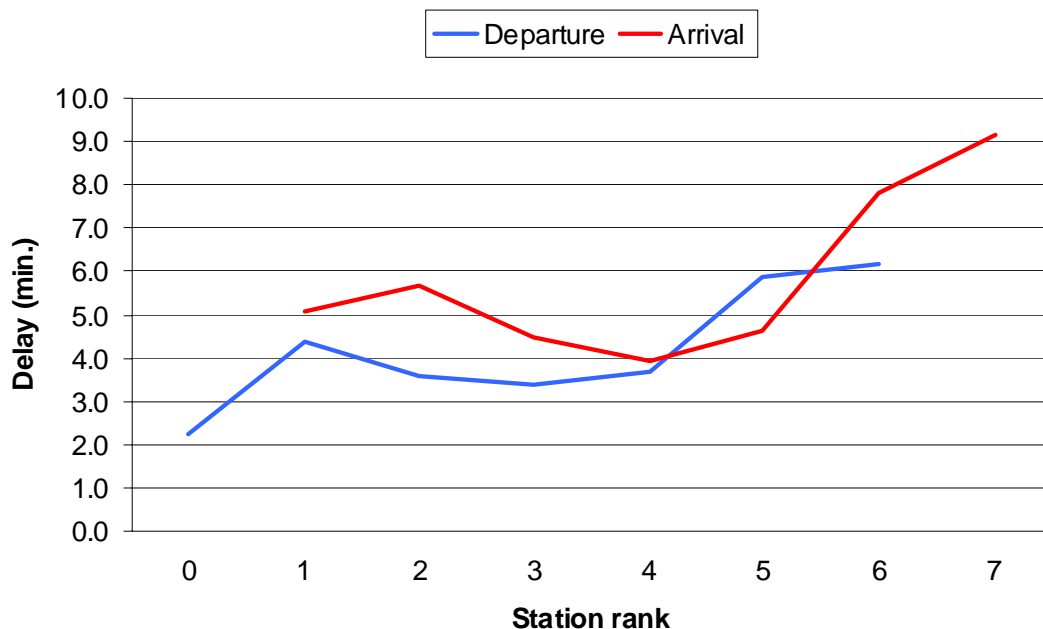
Figure 3-2 : Average delay at station rank

From the average delay at station rank diagram, one can note the following features:

The relationship between the rank of the station and the average arrival or departure delays is not linear. There is a relative peak in the morning followed by an improvement of punctuality in the middle of the day after which delays are continuously increasing up to the last station. It could be noticed that these average delays curves are strongly correlated with the repartition of the overall flight traffic throughout a day.

The curve for the delay on arrival is strongly correlated with the curve of the delay on departure with a double shift :

✓ The average delay on arrival at a station is always greater than the delay on departure from the previous station. This shows that the actual flight duration is, on average, longer than the expected or announced one.

✓ In the same way, the average delay on departure from a station is generally smaller than the delay on arrival at that station. This implies that airport operations do not generate on average additional delay.

However, a deeper analysis demonstrates a very high variability of delays at each station rank. For example, even if as an average airport operations do not generate extra delays, it appears that punctuality may be deteriorated during an airport stop at a specific station.

As a first conclusion, it appears that it is not possible to build a delay propagation model with the station rank as a single explanatory variable. Some more complex propagation models shall be studied and analysed. Those models are described in the following sections.

4 LOCAL MODELS

4.1 Overview

As shown in the previous section, the study of delay propagation needs the development of local delay models for each airport. The aim of this chapter is to define and present, for each airport involved in the itineraries presented in chapter 2, the local models. Those models consist of predictive estimates of arrival and departure delays.

4.1.1 Preliminary approach

A preliminary first approach for the study of the progression of delays through stations consists in comparing average delays values and in particular the values of the average arrival and departure delays for each station. This approach is summarised in the following chart below showing the delays parameters between respectively Paris/Orly and Roissy/CDG airports and the other French airports.

The values in the circles represent the difference between the real average flight duration value and the estimated flight duration value. The values in the rectangular labels under the names of the stations account for the difference between the average delay on arrival and the average delay on departure (nights stop flights being put apart).
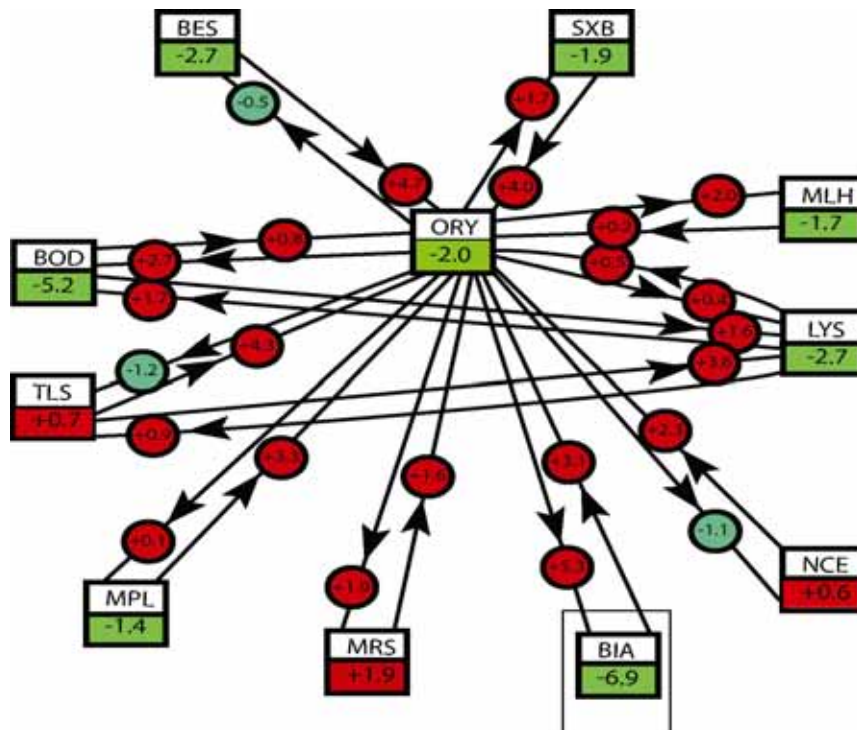

Figure 4-1 : Flights to/from Paris-Orly (ORY)
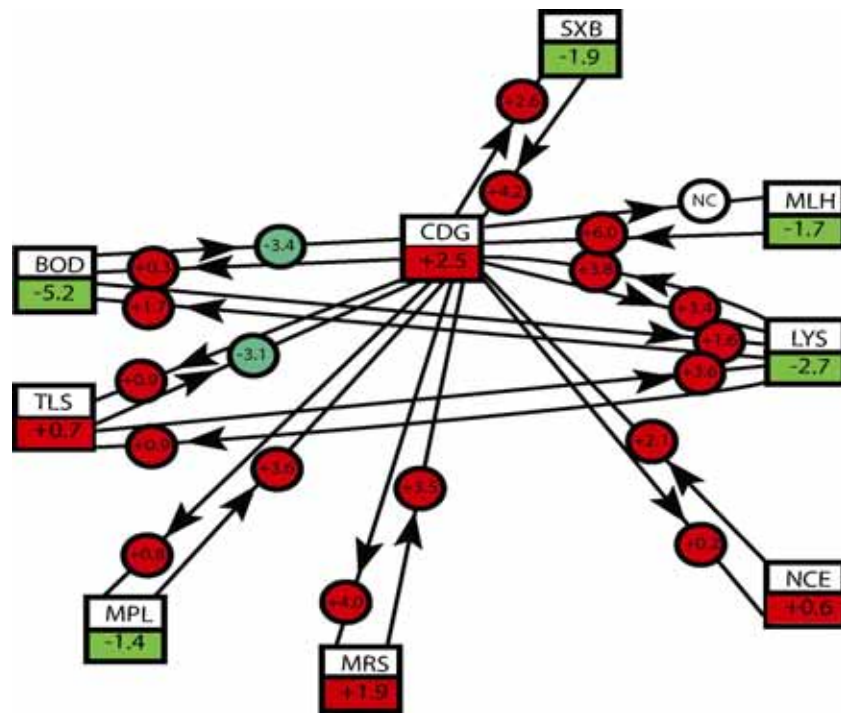
Figure 4-2 : Flights to/from Paris-Charles de Gaulle (CDG)

These charts could be used to find out the average delays on arrival and departure at a specific station. For example, if a plane departs 5 minutes behind its schedule from BOD to CDG, its delays at Roissy airport is on average 2,6 (=5-3,4)min.

If a plane arrives 5 minutes behind its schedule at BOD, it will depart on average 0,2 (=5-5,2) minutes in advance from BOD.

These charts also pinpoint the following results:

- ✓ The influence of the flight duration on the arrival delay as the real flight duration is usually higher (up to 6 minutes) than the expected or announced one. This means that some delays are generated because airlines have on average underestimated the real flight duration. This underestimation could be easily corrected and would cause an automatic decrease in delays.

- ✓ The influence of the station stop time on the departure delay. On average, most airports absorb delays during the aircraft stop. However, airports such as Paris-CDG, Nice, Marseilles, Toulouse do not. It could be noted that, with the exception of Roissy, those airports are the ones from which Air France operates a shuttle from/to Paris/Orly. The generation of delays during the stop of an aircraft indicates :

  - o Either that the average station stop time is not correctly adjusted to handle the airport operations (unboarding, checking-in, boarding) on a proper time.

o Or, if the average stop time is correctly adjusted, that the passengers checking and boarding, aircraft and ramp handling or flight operation and crewing processes generate delays.

When using this first average approach as a model to estimate the delays at stations, it appears that this model is not relevant as a predictive tool. Hence, **local models giving a more accurate assessment of propagation** have to be developed.

4.1.2 Local Models Overview

Based on the previous experience on the Toulouse airport prediction delay model, the local models have been built upon the following parameters:

Arrival models	Departure models
Delay on departure (from previous)	Delay on arrival (at station)
Day of the week	Day of the week
Hour (within day)	Hour (within day)
Origin (station)	Destination (station)
Capacity (*)	Capacity (*)
Load on arrival (**)	Load on arrival (**)
	Load on departure (**)
	Scheduled station stop time

(*) The capacity – explicitly, the local capacity – is a conventional measure of restrictions to runway and approach airport throughput. This value is drawn from recorded events related to airfield disorders (e.g. runway closing due to deterioration of the runway, accident on runway, annual repairs, object on runway, etc.) or to a limitation of the throughput due meteorological events such as storms, low visibility conditions, etc.

(**) The load of the aircraft represents the rate of the actual number of passengers to the number of seats offered. For instance, a plane loaded at full capacity has a 100 (%) load. The load value is 50 for a standard A320, with 159 seats, when boarding 80 passengers. In departure models, we consider separately the load on arrival, with possible incidence on unboarding operations, and the load on departure, which is likely to affect check-in and boarding operations.

4.2 Arrival models

When testing the arrival models against the previous variables, one comes up with some relationships between those variables which are summarised in the following table (Constant terms and minor parameters do not appear in those tables):

	TLS	ORY	CDG	NCE	LYS	MRS	MLH	BOD	SXB	MPL	BES	BIA
N	1924	5677	1743	1285	806	1463	381	826	522	459	351	105
R²	0.850	0.816	0.902	0.884	0.850	0.844	0.874	0.857	0.891	0.821	0.983	0.858
DLD_{-1}	0.949	0.969	0.960	0.998	0.926	0.950	0.983	0.965	1.004	0.974	1.001	0.934
LOAD	0.012	0.034	0.019	-	0.046	-	-	-	-	-	0.025	-

Table 4-1 : Arrival models estimation

The **N line** represents the number of occurrences from which the results are drawn. Those numbers show that local models are being devised on large samples which provide confidence in the overall results. This confidence is corroborated by the R² values (82 to 98 %) which show the relevance of the arrival models.

The lower lines provide the coefficient for the major variables (DLD_{-1} –delay on departure at the previous stations - and LOAD) which represent the delay on departure from the previous station and the load of the aircraft. In order to illustrate the use of the local models, let us consider a flight having undergone a 10-minute delay on departure from any previous station and arriving at Paris-Orly (ORY) loaded at full capacity (100 %). The predicted delay on arrival (DLA) is :

$$DLA = -1.3 + 0.969 \times \mathbf{10} + 0.034 \times \mathbf{100} = 12 \text{ minutes } (^3)$$

This example shows clearly that the real flight duration is not consistent with the scheduled flight duration, as expected from timetables, because, if it were, the delay on arrival would be the same as the one on departure.

This remark applies for all the stations for which the DLD_{-1} coefficient is close to but different from 1.The difference between the delay on departure and the one on arrival accounts for the difference in the scheduled flight duration and the real one.

When developing the analysis, it appears that the delay on departure from the previous station accounts for 80 % of the variability of the delay on arrival .

As shown by the chart here below, the other major variable which impacts the delay on arrival is the load of the aircraft. This variable accounts for 15 % of the variability of the delay on arrival. In the example above, the negative effect of the load exceeds in absolute value the constant term and results in an increase of the flight duration by 2 minutes.

(^{3}) In this example, we do not consider minor effects. Each local model includes a slightly negative constant term : here –1.3 minutes.
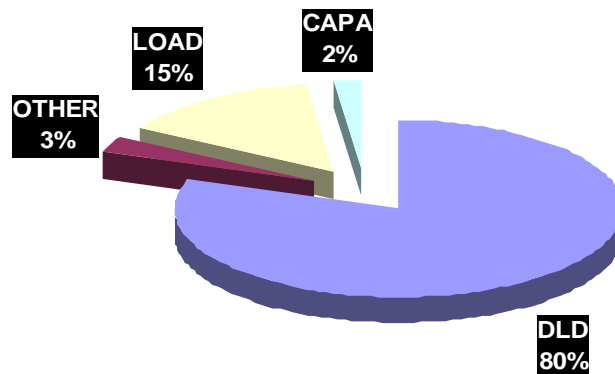
Figure 4-3 : Delays on arrival parameters

The correlation between the delay on arrival and the delay on departure is expected since it is difficult to make up the departure delay during a flight which lasts approximately one hour. Therefore, if airlines do allow any margin between the scheduled and the real flight and taxiing times, the departure delay ends up as an arrival delay of approximately equal value.

The impact of the plane load variable on the arrival delays could be explained in two ways. Firstly, the more a plane is loaded, the more time is needed to handle the airport ground handling (passenger, crew and plane) processes. Secondly, the lighter the plane, the higher the speed and the shorter the flight time.

There is a minor role of the capacity in the arrival delays. It could be explained in two ways. On the one hand, adverse events having caused ATFM regulations, like meteorological events, en-route or terminal restrictions, might have an impact on some non regulated flights. On the other hand, some moderately deteriorated situations do not generate a regulation, especially when the expected traffic is low, but may cause minor disruptions in the flow on arrival.

4.3 Departure models

When testing the sundry variables in order to build the departure model (corrected from slot and unpredictable event ([4])), it appears that the departure delay is sensitive to the delay on arrival. However, the mechanisms whereby the delay on arrival propagates through a station is highly dependent on the station stop time. In order to take into account this particularity, three departure delays models, depending on the station stop time, have been devised :

- ✓ Short station stop time (40-50 minutes)
- ✓ Medium station stop time (55-70 minutes)
- ✓ Long station stop time (>75 minutes)

([4]) About unpredictable events, see the definition of traffic dependency in annex I.

The major parameters of the models are summarised in the tables below. Constant terms and minor parameters do not appear in those tables.

4.3.1 Models

Station	TLS	ORY	CDG	NCE	LYS	MRS	MLH	BOD	SXB	MPL	BES
N	1 031	1 183	138	663	189	627	128	285	191	229	238
R²	0.583	0.625	0.647	0.651	0.708	0.681	0.654	0.792	0.691	0.752	0.498
DLA	0.265	0.246	0.371	0.258	-	0.258	0.186	0.117	0.265	0.250	0.261
DLA²	0.014	0.018	0.023	0.016	0.026	0.020	0.015	0.009	0.013	0.016	0.010
STOP TIME	-0.158	-0.450	-	-0.399	-	-0.148	-	-	-	-0.177	-
LOAD ARR	0.026	0.023	-	-	-	0.046	-	0.023	0.062	-	0.021
LOAD DEP	0.100	0.129	0.100	0.112	0.088	0.124	0.116	0.051	0.116	0.105	0.096

Table 4-2 : Short station stop time

Station	ORY	CDG	LYS	MRS	BOD	BIA
N	1230	68	265	74	106	53
R²	0.392	0.592	0.663	0.496	0.507	0.557
DLA	0.043	-	-	-	-0.158	-
DLA²	0.009	0.020	0.014	0.013	0.020	0.011
LOAD DEP	0.057	0.127	0.054	0.086	0.061	0.077

Table 4-3 : Medium station stop time

Station	TLS	ORY	CDG	NCE	LYS	BOD
N	138	1062	161	41	25	122
R²	0.489	0.260	0.346	0.597	0.452	0.192
DLA	-0.070	0.015	-	-0.209	-	-
DLA²	0.004	-	-	0.010	-	0.002
LOAD DEP	0.080	0.070	0.094	0.047	0.068	0.050

Table 4-4 : Long station stop time

4.3.2 Factors contribution to the average flight delay

Slots and non-traffic dependent causes apart, the delay causes in the studied sample are the load factor which accounts for 80 % (the load on departure alone accounting for 70 %) of the departure delay and the delay on arrival which represents only 20 %. This result shows that if a flight is not lengthily delayed on arrival the most constraining factor is then the load of the aircraft.
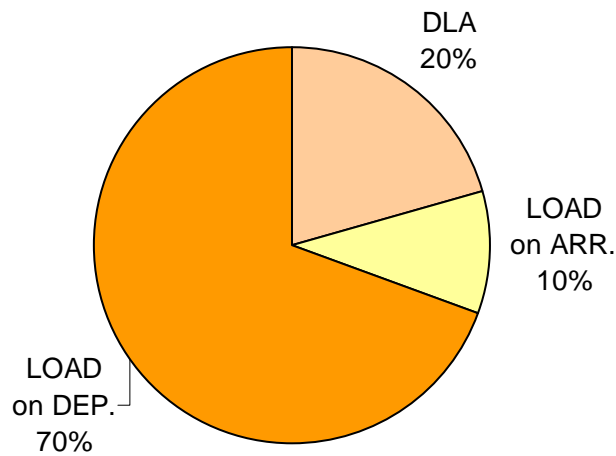
Figure 4-4: Delays on departure parameters

4.3.3 Sensitivity of the delay on departure to factors

When putting together a local model to predict the departure delay from sundry variables, a number of assumptions will have to be made and these assumptions will lead to a particular outcome in terms of delays and indicators. One important feature of the model is its stability and reliability to the variables which could be studied through a sensitivity analysis. Sensitivity analysis involves a systematic variation of the key delay causes and assumptions within a specific range so that the effect on the outcome and the model stability and reliability can be studied.

4.3.3.1   *Sensitivity to the delay on arrival*

The sensitivity analysis is firstly performed on the delay on arrival variable and the result is shown in chart 5. For the calculation of the sensitivity analysis, the load of the aircraft has been fixed 80 %, slightly above the average value (near 70 %).
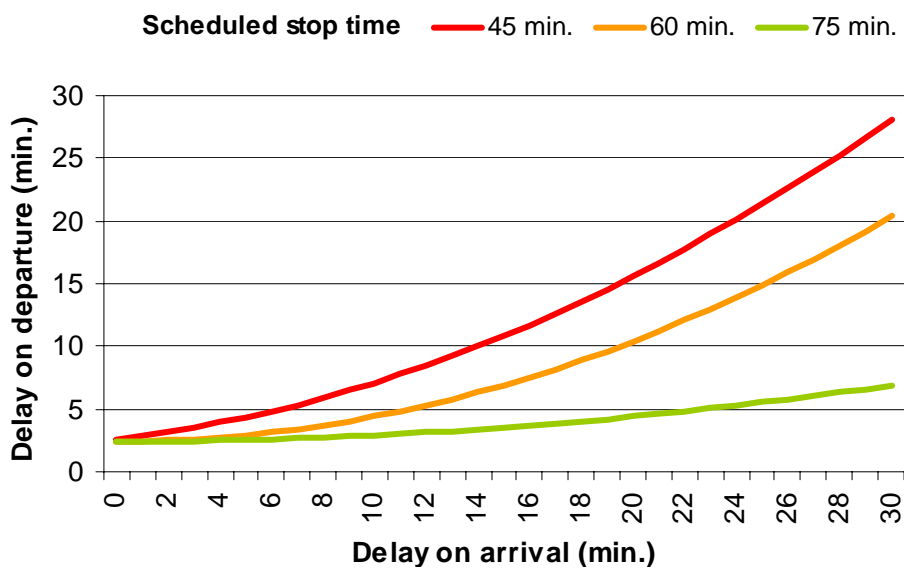

Figure 4-5 : Departure delay sensitivity to the delay on arrival

As expected, the shorter the scheduled station stop time, the greater the sensitivity of the delay on departure to the delay on arrival. In particular, it could be seen that in the case of a station stop time equal to 75 minutes, the arrival delay is almost totally absorbed during stop time as the departure delay varies from 3 minutes to 7 minutes (on 30 minutes arrival delay).

In case of a station stop time equal to 45 minutes, the delay on arrival is all the more absorbed that the delay on arrival is low. As the delay on arrival increases (above 20 minutes), the arrival delay is almost integrally transferred on the departure delay. This result shows that if airlines schedule station stop time with no margin to the expected ground operations time, the station stop time will be too tight to absorb part of the arrival delay and then delays will propagate.

4.3.3.2   *Sensitivity to the load factor*

In order to be as comprehensive as possible, the sensitivity of the departure delay to the load of aircraft has been calculated with several hypothesis on the values of schedule stop time and arrival delay.
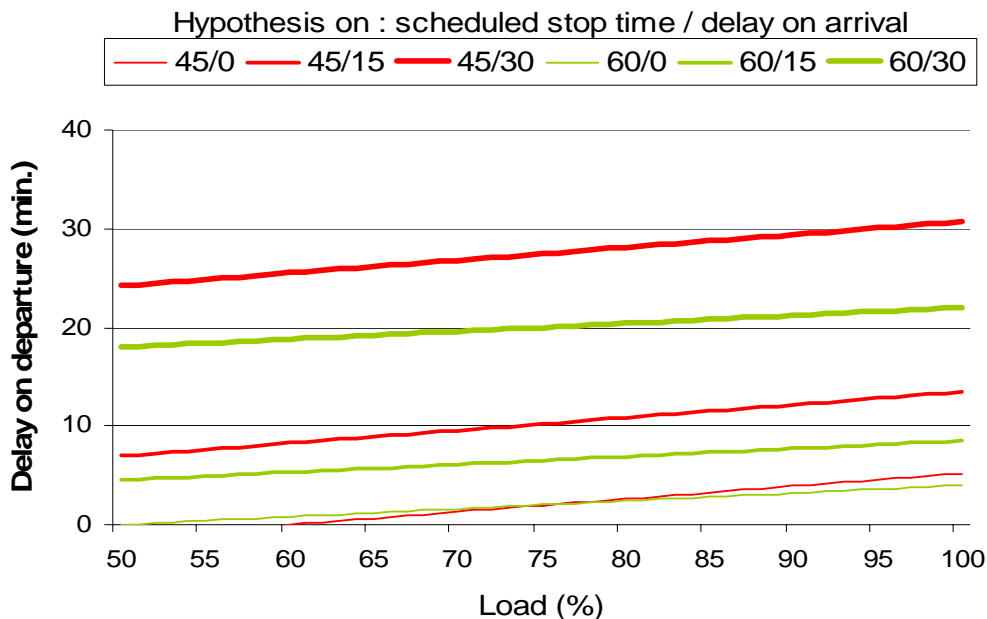

Figure 4-6 : Departure delay sensitivity to the aircraft load

The chart shows that whatever the hypothesis on the scheduled stop time and arrival delay values, the impact of passing from a 50% to a 100% load increases the departure by a maximum of five minutes. The value of this increase is approximately the same whatever the hypothesis which means that the ground processes impacted by the aircraft load are independent from the other ground processes.

5   GENERAL MODEL

5.1   Overview ([5])

The general model aims at explaining the progression of delay through stations. It is built from an algorithm connecting in an iterative manner the local models in order to account for the propagation of delays through stations.The general model is operated for every flight and takes into account all flight delays whose causes range from local causes (the ones studied in the previous chapter: delay on arrival, on departure, load of aircraft, etc ..) to ATFM regulation (ATC slots) and exceptional events (runway closure, aircraft failure or defect, etc.).

The input of the general model are the following:

- ✓ A sequence of stations.

- ✓ A set of local parameters associated to each station (especially the station stop time and the load of the aircraft).

- ✓ A set of possible delays due other causes than local ones. These delays are composed of exceptional delays and ATFM regulations.

The effects of ATFM slots and exceptional events are taken into account by a rule which states that a slot or an exceptional event alters the predicted delay if the resulted delay is higher than the delay caused by the propagation or the local effects. In order to use this model as a predictive model, it is necessary to appraise its quality.

5.2   Model performance

The performance and the quality of the general model are appraised by confronting the difference between real and predicted values of delays along sequences of stations as shown in the figure here below:
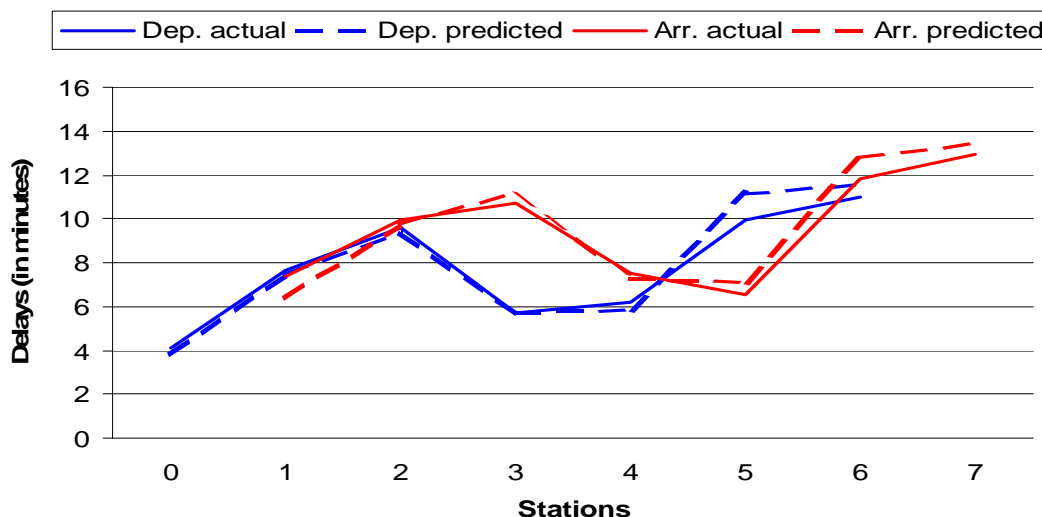

Figure 5-1 : Performance of the general model

([5]) details and explanations about the general models can be found in annex II.

Figure 5-1 represents for all of the 223 daily flights tested in the general model the predicted and the real delays. The predicted profile is based on a simulation from the general model i.e. :

✓ The iterative prediction of the propagated delay based on the actual values of stop time and load of the aircraft.

✓ The actual sequence of slots for each individual flight.

At stations, the predicted values deviate from the real values by 1 to 12 % for departures and by 2 to 14 % for arrivals.

Through the seven stations, the standard deviation is 0.5 minutes for departures and 0.6 minutes for arrival. These results ensure that this model presents a correct accuracy in the prediction of flight delays.

6 STUDY OF THE PROPAGATION

6.1 Preliminary step

As shown in section 3.2, the average delays across stations all over the different itineraries are represented by the following pattern :
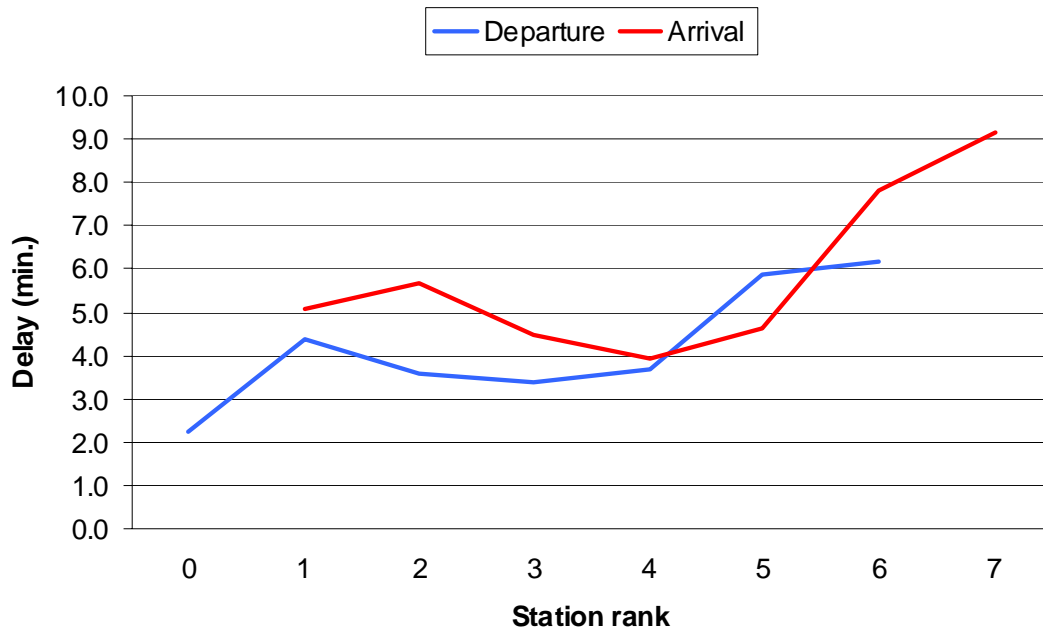

Figure 6-1 : Average delay at station rank

However, in using the general model, it appears that the individual values of delays are very heterogeneous for a specific station rank which leads to different delays propagation profiles. In order to find out the rules of propagation, it was then necessary to identify homogenous profiles regarding the progression of the overall delay through stations. Eventually, six homogenous profiles were identified from the actual delays. Those propagation patterns through stations (0) to (6) correspond to the most widespread patterns and are described and presented below with the sequence of delays on departure.

Even if the patterns are very different from each other, it could be noticed that the sum of these patterns correspond to the curb of the average delays above.

6.2 Categorisation of profiles

a) Increased delays at the end of the day (40,5%)

The patterns n°1 and 5, correspond to situations through which delays increase at the end of the itineraries. However, it could be noted whereas the delays increase continuously throughout the day and successive stations in n°1 pattern, delays skyrocket at the end of the day in the pattern n°5 scenario (13,5 % of occurrence).
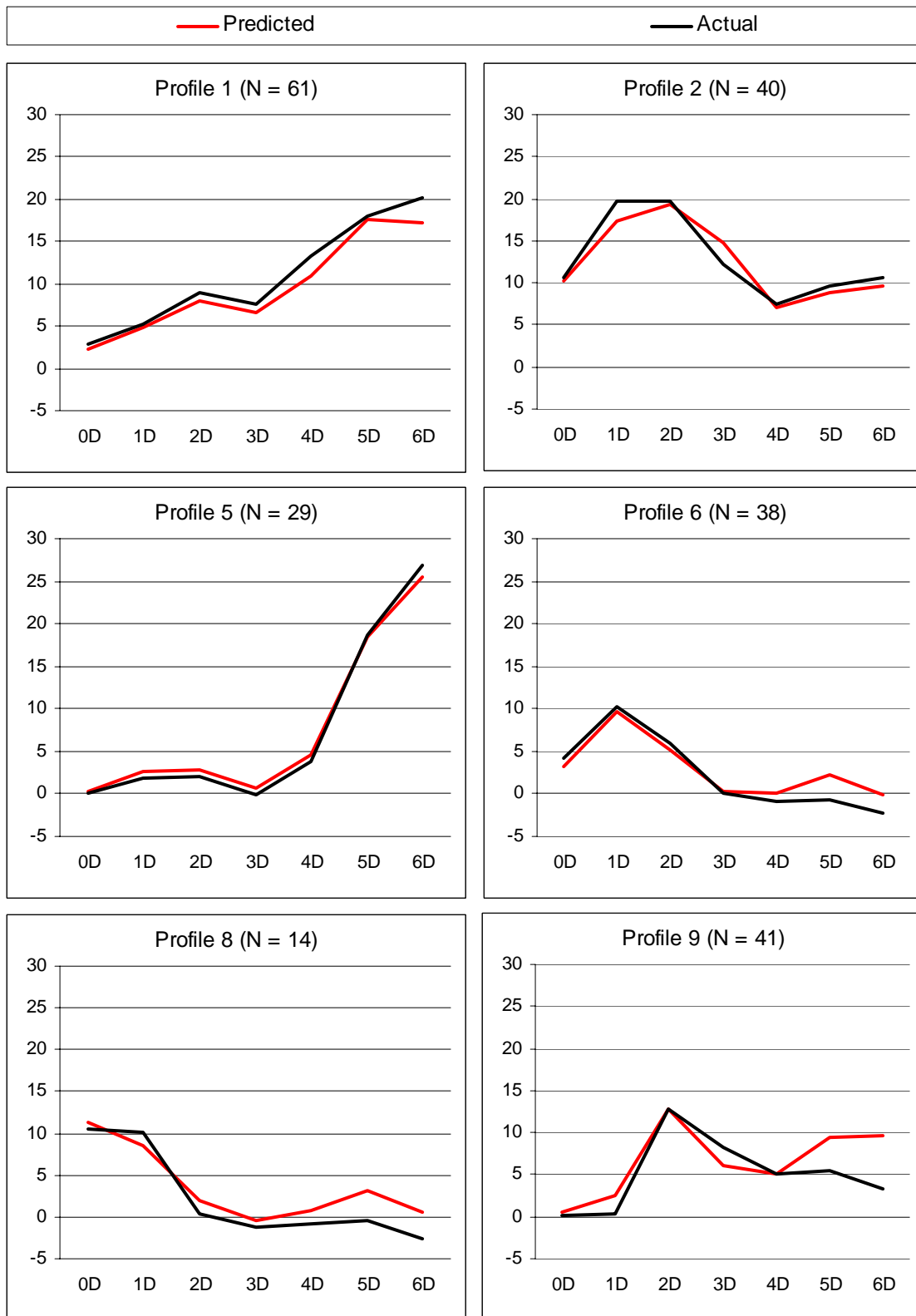
Figure 6-2 : Propagation of delay profiles

b) Peak during the morning rotation

Profiles n°2, 6, 8 and 9 have similar patterns with some delay peaks during the morning. The delay peaks may reach *average* values of 10 to 25 minutes and they differ from one profile to another by their position (first, second or third station) and their amplitude.

In any case, there is no delay peak in the middle of the day (through stations 3 and 4).

6.3   Major contributors to delay

In order to measure the contribution of the propagation effect on the overall delays, the values of each delay contributor have been assessed through the computation of the predicted delay on departure. The delays contributors have been classified as follows:

- ✓ *Propagated :* Drawn from the local model, this contributor represents the part of the departure delays which propagate. This contributor is made of the sum of effects for the delay on arrival and the station stop time.
(usually, $A * DLA + B * DLA^2 - C * STOP\_TIME$) [6]

- ✓ *Local :* This contributor represents the sum of all the other delays effects than the propagation contributor of the local model. In general, the major value of this local effect is the load factor.

- ✓ *Event :* the part of the computed effect of non-traffic dependent events higher than the delay predicted by the local model. Event delays are depending on a restricted number of causes which range from airports causes (mandatory security, airport facilities, etc.) to airlines (non-scheduled maintenance, aircraft defects, etc.) and passenger causes (passenger convenience, VIP, etc.) which seem essentially unpredictable.

- ✓ *Slot :* the part of the computed effect of a slot higher than the delay predicted by the local model.

The diagrams below provide examples of what could be achieved with the general model. Those diagrams represent the progression of delays and the proportion of the major contributors for two individual itineraries.

---

[6] Where DLA is the delay on arrival.
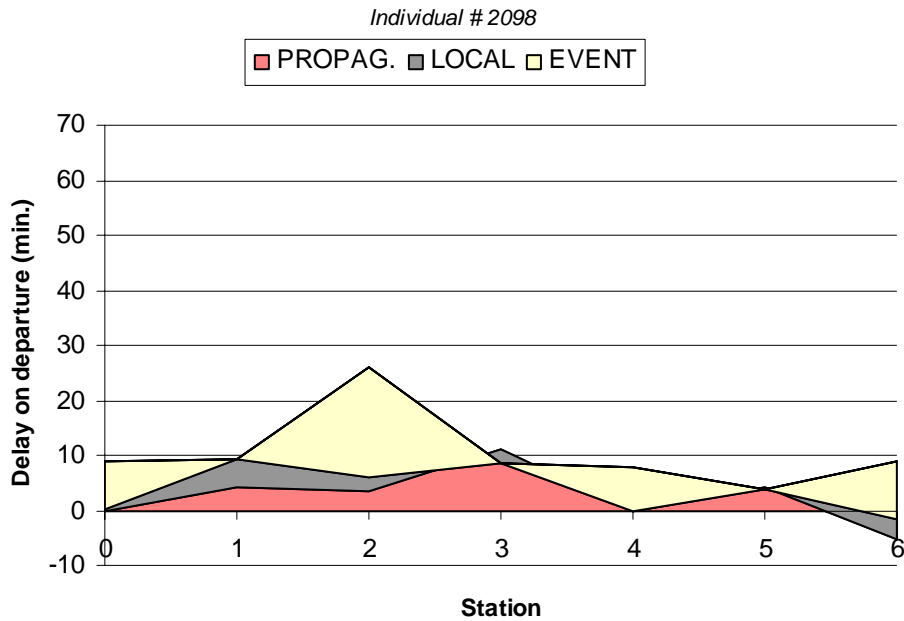
Individual # 2098

Figure 6-3 : Progression of delays – Example n°1

Through this itinerary, the plane undergoes successive events at stations. In that case, the causes are airport facilities, mandatory security (26 minutes), aircraft defects and de-icing. If considering the IATA code, the 10-minute predicted delay at station 3 is correctly imputed to the late arrival of the aircraft.

It should be noted that the propagated and local delays are calculated values (positive or negative), whereas the contributions of events and slots to the overall delay cannot be negative.
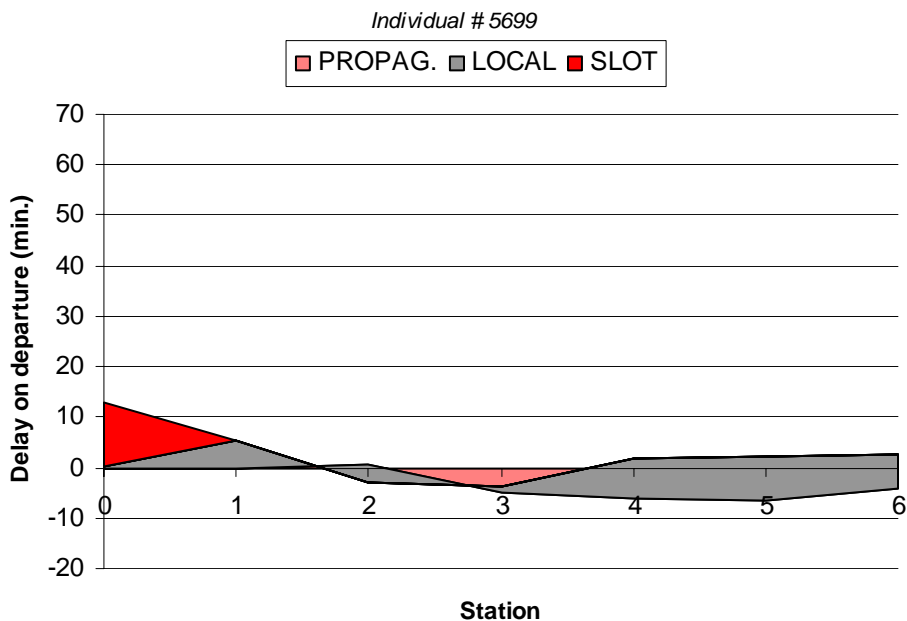
Individual # 5699

Figure 6-4 : Progression of delays – Example n°2

The above itinerary is submitted to a 12-minute slot on the first departure in the morning. Through the next stations, the sum of the propagated and local component stands near zero and the flight is on time.

A few exceptions put apart, a flight cannot be simultaneously submitted to a slot and to an event at the same specific station.

6.4   Case studies

In order to understand the propagation mechanism, some delays cases have been picked up and described below. Those delays cases correspond to the following scenarios :

✓   Continuous delay increase through the day

✓   Inflation of delays at the end of the day

✓   Single-station peak in the morning

✓   Several-stations peak in the morning

6.4.1   Continuous delay increase through the day

Figure 6-5 : Continuous delay increase through the day
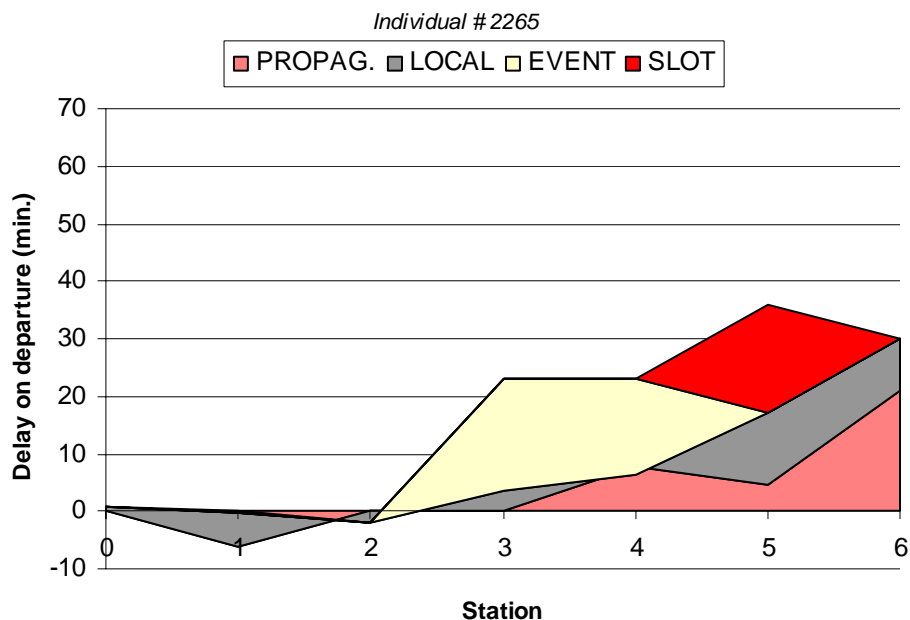
As shown in the diagram above, an event occurred at station 3 : the flight had to wait for the cabin crew from another flight. The flight took off from the 4th station at 14:18 with a 23-minute delay due to aircraft defects.

The arrival at the 5th station, scheduled at 15:05, was delayed by 12 minutes. Thus, 11 minutes were recovered through a 70-minute scheduled flight.

The scheduled stop time being 45 minutes, the flight was not able to achieve its initial EOBT (15:50) and had to declare a change. It was then submitted to a slot at 16:46. Eventually, the flight left blocks from the 5th station at 16:26.

It is to be noted that this flight departed from stations 3 and 4 with approximately the same overall delay but the sum of the local and propagated components at station 5 was more than a two-fold its value at station 4. This illustrates the importance of the local factors (load, scheduled stop time).

The last departure (from station 6) was delayed by some 30 minutes. Propagated and local factors were entirely responsible for this delay. It is noticeable that the actual delay was not 30 minutes (predicted) but 53 minutes : in the hurry, a passenger was missing at boarding !

This case illustrates how the occurrence of two successive events at the third and the fourth station due to airlines and passengers causes has resulted in a constraining slot on departure from the fifth station, because propagation did not allow the flight to depart at its scheduled EOBT. The 35-minute ATFM regulation delay has propagated to an extent of approximately 20 minutes at the sixth station which faced a departure total delay of 30 minutes because of the load factor combined with the propagation effect.

6.4.2   Inflation of delays at the end of the day
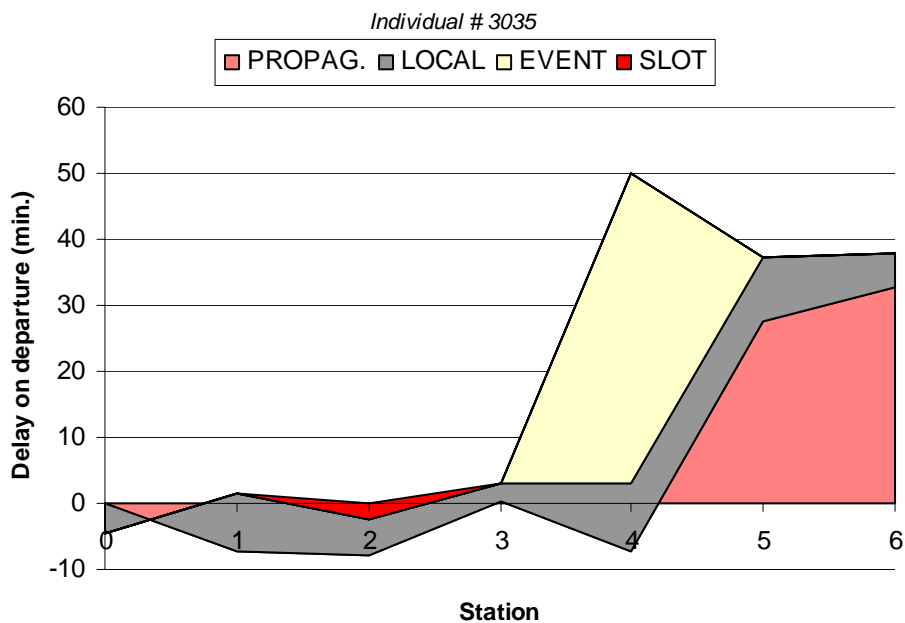

Figure 6-6 : Inflation of delays at the end of the day

This case shows how a specific event delay is maintained during the next two stations because of the propagation effect or local causes. The long delay at station 4 was caused by "commercial publicity or passenger convenience (VIP, press)". At station 5, a lack of staff for loading / unloading is supposed to be responsible for the real delay (close to 60 minutes) but the model gives evidence that a delay close to 40-minute delay was predictable from the propagation and the local conditions of traffic. The long delay at station 6 is, consistently with the model, attributed to the late arrival of the aircraft.

From this special case, it is noticeable that the propagated and local components alone may cause delays through two consecutive stations when following an event (or a slot) which caused a very long delay. This process is particularly true when the event occurs at the fourth or the fifth station since the stop time and the aircraft load conditions combined with the propagated effect do not allow any recovery of delays during evening stations.

6.4.3 Single-station peak in the morning
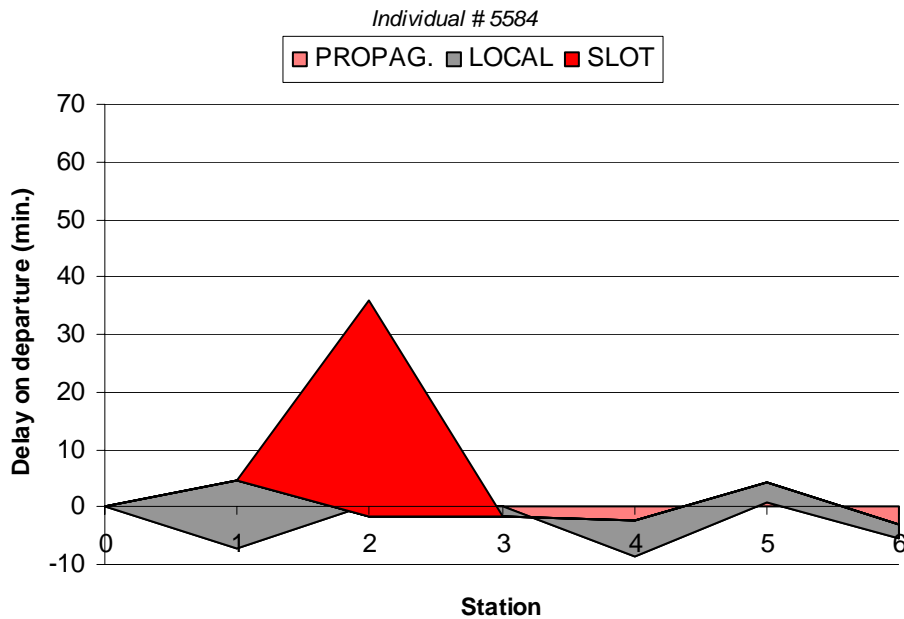
Figure 6-7 : Single-station peak of delay in the morning

This case provides an explanation of the decrease of the average delay at station n°3 on the average delay curve. The flight was submitted to a severe slot at station 2. Nevertheless, it was absolutely on time through all of the next stations.

This is explained by the low load factor on departure (30 %) value combined with the high scheduled stop time (80 minutes) at station 3. This case also demonstrates how airlines can absorb, by the middle of the day, delays generated in the morning itineraries.

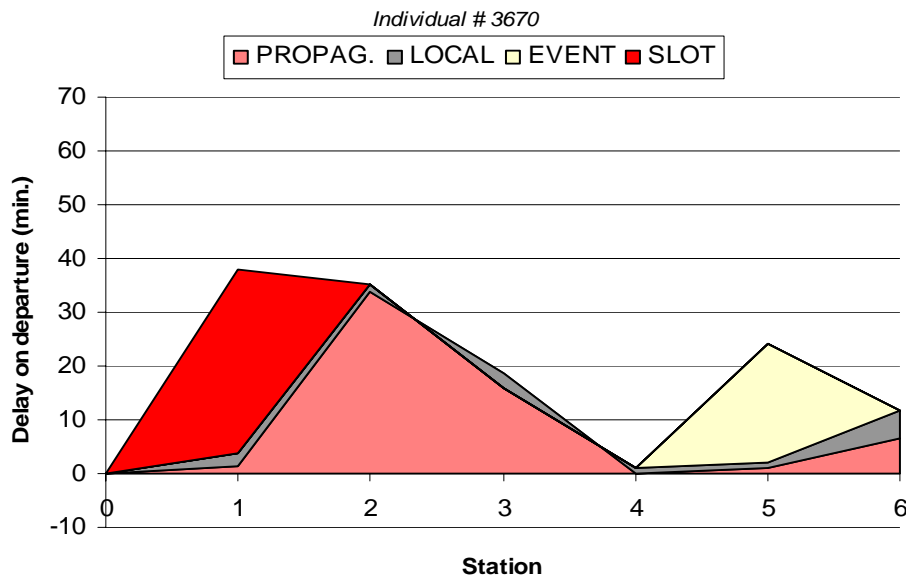### 6.4.4 Several-stations peak in the morning

Individual # 3670

Figure 6-8 : Several-stations peak of delay in the morning

This flight was submitted to a slot at station 1. The propagation resulted in a long delay at station 2. The recorded cause was, consistently with the model, the late arrival of the aircraft. The model over-estimates the delay at station 3 (5 minutes real delay instead of 20 minutes). However, the delay is still attributed to the late arrival in the IATA codification.

At station 4, the load (65 %) and the scheduled stop time (80 minutes) values allow a large recovery of the former predicted delay.

This case shows that, even if an initial delay generated by a slot or any event propagates through the morning, the propagation may be curbed by the middle of the day because of the favourable loads and scheduled stop times values.

6.5 Conclusions on propagation

✓ **Short delays**

Consistently with the COMUTA results, it has been found that some 25 % of flights are delayed by 3 to 14 minutes. The simulations provided by the general model proved that some 60 % of short delays are not attributed to an exceptional event or an ATFM slot. Short delays result mainly from the propagation of a former delay and/or the local conditions (load, scheduled stop time). It has also been proved that propagation and local effects alone cannot reach values up to 15 minutes if they do not result from an event or an ATFM slot.

✓ **Long delays**

Profiles with long delays show the occurrence of an ATFM regulation or a non-traffic dependent event ([7]). Non-traffic dependent events are weakly or not correlated to the local traffic (passengers and/or aircraft) as for instance : aircraft failure or mandatory security.

Nevertheless, the local conditions and the propagation effect play an important role in maintaining a high level of delays. Firstly because the application of the propagation and local effects on an event or an ATFM regulation at a station result in a delay which could trigger the generation of an ATFM slot at the next station. This slot may be the most constraining factor and determine the actual block time but it is a secondary effect of propagation and local factors. Secondly, even if the flight is free from regulation at its new EOBT, it is very likely to undergo a long delay due to the propagated and local contributors alone. This is particularly true when an event or an ATFM regulation occurs during the latest stations of an itinerary where the local conditions (short stop time, high aircraft load) and the propagation effect are at a maximum.

Conversely, by the middle of the day (stations 3 and 4), favourable conditions allow airlines to minimise local and propagated contributors : minimum load, maximum stop time as shown in the following chart.
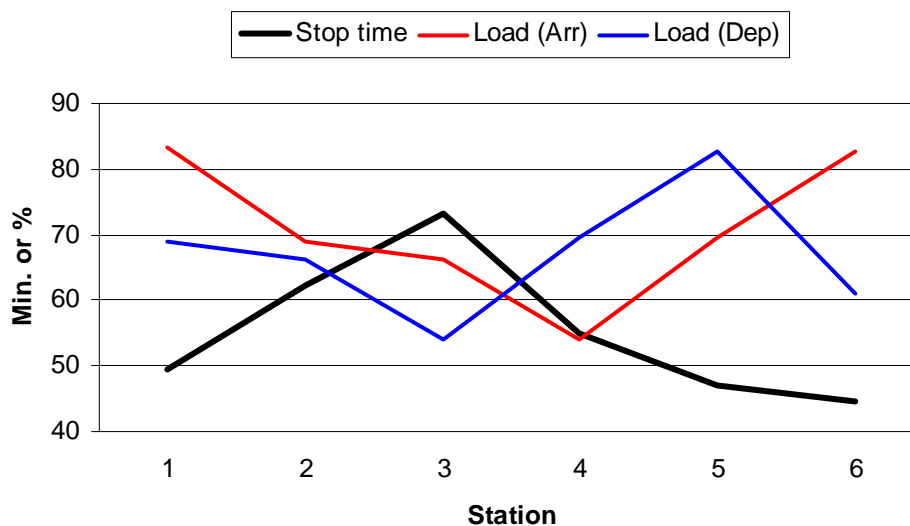

Figure 6-9 : Stop time and aircraft load per station

As a consequence, flights experiencing long delays in the morning may absorb the propagation of delays by the middle of the day (stations 3, 4) whereas the propagation and the local effects sustain the level of event or ATFM delays in the evening until the night stop.

This result explains why 40 % of delayed flights by 15 minutes or more are caused by propagation combined with the local conditions effects.

It should be underlined that the sum of the propagated and local contributors may reach values up to 40 minutes. Since the maximum effect of the local factors does not exceed 10 minutes, it is clear that the propagation plays the major part in the 40% flights delayed by more than 15 minutes.

([7]) In delays above 15 minutes, non-traffic dependent events (35 %) are more frequent than slots (25 %) but slots generate longer delays (35 minutes) than non-traffic dependent events (30 minutes).

7   CONCLUSIONS

7.1   Arrival and departure delays prediction

One of the major causes of airlines operations disruption is the occurrences of delays. The delays can be classified in two categories : ATFM delays and non-ATFM delays.

Non-ATFM delays include exceptional events whose causes seem unpredictable. The exceptional event delays have various sources such as passengers (e.g. VIP), airport (e.g. mandatory security) or airlines (e.g. plane technical breakdown). As a common feature, they are weakly correlated to the overall traffic. In addition to these exceptional events, the organisation of the airlines' ground operations (unboarding, checking, boarding of passengers, plane handling, etc..) could create delays or at least propagate them during a plane daily journey.

In order to study the impact of the ground airline operations, local delay models have been defined. From these models, it has been possible to establish that :

- ✓ Delays on arrival could be predicted essentially from the delay on departure (80%) and to some extent from the load factor (16%).

- ✓ Delays on departure could be predicted essentially from the departure plane load factor (70%) when the plane is not lenghtily delayed on arrival.

- ✓ The airline whose itineraries have been studied allows a very small margin between the scheduled and the real flight and taxiing times, which does not allow the absorption any part of departure delays.

- ✓ The scheduled station stop time plays an important role in absorbing the arrival delays. As an example, the sensitivity analysis shows that a 45-minute scheduled station stop time does not allow a delay recovery if the departure plane is 80% loaded.

- ✓ The passengers checking and boarding operations generate an average departure delay even for long and medium station stops which implies that these operations are not optimized regarding the departure time.

7.2   The propagation of delays

The link between departure and arrival delays indicates that a certain amount of delay is propagated. In order to study the progression of delays through stations, a general model has been built. This model takes into account all flight delays whose causes range from local causes (delay on arrival, on departure, load of aircraft, etc ..) to ATFM regulation (ATC slots) and exceptional events (aircraft failure or defect, mandatory security, etc.). The study of the general model has allowed the following results to established :

✓ As long as exceptional events and ATFM slots are not operating, propagation and local conditions (ground airport operation, plane load, etc.) alone do not reach values up to 15 minutes.

✓ Profiles with long delays show the occurrence of an ATFM regulation or a non-traffic dependent event, whatever its cause : airlines, airports or passengers.

✓ The local conditions (ground airport operation elements) and the propagation effect play an important role in maintaining a high level of delays. As a matter of fact, a flight experiencing a disruptive event or an ATFM regulation at a station is very likely to undergo a long delay due to the propagated and local contributors alone. This effect is especially important during the latest stations of a daily itinerary where the local conditions (short stop time, high aircraft load) and the propagation effect are at a maximum.

✓ The propagated delays could also generate extra delays if the propagated delay does not allow the flight to cope with an EOBT at the next station.

✓ By the middle of the day (stations 3 and 4), favorable conditions allow airlines to minimize local and propagated contributors : minimum load, maximum stop time. This explains why the average delay curve reaches a minimum at station 3 and 4. As a consequence, flights experiencing long delays in the morning may curb the propagation of delays by the middle of the day (stations 3, 4) whereas the propagation and the local effects sustain the level of event or ATFM delays in the evening until the night stop.

ANNEX I : METHODOLOGY

DEFINITIONS

<u>Individual itinerary :</u> Sequence of stations along a single day for a single aircraft.

<u>Standard itinerary (i):</u> Repeated pattern of individual itineraries.

<u>Station :</u> Any stop within an itinerary. Stations are labelled with an index (j). Depending on the itinerary, the (j) station may be any airport but station (i, j) refers to a particular airport. Index (j) starts at zero on every new day. The "0" value refers to the departure of a <u>night stop flight</u> (first departure within a special day).

<u>Segment :</u> Any sequence of two consecutive airports along a particular itinerary.

<u>Non-traffic dependent causes (of delay) :</u>

This term refers to any cause of delay which is weakly or not correlated to the local traffic (passengers and/or aircrafts). For instance : aircraft failure or mandatory security.

In the former study (Toulouse, 2001), these causes had concerned some 4 % flights and thus had been called "exceptional". The current study shows a higher frequency (8 to 10 % flights), reaching at some stations much higher values (36 % at Paris – CDG). Hence, the exceptional characteristic cannot be put forward.

The clustering of flights into "dependent" or "non-dependent" causes is based on the delay IATA cause recorded by the operator for each delayed flight.

A very clear example of "<u>traffic-dependent</u>" cause is "late check-in". First, it is very likely that each passenger, considered separately, may be associated to a given and somewhat constant risk of being late at check-in. Hence, the global risk for a given flight of a delay due to "late check-in" is at some extent proportional to the number of passengers, i.e. to the load. On the other hand, the consequences of a late check-in are probably worse when the plane is heavily loaded, due to congestion of the check-in area, stress and workload for agents and tense check-in to boarding process.

BASIC RELATIONSHIPS

<u>Delay on departure :</u>

$DLD(0) = f_D$(local parameters at station (0))   i.e. night stop flights

$DLD(j) = f_D(DLA(j)$, local parameters at station (j)),   when j>0

<u>Delay on arrival (j $\geq$ 0) :</u>

$DLA(j) = f_A(DLD(j-1)$, local parameters at station (j))

LOCAL MODELS

f_A and f_D are linear functions associated to local models. Each station stop (e.g. Paris-CDG, Paris-Orly, Toulouse-Blagnac, etc.) is characterised by a set of linear functions, usually :

- ✓ one f_A function ;
- ✓ several f_D functions, one for night stop flights and 3 for other departures, depending on the station stop time.

f_A and f_D functions are (local) expressions of the local models i.e. for a given airport.

Regarding to the local parameters, we have two types of variables. On the one hand, variables related to traffic are measurements of positive flows or restrictions associated to air transport, regarding to aircrafts (capacity) or passengers (load) and indicators of the steady-state of flows (delay on the former segment and station stop time). These variables are quantitative.

The local capacity is a conventional measure of restrictions to runway and approach capacity. It is based on the analysis of recorded events related to airfield disorders (e.g. annual repairs) and meteorological events (e.g. low visibility).

On the other hand, the study at Toulouse in 2002 showed that significant corrections should be applied depending on the opposite station (e.g. destination when considering departures), day of the week and scheduled hour ([8]) which are categorical variables.

GENERAL MODEL (RESTRICTED)

From the local models previously introduced, a restricted general model for delays formation along the aircraft trajectory will be built as :

$$^{\wedge}DLA(j) \quad = f_A(\quad ^{\wedge}DLD(j-1), \text{ local parameters at station } (j))$$

$$= f_A(\quad f_D(^{\wedge}DLA(j-1), \text{ local parameters at station } (j-1),$$

$$\text{local parameters at station } (j))$$

$$= G(\quad \text{local parameters at stations } 0,1,\ldots, j-1, j)$$

The G function is an expression of the general model. Under that form, the general model accounts for the variability of the overall delay due to non-ATC and traffic dependent causes. It is a so-called restricted model as it does not.

---

[8] An explanation of delays regarding functional indicators and, especially, flows (traffic level, capacity) is needed but some qualitative factors are expected to affect delays. An in-depth analysis could help identifying quantitative descriptors (e.g. related to staff, aircrafts, etc.) of mechanisms underlying differences between stations, but in this study the opposite station is considered as a "black box". With regard to the day of the week and to the scheduled hour, the possible effects are supposed not to be correlated with traffic and capacity. In other words, they should not be redundant with quantitative variables depicting traffic and capacity. For instance, a correction applied to a particular day or scheduled hour should be borne out by some commercial (type of passengers, e.g. holidays, job) or organisational cause (e.g. meal time) and not by the average level of traffic during this period.

ANNEX II : DATA COLLECTION AND STATISTICAL PROCESS

DATA SOURCES

The following data sources have been analysed :

- ✓ AFR database :
    - o Flight and aircraft identification
    - o Date and hour of the flight (scheduled, actual)
    - o Delay cause when available
    - o Number of passengers
- ✓ French National Archives : Regulations
- ✓ Local ATS on selected airports : data for capacity

SAMPLE DEFINITION

Flight data were provided by a single airline (Air France). In order to avoid variability due to heterogeneous module sizes, information for an homogenous type of aircraft were requested : 102 aircrafts A319 and A320 were involved.

Initially, some 45,000 records were available, between September 16th and December 13th of 2002. Week-ends, legal holidays, days with strikes or any major failure in ATC systems have been discarded.

STANDARD ITINERARIES

Individual itineraries were selected as far as :

- ✓ All the data, at each station stop along the itinerary, were available.
- ✓ They did not include any event altering the predictability of propagation (no setting up of an empty plane).
- ✓ They did not include one or several station(s) out of France (availability of data for the assessment of the local capacity).

This selection provides 2,228 individual itineraries covering the major French airports, with an average of 6.3 stations per itinerary.

In order to assess propagation at late station stops, individual itineraries composed of 7 station stops or more (up to 8) have been considered : hence, 1225 individual itineraries were used.

Out of those 1,225 individual itineraries, 325 "standard" itineraries were identified : no strong regularity in the daily sequences of stations has been observed.

LOCAL MODELS

We were seeking predictive expressions of the delays on arrival and on departure, previously defined as :

Delay on departure :

DLD(0) = f$_D$(local parameters at station (0))

DLD(j) = f$_D$(DLA(j), local parameters at station (j)), j>0

Delay on arrival (j $\geq$ 0) :

DLA(j) = f$_A$(DLD(j-1), local parameters at station (j))

This task requires a statistical calibration, based on recorded parameters and a convenient sample of flights for each considered airport :

> ➢ On arrival : all flights
> ➢ On departure, flights altering the predictability shall be discarded. Thus are selected:
>> ▪ Non regulated flights (no slot) ;
>> ▪ Flights which do not undergo a non-traffic dependent cause of delay (regarding to the IATA code).

The local parameters considered in the models are :

Arrival models	Departure models
Delay on departure (from opposite)	Delay on arrival (at station)
Day of the week	Day of the week
Hour (within day)	Hour (within day)
Origin	Destination
Capacity	Capacity
Load on arrival	Load on arrival
	Load on departure
	Station stop time

GENERAL MODEL

TEST POPULATION

Tests were performed on 223 (aircraft x day) from 9 standard itineraries as described in section 3.

ALGORITHM

For each individual (aircraft x day), we started at station (0) and predicted the delay on departure. Through the next stations, sequentially, we computed the delay on arrival (this computation, through the local model, including the predicted delay on departure from the previous station), then the delay on departure (the latter taking into account the computed delay on arrival).

The delay on departure is derived from the local model. However, as expressed above, the general model did not take into account the effect of slots and non-traffic dependent events. In order to include these effects, we proceeded as follows :

We computed the traffic-dependent delay, including effects of propagation (delay on arrival and station stop time) and other local effects (due to load, local capacity and corrections). This computation was strictly based on the local models.

If the flight was submitted to a slot, we estimated the EOBT involved by the slot (^EOBT) from a standard taxiing time (CFMU data). The gross delay predicted from the slot was then :
        ^EOBT – scheduled time. If this value exceeded that of the actual delay on departure, the latter was considered to be the delay due to the slot. This correction is necessary because a rough analysis of our file confirms that slots allocated by CFMU and reported in the PLN file are usually improved.

If the flight was undergoing a non-traffic dependent cause of delay, the actual delay on departure was considered to be the effect of the reported event. The highest value out of those three was considered to be the predicted delay on departure.