



E.02.06-POEM-D6.2-Final Technical Report

Document information

Project Title	POEM
Project Number	E.02.06
Project Manager	University of Westminster
Deliverable Name	Final Technical Report
Deliverable ID	6.2
Edition	01.01.00
Template Version	03.00.00

Task contributors

University of Westminster
Fundación Instituto de Investigación Innaxis

Abstract

Central to the POEM (Passenger-Oriented Enhanced Metrics) project is the design of new performance metrics and their evaluation through a European network simulation model under novel flight and passenger prioritisation scenarios. A normative day with full passenger itineraries is simulated. Trade-offs between the (new) flight-centric and passenger-centric metrics are explored. The propagation of delay through the network is characterised using classical and complexity science techniques. The importance of using passenger-centric metrics in fully assessing system performance is repeatedly observed, since key changes are not expressed through any of the currently-used flight metrics. Most prioritisation scenarios perform similarly for high cancellation and high delay days, thus demonstrating robustness in terms of their efficacy under increased disruption. Smaller airports are significantly implicated in the propagation of delay. Back-propagation is also an important characteristic of the persistence of delay in the network.

Authoring & Approval

Prepared By - <i>Authors of the document.</i>		
Name & Company	Position & Title	Date
Andrew Cook / University of Westminster	Project Leader	05/08/2013
Graham Tanner / University of Westminster	Consortium Member	
Samuel Cristóbal / Innaxis	Consortium Member	
Massimiliano Zanin / Innaxis	Researcher	

Reviewed By - <i>Reviewers internal to the project.</i>		
Name & Company	Position & Title	Date
Graham Tanner / University of Westminster	Consortium Member	08/08/2013
Samuel Cristóbal / Innaxis	Consortium Member	

Reviewed By - <i>Other SESAR projects, Airspace Users, staff association, military, Industrial Support, other organisations.</i>		
Name & Company	Position & Title	Date
-	-	-

Approved for submission to the SJU By - <i>Representatives of the company involved in the project.</i>		
Name & Company	Position & Title	Date
Andrew Cook / University of Westminster	Project Leader	10/08/2013

Rejected By - <i>Representatives of the company involved in the project.</i>		
Name & Company	Position & Title	Date
-	-	-

Rational for rejection	
-	

Document History

Edition	Date	Status	Author	Justification
01.00.00	10/08/2013	Deliverable	Andrew Cook	New document for review by EUROCONTROL
01.01.00	30/09/2013	Final Deliverable	Andrew Cook	References to other POEM outputs made more explicit

Intellectual Property Rights (foreground)

Foreground owned by one or several Members or their Affiliates.

Table of Contents

EXECUTIVE SUMMARY	7
ACKNOWLEDGEMENTS	9
1 INTRODUCTION.....	10
1.1 PURPOSE OF THE DOCUMENT	10
1.2 INTENDED READERSHIP	10
1.3 EUROPEAN UNION POLICY CONTEXT	11
1.4 INPUTS FROM OTHER PROJECTS	11
2 SUMMARY AND UPDATE OF POEM.....	12
2.1 OVERVIEW AND OBJECTIVES	12
2.2 SUMMARY OF THE METRICS.....	13
2.3 SUMMARY OF THE PRIORITISATION SCENARIOS.....	14
2.4 SUMMARY AND UPDATE OF THE MODEL	16
2.4.1 Core model.....	16
2.4.2 Traffic inputs for three simulation days.....	21
2.4.3 Assigning passengers to flights	30
2.5 MODEL CALIBRATION	37
3 RESULTS OF THE MODEL	39
3.1 THE CHALLENGES OF ESTABLISHING CAUSALITY	39
3.2 CLASSICAL RESULTS.....	43
3.3 FACTOR ANALYSIS (DERIVED METRICS)	54
3.3.1 Review and update of the factor analytical approach.....	54
3.3.2 Applying factor analysis to the model data.....	56
3.4 NETWORK AND COMPLEXITY ANALYSES.....	64
3.4.1 Introduction	64
3.4.2 Reactionary delay and the propagation of delay.....	66
3.4.3 Granger causality analyses.....	73
3.5 COMPARATIVE INSIGHTS FROM THE ANALYTICAL METHODS	81
4 BEYOND POEM – A LOOK AHEAD	83
4.1 DISSEMINATION	83
4.2 FURTHER RESEARCH.....	83
5 REFERENCES.....	84
APPENDIX A KEY MODEL RULES AND VARIABLES	87
APPENDIX B FINAL DATA PREPARATION AND CLEANING.....	115
APPENDIX C FACTOR ANALYSIS METHODOLOGY	120
APPENDIX D PASSENGER VALUE OF TIME.....	129
APPENDIX E BASELINE TRAFFIC DAY DISRUPTION.....	132
APPENDIX F FURTHER ANALYSES OF AIRPORT REACTIONARY DELAY	133

List of tables

Table 1. Summary of prioritisation scenarios.....	15
Table 2. Rules and scenarios.....	15
Table 3. Cancellation event types.....	22
Table 4. Article 5 (cancellation), Regulation 261	23
Table 5. Daily cancellation rates for 2012 (full year c.f. September; 2 s.f.)	25
Table 6. Example allocation of additional cancellations	29
Table 7. Identified passenger allocation issues	31
Table 8. SEP10 passenger breakdown by flight leg and allocation proportion targets	31
Table 9. 17SEP10 passenger allocation targets.....	32
Table 10. Connection options for a three-flight leg itinerary	34
Table 11. Passenger connectivity file – inflexible passengers (extract)	36
Table 12. Key model output calibration values	37
Table 13. Correspondence between rules and case study inputs	38
Table 14. 17SEP10 allocated passengers compared with allocation targets.....	38
Table 15. AEA load factors	38
Table 16. Airport typologies – overview	39
Table 17. Airport typologies – by interaction level	41
Table 18. Core metric set.....	43
Table 19. Summary of scenario and simulation day comparisons	45
Table 20. Core metric results – baseline traffic day.....	47
Table 21. Core metric results – baseline traffic day (main effects).....	48
Table 22. Standard deviation of core metrics – dispersion under scenario A ₂	50
Table 23. Core metric results – high cancellation day	51
Table 24. Core metric results – high delay day.....	52
Table 25. Core metric results – overall summary	53
Table 26. Average departure delay per time period by worst ten airports (baseline traffic day)	57
Table 27. Baseline delay components (highest component loadings; from model input data)	58
Table 28. Baseline delay component correlations (from model input data).....	59
Table 29. Component loading frequencies (from model input data).....	60
Table 30. Key inter-component delay propagation nodes (model input data)	61
Table 31. Component loading frequencies (from model output data)	62
Table 32. Key inter-component delay propagation nodes (from model output data).....	63
Table 33. Highest dozen airport propagation pairs	71
Table 34. Highest dozen airports in terms of back-propagated delay	72
Table 35. Granger causality axioms	73
Table 36. Summary of key complexity metrics by scenario and layer.....	77
Table 37. Percentage of links that are different between network layers	78
Table 38. Node ranking by in-degree and network layer	78
Table 39. Node ranking by out-degree and network layer.....	78
Table 40. Node ranking by eigenvector centrality and network layer	79
Table 41. Comparison of rankings by network layers for eigenvector centralities.....	79
Table 42. Comparison of airport type identification across analytical methods.....	82
Table 43. Minimum turnaround times by aircraft type, AO type and airport size	88
Table 44. Definition of passenger waiting cases	90
Table 45. Passenger reaccommodation - trigger and control events	96
Table 46. Hierarchy of interlining	98
Table 47. Regulation 261 rights afforded with respect to delay.....	99
Table 48. Costs of provisions made by airlines, by delay duration and cost scenario	100
Table 49. Final passenger hard cost assignments by carrier type	100
Table 50. Ticket type by value of time applied.....	101
Table 51. Generic load factors.....	103
Table 52. Passenger soft costs of delay per minute, by three cost scenarios.....	106
Table 53. Summary of soft cost assignments by airline type.....	106
Table 54. Summary of non-passenger cost assignment	107
Table 55. At-gate, non-passenger costs of delay per minute, by three cost scenarios.....	108
Table 56. Taxi, non-passenger costs of delay per minute, by three cost scenarios.....	108
Table 57. En-route, non-passenger costs of delay per minute, by three cost scenarios.....	109

Table 58. Arrival management, non-passenger costs of delay per minute, by three cost scenarios	109
Table 59. Definition of variable types	111
Table 60. Master variable list	111
Table 61. Key correspondences between I-type and D-type variables	113
Table 62. Map of variables and their associated rules	113
Table 63. Tracked flights of A320 EIDSX on 17SEP10	115
Table 64. Example of matching difficult flights with the schedule (local times)	116
Table 65. Summary of flight status for 17SEP10	118
Table 66. Summary of passenger dataset cleaning iterations for SEP10	119
Table 67. Simple data for three airports – perfect off-set	120
Table 68. Unrotated extraction for three airports	121
Table 69. Unrotated extraction for three airports (Frankfurt high delay, one component)	122
Table 70. Unrotated extraction for three airports (Frankfurt high delay, two components forced)	122
Table 71. Varimax extraction for three airports	123
Table 72. Promax extraction for three airports	124
Table 73. Direct oblimin extraction for three airports	124
Table 74. Simple data for three airports – imperfect off-set	125
Table 75. Properties of rotated, <i>oblique</i> components	126
Table 76. Promax extraction for A, B and C airports	127
Table 77. Promax extraction for A, B, C and D airports	127
Table 78. Promax extraction for A, B, C and D (extended) airports	128
Table 79. Value of time per passenger, per hour	130
Table 80. Ticket type by value of time applied	131
Table 81. Flights affected by strike action in France on baseline traffic day	132

List of figures

Figure 1. Project architecture: scenarios and metrics	12
Figure 2. Relationships between metric types	13
Figure 3. Key analysis pathways	14
Figure 4. The POEM event stack in overview	16
Figure 5. The POEM event stack in more detail	17
Figure 6. Extract of model log	20
Figure 7. Airport cancellation rates for 2012	24
Figure 8. Non-zero cancellation rates for 2012, with Gamma fit	25
Figure 9. Non-zero cancellation rates for worst three days in September 2012, with Gamma fit	26
Figure 10. Non-zero cancellation rates for rest of September 2012, with Gamma fit	26
Figure 11. Desired transformation from baseline to target cancellations	27
Figure 12. Fitted difference distribution used to transform baseline to target cancellations	28
Figure 13. Aggregated PaxIS passenger itineraries onto disaggregated PRISME flights	30
Figure 14. Summary of the assignment of target passengers to individual flights	33
Figure 15. Average economy fare per origin-destination GCD	36
Figure 16. Delay propagation and airport typologies	39
Figure 17. Overview of different airport-airport density layers	42
Figure 18. Distribution of delayed passengers under N ₂ scenario	46
Figure 19. Four-airport example of delay patterns	55
Figure 20. Total departure delay by time period and airport (baseline traffic day)	56
Figure 21. Total departure delay by time period, without LEBL and LEMD (baseline traffic day)	57
Figure 22. Inter-component delay propagation nodes	59
Figure 23. Scree plot for inter-component delay propagation nodes (from model input data)	61
Figure 24. Scree plot for inter-component delay propagation nodes (from model output data)	62
Figure 25. Sensitivity of the air transport network to primary delays	64
Figure 26. Simple airport delay amplifiers and attenuators	65
Figure 27. Qualitative reactionary delay propagation tree for AZA_LIRFLFPG01	66
Figure 28. Quantitative reactionary delay propagation tree for AZA_LIRFLFPG01	67
Figure 29. Reactionary and arrival delay distribution	68
Figure 30. Arrival and reactionary delay distributed by scheduled in-block time	69

Figure 31. Time required to yield the same proportions of arrival and reactionary delay	69
Figure 32. Average time until total reactionary delay is absorbed, by airport size	70
Figure 33. Arrival and reactionary delay, by airport size	70
Figure 34. Airport time-line graph	71
Figure 35. Reactionary delay propagated from LFPG	72
Figure 36. Passenger delay by elapsed time	74
Figure 37. Passenger delay causality network for S_0 simulation	75
Figure 38. Passenger delay causality network for A_1 simulation	75
Figure 39. Flight delay causality network for S_0 simulation	76
Figure 40. Flight delay causality network for A_1 simulation	76
Figure 41. Simultaneous propagation causalities	80
Figure 42. Relationships between metric types (repeat of Figure 2)	81
Figure 43. Minimum turnaround time cut-off threshold	88
Figure 44. Minimum turnaround times – heavy aircraft, all airlines	88
Figure 45. Minimum turnaround times – medium aircraft, regional airlines	89
Figure 46. Minimum turnaround times – medium aircraft, charter airlines	89
Figure 47. Minimum turnaround times – medium aircraft, LLCs	89
Figure 48. Minimum turnaround times – medium aircraft, full-service airlines	89
Figure 49. Minimum turnaround times – light aircraft, all airlines	90
Figure 50. Illustrative movements by time block	93
Figure 51. Integral transformation of movements	93
Figure 52. Sequencing over the IAF	94
Figure 53. Minimum connecting times (MCTs) for each ECAC airport	103
Figure 54. Power curve fit of pax hard costs as a function of delay duration and cost scenario	104
Figure 55. Passenger dissatisfaction as a function of delay duration	105
Figure 56. The 0400-0359 operational day across ECAC time zones	116
Figure 57. The restructured raw passenger dataset	119
Figure 58. Four-airport example of delay patterns	126
Figure 59. Value of time in context	129
Figure 60. Arrival and reactionary delay distributed by scheduled in-block time for busiest airports	133
Figure 61. Reactionary delay propagated from busiest airports	134

Executive summary

At the core of POEM (Passenger-Oriented Enhanced Metrics) is the design of new performance metrics and their evaluation through a European network simulation model under novel flight and passenger prioritisation scenarios. Key objectives were to explore the trade-offs between the (new) flight-centric and passenger-centric metrics and to characterise the propagation of delay through the network.

Social and political priorities in Europe are continuing to shift in further support of passenger rights in transportation, as evidenced by high-level position documents such as 'Flightpath 2050' and the European Commission's 'Roadmap to a Single European Transport Area'. To better measure progress in reaching such objectives, passenger-centric metrics are needed. These are largely absent from the metrics currently in place to measure air transport system performance. We also need better models to understand the implications of policy measures and the trade-offs between them. Such models and metrics thus need to reflect the progress of corresponding planned regulatory review, particularly with regard to the underpinning regulatory instrument, Regulation 261.

For the modelling, a baseline traffic day in September 2010 was selected as a busy day in a busy month – without evidence of exceptional delays, strikes or adverse weather. POEM models the busiest 199 European Civil Aviation Conference (ECAC) airports in 2010, having identified that these airports accounted for 97% of passengers and 93% of movements in that year. Routes between the main airports of the EU 27 states (i.e. prior to Croatian accession in July 2013) and airports outside the EU 27 have been used as a proxy for determining the major flows between the ECAC area and the rest of the world. This process allowed the selection of 50 non-ECAC airports for inclusion of their passenger data.

The two principal datasets used to prepare the input data for the POEM model were IATA's PaxIS passenger itineraries and EUROCONTROL's PRISME traffic data. Extensive data cleaning of the source traffic data was required. Departure and arrival times were converted to local times (in addition to UTC) in order to define local 0400-0359 operational days at the modelled ECAC airports and to enable schedule matching (published in local times only), taking into account daylight saving time adjustments. There are approximately 30 000 flights in each day's traffic and around 2.5 million passengers distributed among 150 000 distinct passenger routings. The assignment of passengers to individual flights, with cost characteristics and full itineraries, was a fundamental component of POEM, since the project explores new passenger-centric metrics. All the allocated connections were viable with respect to airline schedules and published minimum connecting times (MCTs).

POEM is a full gate-to-gate model, with passenger connectivities explicitly modelled. Each simulated process is governed by one or more rules. For example, Rule 33 governs realistic decision-making for missed passenger connections due to delays and cancellations (such as dynamic passenger reaccommodation onto aircraft with free seats, using detailed fleet and load factor data) and integrates with the tail-tracked aircraft wait and turnaround (recovery) rules. Core cost estimations in the model are with respect to delay costs to the airline, since it is these which drive airline behaviour. The model represents a normative day and the simulation results thus reflect schedule robustness (e.g. with respect to passenger reaccommodation).

Two airline case studies, including on-site visits and workshops, have focused on developing and testing specific aspects of the model rules as examined in an operational context. Furthermore, key parameters of the model were calibrated against independent data sources, including flight departure, arrival and reactionary delay, and also European load factors.

To measure the effect of increased perturbation, two disrupted days were derived from the baseline traffic day through the application of internal rules in the model. This allowed like-for-like comparisons between the disrupted days and the baseline day. One disrupted day imposed 1 extra minute on the average departure delay (making a new average of 14.9 minutes across all flights). The other disrupted day imposed just under 1% of additional cancellations on morning operations. Comparing the model outputs for the disrupted days showed them to be well modelled in that changes to the core metrics were as expected and reflected operational experience (e.g. with regard to relatively low impacts on flight *punctuality* metrics during periods of high cancellations).

Flight and passenger prioritisation scenarios were applied to the baseline traffic day and the two disrupted days. The prioritisation scenarios provided primary inputs into the network simulation model

and were designed in parallel with the metrics, through which they were assessed in terms of their impact on performance. ANSP and AO modelled scenarios involved decision-making based on reasonable information for that agent to possess in either the current information environment or a future one (e.g. in the context of System-Wide Information Management).

Flight prioritisation scenarios operating during arrival management based simply on the numbers either of inbound passengers or on those with connecting onward flights, were ineffective in improving performance. Such performance was even slightly worse under high delay or increased cancellation rates. These effects were only discernible through the new metrics.

A modelled policy-driven scenario represented the special case where we ran the model under putative conditions not driven by current airline or ATM objectives but which may nevertheless benefit the passenger. This scenario, rebooking disrupted passengers at airports based on minimising delays at their final destination, produced very weak effects when current airline interlining hierarchies were preserved. When these restrictions were relaxed, marked improvements in passenger arrival delay were observed, although at the expense of an increase in total delay costs per flight (due to passenger rebooking costs).

A prioritisation process assigning departure times based on cost minimisation (with due consideration of ATFM delays) markedly improved a number of passenger delay metrics and airline costs, the latter determined by reductions in passenger hard costs to the airline. One of the very few negative outcomes associated with this scenario was an increase of two percentage points in overall reactionary delay. Furthermore, in this trade-off, the additional reactionary delay was manifested through relatively few flights and was introduced purposefully by airlines through the cost model (i.e. waiting for late passengers) such that the overall cost to the airlines decreased. The addition of independent, cost-based arrival management apparently foiled these benefits due to lack of coordination between departures and arrivals.

Factor analysis was undertaken to ascertain the extent to which a derived (data reduction) technique would compare with a complexity science approach in analysing the results. Granger causality is held to be one of the only tests able to detect the presence of causal relationships between different time series. This method was used along with the eigenvector centralities of nodes (airports) to further explore delay propagation under the cost minimisation prioritisation scenario and to compare the passenger- and flight-centric network properties. Remarkably, there was practically no relationship between the role of airports across these different layers, thus demonstrating the difference between the passenger- and flight-centric networks *and* the impact of the cost-based prioritisation process on these networks. This prioritisation scenario also produced a further trade-off: the propagation of delay was contained within smaller airport communities, but these communities were more susceptible to such propagation.

The importance of using passenger-centric metrics in fully assessing system performance was repeatedly observed, since such changes were not expressed through any of the currently-used flight metrics at the common thresholds set. Most prioritisation scenarios performed similarly for the high cancellation and high delay days, demonstrating robustness in terms of their efficacy under increased disruption.

A significant advance on earlier work has been the explicit estimation of reactionary costs (since each flight is individually modelled with its connectivity dependencies) and of the passenger costs of delay to the airline. In previous work these costs were estimated statistically. Passenger value of time has also been quantified as a function of delay at the final destination.

Smaller airports were significantly implicated in the propagation of delay through the network at a level that has hitherto not been commonly recognised. This is probably due to reduced delay recovery potential at such airports (e.g. through flexible or expedited turnarounds, and spare crew and aircraft resources – although the latter are not explicitly modelled in POEM) and whether a given airport has sufficient connectivity and capacity to reaccommodate disrupted passengers. Back-propagation is an important characteristic of the persistence of delay propagation in the network. Paris Charles de Gaulle, Madrid Barajas, Frankfurt, London Heathrow, Zürich and Munich all demonstrated more than one hundred hours of back-propagated delay during the modelled (baseline) day. Most of the delay characterised in the network was indeed distributed between a relatively limited number of airports.

Future dissemination plans and research priorities are summarised.

Acknowledgements

We would like to acknowledge the kind assistance of EUROCONTROL's Performance Review Unit and Central Office for Delay Analysis (CODA) for extensive and patient help during the preparation of this report, plus generous support on airline data provision from Innovata (UK and US offices) and on passenger data provision from IATA's Passenger Intelligence Services (Montreal).

1 Introduction

1.1 Purpose of the document

Central to the POEM (Passenger-Oriented Enhanced Metrics) project is the design of new performance metrics and their evaluation through a European network simulation model under novel flight and passenger prioritisation scenarios. Key objectives were to explore the trade-offs between the (new) flight-centric and passenger-centric metrics and to characterise the propagation of delay through the network.

This report describes how the simulation model was built, delineates the extensive rules and model structure, defines the prioritisation scenarios incorporated into the model and details the creation of a unique European dataset which has passengers assigned to each flight. Analytical techniques and metrics used to explore the results are described, before these are presented with a detailed interpretative commentary.

The effectiveness of the flight and passenger prioritisation scenarios is investigated and the importance of passenger-centric metrics is highlighted. The legislative and political context is established. Delay propagation and the issue of causality are examined in detail. Finally, key next steps beyond the project are identified.

1.2 Intended readership

This report is written for the professional reader and assumes an understanding of air transport and ATM. Without detriment to appropriate referencing and delineation, the text is not cluttered with explanations of common acronyms or principles. Supporting statistical results are given in a practical and unobtrusive manner with interpretative texts and explanations given of less common techniques.

The report has been produced in two versions. This Final Technical Report (POEM Deliverable 6.2) is the fuller text with all appendices. The Final Strategic Report (POEM Deliverable 6.3) is a condensed text intended to convey the key information with less technical detail and without the supporting appendices¹ (although these are still referred to in several places to allow the reader to ascertain where additional material is available in POEM Deliverable 6.2).

Appendices included in this report

Appendix	Coverage
A	Key model rules and variables
B	Final data preparation and cleaning
C	Factor analysis methodology
D	Passenger value of time
E	Baseline traffic day disruption
F	Further analyses of airport reactionary delay

¹ All the appendices present new or extensively updated work, not contained in previous deliverables, with the exception of Appendix D on passenger value of time, which is included here for completeness of reference.

1.3 European Union policy context

Social and political priorities in Europe are continuing to shift in further support of passenger rights in transportation, as evidenced by high-level position documents such as 'Flightpath 2050' (European Commission, 2011a) and the European Commission's 2011 White Paper: 'Roadmap to a Single European Transport Area' (European Commission, 2011b).

To better measure progress in reaching such objectives, passenger-centric metrics are needed. These are largely absent from the metrics currently in place to measure air transport system performance. We also need better models to understand the implications of policy measures and the trade-offs between them. Such models and metrics thus need to reflect the progress of corresponding planned regulatory review, particularly with regard to the underpinning regulatory instrument, Regulation 261 – the European Union's air passenger compensation and assistance scheme (European Commission, 2004).

It has been accepted that there are currently several problems with regard to the implementation and scope of Regulation 261. A roadmap for the revision of the Regulation was published in late 2011 (European Commission, 2011c). The Commission intended to put forward a proposal to make the current Regulation more effective, without imposing undue burdens on operators or passengers.

A consultation on the potential revision was completed in March 2012. There was little consensus on the way forward, with responses from airlines and consumer/passenger organisations often directly opposed (European Commission, 2012). In May 2012, stakeholders met at a conference on air passenger rights to discuss the consultation findings. In March 2013, a memo was released by the Commission (European Commission, 2013) detailing the key proposed changes, which could become law by 2015, subject to approval by member states. In summary, the key changes are to:

- initiate passengers' right to care and assistance after two hours of delay, regardless of the length of the flight (a current determining factor);
- require an airline to re-route passengers onto other carriers (already much commoner in the US) if it cannot re-route onto its own services within 12 hours; this is to obviate the current lengthy periods passengers may need to wait for a re-route;
- offer passengers the same rights for delays relating specifically to connecting flights, and to extend such rights to *compensation* for long delays (including arrival delay) caused by any reason;
- introduce new obligations (currently none exist) regarding information on delayed or cancelled flights;
- better define 'extraordinary circumstances' that exempt carriers from paying passenger compensation (although proposed changes to the compensation rights will make these more complex, allowing the carriers more time to avoid cancelling flights, for example).

A number of important insights into these trade-offs are afforded by the new, passenger-centric metrics examined in the POEM model, as presented in this report. Yet further insights may be gained in future by extending the scenarios simulated.

Whilst the second edition of the SESAR ATM Master Plan (SESAR, 2012) explains that it "provides the best actualised view on the products, technologies and operational procedures, which can be further industrialised and deployed in order to satisfy the needs of the European citizens", it does not refer explicitly to passengers, passenger metrics or Regulation 261.

1.4 Inputs from other projects

N/A.

2 Summary and update of POEM

2.1 Overview and objectives

At the core of POEM (Passenger-Oriented Enhanced Metrics) is the design of new performance metrics and the evaluation of these through a European network simulation model under novel flight and passenger prioritisation scenarios. Key objectives were to explore the trade-offs between the (new) flight-centric and passenger-centric metrics and to characterise the propagation of delay through the network.

Figure 1 shows the model architecture of POEM, with a focus on the scenarios and metrics and the relationships between them. We define a metric as any quantitative measure, particularly one which usefully expresses some output of a system (usually performance), part of the system, or (an) agent(s) within it, usually over an aggregate scale and often as a ratio (e.g. per flight). The central oval represents the ATM system: the main model. The prioritisation scenarios applied in the model may be classified by three central themes according to the agency/orientation of the decision-making. Each is figuratively represented as a 'stream' (horizontal grey band) with its impact 'flowing into' the ATM system.

ATM (ANSP) and AO scenarios involve decision-making based on reasonable information for that agent to possess in either the current information environment, or a future one (e.g. in the context of System-Wide Information Management). A policy-driven scenario represents the special case where we run the model under putative conditions not driven by current airline or ATM objectives but which may nevertheless benefit the passenger. In addition to this, over-arching all the scenario streams, are exogenous factors, such as higher-level network characteristics, economic growth or recession. These may be reflected through changes to the traffic input data, two examples of which that we will model are increased delays and increased cancellations (see Section 2.4.2).

The multiple arrows in each scenario stream represent the different levels therein, as described in Section 2.3. Increasing level numbers represent increasing levels of intervention with respect to the current baseline. Currently, two such levels of prioritisation rules are deployed per scenario in the model. For each of the scenario streams, we may expect a corresponding output effect (on the right-hand side of the figure) on the metrics. For example, if a scenario in the ATM (ANSP) stream prioritises flights according to aircraft delay minutes, we would naturally observe a reduction in metrics such as average flight delay. If a scenario in the AO stream prioritises flights according to passenger cost of delay to the airline, a reduction in passenger metrics associated with delay cost will be observed. More interesting, however, are the quantitative effects observed *between* these streams (represented by the dashed arrows between the horizontal bands). For example, it would be expected that prioritising flights according to aircraft delay minutes (as a scenario in the upper stream) will also reduce passenger delay costs (a metric in the second stream).

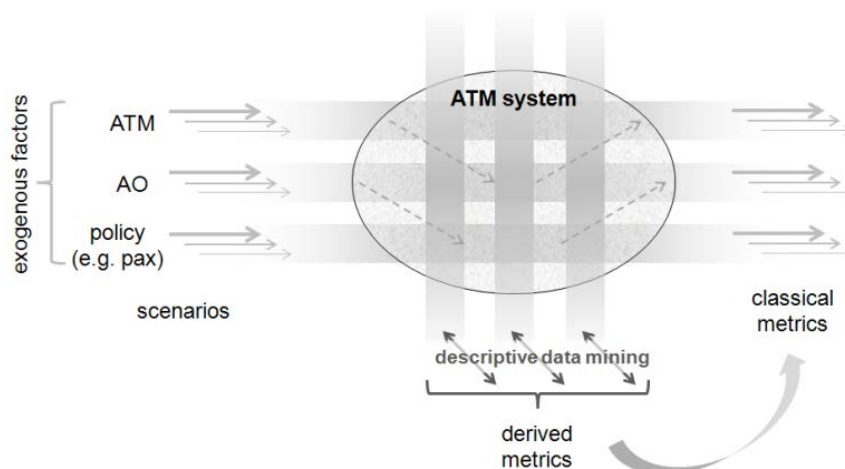


Figure 1. Project architecture: scenarios and metrics

2.2 Summary of the metrics

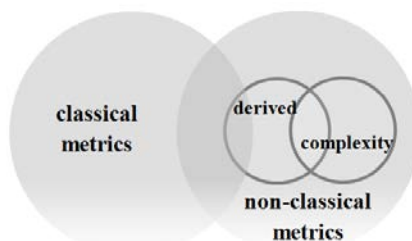


Figure 2. Relationships between metric types

Figure 2 shows a metric classification, which we define in order to more clearly present the manner in which their scope needs to be extended in ATM and to differentiate between the types. The term ‘classical’ metrics is used to denote those that are pre-defined (such as average aircraft delay), are univariate (draw on one variable in the data) and do not use complexity science techniques. Some of these types of metric are already commonly in use (such as, indeed, average aircraft delay) whilst others are not (such as average passenger delay) – and, arguably, thus conspicuous by their absence.

‘Non-classical’ metrics defines both (non-complexity) ‘derived’ metrics, which are in contrast to the classical metrics in that they are not (fully) pre-defined but are derived from the data iteratively and are typically multivariate, and those drawn from complexity science. An example of a derived metric is a factor obtained as the result of factor analysis. An example of a simple complexity metric is the degree of a node.

Figure 2 shows that these relationships are not wholly mutually exclusive. Data mining techniques may be applied not only to generate non-classical metrics but also in topology characterisation, such as identifying complex network communities (groups of densely connected nodes sharing only few connections with nodes outside their group). These techniques are not needed to define classical metrics, however.

Whilst it is thus relatively straightforward to identify some metrics that belong decisively to one of the categories, the *overlap* between the categories is less well defined and is of particular interest to explore. For example, how well do non-complexity metrics and methods capture certain features of ATM system dynamics (such as delay propagation) compared with those of complexity science? In POEM, both types of non-classical metric analysis are employed. We will explore both derived metrics and complexity metrics, in addition to the evaluation of classical metrics, old and new. Metric design was informed by both a literature review² and a stakeholder consultation process³.

Figure 3 summarises the key analysis pathways for POEM. In the data acquisition stage, flight and passenger data were obtained from the PRISME (EUROCONTROL) and PaxIS (IATA) sources, then filtered (to avoid exceptional days), cleaned and integrated (assigning passengers to flights). This then furnished the ‘flight-passenger’ datasets, i.e. wherein each flight has passengers assigned to it. These data were analysed in their own right and also used as inputs into the POEM model. The data preparation and model development processes are detailed in Section 2.4.

² See POEM Deliverable 3.1.

³ See POEM Deliverable 6.1. The project was supported by a consultation and dissemination workpackage. This included an on-line user-requirements survey addressing KPAs. This stakeholder survey secured 157 responses from airlines, airport authorities, air navigation service providers (ANSPs), civil aviation authorities, EUROCONTROL, Regulation 261 national enforcement bodies and researchers/academics. It was followed by a complementary one day seminar and workshop in central London in January 2012, attracting approximately 60 delegates.

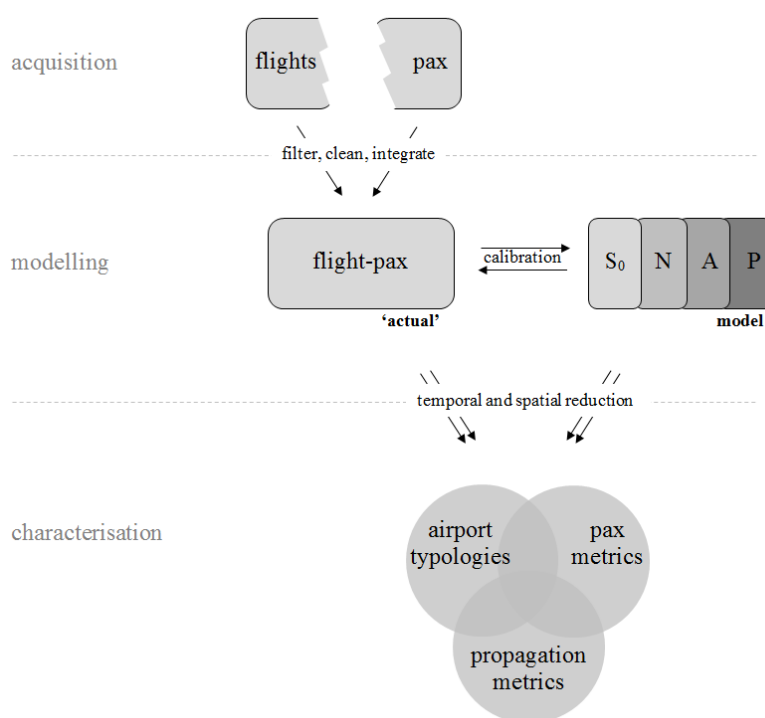


Figure 3. Key analysis pathways

The scenarios are discussed in the next section. In the model, no-scenario (S_0) baselines are modelled in addition to flight and passenger prioritisation scenarios. The active scenario designators in the figure ('N', 'A' and 'P') correspond to ANSP, AO and policy – thus corresponding to the agency/orientation of the decision-making referred to above.

The characterisation identified in the lower part of Figure 3 requires data reduction. This may be achieved by various means, such as through the evaluation of the passenger metrics indicated in the Venn diagram and through the exploration of various airport typologies. These typologies are determined by the role an airport plays in the propagation of delay. There is some broad correspondence within these characterisations, both between the two types of metric in certain cases (e.g. flight and passenger arrival delay⁴) and between the airport typologies and the metrics (e.g. ratios of departure and arrival delay). The analyses are presented in Section 3.

2.3 Summary of the prioritisation scenarios

The prioritisation scenarios (see Table 1) are fundamental drivers of the network simulation and were designed in parallel with the metrics. '(Prioritisation) scenario' refers to the various priority rules applied in the model. We use the metrics to compare the effectiveness of such rules in reducing the costs of delay to the airline, for example. 'Cost scenario' refers to the assumptions made regarding the underlying costs of delay to an airline. These costs are modelled according to 'high', 'base' and 'low' cost scenarios, using models adapted from 'European airline delay cost reference values' (Cook and Tanner, 2011), a standard cost reference used by EUROCONTROL (for example cited in EUROCONTROL's annual Performance Review Reports and Standard Inputs for Cost Benefit Analyses documents). To preserve consistency across documentation but to avoid any confusion regarding terminology, where 'cost scenario' is intended, the word 'cost' will not be omitted, whereas generic references to 'scenario' mean 'prioritisation scenario'.

⁴ Notwithstanding important differences between these metrics, too.

Table 1. Summary of prioritisation scenarios

Type, and level	Designator	Summary description
No-scenario, 0	S ₀	No-scenario baselines (reproduce historical operations for baseline traffic day)
ANSP, 1	N ₁	Prioritisation of inbound flights based on simple passenger numbers
ANSP, 2	N ₂	Inbound flights arriving more than 15 minutes late are prioritised based on the number of onward flights delayed by inbound connecting passengers
AO, 1	A ₁	Wait times and associated departure slots are estimated on a cost minimisation basis, with longer wait times potentially forced during periods of heavy ATFM delay
AO, 2	A ₂	Departure times <i>and</i> arrival sequences based on delay costs – A ₁ is implemented <i>and</i> flights are independently arrival-managed based on delay cost
Policy, 1	P ₁	Passengers are reaccommodated based on prioritisation by final arrival delay, instead of by ticket type, but preserving interlining hierarchies
Policy, 2	P ₂	Passengers are reaccommodated based on prioritisation by final arrival delay, regardless of ticket type, and also relaxing all interlining hierarchies

Table 2 shows under which rules the various scenarios are implemented. Note that these are run independently. For example, P₁ and P₂, prioritising passenger reaccommodations at airports, are not coordinated with arrival management rules. This allows us to assess independently the impacts of the various scenarios. A partial exception to this is under A₂ in Rule 13, whereby flights follow prescribed rules for waiting for delayed passengers *and* are then given arrival priority: both rules are based on delay costs, but without any gate-to-gate coordination (see also Section 2.4). Full details of the way in which each scenario is implemented within a rule, are given in the individual descriptions of rules 13, 26, and 33 (see Appendix A).

Table 2. Rules and scenarios

Designator	Rule 13 Wait for boarding	Rule 26 Airborne arrival management	Rule 33 Passenger reaccommodation
N ₁		●	
N ₂		●	
A ₁	●		
A ₂	= A ₁	●	
P ₁			●
P ₂			●

● = implemented

2.4 Summary and update of the model

2.4.1 Core model

2.4.1.1 Overview of the core model

This section presents some of the key facets of the model as a stand-alone text. The fully updated model rules are to be found in Appendix A. Through POEM's Case Study 2 (see POEM Deliverable 7.2), feedback was received from four airlines (one LCC, one regional airline and two flag carriers) regarding Rule 33 and from another flag carrier regarding the use of flexible fares.

Note that each simulation day of the model runs in a *local* time window from 0400 to 0359, which took considerable effort to assign to each airport, converting from the more commonly employed UTC (see Appendix B).

2.4.1.2 Event stack

Instead of traditional sequential execution, the POEM model follows an event-driven paradigm. The execution flow of the model is driven by the 'event stack'. This stack consists of an ordered sequence of events. Each event has a time stamp associated with it, which determines its position in the stack. Each event is defined by the following:

- **time stamp** – indicates the time at which the action takes place, e.g. 1202220812 (or 12:02 on 22 August 2012);
- **event type register** – e.g. 'ask for departure clearance';
- **identifier** – of the class type and instance of the object involved, e.g. flight SN377;
- **log** – a register to keep track of the processes.

Usually, events are given in the following format: *identifier::type(time stamp)*.

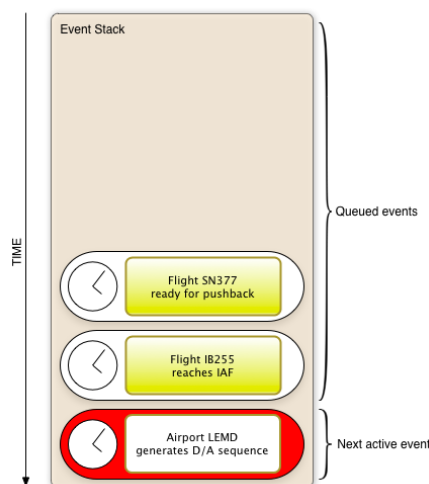


Figure 4. The POEM event stack in overview

The event stack is defined as an object class and is controlled by a number of methods. These methods include insertion, deletion, extraction, search and integrity checks. By using the event stack, one can abstract to a design layer defining the events and their relationships, letting the model flow naturally and be managed automatically by the event stack.

Execution follows in a series of steps. First, the event stack is initialised by a constructor method at the beginning of the simulation. If the stack is not empty the event with the smallest time stamp becomes active. Next, the active event calls the method defined by the type of the event.

This method follows an algorithm and ultimately introduces new events, which are inserted into the stack, preserving the ordered structure. Once an event has been completely processed it is removed from the stack. The next event then becomes active and the process repeats until the stack is empty, in which case the simulation is over.

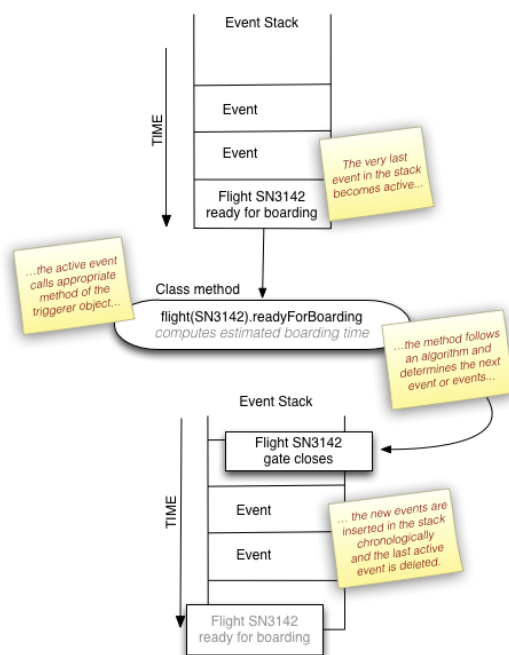


Figure 5. The POEM event stack in more detail

2.4.1.3 Defined events

We present here the defined events and the corresponding rules they implement. The event stack is initialised by introducing, for each flight from the initial traffic sample, an event of type *airline::flightLegStart*, with time stamp SOBT.

airline::flightLegStart

This event triggers the flight process. Initially, the event looks for the location and current status of the object aircraft, given by the aircraft registration. If the aircraft is still flying a previous segment or has already landed but the minimum turnaround time (computed as described in Rule 10 and Rule 31) has not yet elapsed, then a new *airline::flightLegStart* (with corresponding time stamp plus minimum turnaround time) is inserted into the event stack. Otherwise, the aircraft is considered ready for boarding, the event *flight::readyForPushback* is introduced into the event stack and the aircraft ready time is set to the current time.

flight::readyForPushback

Once the aircraft is considered ready for boarding, the next step is passenger boarding. Following the wait for boarding rule (Rule 13), the estimated gate-arrival times of passengers from inbound connecting flights are calculated, taking into account the minimum connecting times (MCTs) of the airport, and the airline decides whether to wait for certain passengers.

In the case that the airline decides not to wait, these passengers would be included in the unaccommodated passenger list under Rule 33, and one event of type *airline::reallocatePax* would be created to assign these passengers new flights. If the airline decides to wait for (a) connecting passenger(s), then a new *flight::readyForPushback* event is introduced into the event stack at the latest expected gate-arrival time of the waited-for passengers. If the airline is not waiting for any more passengers, then 'aircraft ready for pushback' is set to true and an event *airport::askForDepartureSlot* is introduced into the stack.

airline::reallocatePax

This method implements the passenger reaccommodation algorithm described in Rule 33.

flight::askForDepartureSlot

As soon as a flight is flagged as ready for pushback, a departure request is sent to the departure manager of the origin airport so that the departure manager can check, following Rule 16, whether there are any (further) network capacity restrictions. If so, the flight becomes pseudo-ATFM regulated and some departure delay is applied (see Rule 118) and a new *flight::askForDepartureSlot* event is introduced into the event stack.

Additionally, if the current runway queue exceeds 5 minutes of queuing the flight is delayed at-gate and a new *flight::askForDepartureSlot* is introduced into the stack. Otherwise, the next available departure slot is assigned. Once the slot is assigned, the flight is considered off-blocks and Rules 19 and 20 are used together to compute the actual taxi-out time, which in turn determines the time stamp of the event *flight::runwayHit*, which is then introduced into the event stack.

flight::runwayHit

Once the flight has been granted a departure slot and taxi-out is completed, the flight reaches the runway and it is then tagged as 'waiting for take-off'. This does not mean an imminent take-off, for it has to be selected from all aircraft in the 'waiting for take-off' and 'waiting for landing' airport lists. This selection is made by the airport when the event *airport::manageRunway* becomes active. This event checks if there is an *airport::manageRunway* event already in the stack. If not, such an event is introduced into the stack with the appropriate time stamp.

airport::manageRunway

This event is repeated whilst there is any flight tagged as ready to take-off or to land at a given airport. The airport chooses a flight from the corresponding lists as the next flight to take-off or land. If a flight is selected to take-off, an event *flight::takeOff* is introduced into the stack. Otherwise, a *flight::landing* event is introduced.

flight::takeOff

As soon as the *flight::takeOff* event becomes active, a modelled take-off time is set a time stamp, meaning that the aircraft has already departed. Using the planned en-route time (modified accordingly following Rule 23) a passing time for the IAF is calculated and an event type *flight::askForArrivalSlot* with a time stamp of the estimated time over is added to the event stack.

flight::askForArrivalSlot

Once the aircraft has reached the IAF, it requests an arrival slot from the arrival manager of the destination airport. The manager, following Rule 26 (airborne arrival management), designates an available arrival slot. Note that the way in which arrival management handles the available slots varies substantially with the scenario (Rule 26). Once the arrival slot is granted an event *flight::hitPTI* is introduced into the stack.

flight::hitPTI

Similar to the *flight::runwayHit* event, this event tags the calling flight as 'waiting for arrival' and waits until the *airport::manageRunway* event chooses it for landing. When this happens, an event *flight::landing* is introduced into the stack.

flight::landing

The landing time determines the modelled time on and concludes the airborne phase of the flight. It also introduces the *flight::flightLegEnd* event. Using Rule 29, the taxi-in time is calculated and added to the arrival time to determine the final in-block time – thus ending the flight process.

flight::flightLegEnd

When this event becomes active it means that the flight is in-block. The flight process is therefore ended and with it all passengers that are not connecting to another flight have reached their final destination. For all connecting passengers, the actual connection time is computed using Rule 35 and a new event *pax::gateArrival* is introduced into the stack, with its corresponding time stamp. If the aircraft is intended to start a new flight leg, a new *flight::flightLegStart* event is introduced. If there is sufficient time for an early departure, Rule 118 permits this.

pax::gateArrival

Upon activation, this event tags the passengers as ready for boarding, if the connecting flight has not departed yet. Otherwise, it introduces an event of type *airline::reallocatePax*, so that such passengers can (usually) continue their trip on a different flight.

2.4.1.4 Extract of model log

Figure 6 shows some of the activities of the first elapsed second from 12:25:00 UTC during a model run, with examples of wait rules, cost estimation and passenger reaccommodation in progress.

```
[...](17-Sep-2010 12:25:00) 47 out of 49 of pax (95.92 pct.) of DLH_EDDLEGBB02:15877 were ready, flight over 80 pct. occupancy, no more delay added

(17-Sep-2010 12:25:00) Total cost of flight DLH_EDDLEGBB02:15877 departing at 17-Sep-2010 12:25:00 now estimated at 127.15 euros

(17-Sep-2010 12:25:00) No further pax delay will be introduced, thus flight DLH_EDDLEGBB02:15877 is now pushback ready, reaccommodating connecting pax

(17-Sep-2010 12:25:00) Pax group DLH1815:37550 of 2 inflex pax coming from DLH_EDDHEDDL06:12246 to EGBB did not make it to DLH_EDDLEGBB02:15877 (no more connections afterwards) and need to be reaccommodated

(17-Sep-2010 12:25:00) 2 inflex pax of group DLH1815:37550 of DLH_EDDHEDDL06:12246 that missed DLH_EDDLEGBB02:15877 were successfully reaccommodated in DLH_EDDLEGBB03:23396 same alliance, DLH1815/1:145607 Arrival: 17-Sep-2010 17:50:00 delay: 04:00'00" (airport wait 03:01'51")

(17-Sep-2010 12:25:00) Trying to reaccommodate the 80 pax waiting at EDDL:10

(17-Sep-2010 12:25:00) A total of 2 pax of DLH_EDDLEGBB02:15877 were left behind and all of them were successfully reaccommodated

(17-Sep-2010 12:25:00) Flight SAS_ENKBENGM03:15843 loading 67 pax and all of the 67 pax are not coming from a previous flight. There are NO connecting pax

(17-Sep-2010 12:25:00) There are 29 pax groups in SAS_ENKBENGM03:15843 connecting with another flight afterwards (SAS3310:87574, SAS3311:87575, SAS3312:87576, SAS3313:87577, SAS3314, [...])
```

Figure 6. Extract of model log

2.4.1.5 Processing time

There are approximately 30 000 flights in each day's traffic and around 2.5 million passengers distributed among 150 000 distinct passenger routings. (See also Section 2.4.3 on passenger allocation to flights and Appendix B on data cleaning and flight demotion.) In each simulation, the number of processed events depends greatly on the scenario being simulated, from around 300 000 to almost one million. On a single-processor machine (Intel i7 1.6GHz with 4GB RAM) the average run took around 20 – 40 minutes, plus five to ten additional minutes to compute the metrics. However, simulations were carried out using a cloud-computing platform. Using this technology one can execute a large number of runs simultaneously. In fact, all three scenarios and days, a total of 750 runs, were executed in parallel over 24 hours, resulting in a (crude) average of less than two minutes per scenario.

2.4.1.6 Pre-computed cost functions

During initial runs of the model, it was found that scenarios A_1 and A_2 were highly demanding in terms of computational power and, moreover, produced some unexpected cost deteriorations. This was due to negative impacts of rules on subsequent rotations of aircraft. This confirmed that an instantaneous optimisation based on first-order effects produces a local optimum only, whereas the network effects may worsen.

In order to improve the cost optimisation for the airlines, without running the entire model to estimate the implication of each decision, pre-computed cost functions were developed. These were implemented as complementary procedures in the scenario modules by calculating delay propagation costs based on scheduled times, i.e. without dynamic data or stochastic assessment. These functions work recursively (i.e. backwards from the end of the simulation day) using entire propagation cost trees based on discrete delay values (0, 5, 10, 15 ... minutes of delay, up to 6 hours).

Dynamic cost functions are (still) used to determine costs for a given flight based on the current status of the network. For example, these drive the airline passenger wait rules (as applied in scenarios A_1 and A_2 ; see also Rule 13). The pre-computed functions drive the arrival queuing determined by ANSPs under scenario A_2 (see Rule 26), for example. In certain other circumstances the cost functions work mutually, whereby the dynamic cost functions call the pre-computed functions to assess second-order effects.

2.4.2 Traffic inputs for three simulation days

2.4.2.1 Overview of baseline and disrupted traffic days

The POEM baseline traffic day of Friday 17SEP10 was selected as a busy day in a busy month, without any evidence of atypical delays or weather. Relevant CFMU monthly and daily ATFCM summaries were consulted to check for days with unusually high disruption – primarily weather events and strikes⁵. The integration of these flight data from EUROCONTROL (PRISME) with the passenger data from IATA (PaxIS) is described in Section 2.4.3. (See also Appendix B for data cleaning.)

September was selected⁶ as the busiest non-holiday month for traffic and passengers, after the holiday months of August and July. The busiest 199⁷ European Civil Aviation Conference (ECAC) airports are modelled, having previously identified⁸ that these airports accounted for 97% of passengers and 93% of movements in 2010. Routes between the main airports of the EU 27 states (i.e. prior to Croatian accession in July 2013) and airports outside the EU 27 have been used as a proxy for determining the major flows between the ECAC area and the rest of the world. This process allowed the selection of 50 non-ECAC airports for the inclusion of their passenger data.

The baseline traffic day includes delays and cancellations. The delays, however, are reset and randomised on each model run to reflect more typical delay patterns, as explained through the rules of Appendix A and through the use of schedule data. It is thus important to note that the model represents a normative day, to the greatest extent possible, and that the simulation results thus reflect schedule robustness (e.g. whether a given airport has sufficient connectivity and capacity to reaccommodate disrupted passengers) in addition to the rules formulated in Appendix A.

Cancellations are already included in the baseline traffic day (Friday 17SEP10) and cannot be restored. Even if such flights could be somehow restored, it would be wholly impractical to reset all the other flights (and their properties) that may have been affected. The effect of these cancellations are thus part of the baseline traffic day's properties, although we already know that these days were not atypical, due to the selection and filtering process.

We had originally planned in POEM to build other traffic-passenger days from first principles but due to the extremely labour intensive process and the inevitable degree of stochasticity in the resulting traffic-passenger dataset used as input into the model, it was decided instead to derive two disrupted days from the baseline traffic day through the application of internal rules.

Critically, this allows us to make like-for-like comparisons across the three days, instead of incurring the risk that observed differences in the metrics could result from (stochastic) differences in the input data, particularly as the metrics are designed to be sensitive to passenger allocation characteristics.

In order to assess the effects of increased perturbation on the system, the internal rules employed in POEM to create the two disrupted days will impose:

- (i) increased flight departure delays;
- (ii) additional flight cancellations.

It is of course known that departure delays and cancellations often occur together, during various types of disruption, although airline rules on when and where to cancel a flight based on its accumulated delay will vary by operator and (legislative and cost) context. These effects will here be modelled separately: the types of additional cancellations applied are not a function of prevailing delay types or magnitudes. Although one disadvantage of this approach is that this independency is somewhat less realistic than empirical observations, a concomitant advantage is that it will permit independent quantification of increased departure delay and increased cancellations.

⁵ See POEM Deliverable 4.2 for full details. See Appendix E for a summary of strike activity during the baseline traffic day.

⁶ See POEM Deliverable 1.1 for full statistics.

⁷ Doncaster Sheffield/Robin Hood airport (DCS, EGCI) was included in the top 200 airports originally specified, but there were no corresponding traffic data in the PRISME traffic database, so this airport was subsequently omitted.

⁸ See POEM Deliverable 1.2 for full details.

Although we are modelling September 2010 in POEM, cancellation data are not available in any detail for this year, nor are *daily* departure delay data. Such data are available for 2012, however, due to the implementation of Regulation 691/2010. We will draw on such data kindly provided by EUROCONTROL's Central Office for Delay Analysis (CODA) with regard to 29 reporting European airports, to infer rubrics for September operations.

2.4.2.2 High delay day

In Rule 118 of the POEM model (see Appendix A), it is explained that stochastic values are assigned to the departure delays in order to match the target calibration value (see Section 2.5) of **13.9 minutes** for the average departure delay for a flight (across all flights), and 13.6 minutes for the average arrival delay (across all flights). Rule 118 also allows some early departures, all departures being subject to the other constraints and rules of the model (not least, waiting for connecting passengers).

Daily CODA delay data for the month of September 2012 were used in lieu of available daily data for September 2010. The average departure delay for the worst three days⁹ of September 2012 (14.7 minutes) was 1.43 times greater than that of the rest of the month. Applying this simple factor to the September 2010 baseline traffic day ($\mu = 13.9$ minutes), would yield a high delay target value of $\mu = 19.9$ minutes. However, this value exceeds the highest average delay for any day in September 2012, even during strike peaks, such that this value is discarded as being too high.

The average value for the worst three days in September 2012 is 14.7 minutes. By adding one minute onto the baseline September 2010 average value (also through Rule 118), to give 14.9 minutes, a value is obtained that is very close to that of the worst three days in September 2012 and just below the commonly used threshold of 15 minutes used to classify departures as 'on time'. We therefore set the high delay target at an average delay of **14.9 minutes** (across all flights), adding one minute to the baseline traffic day average. This is achieved by increasing the indiscriminate, randomised delays applied in Rule 118, whereby we do not attempt to generate specific spatial effects (e.g. national strikes) or to impose temporal effects (e.g. delays peaking in the evening).

2.4.2.3 High cancellation day

Since the underlying cancellation rates (the percentages of flights cancelled per airport) in the baseline traffic data are not known, and it is further necessary to separately consider the costs associated with cancellations, the methodology for the implementation of the high cancellation day is considerably more involved than that of the high delay day. Furthermore, we have here set out a full methodology, which may be employed under future modelling conditions whereby the underlying cancellation rates are known. Table 3 shows four types of cancellation pattern that could be simulated.

Table 3. Cancellation event types

Reference	Event type	Intensity	Example
1	widespread, multinational	high	ash cloud
2	National	high	ATC strike
3	regional (multinational)	high	sustained winter weather
4	regional (multinational)	low	transient (winter) weather

The types presented are not exhaustive, but illustrate the potential range. The high intensity types, by which is meant high rates of cancellations for a sustained period, present particular problems for the model. Since POEM models the associated costs of disruption, types 1 – 3 are difficult to assess. These events would be known in advance, to a certain extent, and partly handled in advance by airlines, again – to certain and variable extents. This forewarning causes numerous problems in terms of the cost implications for the airlines, for which either extensive further research or assumptions would have to be made. Furthermore, such types of event would involve airports being (effectively)

⁹ Two of these three days followed peak strike action in France and the other one coincided with a day of strike action in Spain, although the latter was notified in advance, thus enabling operators to plan ahead to a considerable extent and with a decrease in traffic in Spain on that day.

closed for sustained periods of time, which would generate numerous modelling difficulties in terms of flights inbound to these destinations and the likely passenger behaviour. For example, if the London airports were closed, which passengers would cancel, defer, substantially re-route and/or change modes with respect to their planned itineraries? Both short- and long-haul alternatives would require extensive socio-behavioural modelling to support the model.

The type 4 event, for example with regional snow or fog causing difficulties for morning operations, would be somewhat easier to model. (Although fog affects certain airports less frequently than others, and snow is not likely in the modelled month of September, a model of transient weather problems will suffice for the purposes of the model, i.e. to assess the impact of increasing cancellation rates).

It is necessary to explore the cost implications of the type of cancellations modelled. Under the Commission's Regulation 261 (European Commission, 2004) the rules regarding cancellation are governed by Article 5¹⁰ (the emboldening and marginal notes have been added). The table shows that the type of cancellations to be assessed are thus already covered by the POEM cost models. (Evidence from a Danish survey suggests that only 2 – 4% of passengers entitled to financial compensation actually received it - European Commission, 2013).

Table 4. Article 5 (cancellation), Regulation 261

1. In case of cancellation of a flight, the passengers concerned shall:

(a) be offered assistance by the operating air carrier in accordance with **Article 8**; and

(b) be offered assistance by the operating air carrier in accordance with **Article 9(1)(a) and 9(2)**, as well as, in event of re-routing when the reasonably expected time of departure of the new flight is at least the day after the departure as it was planned for the cancelled flight, the assistance specified in **Article 9(1)(b) and 9(1)(c)**; and

(c) have the right to compensation by the operating air carrier in accordance with **Article 7**, unless:

(i) they are informed of the cancellation at least two weeks before the scheduled time of departure; or

(ii) they are informed of the cancellation between two weeks and seven days before the scheduled time of departure and are offered re-routing, allowing them to depart no more than two hours before the scheduled time of departure and to reach their final destination less than four hours after the scheduled time of arrival; or

(iii) they are informed of the cancellation less than seven days before the scheduled time of departure and are offered re-routing, allowing them to depart no more than one hour before the scheduled time of departure and to reach their final destination less than two hours after the scheduled time of arrival.

2. When passengers are informed of the cancellation, an explanation shall be given concerning possible alternative transport.

3. An operating air carrier shall not be obliged to pay compensation in accordance with **Article 7**, if it can prove that the cancellation is caused by extraordinary circumstances which could not have been avoided even if all reasonable measures had been taken.

4. The burden of proof concerning the questions as to whether and when the passenger has been informed of the cancellation of the flight shall rest with the operating air carrier.

Note. **Article 8** refers to the reimbursement and re-routing rights, as covered already in POEM Rule 33 (see Appendix A), according to the Regulation.

Note. **Article 9** refers to right of care (meals, accommodation, etc.) as covered already in POEM Rule 33 (see Appendix A), according to the Regulation.

Note. See note below regarding **Article 7**.

Note. By design, **Article 7** therefore does not apply to the specific types of weather-induced cancellations applied in the model.

¹⁰ Proposed changes to Regulation 261 are discussed in Section 1.3.

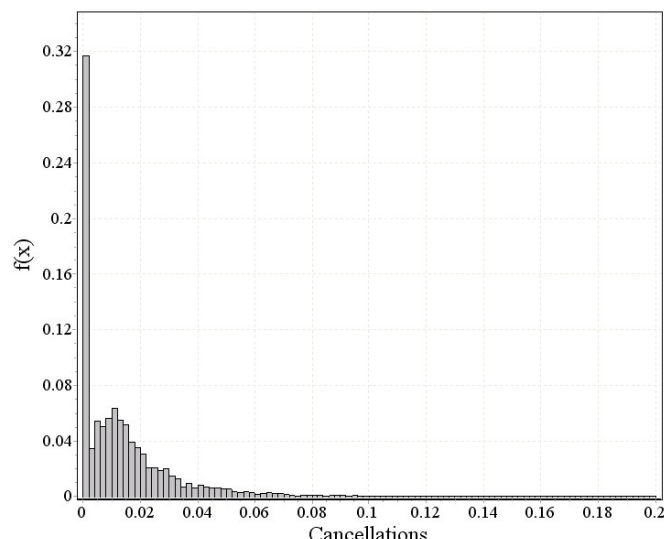


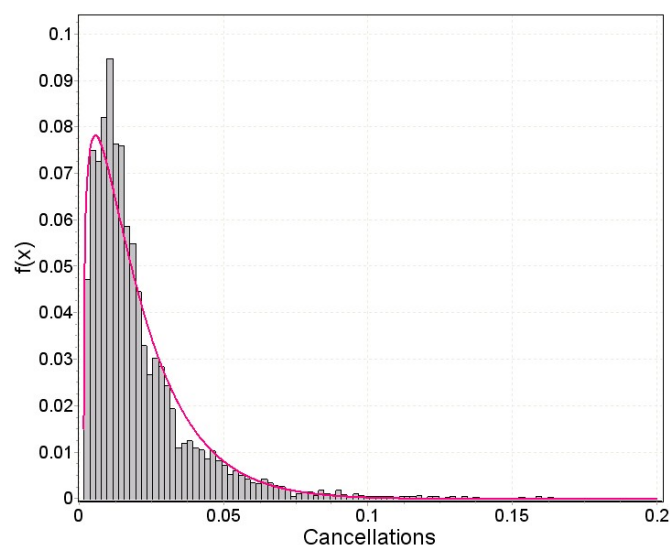
Figure 7. Airport cancellation rates for 2012

Figure 7 shows the full distribution of daily cancellation rates for the 29 airports reporting to CODA, for year 2012 ($n = 10\,496$; 0.8% missing observations; $\mu = 0.015$, $\sigma = 0.024$). A small number (0.2%) of these values are greater than 0.20 (i.e. cancellation rates of over 20%), which have been omitted from the plot (but included in the summary data calculations) to give a better illustration of the shape of the distribution. Notably, 31% of the rates are exactly zero, which makes fitting a distribution to the bimodal curve difficult. In Figure 8, the zeros have been removed and a three-parameter Gamma distribution fitted, with probability density function:

$$f(x) = \frac{(x - \gamma)^{\alpha-1}}{\beta^{\alpha} \Gamma(\alpha)} e^{-(x-\gamma)/\beta} \quad [1]$$

whereby $\alpha > 0$ is the shape parameter, $\beta > 0$ is the scale parameter and $\gamma \leq x$ is the location parameter. Maximum likelihood estimation (MLE) fits¹¹ were performed using EasyFit (Professional) from MathWave Technologies. The more common, two-parameter Gamma distribution is obtained with $\gamma = 0$, but this did not model the leading shoulder of the curve well, giving Pareto-like fits instead (i.e. simply falling away from left to right). With the zero-rate values removed, the mean increases and the standard deviation falls ($n = 7198$; $\mu = 0.021$, $\sigma = 0.020$).

¹¹ Inference problems are sometimes encountered with small sample MLE fits on three-parameter Gamma distributions, but this is out of the current scope of discussion. A reserve solution is to revert to 2-parameter fits, but we have found this unnecessary. Simple moment methods also gave inferior fits.



$$\Gamma(\alpha, \beta, \gamma): \alpha = 1.27, \beta = 0.0151, \gamma = 0.00172 \text{ (3 s.f.)}$$

Figure 8. Non-zero cancellation rates for 2012, with Gamma fit

Table 5. Daily cancellation rates for 2012 (full year c.f. September; 2 s.f.)

Daily cancellation rate metric	Full year (n = 10 496)	Worst three September days (n = 87)	Rest of September (n = 783)
Proportion of zero-rated values	31%	20%	31%
Mean rate per airport*	2.1%	2.2%	1.6%
Standard deviation per airport *	2.0%	1.5%	1.4%
Maximum rate	73% [†]	8.1%	9.5%

* Excluding zero-rated values

[†] Although only 0.2% of values > 20%

Table 5 shows similar data for the worst three days¹² of September 2012 (defined as the three days with the highest average cancellation rates across the 29 reporting airports) and for the rest of the month. The worst three days have an average cancellation rate closer to the annual average (which includes the worse winter months) and lower proportions of zero-rated days, as expected.

¹² These three days did not coincide with any of the three worst delay days, discussed in the previous section, although they were either close to, or included in, peak strike action days in France.

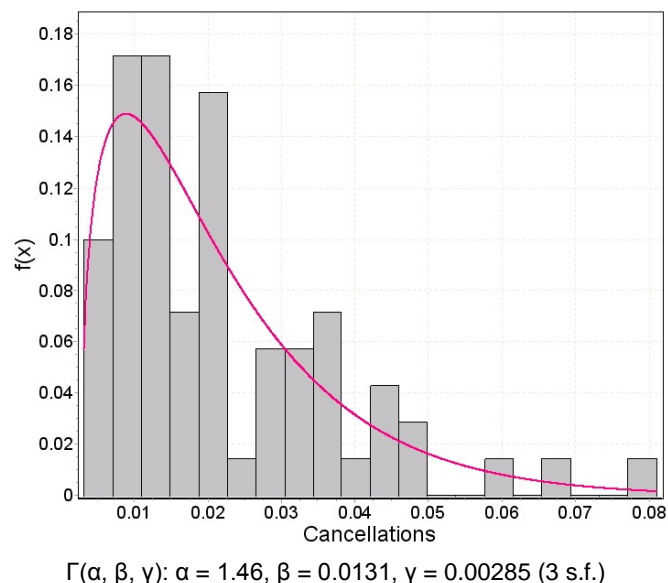


Figure 9. Non-zero cancellation rates for worst three days in September 2012, with Gamma fit

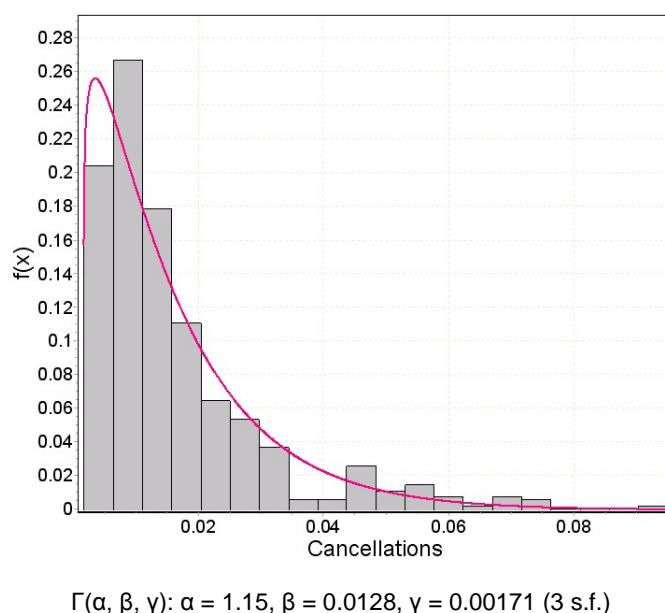


Figure 10. Non-zero cancellation rates for rest of September 2012, with Gamma fit

Figures 9 and 10 show that the three-parameter Gamma distribution also provides a good fit for these two selected periods. As referred to at the start of this discussion, the actual baseline distribution of cancellations, and at which airports they occurred, is unknown in the POEM traffic data. Assuming the distribution follows something fairly close to the ‘rest of September’ plot (Figure 10), we essentially wish to move from this distribution, to the ‘worst three days’ distribution (Figure 9), i.e. to add the *difference* between these shapes (notably adding more right-hand tail) to the second figure. This desired transformation is shown in Figure 11.

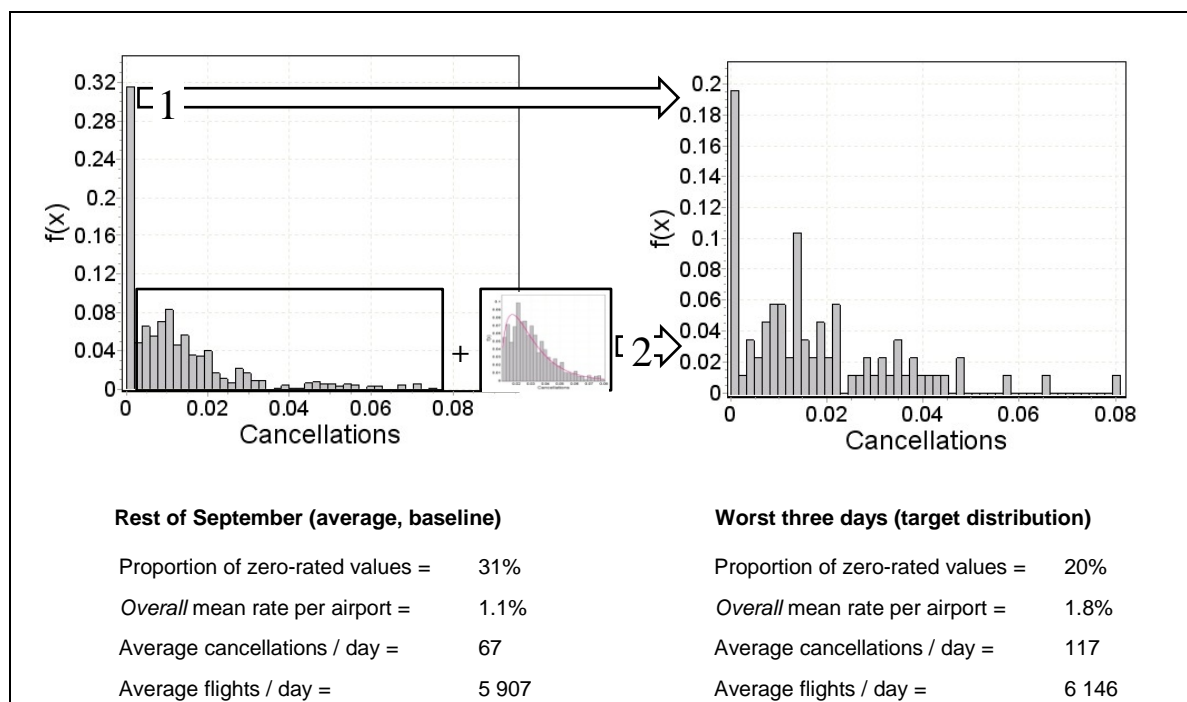


Figure 11. Desired transformation from baseline to target cancellations

Simulating ($n = 10\,000$) these distributions as random variables, using their zero-excluded Gamma distribution fits (as shown in the previous two figures), the *difference* between the two curves is fitted well by several distributions – including, as expected, the Gamma (adopted fit: shown inset, and full size in the following figure) and the Normal (the inverse Gaussian also offering a good fit). With *given* model baseline traffic data, the two steps illustrated in the figure for transforming such average data to one of the worst cancellation days would be:

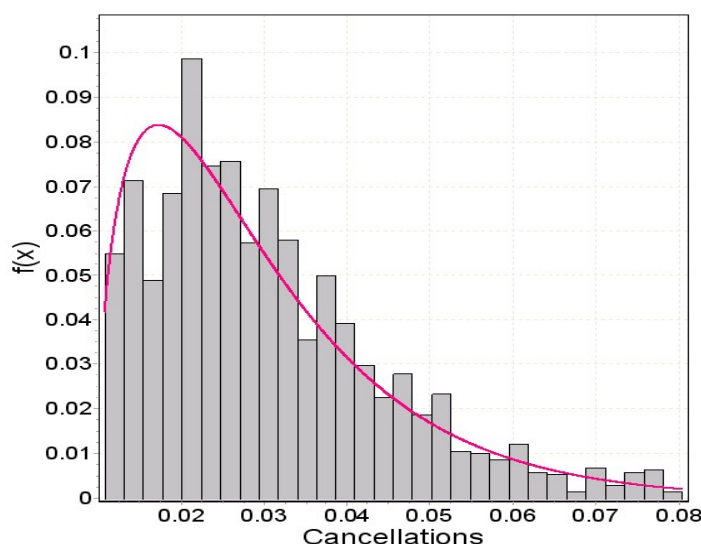
(1) replace some of the zero-rated airports with high-rated cancellations

i.e. reduce the proportion of (baseline) zero-rated airports from 31% (left of figure) to 20% (right of figure), and distribute this 11% as per the Gamma fit obtained for the worst three days, protecting the remaining 20% (right-hand side) as zero-rated;

(2) increase the cancellation rates at the rest of the airports

i.e. adding the *difference* distribution to the boxed area (non-zero-rated) of the left-hand side distribution, thus increasing the tail of the distribution (right-hand side), for example.

Such a method may be employed in future when it is possible to identify the baseline cancellation rates at each airport. However, currently, there is no such information, such that neither the zero-rated airports can be identified, nor the non-zero rates at any other airports, to which additional cancellations can be added to produce the desired distribution. Instead, the available option is to use the fitted Gamma difference distribution shown in Figure 12.



$$\Gamma(\alpha, \beta, \gamma): \alpha = 1.59, \beta = 0.0126, \gamma = 0.00985 \text{ (3 s.f.)}$$

Figure 12. Fitted difference distribution used to transform baseline to target cancellations

The final method used is therefore as follows:

- select an airport at random;
- sample a random value, $0\% < x \leq 8\%$, from $\Gamma(\alpha, \beta, \gamma): \alpha = 1.59, \beta = 0.0126, \gamma = 0.00985$, with re-sampling if $x > 8\%$;
- cancel (an additional) $x\%$ of flights with valid metrics from 0400 to 0900 (inclusive) at the airport selected;
- repeat these steps until the total desired number of additional flights has been cancelled (since this involves discrete airport selection with an applied rate of x_i at the final airport, this will produce a very slightly higher overall number of cancellations, rather than any bias towards a lower number, as illustrated below).

In general, cancellation rates are higher after 1900 UTC for all months and they are also rather higher during the first few hours of many days over the winter months¹³. Although we are here hypothetically modelling transient weather in autumn (such as morning fog), the intention is to more generically assess the impact of an increase in cancellations during the start of the operational day. We therefore judgmentally apply the $x\%$ of cancellations to the first part of the morning, i.e. from first operations (defined as from 0400 local) up to and including 0900. The random selection of airports does not take account of the likely spatial distribution of such weather. Instead, it is assumed to occur spatially at random throughout Europe during this disrupted day. Only flights with valid metrics (see Appendix B) are cancelled, as there is no point in cancelling flights for which the metrics are turned off, as this impact would not be measured in the model.

This method is the statistically best option available to assign additional cancellations to unknown baseline data, preserving some of the unevenness of the observed distributions of cancellation rates across the airports. It statistically (but blindly) reduces some of the zero-rated allocations until the total desired number of cancellations is reached. To minimise the bias associated with assigning the additional cancellations to unknown baseline data, these extra cancellations are allocated randomly in each run of the model. In the September 2012 data, there is actually a statistically significant relationship between airport size and cancellation rates, such that larger airports are less likely to have zero rates. However, the largest airports in the sample did experience zero (and very low) rates on certain days in September 2012, such that we cannot exclude them from assignments of zero rates. Due to the randomisation, however, all runs could not be impacted by a single random assignment of zero rates to the same one or two large airports (e.g. running every simulation with no (additional) cancellations at Heathrow and Frankfurt).

¹³ Personal communication from CODA.

Table 6. Example allocation of additional cancellations

(a) Sequence number of randomly selected airport	(b) Airport ID	(c) Additional cancellation rate	(d) Number of baseline flights	(e) No. of new cancellations [(c) x (d)]	(f) Cumulative total of new cancellations
1	L	2.2%	289	6	6
2	K	3.2%	431	14	20
3	P	5.3%	222	12	32
4	B	2.1%	393	8	40
5	H	1.8%	149	3	43
6	J	5.2%	389	20	63
7	N	3.1%	95	3	66
8	A	2.0%	106	2	68
9	G	3.5%	98	3	71
10	D	5.0%	360	18	89

Table 6 shows an example test allocation (using real data from September 2012) whereby the target was to add another 86 cancellations for the whole day. The values in column (c) are generated from the Gamma difference distribution. By airport number 10 in the sequence, (a), of randomly selected airports, (b), a total of 89 additional cancellations were made, (f), slightly over-allocating compared with if the algorithm had stopped at airport number 9 in the sequence. In the September 2012 data examined above, it was observed that an average baseline day had 67 cancellations, whereas a typical worst day had 117. An extra 50 cancellations would thus need to be added, on average, to a baseline day to create one of the worst days. The number of airports implicated in such an addition of cancellations depends on the:

- number of additional cancellations to be added;
- variance of the number of flights amongst the airports (σ_f^2);
- variance of the Gamma difference distribution used for sampling cancellation rates (σ_G^2).

Let us consider the additional cancellations required to be fixed at 50. Of the 29 airports in the 2012 sample, the average number of flights is 204 with $\sigma_f = 138$. For the Gamma distribution ($\alpha = 1.59$, $\beta = 0.0126$, $\gamma = 0.00985$) the average cancellation rate is 3% with $\sigma_G = 1.5\%$. It is thus expected that on average, over numerous runs, $50/(0.03 \times 204) = 8$ airports would be required to effect the additional cancellations, although this would vary quite substantially from run to run, particularly due to the high value of σ_f . (As σ_f and $\sigma_G \rightarrow 0$, the number of airports in each run would tend to exactly 8.)

Where proportions \bar{p}_w and \bar{p}_b are the simple, daily average cancellation rates (number of cancelled flights per day over the average number of flights per day, \bar{n}_w and \bar{n}_b) for the September 2012 worst days and baseline days, respectively, the estimator of the **proportion of extra cancellations** to be added to the baseline days is given by:

$$\hat{p}_e = \bar{p}_w \left(\frac{\bar{n}_w}{\bar{n}_b} \right) - \bar{p}_b = 0.0085 \quad (\text{i.e. } 0.85\%)$$

[2]

For the September 2012 data this yields $0.0085 \times 5\,907 = 50$ extra cancellations per day. For the POEM baseline traffic day, applying the cancellations only to flights with valid metrics departing from the 199 ECAC airports, this yields $0.0085 \times 18\,526 = 157$ extra cancellations. In the POEM model, with some fairly small airports included in the sample, there is an average of 93 ECAC departures per day with valid metrics per airport in the baseline traffic day. It is thus expected that on average over numerous runs, $157/(0.03 \times 93) = 56$ airports would be required to effect the additional cancellations, although this will again vary quite substantially from run to run. The simplifying assumption was made that subsequent flights involving a previously cancelled aircraft were operated by replacement aircraft (i.e. we did not model the cancellation chains), although in future modelling the cancelled rotations will be specifically propagated.

2.4.3 Assigning passengers to flights

This section updates the passenger allocation process. Major developments were made regarding the final process compared with the methods previously reported in POEM Deliverable 4.2. The assignment of passengers to flights, with cost characteristics and full itineraries, was a fundamental component of POEM, since the project explores new passenger-centric metrics. This process required a very sizeable amount of resource.

The two principal datasets used to prepare the input data for the POEM model were sourced from IATA and EUROCONTROL. IATA's PaxIS dataset provides passenger itineraries aggregated by month, captured through ticket sales; EUROCONTROL's PRISME dataset details European flight movements. (Both datasets were described in detail in POEM Deliverable 1.1 and Deliverable 1.2.)

Passenger allocation algorithms have been developed to assign aggregated passengers for the month onto individual flights – Figure 13 illustrates the challenge, highlighting a sample of passenger itineraries that include an Amsterdam Schiphol to Rome Fiumicino flight leg and the nine passenger flights that flew on the baseline traffic day (17SEP10) that need to be allocated passengers.

Dom_AI	Mar_AI1	Mar_AI2	Mar_AI3	Orig	Connect_2	Connect_3	Dest	Class	Est_Pax	Avg_Fare
KL	KL	KL	KL	ABZ	AMS	FCO	AOI	ECON DISC	4	153.5
KL	KL	KL	AZ	ABZ	AMS	FCO	BRI	ECON DISC	2	180.4
KL	KL	KL	AP	ABZ	AMS	FCO	CAG	ECON DISC	2	167.9
KL	KL	KL	KL	ABZ	AMS	FCO	PMO	OTHER	9	94.9
KL	KL	KL	KL	ABZ	AMS	FCO	TRS	BUSINESS	5	443.7
KL	KL	KL	KL	ACA	MEX	AMS	FCO	ECON DISC	4	223.9
KL	KL	KL	KL	ADL	KUL	AMS	FCO	ECON DISC	8	623.3
AZ	AZ	AZ		AMS	FCO		ACC	ECON DISC	3	344.4
AZ	AZ	AP		AMS	FCO		AHO	ECON FULL	11	105.2
AZ	AZ	AZ		AMS	FCO		AMM	ECON DISC	15	209.5
AZ	AZ	AZ		AMS	FCO		ATH	ECON DISC	100	125
AZ	AZ	AZ		AMS	FCO		ATH	ECON DISC	122	127.2
AZ	AZ	AZ	PZ	AMS	FCO	EZE	CBB	ECON DISC	6	357.6
KL	LP	KL	KL	AQP	LIM	AMS	FCO	ECON DISC	3	425.3
AZ	AZ	AZ	AZ	ARN	AMS	FCO	BDS	ECON DISC	3	180.8
KL	KL	KL	KL	ARN	AMS	FCO	BDS	ECON DISC	3	167.8
KL	KL									
KL	KL									
KL	PZ									
KL	KL									

Aircraft_Operator	Aircraft_Type_ICAO_ID	Conn_Registration	Seats	ADEP	ADES	AOBT_3	ARVT_3	FltNum
KLM	B738	PHBXF	171	EHAM	LIRF	17/09/2010 05:03	17/09/2010 07:04	KLM_EHAMLIRF01
KLM	B738	PHBGB	171	EHAM	LIRF	17/09/2010 07:55	17/09/2010 09:50	KLM_EHAMLIRF02
AZA	A320	EIDSC	159	EHAM	LIRF	17/09/2010 11:29	17/09/2010 13:30	AZA_EHAMLIRF01
EZY	A319	GEZBH	156	EHAM	LIRF	17/09/2010 11:56	17/09/2010 14:00	EZY_EHAMLIRF01
KLM	B738	PHBXF	171	EHAM	LIRF	17/09/2010 11:49	17/09/2010 13:51	KLM_EHAMLIRF03
KLM	B739	PHBXR	189	EHAM	LIRF	17/09/2010 14:31	17/09/2010 16:34	KLM_EHAMLIRF04
AZA	A320	EIDSA	159	EHAM	LIRF	17/09/2010 15:07	17/09/2010 17:08	AZA_EHAMLIRF02
AZA	A320	IBIKU	159	EHAM	LIRF	17/09/2010 17:13	17/09/2010 19:24	AZA_EHAMLIRF03
KLM	B738	PHBXM	171	EHAM	LIRF	17/09/2010 18:41	17/09/2010 20:37	KLM_EHAMLIRF05

Figure 13. Aggregated PaxIS passenger itineraries onto disaggregated PRISME flights

2.4.3.1 Recap of passenger allocation issues

The previous iteration of allocating passengers to flights on 500 routes identified a number of allocation problems to be addressed. Table 7 summarises these.

Table 7. Identified passenger allocation issues

Allocation problem	Problem description	Outcome
Passenger itinerary data structure	Raw passenger itineraries contained the details of one to potentially thousands of passengers, causing problems when allocating passengers to flights from itineraries with multiple passengers	Itineraries restructured so that each record consists of one passenger (refer to Appendix B)
Code share and airline alliance passengers not considered	Passengers were only allocated to flights if the airline-issued ticket record matched the operator (e.g. Lufthansa flights were only allocated passengers with Lufthansa tickets, thereby excluding passengers from other Star Alliance airlines)	Range of passengers now allocated to flights, based on tickets issued by code sharing, alliance and partner airlines
Allocated passengers per airline-route should remain constant	Partly caused by the itinerary data structure, overall passengers per airline-route (e.g. KLM: London Heathrow to Amsterdam Schiphol) reduced as allocated passengers were balanced iteratively to resolve overloaded aircraft	Itineraries have been restructured and improved allocation greatly reduces any required balancing
Unsophisticated distribution of passengers to multiple flights per airline-route	Sequential distribution of passengers per flight per airline-route resulted in full flights in the morning with load factor tailing off over the course of the day	New stochastic distribution of passengers per flight per airline-route
Allocated passengers given unachievable connections	Connecting passengers were allocated to flights by flight numbers rather than considering arrival, connecting and departure times	Connections now allocated using published schedule times and appropriate MCTs at each airport

2.4.3.2 Passenger allocation targets

After data cleaning and recoding (summarised in Appendix B), 58.5 million passengers for SEP10 were available for allocation onto individual flights. Table 8 shows the breakdown by flight leg for the month and the finalised passenger allocation proportion targets per flight leg (right hand column). Overall, 80% of passengers should be allocated to single/direct flights with the remaining 20% of passengers making one or two connections. A target was not set to allocate passengers to itineraries consisting of three connections (four flight legs) – these itineraries only accounted for 0.02% of passengers overall and would be difficult to achieve within the 24-hour window of the model.

Table 8. SEP10 passenger breakdown by flight leg and allocation proportion targets

Flight legs	Passenger itinerary	Total passengers	Passenger proportions	Passenger proportion targets
1	direct flight	47 060 468	80.36%	80%
2	one connection	10 736 080	18.33%	18.75%
3	two connections	753 341	1.29%	1.25%
4	three connections	12 327	0.02%	0%
Totals		58 562 216	100%	100%

By connecting and using multiple flight legs, the overall number of available passenger *movements* increases from 58.5 million to almost 71 million passengers for the month, i.e. $(1 \times 47m) + (2 \times 11m) + (3 \times 750k) + (4 \times 12k)$. With the target breakdown between flight legs established, the target number of passengers for 17SEP10 was estimated from actual seat capacities – as there are more flights on busy days, seat capacity is the basis for estimating the daily passenger target from the aggregated monthly data.

Having excluded non-commercial passenger flights, 14 272 unique airline-routes were available on 17SEP10 from 25 797 flights. The summation of seating capacity from the contributing aircraft on each 17SEP10 airline-route was compared with the equivalent available seating capacity per airline-

route for SEP10. Taking the airline-route 'AFR_LFMNLFPO' (Air France Nice Côte d'Azur to Paris Orly) as an example, on 17SEP10, 3 217 seats were available on the 20 Air France flights, accounting for 3.6% of the monthly capacity (88 000 seats) in this direction. In SEP10, the PaxIS dataset has 68 632 passengers (all ticket fares and classes) on the AFR_LFMNLFPO airline-route, contributing the majority of the 92 700 passengers on the Nice Côte d'Azur to Paris Orly route (i.e. including passengers from all airlines). Applying the 3.6% proportion to the 68 632 AFR_LFMNLFPO passengers predicts 2 502 passengers for 17SEP10 with a load factor of 78%. This estimation gave the initial target passengers per airline-route before adjustments were made for:

- **low load factors:** a minimum load factor could not be breached for each airline type (10% below the generic load factor for full-service, regional, low-cost carrier or charter – refer to Rule 36 (Appendix A) for generic load factors;
- **high load factors:** airline-routes with 100% load factors (i.e. every seat occupied on all flights flown by an airline on the route) had their passengers reduced to a capped maximum load factor (95%).

Taking into account these load factor adjustments, the overall target passenger total for 17SEP10 was calculated to be 2 868 522 passengers (from the 14 272 unique airline-routes). This suggested the overall load factor for 17SEP10 to be 75-77%, depending on whether maximum or average seats per aircraft are considered. Table 9 applies the target breakdown between flight legs to this overall daily total.

Table 9. 17SEP10 passenger allocation targets

Flight legs	Overall passenger target	Passenger proportion targets	Passenger allocation targets
1	2 868 522	80%	2 294 818
2		18.75%	537 848
3		1.25%	35 856
4		0%	0
Totals		100%	2 868 522

This overall passenger target and its proportional breakdown by flight leg were used to monitor progress through the ensuing passenger allocation iterations.

Next, with each airline-route allocated a target total for the day, the passengers were distributed among the available flights. Airline-routes with a single flight were allocated all the corresponding passengers. Airline-routes with multiple flights were allocated passengers stochastically, respecting the maximum seating capacity per flight and the overall number of target passengers for the airline-route. This was achieved through a series of distribution parses:

- allocating passengers to each flight according to random sampling from the Normal distribution, with load factor as mean;
- assigning the total passengers in proportion to parse 1;
- allocating the shortfall (due to rounding) from parse 2, randomly across the flights;
- removing the overcapacity passengers by re-allocating them sequentially to the flights, using only 50% of the remaining capacities (in order to try to avoid simply filling up all the early flights in the sequence);
- allocating any remaining required passengers, using all remaining capacity.

Figure 14 summarises the allocation of actual monthly passenger totals to the individual flights on 17SEP10. The Air France Nice Côte d'Azur to Paris Orly airline-route example (AFR_LFMNLFPO) is shown in bold.

Actual total passengers	58 562 216
Actual total passenger movements (including multiple flight legs)	70 841 959
Actual total passengers per airline-route	
AAF_DAAELFML	495
AFR_LFMNLFPO	68 632
XLF_LICJLFRS	363
Total	70 841 959
SEP10	
Target total passengers per airline-route (predicted daily capacity as a proportion of the monthly capacity)	
AAF_DAAELFML	(1 flight) 99
AFR_LFMNLFPO	(20 flights) 2 502
XLF_LICJLFRS	(1 flight) 113
Total	(25 797 flights) 2 868 522
Target total passengers per flight leg	
Single-flight leg	2 294 818
Two-flight leg	537 848
Three-flight leg	35 856
Four-flight leg	0
Total	2 868 522
Target total passengers per flight (stochastically distributed)	
AAF_DAAELFML01	(144 seats) 99
AFR_LFMNLFPO01	(212 seats) 120
AFR_LFMNLFPO02	(142 seats) 99
AFR_LFMNLFPO03	(212 seats) 129
AFR_LFMNLFPO04	(178 seats) 95
AFR_LFMNLFPO05	(178 seats) 118
AFR_LFMNLFPO06	(212 seats) 136
AFR_LFMNLFPO07	(142 seats) 112
AFR_LFMNLFPO08	(142 seats) 112
AFR_LFMNLFPO09	(142 seats) 110
AFR_LFMNLFPO10	(178 seats) 169
AFR_LFMNLFPO11	(142 seats) 114
AFR_LFMNLFPO12	(142 seats) 103
AFR_LFMNLFPO13	(212 seats) 157
AFR_LFMNLFPO14	(142 seats) 100
AFR_LFMNLFPO15	(212 seats) 146
AFR_LFMNLFPO16	(178 seats) 131
AFR_LFMNLFPO17	(178 seats) 129
AFR_LFMNLFPO18	(212 seats) 144
AFR_LFMNLFPO19	(178 seats) 122
AFR_LFMNLFPO20	(178 seats) 156
XLF_LICJLFRS01	(189 seats) 113
Total	2 868 522
17SEP10	

Figure 14. Summary of the assignment of target passengers to individual flights

2.4.3.3 Passenger itinerary pre-allocation

Having determined the target totals per flight, actual passenger itineraries now needed to be assigned. Overall, allocated passengers needed to be distributed proportionally between flight legs, so that 80% of passengers were allocated to a single flight, 18.75% onto two connecting flights and 1.25% onto three connecting flights.

Passengers in scope for allocation to flights on 17SEP10 were prepared as one-, two- and three-flight leg passengers. To accommodate load balancing, more passengers were prepared than required by the flight leg targets to allow flexibility when allocating passengers among flights – with excess passengers discarded.

As single leg passengers (without connections) could be allocated more flexibly, unconstrained by the requirement to connect with other flights, the more complex three-flight leg passengers with two connections were pre-allocated.

(i) Three-flight leg passenger itineraries

These were the most constrained passengers, requiring viable connections at two airports. Three-flight leg passengers were augmented with four-flight leg passengers (less the first or last leg) with

fares appropriate and available for the reduced itinerary (i.e. the longest flight leg must be retained after dropping the first or last leg).

All combinations of connections were built for each three-flight leg passenger itinerary using schedule (at-gate) times. Average MCTs per airport were used – with the average ‘domestic’ MCT applied to connections for arriving flights that were ≤1500NM and the average ‘international’ MCT applied to connections for arriving flights that were >1500NM.

Table 10 illustrates the connection options available to a viable three-flight leg itinerary, for a passenger flying from Aarhus (EKAH) to Alicante (LEAL) via Copenhagen (EKCH) and Madrid Barajas (LEMD).

Table 10. Connection options for a three-flight leg itinerary

Flight leg	Airline-route	Options	Flight number	SOBT	SIBT	SIBT+MCT
1	SAS_EKAHEKCH	5	SAS_EKAHEKCH01	07:20	08:00	08:30
1	SAS_EKAHEKCH	5	SAS_EKAHEKCH02	09:05	09:40	10:10
1	SAS_EKAHEKCH	5	SAS_EKAHEKCH03	13:40	14:15	14:45
1	SAS_EKAHEKCH	5	SAS_EKAHEKCH04	15:55	16:30	17:00
1	SAS_EKAHEKCH	5	SAS_EKAHEKCH05	18:10	18:45	19:15
2	SAS_EKCHLEMD	1	SAS_EKCHLEMD01	09:00	12:20	14:04
3	JKK_LEMDLEAL	3	JKK_LEMDLEAL01	07:20	08:20	08:50
3	JKK_LEMDLEAL	3	JKK_LEMDLEAL02	12:40	13:45	14:15
3	JKK_LEMDLEAL	3	JKK_LEMDLEAL03	18:55	19:55	20:25

Note. All schedule times are local; viable connecting flights in bold; unviable connecting flights in grey.

The constraining connection in this example is the second flight leg – with just one flight available there are no alternatives to SAS_EKCHLEMD01. If none of the first or third flights connect with sufficient time¹⁴ then the itinerary is discarded (in such instances the itinerary may be more suited to other days in SEP10 with different schedule times, or an overnight stopover would be required). The two right-hand columns give the scheduled in-block time per arrival and the scheduled arrival time with the appropriate MCT applied. In this example, a valid flight itinerary would be:

First flight leg: **SAS_EKAHEKCH01** dep EKAH 0720; arr EKCH 0800; 30 minutes MCT
Second flight leg: **SAS_EKCHLEMD01** dep EKCH 0900; arr LEMD 1220; 104 minutes MCT
Third flight leg: **JKK_LEMDLEAL03** dep LEMD 1855; arr LEAL 1955

Itineraries containing flight legs with no corresponding PRISME flights were also discarded. This process was repeated for other three-flight leg itineraries, pre-allocating passengers with tickets that matched the operating airline. In cases with a choice of viable connections (unlike the above example) the connecting flights were randomly selected from the three shortest connection times.

The next iteration repeated this process, focusing on passenger itineraries with ticketing across code shares, alliances and partner airlines. Partner examples include shuttle flights feeding hub airports – the PRISME flight record reports the *operating* airline such as British Airways Shuttle (‘SHT’) or CityJet (‘BCY’), whereas the passenger ticket reflects the *marketing* airline (British Airways, ‘BAW’ or Air France, ‘AFR’). Alliance examples include allowing Scandinavian Airlines or Spanair ticketed passengers to be pre-allocated to Lufthansa operated flights.

Passengers with valid three-flight leg itineraries were then randomly selected and pre-allocated to their corresponding flights.

¹⁴ I.e. the arriving SIBT + MCT is later than the departing SOBT.

(ii) Two-flight leg passenger itineraries

The same process was repeated for two-flight leg passengers, although with just one connecting airport. Only passengers with two-flight leg itineraries were considered – there was no requirement to augment this target group with passengers from three- or four-flight itineraries.

As with the pre-allocation of three-flight leg passengers, MCTs were applied to ensure connections were achievable. Iterations considered passengers with tickets matching the operating airline, as well as itineraries based on code share, alliance and partner airlines.

Passengers with valid two-flight leg itineraries were then randomly selected and pre-allocated to their corresponding flights.

(iii) Single-flight leg passenger itineraries

Although single-flight itineraries are far simpler, with no need to synchronise connecting flights, the volume of records (over 42 million) to be processed placed considerable burden on a typical desktop PC's resources.

The first iteration (again) pre-allocated passengers with tickets matching the operating airline, with further iterations pre-allocating passengers with code share, alliance and partner tickets. Single-flight leg passengers were thus pre-allocated to flights, in most cases increasing the total number of passengers to satisfy the target. Any excess passengers were discarded.

However, not all PRISME airline-routes could be matched with passenger itineraries, or could only be partially allocated passengers: some airline-routes had either no corresponding passengers with no alliance or other alternative passengers available to allocate to the flights, and some airline-routes had insufficient corresponding passengers to meet the target(s). These shortfalls are discussed in the next section.

2.4.3.4 Passenger itinerary final allocation

Before addressing the problem of flights with missing or insufficient passengers, it was important to ensure that the three- and two-flight leg passenger totals were as near as possible to the targets.

Checks were made to ensure that flights were not overcapacity. This was possible at the pre-allocation stage since multi-flight leg passengers affected several flights – in some cases a common passenger itinerary allocated passengers onto the same flights (particularly if alternative flights failed to connect at key airports), resulting in the smaller aircraft on one of the flight legs becoming overloaded. Reducing the number of two- and three-flight leg passengers to correct an overcapacity flight caused other flights to fall under the previously met target. In some instances, discrepancies such as these could only be remedied by substituting multi-flight leg passengers with single leg passengers (on the affected flights). For a limited number of flights, a small number of pre-allocated passengers had to be duplicated to satisfy the targets of flights with insufficient passengers.

On completion of the checks for overloaded flights, that multi-leg passengers were balanced between flights and that multi-flight leg targets were achieved overall, the final allocation step was to impute passengers for the 14.6% of flights with *no* matching PaxIS itineraries. In addition to missing passengers at the airline-route level (e.g. 'ADR_ENALLJLJ' Adria Airways Alesund Vigra to Ljubljana) these flights were also missing passengers at the route level (e.g. ENALLJLJ Alesund Vigra to Ljubljana – all operators) and thus could not employ code share, alliance or partner airlines. Flights without passengers included unidentified flights for which flight details were known apart from the operator (coded as 'ZZZ'), and overall, approximately half of the flights without any passengers were categorised as either charter or unknown operation types¹⁵.

Imputed passengers were simply the (stochastically distributed) target number of passengers per flight. Taking the Adria Airways flight 'ADR_ENALLJLJ01' as an example, this had a target of 56 passengers (Canadair Regional Jet 900 with 86 seats) – with no matching passenger itineraries available, the target of 56 passengers was simply allocated.

These imputed passengers need to be treated as 'inert', i.e. having no effect on the metrics. Flights with passengers allocated in this manner were demoted (see Appendix B) but retained in the model to

¹⁵ 3 769 flights with unmatched passengers; operator type categorised as: full-service (9%), low-cost (16%), regional (26%), charter (33%) and unknown (16%) – with unknown treated as regional by the model.

offer reaccommodation options. Imputed passengers were only allocated to single leg itineraries and were assigned average economy fares based on origin-destination distances¹⁶. Figure 15 plots average economy fares against the great circle distance of the itinerary origin-destination (rather than per flight-leg), giving a reasonable linear fit ($r^2 = 0.70$) with 50 outliers removed.

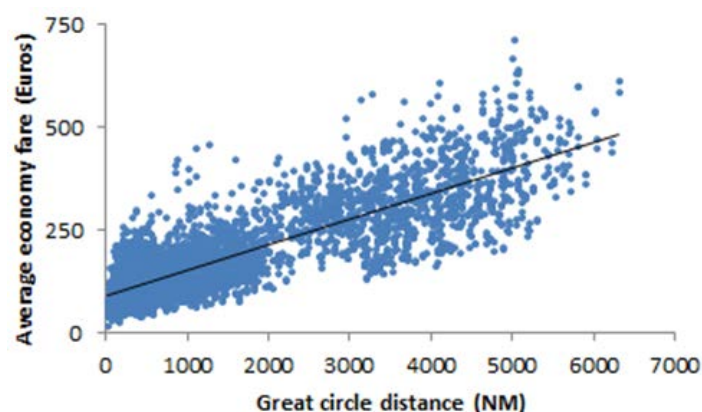


Figure 15. Average economy fare per origin-destination GCD

2.4.3.5 Passenger connectivity files

With passengers allocated to flights, two passenger connectivity files were required by the model to track passengers' intentions when reaccommodation was necessary. Following Rule 33 (see Appendix A), passengers with first class tickets plus those having paid the highest fares of all other fare types were classified as flexible (10% overall), with the remaining 90% of passengers assigned inflexible tickets. Note that all inert passengers were assigned to the inflexible group.

Table 11. Passenger connectivity file – inflexible passengers (extract)

Flight leg 1	Flight leg 2	Flight leg 3	Fare (Euros)
AEA_KJFKLEMD01	JKK_LEMDLEAL03		181.0
AZA_LICALIRF02	AZA_LIRFLEMD01	JKK_LEMDLEAL03	171.5
DLH_LKPREDDF02	JKK_EDDFLEMD01	JKK_LEMDLEAL03	159.8
DLH_LKPREDDF02	JKK_EDDFLEMD01	JKK_LEMDLEAL03	159.8
DLH_USSSEDDF01	DLH_EDDFLEMD02	JKK_LEMDLEAL03	233.9
DLH_USSSEDDF01	DLH_EDDFLEMD03	JKK_LEMDLEAL03	233.9
JKK_LEMDLEAL03			104.1
JKK_LEMDLEAL03			104.1
JKK_LEMDLEAL03			89.0
JKK_LEMDLEAL03			41.5
JKK_LEMGLEMD01	JKK_LEMDLEAL03		51.3
JKK_LEMGLEMD01	JKK_LEMDLEAL03		51.3
JKK_LYBELEMD01	JKK_LEMDLEAL03		106.0
SAS_EKAHEKCH01	SAS_EKCHLEMD01	JKK_LEMDLEAL03	75.9
SAS_EKAHEKCH01	SAS_EKCHLEMD01	JKK_LEMDLEAL03	75.9
SAS_EKAHEKCH01	SAS_EKCHLEMD01		119.9
SAS_EKAHEKCH01	SAS_EKCHLEMD01		100.2
SAS_EKAHEKCH01	SAS_EKCHLEMD01		100.1
SAS_EKAHEKCH01	SAS_EKCHLEMD01		100.1
SAS_EKAHEKCH01	SAS_EKCHLEMD01		224.7

¹⁶ Economy fare = $(0.062 \times \text{GCD}) + 89.6$.

Table 11 provides an extract of the inflexible passengers' connectivity file, which includes two passengers with the same three-flight leg itinerary illustrated previously (Aarhus to Alicante via Copenhagen and Madrid Barajas – shown in bold). Each row contains the itinerary of one passenger, with flight legs and fare paid for the sequence of flights (i.e. one way fare). Based on the schedules and MCTs, all allocated connections were viable.

The outcome of the passenger allocations is summarised in the next section, on model calibration.

2.5 Model calibration

Before presenting the results of the model, we summarise in Table 12 the key model output calibration values (finalising work first presented in Table 17 of POEM Deliverable 5.2), demonstrating good correspondence between the calibration targets and model values. Additional notes regarding calibration are given in the corresponding rules (Appendix A). Table 13 summarises the relationships between the rules and the case studies.

Table 12. Key model output calibration values

Delay value	Definition	Target mean	Target SD	POEM mean	POEM SD	Notes on target values
Departure delay	AOBT - SOBT	13.9 ^Δ	35.9	13.8 ^{¶Δ}	30.8	sourced from CODA data, SEP10, 457 571 flights*
Taxi-out delay	actual - filed taxi-out	unknown	unknown	-	-	no information available on <i>planned</i> times [‡]
En-route delay	filed - actual duration	0.6	7.0	-0.4	8.7	based on PRISME data, 17SEP10, 21 552 flights (valid metrics only)
Taxi-in delay	actual - filed taxi-in	unknown	unknown	-	-	no information available on <i>planned</i> times [‡]
Arrival delay	AIBT - SIBT	13.6	38.8	13.5 [¶]	28.4	sourced from CODA data, SEP10, 457 571 flights*
Reactionary delay	proportion of dep. delay that is secondary [†]	46%	N/A	49% [¶]	-	sourced from CODA data for SEP10**

* 199 ECAC airports modelled in POEM; early values recoded as zero; SDs calculated using simple weighted variances.

** See: Central Office for Delay Analysis (2010).

[†] Includes IATA delay codes 91-96: load connection (91), through check-in of passengers/baggage (92), aircraft rotation (93), (cabin) crew rotation (94 and 95 – these tend to be grouped together by airlines: see Jetzki (2009)), re-route/diversion/consolidation (96). The same value (46%) applies to the whole year: EUROCONTROL (2011a).

[‡] See Rules 19 and 20 (taxi out) and Rule 29 (taxi in), Appendix A, for further information.

[¶] As cited in Table 20 for baseline, no-scenario model run.

^Δ Approximately 30% of departures are early (driven in the model by Rule 118 – see Appendix A).

Table 13. Correspondence between rules and case study inputs

Rule number	Case study link
23 Modelled en-route time variations and recoveries	Draws on the results of Case Study 2 (see POEM Deliverable 7.2)
33 Passenger reaccommodation - protocol at airports for rebooking onto onward destinations	
35 MCTs for airports - possible passenger connections	Draws on the results of Case Study 1 (see POEM Deliverable 7.1)

In Section 2.4.3, the methodology for allocating aggregated passengers to individual flights was presented. Although imputed data were required to supplement real passenger itineraries, Table 14 shows that the overall flight-leg proportion targets were reached.

Table 14. 17SEP10 allocated passengers compared with allocation targets

Flight legs	Targets			Allocated (actual)			
	Overall passengers	Passenger proportion	Passenger allocation	(Active) passengers	Inert passengers	Total passengers	Passenger proportions
1	2 868 522	80%	2 294 818	1 935 172	361 144	2 296 316	80.06%
2		18.75%	537 848	536 060	0	536 060	18.69%
3		1.25%	35 856	35 997	0	35 997	1.25%
4		0%	0	0	0	0	0%
Totals		100%	2 868 522	2 507 229	361 144	2 868 373	100%

No aircraft was overcapacity and, as noted, based on the schedules and MCTs, all allocated connections were viable. The overall passenger load factor for the baseline traffic day (17SEP10) was in the range 75 – 77%, depending on whether maximum (75%) or average (77%) seats per aircraft were considered. Although daily load factors are not reported, the modelled load factors for 17SEP10 compare well with weekly and monthly statistics published at the European level by the Association of European Airlines (Table 15).

Table 15. AEA load factors

Load factor	AEA time period	AEA description
75.0%	SEP10	Europe total – sum of Cross-Border Europe and Domestic
76.2%	SEP10	Cross-border Europe – includes all cross border/international routes originating and terminating within Europe (including Turkey and Russia up to 55°E), Azores, Canary Islands, Madeira and Cyprus
77.5%	13SEP10-19SEP10 (week 38)	Cross-border Europe – includes all cross border/international routes originating and terminating within Europe (including Turkey and Russia up to 55°E), Azores, Canary Islands, Madeira and Cyprus

Load factors and descriptions from AEA Monthly and Weekly Monitors (AEA 2013)

3 Results of the model

3.1 The challenges of establishing causality

In Section 2.2 we overviewed the metric types to be used in POEM, also referring to an objective to characterise the propagation of delay through the network. We here build on groundwork established in POEM Deliverable 5.2, exploring the associated difficulties of establishing causality. For these introductory illustrations we will use the baseline traffic day's input data.

Airport typologies may be determined by the role played in the propagation of delay, as outlined in Table 16.

Table 16. Airport typologies – overview

Airport type	Summary description
Delay source	Outbound delay is non-zero, whereas inbound delay is zero (e.g. during the first departures of the day)
Delay sink	Outbound delay is zero, whereas inbound delay is non-zero (e.g. during the last arrivals of the day)
Delay transferor	Outbound delay = inbound delay (e.g. arrival delays and departure delays are equal, with neither recovery nor exacerbation during turnaround)
Delay amplifier	Outbound delay > inbound delay (e.g. congestion at the airport and/or ATFM departure delays exacerbate arrival delays)
Delay attenuator	Outbound delay < inbound delay (e.g. airline buffers and/or expedited turnaround times enable arrival delays to be partially recovered)

An airport may act in one or more of these roles at different periods during the day. The first three types are included for completeness, although they will be relatively rare compared with the final two, as they require very specific operational characteristics to be manifested (i.e. no inbound delay, no outbound delay, or an exact equality of the two).

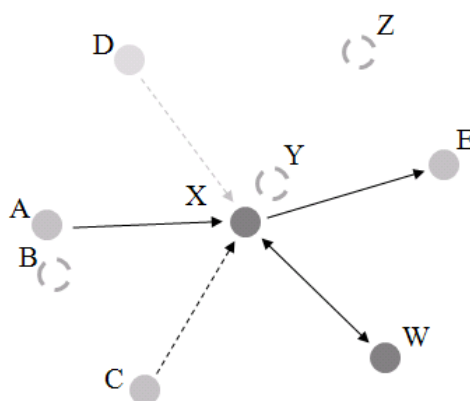


Figure 16. Delay propagation and airport typologies

Figure 16 shows a central node, X, in an airport network. We may readily propose various types of generic delay propagation effects through such a network. Firstly, suppose some aircraft operate at X several times during the same day. Delay incurred at X may be back-propagated to X later during the day, experiencing either amplifier or attenuator effects at outstations such as W. This is addressed by a specific back-propagation metric in POEM. Taking Frankfurt as an example and examining all movements during the baseline traffic day, the average number of times that a specific aircraft arrives

or departs is 2.9. (The most common numbers of movements were one ($n = 120$) and two ($n = 195$); the maximum was ten ($n = 1$; a Lufthansa CRJ-700 mostly shuttling to/from Basel-Mulhouse-Freiburg)).

Another simple example of delay propagation starts with late departures from A arriving late at X. Again, taking Frankfurt as an example, the 17 arrivals from London Heathrow had an average departure delay of 6.2 minutes and an average arrival delay of 3.1 minutes (implying en-route recovery and/or schedule buffering). Inbound flights may cause onward delay due to direct factors such as connecting passengers, crew dependencies or the same aircraft being used for the onward destination, or indirect factors such as gate occupancies. Taking A-X-E as an example, again with Frankfurt as X, the most numerous departures during the baseline traffic day were to Berlin Tegel (22 flights). Of course, such dependencies may be multiple (such as W-X-E) and some may be a subset of $X \leftrightarrow W$ (e.g. A-X aircraft returning from Frankfurt to Heathrow, which most in fact did).

We next examine more subtle effects. For example, airport pairs where one end of the route is a relatively busy airport, C, with a significant number of passengers flying to X, but with no direct flights serving C-X. It may thus be envisaged that late departures from C may cause late arrivals at X, due to delays propagated through the intermediate (connecting) airports. This emphasises the importance of considering the passenger connectivities in these analyses. An example of airport C is Valencia, with the PaxIS data indicating approximately 2 000 passengers flying this itinerary in September 2010, none of them on direct flights. Indeed, during September 2010 some 325 000 passengers arrived at Frankfurt having made one or more pre-connecting flights. Such effects, summed over the whole network for hubs and non-hubs alike, are clearly very large.

Consider also airports such as D with a significant number of departures during the day but with no direct flights, and rather fewer connecting passengers, for X. Of course, the threshold could be arbitrarily defined for 'rather fewer' passengers. Both Trondheim and Valencia had around 90 departures on the sample day, whereas the former carried only 700 connecting passengers for Frankfurt during that month.

Whether this *particular* flow is below the threshold for observable effects on Frankfurt is not the point in question, as we would be certain to find such instances elsewhere. However, by the time we find origins with almost no connecting passengers into X, e.g. from Z, these airports may themselves be very small and not play a large role in network delay propagation. We may also more readily find such Z-X relationships where X is also smaller, particularly where X is not a hub. Furthermore, many pairs of airports may demonstrate apparent relationships such as sharing common delay peaks from 0600 to 1000 simply because these are commonly congested periods, i.e. with no causal relationships at all.

Finally, we return to A. Often, we may expect that neighbouring airports, such as B, have no direct flights to X. Thus A and B will often be a primary-secondary airport pair in the same city. An example of B (if we take A to be London Heathrow) is London Gatwick – indeed neither Gatwick, Luton nor Stansted had any direct flights to Frankfurt during the baseline day. (Other such pairing examples for Frankfurt include Paris Orly, Berlin Schönefeld, Milan Bergamo and Stockholm Bromma). Focusing on Gatwick, the PaxIS data did show some passengers (averaging one per day) flying from Gatwick direct to Frankfurt in September 2010, although there were no such direct flights in the PRISME data for the whole month. This may be due to a ticketing nuance but the number is negligible, in any case. Around 120 passengers flew from Gatwick to Frankfurt indirectly.

More importantly for network effects, however, due to the proximity of Gatwick and Heathrow, is that dependent airspace factors may prevail. We may find that delays at Heathrow are associated with delays at Gatwick, and hence that both are associated with delays at Frankfurt, even though Gatwick has no direct flights and negligible passenger connectivities with Frankfurt. There may be other types of A-B dependencies, such as B being en-route between A and X or sharing some other type of airspace dependency. Whereas some of these A-B relationships may be transparent, others may be very difficult to ascribe any cause and effect relationships to. Some may indeed be just statistical noise. The A-B relationships prompt a corresponding type of relationship for X, with a proximal airport Y (e.g. Cologne-Bonn). Multicollinearities may also be expressed through relationships as complicated as B-(A-X)-Y. These effects are summarised in Table 17 as interaction levels.

Table 17. Airport typologies – by interaction level

Interaction level	Example	Same airport	Connected by flights	Connected by pax	Indirect effects only	No causal relationship
4	X<>W	✓	✓	✓		
3	(A-)X(-E)		✓	✓		
2	C-X, D-X [†]			✓		
1	B-X, Y-X [‡] ; B-Y				✓	
0	Z-X					✓

[†] Strong and weak, respectively

[‡] Remote and proximal, respectively

It is thus important that we understand the underlying context of various metrics analysed in POEM, particularly the difference between the passenger layer and the flight layer. This applies to the factor analyses, the complexity metrics and the classical metrics. Some of the classical metrics play an important role in characterising special properties of the system, which we cannot detect using the higher-level approach taken through the factor analyses and the complexity metrics. An example is the back-propagation metric, which specifically measures the ratio of delay from an airport that later propagates back to the same airport, as discussed above.

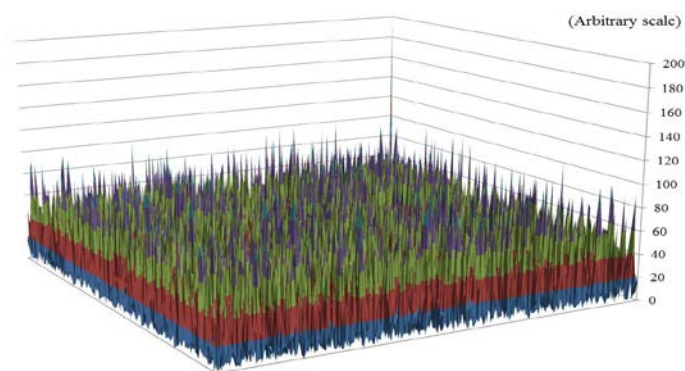
Taking a preview of some key network characteristics, Figure 17 shows (panel 1) that a simple examination of delay correlations between airports¹⁷ produces many instances of cotemporaneous delay between the 199 ECAC airports. This underlines the importance of a rather more sophisticated approach to understanding these properties, as will be explored in detail. Panel 2 shows the number of flights between each airport for the baseline traffic day, with the direct *and* indirect passengers between these pairs shown in panel 3. Although these two patterns are broadly similar, panel 4 emphasises the importance of considering *indirect* (connecting) passengers, i.e. many of which correspond to airport pairs between which no flights operate: such as LGAV (Athens) and EDDH (Hamburg).

These principles and the issue of establishing causality in the context of delay propagation will feature prominently in the results that follow.

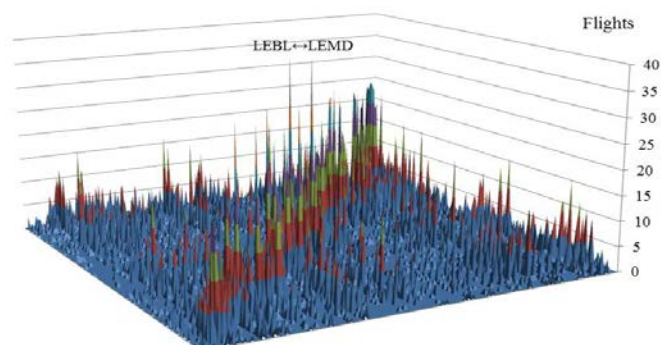
Note.

In POEM Deliverable 5.2, three basic ways of treating early departures (or arrivals), relative to scheduled / planned times were presented: (i) as they appear in the data, i.e. with negative values subtracted from aggregate delays; (ii) recoding them to zero delay; (iii) recoding them to absolute differences. Different rubrics suit different methodological requirements. Throughout this report, method (ii) is adopted, which is the most common approach in other research.

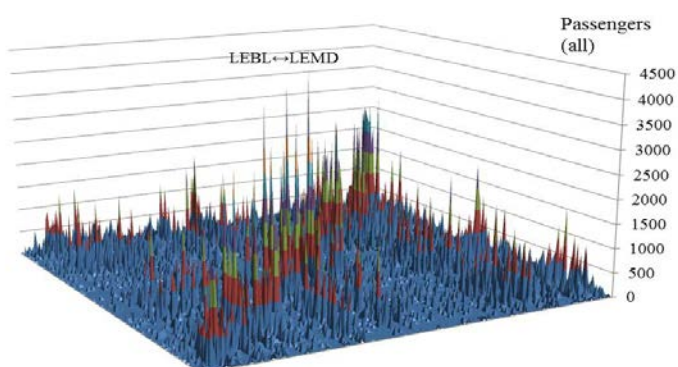
¹⁷ Using the hourly time windows discussed later (Section 3.3).



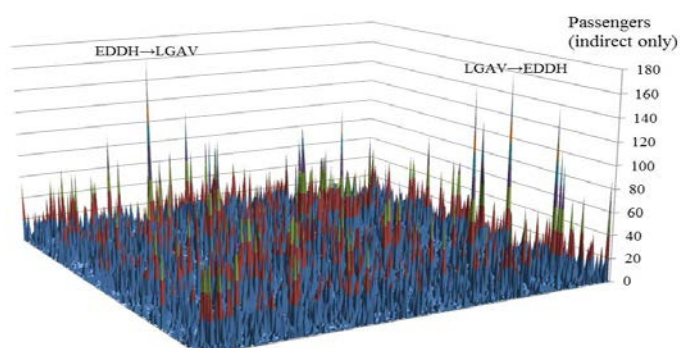
Panel 1.
Delay correlation coefficients



Panel 2.
Flights



Panel 3.
Passengers
– direct and indirect



Panel 4.
Passengers
– indirect only

Figure 17. Overview of different airport-airport density layers

3.2 Classical results

Table 18. Core metric set

Core metric*	Units	Definition	Threshold
Flight departure delay	mins / flight	Delay from the gate relative to schedule	0.2
Flight arrival delay	mins / flight	Delay at the gate relative to schedule	0.2
Departure delay of departure-delayed flights [^]	mins / flight	Delay from the gate relative to schedule	1.0
Arrival delay of arrival-delayed flights [^]	mins / flight	Delay at the gate relative to schedule	1.0
Pax departure delay [†]	mins / pax	Delay from the gate relative to schedule	0.2
Pax arrival delay [†]	mins / pax	Delay at the gate relative to schedule	0.2
Departure delay of departure-delayed pax [^]	mins / pax	Delay from the gate relative to schedule	1.0
Arrival delay of arrival-delayed pax [^]	mins / pax	Delay at the gate relative to schedule	1.0
Passenger hard cost	Euros / pax	Hard costs (see Appendix A) averaged per passenger	0.2
Passenger soft cost	Euros / pax	Soft costs (see Appendix A) averaged per passenger	0.2
Passenger value of time	Euros / pax	Pax value of time (see Appendix A) averaged per passenger	0.2
Non-passenger costs	Euros / flight	Fuel, crew and maintenance costs averaged per flight	10
Per-flight pax hard cost	Euros / flight	Passenger hard costs to airline averaged per flight	10
Per-flight pax soft cost	Euros / flight	Passenger soft costs to airline averaged per flight	10
Total flight cost [‡]	Euros / flight	Passenger plus non-passenger costs per flight	10
Total flight cost per minute of departure delay [¶]	Euros / min	Pax plus non-pax costs per minute of departure delay	2.0
Reactionary delay ratio	ratio	Reactionary delay (see Section 2.5) / flight departure delay	n/a
Arrival-delayed passenger / flight ratio	ratio	Arrival delay of: arrival-delayed pax / arrival-delayed flights	n/a

* All values are means, except where "ratio" is indicated in the units column.

[^] Calculated for delays > 15 minutes, as a differential metric from those including all delay.

[†] Two versions of these metrics are used: measured for the first leg only (for departures) and last leg only (for arrivals), and also as accumulated metrics, whereby a late passenger on a multi-leg journey contributes to the delay measured on each leg, relative to their original (scheduled) itinerary. The threshold for the accumulated metrics is 0.5 minutes.

[‡] These costs are calculated for all flights, such that even an on-time flight may have secondary passenger costs (from further along the propagation tree) assigned equally back to it. The alternative would be to set a threshold for such retrospective assignment. If this threshold were to be any delay, a 1 minute delay on a flight could attract half of a later-incurred €500 cost, whereas zero minutes of delay would attract no cost at all. If the threshold were set higher, e.g. at 15 minutes, then the cost allocations would become unevenly distributed only across flights above the threshold set. Only passenger-related costs are distributed in this way, i.e. equally, back along the original itinerary. Non-passenger costs are assigned directly to the individual flights concerned. In future, other attribution methods may be explored, but even assignments based on the relative delay of each leg are problematic due to the non-linearities of the cost functions. Notwithstanding these arguments, it is nevertheless less likely that a low-delay flight will be associated with propagated costs later on, so these attribution effects should to some extent be self-adjusting. (See also following note.)

[¶] Calculated for each flight with departure delay > 0, with corresponding statistics (such as the mean) computed across the distribution of all such flights. Retrospective cost assignments otherwise as per preceding note: these two metrics should be considered as primarily indicative and relative, their main use being to compare the performance of the scenarios in conjunction with other metrics.

At the outset of these analyses, we wish to stress the importance we have placed on presenting and exploring key results, distilling key conclusions from the model outputs, rather than presenting multiple tables of outputs in which the most pertinent findings may be lost. In particular, we wish to select one combination of simulation day and scenario of special interest from the perspective of the core classical results, to explore in more detail through the network and complexity analyses of Section 3.4. This ‘focus’ combination will be contrasted with the baseline traffic day under no applied scenarios.

Table 18 shows the core set of classical metrics that will be used to explore the outputs of the model runs for the baseline traffic day, the high delay day and the high cancellation day (described in Section 2.4.2) under the scenarios described in Section 2.3. In addition to the core metrics, a greater number of complementary metrics has been output for each model run, which will be used as additional diagnostics during the discussion, rather than being presenting in full in each table and thus cluttering the presentation. Full tabulations of any model run and scenario combination are available from the authors in electronic format for the interested reader, for non-commercial purposes.

Assigning airline costs to the passengers and deriving averages therefrom is, in theory at least, relatively straightforward since each passenger has their costs tracked throughout the model runs. However, it is important to note that costs cannot be trivially assigned to some of the metrics in the table, as discussed in the footnotes. Since most of the costs involve non-linear passenger cost components, the average values produced are a function of the distribution of the delays (which act as inherent cost weights in such calculations). Therefore, the total flight cost per minute of departure delay is not the same as the total cost of all departure delays divided by the total number of departure minutes, for example.

Unless otherwise stated, when comparing all mean values between the model runs, a z-test is applied to check the statistical significance of any differences, in each case using $p < 0.05$ to determine such significance. All comparisons are based on 50 runs of the model for each simulation day, under each scenario¹⁸. In addition to this, in the final column of the core metrics table, a threshold is presented. Where differences (Δ) are detected below this threshold, we do not usually report the difference. For example, a difference of less than 0.2 minutes between average delays per flight or per passenger would not be considered worthy of attention in the context of the modelling fidelity, whether statistically significant or not.




















These thresholds were derived judgementsally and through co-inspection of the p-values through numerous phases of test runs of the model under all simulation days and scenarios, to avoid reporting on artefactual results, although they are often close to 2% of the baseline no-scenario mean value. Thresholds are not applied to the ratio metrics. Differences (Δ) are given as the (scenario) value minus the corresponding no-scenario (baseline) value, with explicit signs for clarity.

The summary (Table 19) of the comparisons shows which outputs are compared with which corresponding no-scenario baselines. Note that for each simulation day, a dedicated no-scenario baseline was run. These three baselines need to be compared with each other (as represented in the first row of the table), in order to check the integrity of the two disrupted days (e.g. that they produce appropriately worse metrics than the baseline traffic day).

Even this high-level comparison would entail the presentation of 20 tables of output, with 360 metric comparisons. Although these comparisons have been computed, we focus on the key results, summarising more general conclusions where possible (notably with regard to the N_1 and N_2 results).

¹⁸ Except for some earlier evaluation runs used to identify the more important scenarios to pursue, as explicitly detailed below.

Table 19. Summary of scenario and simulation day comparisons

Scenario			Simulation day		
Type, and level	Designator	Summary description	Baseline traffic	High delay	High cancellation
No-scenario, 0	S ₀	No-scenario baselines			
ANSP, 1	N ₁	Prioritisation of inbound flights based on simple passenger numbers			
ANSP, 2	N ₂	Inbound flights arriving > 15 minutes late are prioritised based on the no. of onward flights delayed by inbound connecting pax			
AO, 1	A ₁	Wait times and associated departure slots are estimated on a cost minimisation basis, with longer wait times potentially forced during periods of heavy ATFM delay			
AO, 2	A ₂	Departure times and arrival sequences based on delay costs – A ₁ is implemented and flights are independently arrival-managed based on delay cost			
Policy, 1	P ₁	Pax reaccommodated based on prioritisation by final arrival delay, instead of by ticket type, but preserving interlining hierarchies			
Policy, 2	P ₂	Pax reaccommodated based on prioritisation by final arrival delay, regardless of ticket type, and also relaxing all interlining hierarchies			

Comparing the high cancellation day no-scenario model runs with the no-scenario baseline traffic day, the former demonstrated no significant change in any average aircraft departure or arrival delays. This was reflected in practice, on a wider scale for 2010 in Europe, which was a year of high cancellation rates (especially due to the Eyjafjallajökull ash cloud in April and May) without a large impact on the punctuality of flights that operated (EUROCONTROL, 2011a). However, *all* the passenger delay metrics in the model worsened significantly, especially the (simple) departure delay of departure-delayed passengers ($\Delta = +23$ minutes) and the arrival delay of arrival-delayed passengers ($\Delta = +17$ minutes). As expected, a number of costs also increased significantly, particularly the average total flight cost, which increased by EUR 39 (all $p = 0.00$). This again underlines the importance of using passenger-centric metrics in measuring true system performance, since such changes are not expressed through any of the flight metrics.

Comparing the high delay day no-scenario model runs with the no-scenario baseline traffic day, all flight and passenger metrics worsened significantly, this time thus including the average flight departure and arrival delays ($\Delta = +1.3$ minutes, $\Delta = +1.2$ minutes; both $p = 0.00$). Whilst all passenger delay metrics also worsened significantly (all $p = 0.00$), none of the changes were as large as those during the high cancellation day, all being in the range $\Delta = +1$ to $+3$ minutes. Again, a number of costs also increased significantly, particularly the average total flight cost, which increased by EUR 33 per flight ($p = 0.00$) – close to the value for the high cancellation day.

The average arrival-delayed passenger / arrival-delayed flight ratio was 1.5 (in alignment with the literature values reported in Section 2.5 on model calibration) for the no-scenario runs for both the baseline day and the high delay day, but increased to 1.9 for the high cancellation day – as would be expected from the preceding discussion. These core metric results for the no-scenario runs for the high cancellation day and the high delay day are presented fully later, with their corresponding active

scenario results. These results have demonstrated that the disruption days have been well-modelled in so far as the changes in the core metrics are all as expected.

Considering the results from the N_1 and N_2 scenarios applied to the baseline traffic day, N_1 produced no differences relative to the no-scenario baseline and N_2 only produced an *increase* in the average delay of arrival-delayed passengers (i.e. as measured for the final leg; $\Delta = +2.2$ minutes, $p = 0.00$).

These ANSP scenario rules, by design, explored the effectiveness of information and control within the scope of the ANSP: numbers of inbound passengers and arrival management (see Rule 26 in Appendix A for details). The lack of any difference in the core metrics for N_1 demonstrates that a prioritisation of inbound flights, based only on simple passenger numbers (i.e. minimising the product of passengers and delay minutes in the delay sequence) is ineffective in improving performance.

Under N_2 , flights arriving more than 15 minutes late were prioritised based on the number of onward flights that would be delayed by inbound connecting passengers (with N_1 resolving any ties). Although N_2 produced an increase in the average delay of arrival-delayed passengers of 2.2 minutes, this was associated with a relatively small *decrease* (of 2%) in the *number* of arrival-delayed passengers. The change in the number of missed connections was very small, albeit an increase of 0.6% (representing only around 130 passengers). There were no significant changes in any flight or cost metrics.

These changes are thus probably explained by some flights having modest arrival delay reductions, and a relatively smaller number suffering slightly worse delay (perhaps larger aircraft and higher load-factor LCCs with passenger forward connections onto relatively fewer/no flights – thus not being prioritised by the scenario rules). The net result is evidently only a small change in the distribution of the number of arrival-delayed passengers (as shown in Figure 18). Compared to the baseline situation (with no-scenario), under N_2 the skewness becomes less negative (from -4.9 to -3.4) and the kurtosis doubles (from 0.45 to 0.91), suggesting a heavier upper tail.

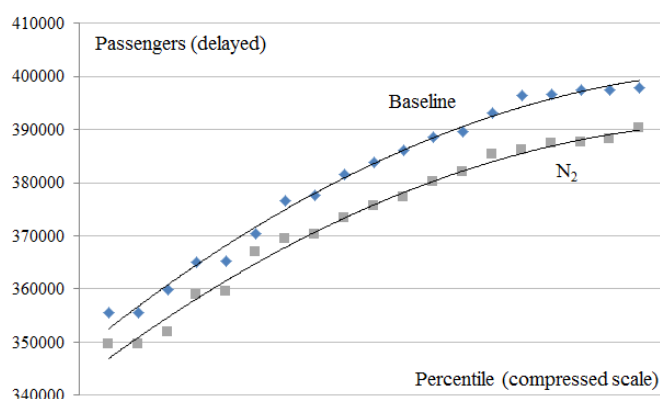


Figure 18. Distribution of delayed passengers under N_2 scenario

N_2 also produced statistically significant ($p = 0.00$) increases in the average delay of arrival-delayed passengers under the high delay and high cancellation day simulations (both between 2 and 3 minutes, based on earlier evaluation runs of the model¹⁹). Here, both the N_1 and N_2 scenarios also produced some above-threshold (and significant; $p = 0.00$) but often very modest worsening of passenger departure metrics, but with no changes in any flight or cost metrics. All in all, however, these effects are small and negative in terms of performance change and were not investigated in any further detail.

In conclusion, prioritisation scenarios instigated under arrival management, based on the *numbers* either simply of inbound passengers or on those with connecting onward flights, were ineffective in improving performance. Such performance was even slightly worse under high delay or increased cancellation rates but was only discernible through the use of passenger-centric metrics. No significant changes in any flight punctuality or cost metrics were observed.

¹⁹ Each running the test day 20 times. Differences in the standard deviations of the average delay of arrival-delayed passengers were less than 0.1% between these runs and the full runs of the baseline traffic day.

Table 20 presents the results of the core metrics under the remaining scenarios for the baseline traffic day. Mean values are shown, preceded by the corresponding standard deviations in parentheses (this reversal of convention facilitates visual inspection of the means). Significant improvements in metrics, whereby the scenario mean is lower than the no-scenario value, are highlighted in a white font and shaded background. Significant deteriorations are shown underlined.

Table 20. Core metric results – baseline traffic day

Core metric	Units	S ₀	P ₁	P ₂	A ₁	A ₂
Flight departure delay	mins / flight	(31) 13.8	(31) 13.8	(31) 13.8	(31) 14.3	(39) <u>15.8</u>
Flight arrival delay	mins / flight	(28) 13.5	(29) 13.5	(28) 13.5	(29) 13.9	(37) <u>15.7</u>
Departure delay of departure-delayed flights	mins / flight	(45) 52.3	(45) 52.6	(45) 52.5	(45) 52.7	(62) <u>61.4</u>
Arrival delay of arrival-delayed flights	mins / flight	(45) 50.3	(45) 50.6	(45) 50.6	(45) 50.5	(62) <u>60.0</u>
Pax departure delay (accumulated)	mins / pax	(33) 12.8 (200) 24.2	(31) 12.8 (142) <u>22.1</u>	(30) 12.7 (108) <u>20.6</u>	(34) <u>13.2</u> (193) 23.7	(41) <u>14.0</u> (172) 24.2
Pax arrival delay (accumulated)	mins / pax	(84) 16.3 (224) 23.9	(84) 16.4 (157) <u>21.5</u>	(81) <u>15.9</u> (120) <u>19.9</u>	(70) <u>14.6</u> (217) <u>23.1</u>	(81) <u>16.7</u> (198) 23.9
Departure delay of departure-delayed pax	mins / pax	(54) 53.0	(49) 52.7	(47) 52.4	(56) 53.5	(74) <u>61.2</u>
Arrival delay of arrival-delayed pax	mins / pax	(187) 77.8	(188) 78.4	(180) <u>75.6</u>	(153) <u>68.0</u>	(179) <u>82.4</u>
Passenger hard cost	Euros / pax	(23) 1.3	(23) 1.3	(26) <u>1.6</u>	(19) <u>0.9</u>	(22) 1.3
Passenger soft cost	Euros / pax	(3) 0.8	(3) 0.9	(3) 0.8	(3) 0.8	(4) 1.0
Passenger value of time	Euros / pax	(46) 11.0	(46) 11.0	(44) <u>10.7</u>	(39) <u>10.2</u>	(45) <u>11.7</u>
Non-passenger costs	Euros / flight	(428) 327	(429) 328	(428) 328	(430) 331	(518) <u>346.7</u>
Per-flight pax hard cost	Euros / flight	(723) 133	(720) 134	(817) <u>159</u>	(545) <u>94</u>	(730) 131.6
Per-flight pax soft cost	Euros / flight	(298) 84.4	(301) 85.1	(300) 83.9	(299) 82.0	(374) <u>103.8</u>
Total flight cost	Euros / flight	(1115) 545	(1117) 546	(1191) <u>571</u>	(977) <u>506</u>	(1311) <u>582.1</u>
Total flight cost per minute of departure delay	Euros / min	(424) 93.6	(417) 94.0	(452) 99.8	(348) <u>85.9</u>	(441) 99.1
Reactionary delay ratio	ratio	49%	49%	49%	51%	58%
Arrival-delayed passenger / flight ratio	ratio	1.5	1.5	1.5	1.3	1.4

To simplify the table further, non-significant changes are removed; significant improvements are represented as '✓' and significant deteriorations as '✗'. Changes of more than 3 minutes, of more than 20 Euros or more than 3% in reactionary delay ratios are shown as doubled symbols. These are presented in Table 21 and although the thresholds are arbitrary, they allow us to focus on the larger effects, comparatively across the scenarios. It is to be stressed that the smaller changes could still represent substantive changes over the network, which we will explore in future work, restricting the focus here to the larger comparisons.

Table 21. Core metric results – baseline traffic day (main effects)

Core metric	Units	P ₁	P ₂	A ₁	A ₂
Flight departure delay	mins / flight				✗
Flight arrival delay	mins / flight				✗
Departure delay of departure-delayed flights	mins / flight				✗✗
Arrival delay of arrival-delayed flights	mins / flight				✗✗
Pax departure delay (accumulated)	mins / pax	(✓)	(✓✓)	✗	✗
Pax arrival delay (accumulated)	mins / pax	(✓)	(✓✓)	(✓)	✗
Departure delay of departure-delayed pax	mins / pax				✗✗
Arrival delay of arrival-delayed pax	mins / pax		✓	✓✓	✗✗
Passenger hard cost	Euros / pax		✗	✓	
Passenger soft cost	Euros / pax				
Passenger value of time	Euros / pax		✓	✓	✗
Non-passenger costs	Euros / flight				✗
Per-flight pax hard cost	Euros / flight		✗✗	✓✓	
Per-flight pax soft cost	Euros / flight				✗
Total flight cost	Euros / flight		✗✗	✓✓	✗✗
Total flight cost per minute of departure delay	Euros / min			✓	
Reactionary delay ratio	ratio				✗✗

Key results apparent from the table are:

- P_1 demonstrates some weak improvements;
- P_2 demonstrates a more substantial improvement in passenger arrival delay, at the expense of incurring some extra costs to the airline;
- A_1 markedly improves a number of passenger delay metrics and airline costs;
- A_2 fails to improve the metrics;
- the improvements in passenger metrics and costs are not associated with concomitant changes to the flight delay metrics.

Under P_1 , passengers are reaccommodated based on prioritisation in order of delay saving at the final destination, instead of by ticket type, although ticket type still determines which carriers passengers are rebooked onto and interlining hierarchies are preserved (see Rule 33, Appendix A). If reaccommodation 'A' would reduce a final delay from three hours to one hour, it is favoured over reaccommodation 'B', reducing a final delay from four hours to two hours, since 'A' has the smaller residual. This is not a global optimisation, since the priorities are determined locally and sequentially: better solutions arising by jointly considering other airports, or arising later, are not taken into account. The only improvement detected through the core metrics is through the accumulated passenger delay metrics, whereby (as explained above) a late passenger on a multi-leg journey contributes to the delay measured on each leg, relative to their original (scheduled) itinerary. Indeed, the sensitivity of these metrics is echoed under P_2 , where the effects are larger still.

Under P_2 , passengers are similarly reaccommodated by arrival delay, but now regardless of ticket type and also relaxing all interlining hierarchies. Passengers are reaccommodated regardless of ticket type onto the next available flight operated by any carrier. This has rather more marked effects for improvements in passenger arrival delay (e.g. average passenger arrival delay $\Delta = -0.4$ minutes, $p = 0.00$; and for concomitant passenger value of time savings, as would be expected). However, the total cost per flight has increased ($\Delta = +25.7$ Euros, $p = 0.02$) – almost entirely driven by increased passenger costs to the airline. In turn, these will be dominated by reaccommodation costs, with P_2 demonstrating a 6% increase ($p = 0.03$) in airline rebookings and a 4% decrease ($p = 0.00$) in passenger time spent at airports or hotels, relative to the baseline traffic day no-scenario simulation (latter results not tabulated elsewhere – from supplementary metrics).

To address the poor performance of scenario A_2 , we first remind ourselves that under A_1 (as described under Rule 13, waiting for boarding: see Appendix A):

- when ATFM delays are 60 minutes or greater, the flight will wait another 60 minutes if the net cost of doing so is estimated as less than that associated with not waiting;
- when ATFM delays are less than 60 minutes, cost estimates for waiting by increments of 15 minutes are estimated, with the minimum cost alternative chosen.

Additionally, under A_2 , A_1 is implemented as described above and flights are also *independently* arrival-managed based on delay cost (as described under Rule 26: see Appendix A).

Table 22 shows that scenario A_2 has caused increased dispersion on all of the key metrics shown, relative to A_1 . Most of the standard deviations are also higher than those under the no-scenario simulation for the high delay day (right-hand column). Further comparing the results of A_2 with the no-scenario run for the high delay day, *none* of the costs are significantly different, nor is the (simple) average passenger departure delay or the average flight departure delay ($p > 0.05$). However, the departure delay of departure-delayed flights ($\Delta = +6.4$ minutes, $p = 0.00$) and the departure delay of departure-delayed passengers ($\Delta = +5.6$ minutes, $p = 0.00$) are both significantly worse under A_2 . Similarly, the arrival delay of arrival-delayed flights ($\Delta = +7.3$ minutes, $p = 0.00$) and the arrival delay of arrival-delayed passengers ($\Delta = +2.3$ minutes, $p = 0.00$) are worse under A_2 . Noting that the arrival delay for arrival-delayed passengers is only relatively modestly worse under A_2 , it is further observed that the average arrival delay of all passengers is actually significantly better under A_2 , albeit by only a small margin ($\Delta = -0.9$ minutes, $p = 0.00$). The corresponding flight metric is comparably worse ($\Delta = +1.0$ minutes, $p = 0.00$). A_2 performs similarly poorly on the high cancellation day.

Table 22. Standard deviation of core metrics – dispersion under scenario A₂

Core metric	Units	Baseline traffic day		High delay day
		A ₁	A ₂	S ₀
Flight departure delay	mins / flight	31	39	33
Flight arrival delay	mins / flight	29	37	31
Departure delay of departure-delayed flights	mins / flight	45	62	48
Arrival delay of arrival-delayed flights	mins / flight	45	62	47
Pax departure delay	mins / pax	34	41	35
Pax arrival delay	mins / pax	70	81	87
Departure delay of departure-delayed pax	mins / pax	56	74	55
Arrival delay of arrival-delayed pax	mins / pax	153	179	187
Total flight cost	Euros / flight	977	1311	1180
Reactionary delay ratio	ratio	51%	58%	48%

Values shown are standard deviations, except for reactionary delay ratio percentages.

With regard to the A₂ scenario, we may conclude that this has caused significant disruption to performance. There is increased dispersion of many metrics, relative even to the high delay day, with a resulting loss in performance on all core metrics relative to baseline no-scenario performance. Hence, whilst A₁ has proven to be remarkably successful, the addition of *independent*, cost-based arrival management has apparently foiled these benefits due to lack of coordination between departures and arrivals – an important finding in itself, worthy of further investigation beyond this deliverable. Part of this failure is reflected in the high reactionary delay ratio under A₂, further indicating a lack of optimisation between inbound and outbound cost functions and wait rules. It is noteworthy that the loss of performance through A₂ compared with the high delay day was not expressed through worsening cost metrics, reflecting some degree of successful outcome of the cost-oriented algorithms of A₂. This was also evidenced in a minor positive outcome for arrival delay averaged across all passengers – again emphasising the need to consider the full range of performance metrics in any such outcome assessment.

Scenario A₁ has markedly improved a number of passenger delay metrics and airline costs, one of the few negative effects being a small increase in passenger departure delay ($\Delta = +0.4$ minutes, $p = 0.00$). Passenger arrival delay, both as a simple average and for arrival-delayed passengers, is significantly improved, the latter improvement being sizeable (respectively: $\Delta = -1.6$ minutes, $p = 0.00$; $\Delta = -9.8$ minutes, $p = 0.00$). Thus, whilst there is an *overall* small increase in average passenger departure delay, the *distribution* of which passengers are more departure delayed, and which are not, is such that the net effect on arrival delay and airline costs is beneficial. Small but significant improvements in passenger value of time and hard costs are made per passenger (see Table 20), with substantial cost reductions in the average total flight cost ($\Delta = -38.6$ Euros, $p = 0.00$) and total flight cost per minute of departure delay ($\Delta = -7.8$ Euros, $p = 0.04$). Both of these effects are driven by reductions in hard costs, with a 6% reduction in airline rebookings (although $p = 0.07$) and an 11% decrease in passenger time spent at airports or hotels ($p = 0.00$) – again relative to the baseline traffic day no-scenario simulation (from supplementary metrics). Finally, we note that A₁ is accompanied by an increase of 2 percentage points in overall reactionary delay, from 49% to 51% (per flight: $\Delta = +0.6$ minutes, $p = 0.00$). These discussions will be taken forward in Section 3.4.

At the start of this section we explained that we intended to select one combination of simulation day and scenario of special interest from the perspective of the core classical results, to explore in more detail through the network and complexity analyses of Section 3.4. Following the foregoing discussion on the successful outcomes of scenario A₁, this selected combination will indeed be A₁ applied to the baseline traffic day, compared with the no-scenario baseline traffic day. This allows us to make a good like-for-like comparison, in that the two selected days both relate to the common source of the baseline traffic day, rather than comparing combinations of scenarios and days across different types of simulation day – such as the high cancellation or high delay day. For these two sets of simulation days, respectively, Table 23 and Table 24 present the core metric results under the successful scenarios²⁰.

Table 23. Core metric results – high cancellation day

Core metric	Units	Baseline traffic day	High cancellation day			
		S ₀	S ₀	P ₁	P ₂	A ₁
Flight departure delay	mins / flight	(31) 13.8	(30) 13.4	(30) 13.6	(30) 13.6	(31) 14.0
Flight arrival delay	mins / flight	(28) 13.5	(28) 13.2	(28) 13.4	(28) 13.3	(28) 13.6
Departure delay of departure-delayed flights	mins / flight	(45) 52.3	(45) 51.9	(44) 52.0	(44) 52.0	(44) 52.4
Arrival delay of arrival-delayed flights	mins / flight	(45) 50.3	(44) 50.0	(44) 49.9	(44) 49.9	(43) 50.1
Pax departure delay (accumulated)	mins / pax	(33) 12.8 (200) 24.2	(300) 17.6 (641) 37.9	(281) 17.5 (604) 37.1	(58) 14.5 (124) 23.1	(295) 17.9 (520) 33.4
Pax arrival delay (accumulated)	mins / pax	(84) 16.3 (224) 23.9	(101) 19.5 (784) 38.6	(101) 19.6 (735) 37.5	(94) 18.5 (136) 20.5	(91) 18.0 (660) 33.5
Departure delay of departure-delayed pax	mins / pax	(54) 53.0	(655) 76.2	(608) 74.8	(115) 60.3	(634) 75.6
Arrival delay of arrival-delayed pax	mins / pax	(187) 77.8	(224) 95.1	(223) 94.8	(206) 89.1	(199) 85.8
Passenger hard cost	Euros / pax	(23) 1.3	(24) 1.6	(24) 1.6	(28) 2.0	(20) 1.2
Passenger soft cost	Euros / pax	(3) 0.8	(4) 1.0	(4) 1.0	(4) 1.0	(4) 1.0
Passenger value of time	Euros / pax	(46) 11.0	(56) 12.8	(56) 12.9	(52) 12.3	(51) 12.2
Non-passenger costs	Euros / flight	(428) 327	(424) 322	(424) 324	(423) 323	(424) 326.0
Per-flight pax hard cost	Euros / flight	(723) 133	(834) 162	(845) 163	(960) 197	(682) 122.9
Per-flight pax soft cost	Euros / flight	(298) 84.4	(399) 100.0	(397) 100.5	(391) 98.1	(397) 97.8
Total flight cost	Euros / flight	(1115) 545	(1264) 584	(1271) 587	(1362) 618	(1136) 546.7
Total flight cost per minute of departure delay	Euros / min	(424) 93.6	(417) 94.6	(424) 94.9	(447) 100.9	(346) 87.1
Reactionary delay ratio	ratio	49%	49%	49%	49%	52%
Arrival-delayed passenger / flight ratio	ratio	1.5	1.9	1.9	1.8	1.7

²⁰ The no-scenario (S₀) results for the high cancellations and high delays are statistically compared with S₀ for the baseline traffic day, as explained above.

Table 24. Core metric results – high delay day

Core metric	Units	Baseline traffic day	High delay day			
		S ₀	S ₀	P ₁	P ₂	A ₁
Flight departure delay	mins / flight	(31) 13.8	(33) <u>15.1</u>	(33) 15.2	(33) 15.0	(34) 15.7
Flight arrival delay	mins / flight	(28) 13.5	(31) <u>14.7</u>	(31) 14.8	(31) 14.6	(31) 15.2
Departure delay of departure-delayed flights	mins / flight	(45) 52.3	(48) <u>54.9</u>	(48) 55.1	(48) 54.7	(48) 55.6
Arrival delay of arrival- delayed flights	mins / flight	(45) 50.3	(47) <u>52.6</u>	(47) 52.9	(47) 52.5	(47) 53.0
Pax departure delay (accumulated)	mins / pax	(33) 12.8 (200) 24.2	(35) <u>14.0</u> (185) <u>25.4</u>	(33) 14.0 (134) <u>23.6</u>	(32) <u>13.8</u> (111) <u>22.3</u>	(37) <u>14.6</u> (201) <u>26.1</u>
Pax arrival delay (accumulated)	mins / pax	(84) 16.3 (224) 23.9	(87) <u>17.6</u> (207) <u>24.8</u>	(87) 17.7 (148) <u>22.8</u>	(83) <u>17.0</u> (123) <u>21.4</u>	(74) <u>16.1</u> (227) <u>25.4</u>
Departure delay of departure-delayed pax	mins / pax	(54) 53.0	(55) <u>55.6</u>	(51) 55.2	(49) 54.7	(59) 56.4
Arrival delay of arrival- delayed pax	mins / pax	(187) 77.8	(187) <u>80.1</u>	(188) 80.5	(179) <u>77.5</u>	(156) <u>71.0</u>
Passenger hard cost	Euros / pax	(23) 1.3	(23) 1.4	(24) 1.4	(27) <u>1.7</u>	(19) <u>1.0</u>
Passenger soft cost	Euros / pax	(3) 0.8	(4) 1.0	(4) 1.0	(4) 0.9	(4) 0.9
Passenger value of time	Euros / pax	(46) 11.0	(47) <u>11.9</u>	(48) 11.9	(46) <u>11.5</u>	(41) <u>11.2</u>
Non-passenger costs	Euros / flight	(428) 327	(451) <u>342</u>	(451) 343	(449) 341	(453) 347
Per-flight pax hard cost	Euros / flight	(723) 133	(743) 141	(762) 143	(847) <u>169</u>	(563) <u>101</u>
Per-flight pax soft cost	Euros / flight	(298) 84.4	(327) <u>95.4</u>	(329) 96.4	(327) 93.8	(331) 94.2
Total flight cost	Euros / flight	(1115) 545	(1180) <u>578</u>	(1195) 583	(1257) <u>604</u>	(1044) <u>542</u>
Total flight cost per minute of departure delay	Euros / min	(424) 93.6	(422) 94.3	(444) 95.2	(463) 101.4	(357) 87.7
Reactionary delay ratio	ratio	49%	48%	48%	48%	51%
Arrival-delayed passenger / flight ratio	ratio	1.5	1.5	1.5	1.5	1.5

These scenario results are fairly similar for the high cancellation and high delay days, showing a degree of robustness in terms of their efficacy under increased disruption. We note again that the improvements in passenger metrics and costs are not associated with concomitant changes in the flight delay metrics.

A summary of the key changes discussed in this section is presented in Table 25. The trade-off under scenario A₁ between improved passenger and cost metrics, and increased reactionary delay, will be explored in detail in Section 3.4.

Table 25. Core metric results – overall summary

Core metric	Units	N ₁ & N ₂	P ₁	P ₂	A ₁
		Inbound prioritisation based on: simple pax numbers, or on onward flights delayed	Passenger reaccommodated based on delay at final destination preserving interlining hierarchies	... relaxing interlining hierarchies	Departures times based on cost minimisation (& consideration of ATFM delay)
Flight departure delay	mins / flight	no significant changes in current flight-centric metrics: stresses need for passenger-centric metrics			
Flight arrival delay	mins / flight				
Departure delay of departure-delayed flights	mins / flight				
Arrival delay of arrival-delayed flights	mins / flight				
Pax departure delay	mins / pax	no significant changes under simple inbound scenarios driven by passenger numbers, or by numbers of delayed onward flights	revised passenger re-booking rules produce only weak improvements whilst current airline interlining rules are preserved, c.f. →	=	+0.4
Pax arrival delay	mins / pax			-0.4	-1.6
Departure delay of departure-delayed pax	mins / pax			=	=
Arrival delay of arrival-delayed pax	mins / pax			-2.2	-9.8
Passenger value of time	Euros / pax			-0.2	-0.7
Non-passenger costs	Euros / flight			=	=
Per-flight pax hard cost	Euros / flight			+26	-40
Per-flight pax soft cost	Euros / flight			=	=
Total flight cost	Euros / flight			+26	-39
Total flight cost per minute of departure delay	Euros / min			=	-7.8
Reactionary delay ratio	ratio			49%	51%

Values to 2 s.f.; all changes indicated are significant ($p < 0.05$); “-” implies metric improvement, “+” implies metric deterioration.

3.3 Factor analysis (derived metrics)

3.3.1 Review and update of the factor analytical approach

The objective here is to use factor analysis to describe delay propagation by reducing the 199 airports modelled to a much smaller number of groups ('components') of airports sharing common characteristics. In particular, we wish to identify delay propagation nodes, i.e. airports with fairly high loadings across these components.

In order to be able to observe patterns in complex data, it is necessary to carefully select the best options available from the range of factor analytic methods. A detailed description of how these options are selected is presented in Appendix C, with a briefer overview given in this section.

We introduced factor analysis in POEM Deliverable 5.1, where we explained that this technique attempts to express a set of observed, independent variables, as a new set of independent variables – these 'factors' are always linear combinations of the original variable set²¹. The technique was originally developed in psychology to simplify the description of behavioural traits, for example. It shares its underlying principles with multivariate, linear regression.

One of the key differences is that factor analysis (usually) deals with the issue of (multi-)collinearity associated with the independent variables. This is true especially of the technique known as principal components, usually considered a variant of factor analysis. There cannot be more factors (also known as components) in the 'solution' than there were variables in the original set and it is obviously preferable that there will be rather fewer of them.

A key indication of the quality of the solution is the percentage of the original variance between the original variables²², which is described by the (fewer) factors. It is not acceptable to obtain a purely 'mathematical' solution in the analysis, i.e. whereby the analyst is not able to assign real meaning to the factors. The analyst often 'rotates' the factors, to increase loadings on some of the original variables, and decrease them on others, in order to improve the interpretability and simplicity of the solution. The 'initial' solution produces orthogonal factors, with sequential factors each accounting for as much of the residual variance as possible. These two restrictions are, however, arbitrary and one or both may be relaxed during the production of rotated solutions. Rotated solutions do not improve the fit between the factors and the source data – the total amount of variance in the data explained is just the same as for the unrotated solution, although the amount of variance explained by each factor may well be different.

Principal components analysis (PCA) is a widely used type of factor analysis method (Tull and Hawkins, 1993). Principal components are mathematical functions of the *observed* variables whereby we need not assume the existence of underlying, hypothetical factors. The objective is to explain as much *variance* as possible in the data. In contrast, common factors (in 'general' factor analysis) are not expressible as combinations of the observed variables and this method focuses on accounting for correlation between the variables, i.e. the covariance (Kim and Mueller, 1978b). Here, without the (underlying) common factor, there is no correlation between the observed variables.

Our analysis of the (modelled) delay patterns at the airports is not premised on underlying, hypothetical factors, but is rather based on the relationships between the observed delays. We therefore favour the PCA approach, although both this and general factor analysis will often lead to similar insights into the underlying data. With either general (unweighted least squares) factor analysis or PCA, if all the variables are perfectly correlated, then only one factor (component) will be extracted, and it will explain 100% of the variance. We show in Appendix C, through the exploration of structured examples, how such general factor analysis may sometimes not yield suitable results, however.

From Appendix C, the following methodological observations and conclusions are drawn:

²¹ For good introductions to the methodology see Stopher and Meyburg (1979) and Aaker *et al.* (2001). Kim and Mueller (1978a) provides further very helpful practical guidelines.

²² See Appendix C. Not suitable for oblique rotations, however.

- extraction of sufficient factors²³ helps to ensure that all of the variance is explained;
- rotated solutions give more intuitive factor loadings;
- oblique rotation is favoured so that component orthogonality is not a methodological artefact²⁴;
- promax rotation gives intuitive factor loading signs.

The favoured method is thus principal components analysis with promax rotation, whereby:

- the covariance matrix will be used since we wish to compare results across different samples;
- a default value of $\kappa = 4$ will be set for the promax;
- factors with more than 0.5 times the mean eigenvalue are retained as the default setting.

In Appendix C we also considered a 24-hour period covering four airports, A through D, with hours during which heavy delays occur shaded as grey blocks (see Figure 19). (In fact, the corresponding main POEM analyses run in a window from 0400 to 0359 local time, as discussed previously).

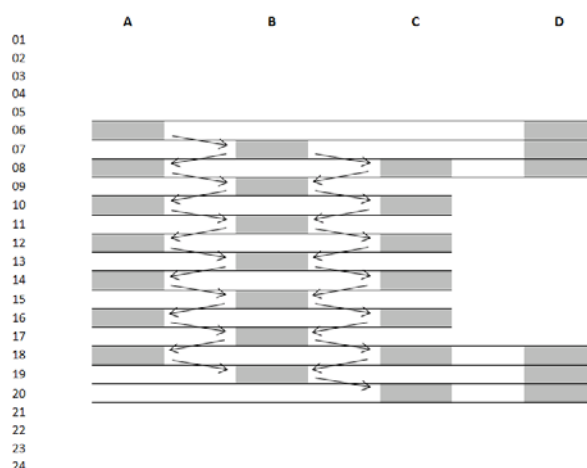


Figure 19. Four-airport example of delay patterns

It was noted that, due to the method of component extraction selected (i.e. promax rotation, as further defined above), such an extraction could distinguish patterns such as D from the other three airports well, with clean components produced for: (i) A with C; (ii) B and (iii) D.

However, if airport D showed high levels of delay throughout the day, it loaded highly on the other components, with potentially less interpretability with respect to D's *relative* loadings across such components. It was also noted that higher delays at some airports may produce solutions with fewer factors, and thus possibly less variance explained.

²³ This is rather less of an issue with more complex datasets, where more factors are typically extracted.

²⁴ After component extraction, we do not expect one group of airports (factors/components) to have delay characteristics independent of other groups, i.e. we want to allow residual correlations between the factors.

3.3.2 Applying factor analysis to the model data

A matrix of delays ($M_{a,t}^d$) has been built such that $d_{a,t}$ is the sum of all departure delays for each airport, a , and for time windows, t , of period one hour each:

$$M_{a,t}^d = \begin{pmatrix} d_{1,1} & d_{2,1} & \dots \\ d_{1,2} & d_{2,2} & \dots \\ \dots & \dots & \dots \end{pmatrix}$$

[3]

These data relate to the model input data (as opposed to the model output data from the simulations, which we will consider later). The columns thus comprise airports, with time periods in each row. If a flight was scheduled to depart at 1015 (i.e. in time window 1000 – 1059) but actually departed at 1115, the 60 minutes of delay are allocated to the scheduled departure window, rather than the window in which the flight actually departed. Delay is included from the first minute of delay and early departures are recoded to zero delay. In these matrices, all times are in UTC to prevent the appearance of delay propagating from one airport to another, e.g. apparently one hour later, when in fact this was only due to a local time difference. All flights with demoted metrics are excluded from these analyses although in the input delay matrix for the baseline traffic day delays for non-ECAC destinations are included.

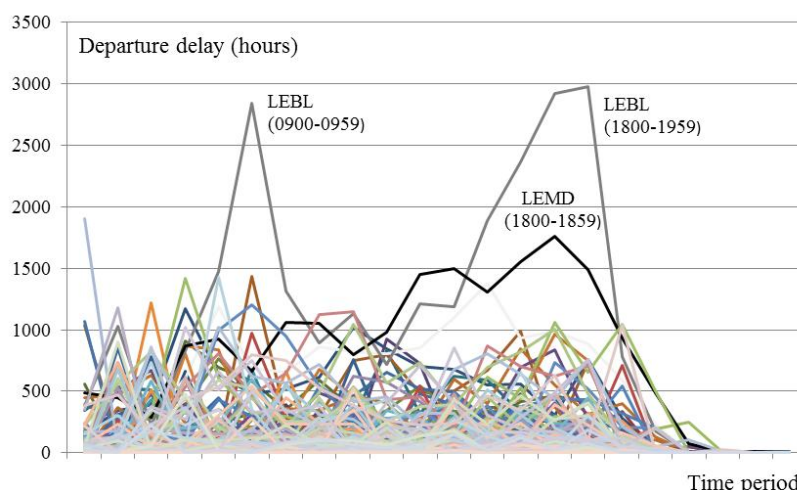


Figure 20. Total departure delay by time period and airport (baseline traffic day)

Figure 20 shows the delays by time period for each airport. Whereas pronounced peaks may have been expected for the morning and evening, the distributions are only markedly peaked for LEBL (Barcelona) and LEMD (Madrid Barajas). Removing these two airports and rescaling does not cause new pronounced peaks to become apparent. (LIML (Milan Linate) is the airport on the left with high delays in the period 0400-0459.)

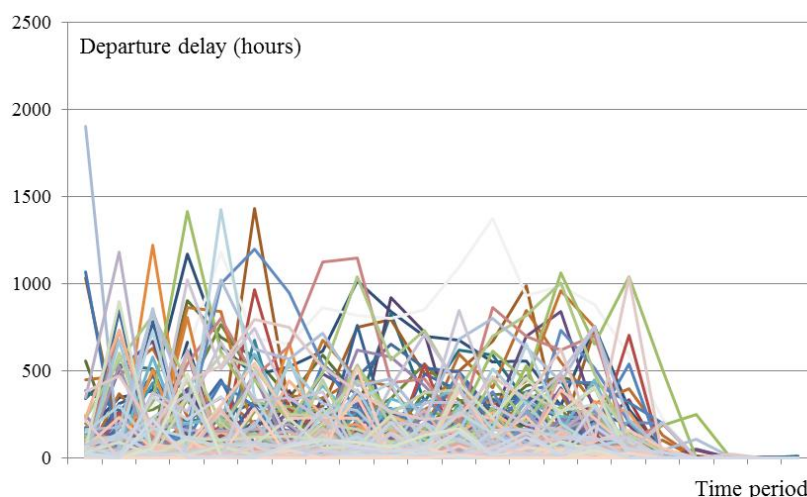


Figure 21. Total departure delay by time period, without LEBL and LEMD (baseline traffic day)

Table 26 shows the worst ten airports in terms of average departure delay per time period. It is noteworthy that over all 199 airports there was a good correlation ($\sigma = 0.8\mu + k$; $r^2 = 0.87$) between the standard deviation (SD) of the delays and the mean delays, whereby the standard deviation increases slower than the mean, as the mean increases.

Table 26. Average departure delay per time period by worst ten airports (baseline traffic day)

	Airport	Mean	SD
LEBL	Barcelona	1107	967
LEMD	Madrid Barajas	824	549
EGLL	London Heathrow	595	434
EDDF	Frankfurt	457	339
LFPG	Paris CDG	449	371
LFPO	Paris Orly	420	337
LIRF	Rome Fiumicino	420	311
EDDM	Munich	417	407
LEPA	Palma de Mallorca	397	388
LTBA	Istanbul Atatürk	355	287

Barcelona, Madrid and Palma de Mallorca had the greatest share of high departure delays (> 60 minutes) on the sample day²⁵, which has driven their mean values up. The former two also have high arrival delays²⁶. With regard to 2010 delay data for thirty 'core' selected European airports from CODA (also see POEM Deliverable 5.2), these high delay flights at Barcelona and Madrid represented only 0.6% of *all* flights on the sample day but contributed 4.7% of the total arrival and departure delay. Extending this analysis, 2.0% of all arrivals for these thirty airports had an arrival delay of more than 60 minutes but these flights contributed 14.7% of all the corresponding arrival delay minutes. Likewise, 1.2% of the high delay departures contributed 10.0% of the total departure delay.

²⁵ As we observed in POEM Deliverable 5.2.

²⁶ For Madrid, this involved 58 flights, with 19 departures contributing 1 543 minutes of departure delay between them and 39 arrivals with a total of 3 006 minutes of arrival delay. 46 of these movements were single rotations, i.e. not aircraft arriving with over 60 minutes of delay and departing with over 60 minutes of delay. One aircraft left Madrid with 75 minutes of delay in the morning, and contributed another four movements during the day, each with over 60 minutes of delay: most of this was recorded as ATFM delay (due to the ATC industrial action in France and en-route weather). For Barcelona, these delays involved 82 flights: 18 departures (with 1 354 minutes of delay) and 64 arrivals (with 5 488 minutes of delay). 68 of these flights were single rotations, 7 were double rotations (two movements).

The most common regulation reason for the high delay instances in Barcelona (arrivals and departures) was weather (affecting 47 flights); nine such flights (only) were affected by the French strike²⁷. For Madrid (arrivals and departures) the most common regulation reason was aerodrome capacity (24 flights), with fourteen flights affected by the strike. For the fourteen high departure delay flights from Palma de Mallorca, eleven had a regulation due to weather, and none were affected by the strike. As demonstrated later, the strike action is in any case not a factor for the simulations since such delays are reset in the model.

Table 27. Baseline delay components (highest component loadings; from model input data)

1	2	3	4	5	6	7	8	9	10
BKPR	BIKF	EFOU	EFHK	EDDC	EGAA	EDDB	EDDM	EDDC	EGLC
EDDG	EGJB	ENCN	EGJJ	EGHH	EINN	EGAC	EGNX	EHRD	ESSA
EDLW	EGNM	ESGG	EGTE	ENAL	ENZV	LBBG	GCTS	EKBI	GCLP
ESMS	EGPD	EYVI	ENTC	ENBR	ESNU	LECO	GCXO	LFBZ	LDZA
LATI	ENBO	GCLA	EPKT	LEBB	ESSB	LEMH	LFKB	LFQQ	LFBO
LDSP	EPKK	LCPH	ESGP	LEST	LEIB	LEVC	LFMN	LGRP	LFRS
LEXJ	LGKR	LEAM	LGTS	LFBZ	LELC	LEVX	LKPR	LIPZ	LMML
LGAV	LIBD	LGKO	LGZA	LFLL	LFOB	LICJ	LOWS	LIRQ	LRBS
LRCL	LROP	LIBR	LIEE	LFRB	LGSR	LPPR	LTAF	LJLJ	LTCG
LZIB	LYBE	UGTB	LIPQ	LIPQ	LTFE	LTAC	LTBS	LPMA	UKBB

Running a factor analysis on ($M_{a,t}^d$) using the settings summarised above produced eighteen components, with 89% of the total variance explained (in the unrotated solution) by the first ten components. Table 27 shows the ten highest loading airports for each of these components, sorted alphabetically. Airports appearing in more than one component are shown in bold. The relative uniqueness of the components is noteworthy, in that there is very little overlap between them.

We observe also that all components except component 3 have at least one pair of airports in the same country, in many components rather more so. As expected for the highest loadings on unique components, a number of the airports are quite small, albeit with some exceptions. A full list is given below²⁸. Of these airports, only two (Frankfurt and Munich) also appear in the high mean delay airports presented in Table 26.

Table 28 (further) shows that these ten components display remarkable orthogonality, compared with the correlations that may have been expected between only eighteen components derived from 199 airports (some of the other eight components had higher inter-component correlations, some had none at all; the maximum correlation coefficient was only 0.44).

²⁷ See Appendix E for a summary of strike activity during the baseline traffic day.

²⁸ The airports shown in the table, in the same order per component, are as listed below:

- 1) Pristina, Munster, Dortmund, Malmo, Tirana, Split, Santander, Athens, Cluj, Bratislava;
- 2) Keflavik, Guernsey, Leeds Bradford, Aberdeen, Bodo, Krakow, Kerkira, Bari, Bucharest Henri Coanda, Belgrade;
- 3) Oulu, Kristiansand, Goteborg Landvetter, Vilnius, La Palma, Paphos, Almeria, Kos, Brindisi, Tbilisi;
- 4) Helsinki, Jersey, Exeter, Tromso, Katowice, Goteborg Save, Thessaloniki, Zakynthos, Cagliari, Trieste;
- 5) Dresden, Bournemouth, Alesund, Bergen, Bilbao, Santiago, Biarritz, Lyon, Brest, Trieste;
- 6) Belfast Aldergrove, Shannon, Stavanger, Umea, Stockholm Bromma, Ibiza, Murcia, Beauvais, Santorini, Milas;
- 7) Berlin Schönefeld, Belfast City, Burgas, A Coruna, Menorca, Valencia, Vigo, Palermo, Porto, Ankara;
- 8) Munich, East Midlands, Tenerife Sur, Tenerife Norte, Bastia, Nice, Praha, Salzburg, Adana, Dalaman;
- 9) Dresden, Rotterdam, Billund, Biarritz, Lille, Diaporas, Venice, Florence, Ljubljana, Madeira;
- 10) London City, Stockholm Arlanda, Gran Canaria, Zagreb, Toulouse, Nantes, Malta, Bucharest Baneasa, Trabzon, Kiev.

Table 28. Baseline delay component correlations (from model input data)

Component	1	2	3	4	5	6	7	8	9
2	0.07								
3	0.16	0.11							
4	0.13	0.07	0.04						
5	0.10	0.13	0.09	0.08					
6	0.14	0.10	0.10	0.20	0.16				
7	0.04	0.13	0.16	0.15	0.19	0.22			
8	0.16	0.10	0.11	0.13	0.21	0.15	0.16		
9	0.21	0.07	0.13	0.12	0.15	0.24	0.12	0.09	
10	0.16	0.14	0.12	0.26	0.20	0.26	0.18	0.07	0.19

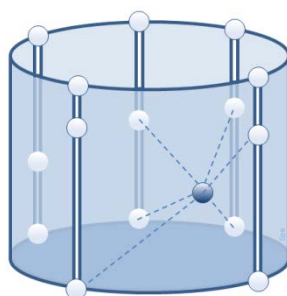


Figure 22. Inter-component delay propagation nodes

Figure 22 shows (in a simplified way) the approach to be taken in the next steps. We may imagine the components of the factor analysis to be represented as strips of airports (spheres) on the surface of a cylinder. Each component represents a 'backbone' of delay with its own temporal characteristics, such that we can consider these vertical strips to be 'waves' of delay. They need not be chronological – e.g. component 6 could 'occur before' component 5. (We cannot assign actual sequencing to the components in practice, but the concept is represented by putting the strips on a closed surface, such that a given time may be (approximately) repeated as we move from one component to the next). The darker sphere in the centre of the cylinder represents an airport with a relatively high loading on several components: it may be considered as a node through which the components are connected, and through which delay is propagated. As we have observed, in each component there is a group of airports that load particularly highly onto it. Since these airports are, by definition, particular (even unique) to the component, they tend to be smaller airports. Lower down the component loading lists, there will clearly be airports with low coefficients on all components. We wish to find airports somewhere in between these two characteristics, with fairly high loadings across the (unique) components – such as the propagation node denoted in the centre of the cylinder.

To identify such nodes, for each of the ten extracted components, coefficients lower than 0.1 were recoded to zero and the number of components onto which each of the 199 airports loaded was counted. We found that counting the number of components on which airports had substantive loadings in this manner was a better way of differentiating the airports than using a simple sum of the coefficients for each one (e.g. since several airports would have multiple low coefficients, driving them relatively high up a ranked sum list).

The frequencies of such loadings are shown in Table 29. The 24 airports²⁹ loading substantively on five or more components were considered to be candidates for propagation nodes.

Table 29. Component loading frequencies (from model input data)

Components loaded upon	Frequency
6	5
5	19
4	34
3	41
2	53
1	40
0	7

For each pairing of these 24 airports a measure of delay asynchrony was calculated. Bivariate correlation coefficients were computed between the airports across the time periods of $M_{a,t}^d$. Coefficients nearer to +1 indicate a high degree of delay alignment between the time periods. Values nearer to zero indicate many periods of zero delay, especially if common for both airports in the pairing or for one airport and with *variable* non-zero delays for the other. Mixed patterns, such as several high-delay – zero-delay correspondences amongst many zero-delay – zero-delay correspondences, give intermediate correlations (e.g. of around 0.5). Perfectly off-set patterns (such as A and B in Figure 19) give correlations approaching -1.

We therefore use the simple linear transformation $c \rightarrow (1 - c)/2$ to convert these coefficients from $+1 \rightarrow 0$, $0 \rightarrow 0.5$ and $-1 \rightarrow +1$, such that instead of representing cotemporaneous delay with higher coefficients, we now have values ranging from +1 (perfect off-set) to 0 (cotemporaneous). In other words, the new coefficients may be used as weights describing asynchrony, thus weighting down component interactions which occur in the same time periods and are thus less likely to have been propagated from one airport to another. This a crude method to weight for causality and diminish artefactual delay relationships, such as high delay occurring everywhere at peak hours.

Finally, the number of passengers travelling between each airport pair (either directly or indirectly) was calculated and where this totalled fewer than 20 (judgementally assigned) the asynchrony weight was (re)set to zero for the corresponding airport pair. This weights-out airport pairs with no passenger itineraries, considering these delay relationships to be non-causal: again, in a fairly crude manner since other causalities may be in effect, not least aircraft dependencies, whereas this overall method is at least superior to a simple delay correlation approach based on flights only.

Other methods, such as applying flight or passenger weights directly (instead of using the passenger itineraries effectively as a filter) would simply produce final results that were dominated by high volume airport pairs in terms of flights and/or passengers.

²⁹ In alphabetical order: EBBR, EDDH, EDDL, EDDM, EDDT, EDLW, EGGP, EGGW, EGHI, EGLL, EGPH, EHEH, ESSA, LEAL, LEBL, LECO, LEMD, LFBO, LFL, LFMN, LFPO, LIMC, LIME, LTAC.

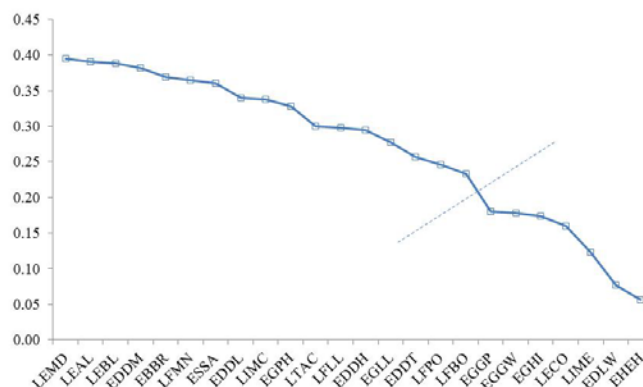


Figure 23. Scree plot for inter-component delay propagation nodes (from model input data)

This final filtering process produced (arbitrary) weights between 0.40 (LEMD, Madrid Barajas) and 0.06 (EHEH, Eindhoven), such that retaining airports with weights of at least half the maximum (the lowest of these being LFBO, Toulouse, at 0.23) yielded the final list shown in Table 30. This filtering process has thus removed seven lesser-ranked airports³⁰. Figure 23 shows the scree plot with the boundary between Toulouse and these seven airports. Table 30 shows the final listing in alphabetical order, rather than by the weights, since these were used more as discriminatory filters.

Table 30. Key inter-component delay propagation nodes (model input data)

EBBR	Brussels
EDDH	Hamburg
EDDL	Düsseldorf
EDDM	Munich
EDDT	Berlin Tegel
EGLL	London Heathrow
EGPH	Edinburgh
ESSA	Stockholm Arlanda
LEAL	Alicante
LEBL	Barcelona
LEMD	Madrid Barajas
LFBO	Toulouse
LFLL	Lyon
LFMN	Nice
LFPO	Paris Orly
LIMC	Milan Malpensa
LTAC	Ankara Esenboğa

(Alphabetical order)

Although the objective here was to identify delay propagators, it is not surprising that based on the model input data, several airports (Munich, London Heathrow, Barcelona, Madrid and Paris Orly) have already been identified in the list of the worst ten airports for average departure delay per time period on the baseline traffic day. Four of the airports (including Paris Orly again) are in France, all of which feature in the discussion of French strike action in Appendix E. The role of smaller airports in delay propagation, and a more detailed examination by the flight- and passenger-oriented layers, will be taken up in Section 3.4.

³⁰ Viz.: EGGP, Liverpool; EGGW, London Luton; EGHI, Southampton; LECO, A Coruna; LIME, Bergamo; EDLW, Dortmund; EHEH, Eindhoven

The whole process was then repeated for the modelled (*) *output* data, specifically, for arrivals for the S_0 baseline traffic day. A matrix of arrival delays ($M_{a,t}^{a*}$) was defined such that $a_{a,t}^*$ is the sum of all arrival delays for each airport, a , and for time windows, t , of period one hour each:

$$M_{a,t}^{a*} = \begin{pmatrix} a_{1,1}^* & a_{2,1}^* & \dots \\ a_{1,2}^* & a_{2,2}^* & \dots \\ \dots & \dots & \dots \end{pmatrix}$$

[4]

The method applied was exactly as per that described for the model input data, except that the delay asynchrony bivariate correlation coefficients were all recalculated based on the modelled *arrival* delays. The final filtering process produced (arbitrary) weights between 0.60 (EDDF, Frankfurt) and 0.02 (LEAM, Almería), such that again retaining the airports with weights of at least half the maximum (the lowest of these being LSGG, Geneva, at 0.31) yielded the final list shown in Table 32. Figure 24 shows the scree plot with the boundary between Geneva and the nine lesser-ranked³¹ airports.

Table 31. Component loading frequencies (from model output data)

Components loaded upon	Frequency
8	1
7	2
6	20
5	36
4	46
3	47
2	24
1	14
0	9

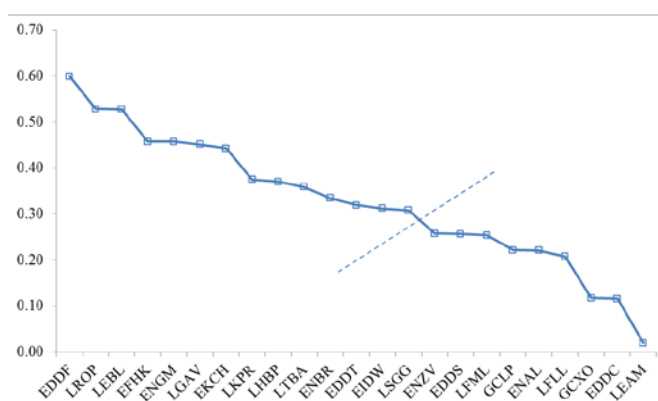


Figure 24. Scree plot for inter-component delay propagation nodes (from model output data)

³¹ Viz.: ENZV, Stavanger; EDDS, Stuttgart; LFML, Marseille; GCLP, Gran Canaria; ENAL, Alesund; LFLL, Lyon; GCXO, Tenerife Norte; EDDC, Dresden; LEAM, Almería.

Table 32. Key inter-component delay propagation nodes (from model output data)

EDDF	Frankfurt
EDDT	Berlin Tegel
EFHK	Helsinki
EIDW	Dublin
EKCH	Copenhagen
ENBR	Bergen
ENGM	Oslo Gardermoen
LEBL	Barcelona
LGAV	Athens
LHBP	Budapest
LKPR	Prague
LROP	Bucharest Henri Coanda
LSGG	Geneva
LTBA	Istanbul Atatürk

(Alphabetical order)

Apparent from these results is that more airports loaded onto a higher number of components in the model output data (see the frequency table) and that the gradients of the scree plots are similar. The key airports finally selected (above) differ from those based on the analysis of the model input data, with only Barcelona and Berlin Tegel being common to both lists. This is to be expected to a considerable extent, as the model resets the actual delays of the baseline traffic data and simulates a normative day based on typical delay patterns. The difference in the two lists thus goes some way to indicating that this has been successful. The French airports affected by the French strikes do not appear in the list above, for example.

The next section will examine the network perspectives using a further set of analytical techniques, after which the results will be reviewed in Section 3.5.

3.4 Network and complexity analyses

3.4.1 Introduction

Section 3.4 focuses in particular on delay propagation. These investigations are from the perspective of the flight, the passenger and the airport. The study of delay in terms of its distribution between its constituent components, viz. primary and reactionary delay, has for some time been a focal point of research and recognised as an area requiring further exploration. As we opened POEM Deliverable 1.1:

Despite [...] the large share of almost 50% of reactionary delay, there is presently only a limited knowledge of how airline, airport and ATM management decisions affect the propagation of reactionary delay throughout the air transport network.

EUROCONTROL (2011a)

Reactionary (higher order³²) delay is any departure delay that is causally linked to a previous rotation of the same, or a different, aircraft. Such dependencies may be due to connecting passengers or crew, or may reflect indirect effects such as gate occupancies as discussed in Section 3.1. Primary (first order) delay arises due to such issues as weather, an aircraft technical problem or capacity restrictions (e.g. ATFM delay). In POEM, the causality of *primary* delay is less of an issue, as we wished to focus more on the generation of reactionary delay and its propagation, as examined in this section³³.

In 2010, as observed in Section 2.5, reactionary delay accounted for 46% of all departure delays in Europe (EUROCONTROL, 2011a). The value reported for 2012 was the same, to the nearest percentage point (EUROCONTROL, 2013a). Figure 25 shows the sensitivity of the air transport network to primary delays, through the ratio of reactionary to primary delay. In 2010, the ratio was approximately 0.9, such that on average, one minute of primary delay resulted in approximately 0.9 minutes of reactionary delay. Peaking in 2010, the ratio improved somewhat in 2011 and 2012.

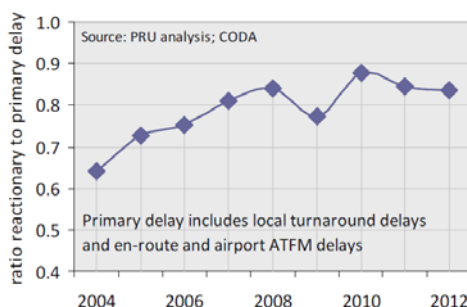


Figure 25. Sensitivity of the air transport network to primary delays

Source: EUROCONTROL (2013a).

However, as Jetzki (2009) points out, most delay studies have traditionally focused on the analysis of primary delays at departure airports. Indeed, this work (*ibid.*), using European data covering a period from 28 October 2007 to 25 October 2009, is a rare example of research exploring reactionary delays in Europe. Here, it is cited that the main share of reactionary delay is due to rotational reactionary delay (IATA code 93 – see Section 2.5), accounting for 89% of all reactionary delay reported during the analysed period. There are degrees of variability regarding the extent to which this type of reactionary delay dominates: according to the airline business model, airport type and time of day, for example. In Case Study 2 (POEM Deliverable 7.1) we reported values of 25 – 30% relating to passenger connections (code 91) for a flag carrier at its European hub.

³² Second order refers to the first reactionary delay in a chain, third order to the second reactionary delay in the sequence, etc.

³³ The calibration of primary delay was nevertheless important, as summarised in Rule 118 (Appendix A) and under model calibration (Section 2.5).

The precise attribution and recording of reactionary delay codes is not without its problems. Consider for example the dispatcher at the end of the operational day whereby an aircraft has accumulated delays during multiple rotations and pushes back 60 minutes late. It is by no means trivial to accurately assign these 60 minutes back through the propagation tree.

In Section 3.4, a number of techniques are applied to analyse different types of delay. Reactionary delay is analysed using random trees and random graphs. Since the causes for the reactionary delay can be attributed in POEM, a reactionary delay propagation tree can be constructed. However, this is not a deterministic tree, for the model is of a stochastic nature and the causes may be different across the simulation runs. The random tree is then approximated using an adapted Monte Carlo method. Primary delays are analysed using Granger causality in a complexity network context, also exploring nodal connectivities and propagation. Primary delays tend to be distributed more evenly during the operational day, whereas reactionary delays impact the network more heavily as the day progresses – although the causal origins of these may have been much earlier, as will be observed. Propagation chains clearly have longer to develop (or perhaps to be mitigated) through the day. Higher order delays may be experienced as both arrival and departure delay.

Figure 26 shows total departure delay plotted as a function of total arrival delay for given airports (circles). The diameter of each circle expresses the airport's size (in terms of movements) whilst the sequentially graded colours reflect the simple departure / arrival delay ratios, r : red for $r \gg 1$ (delay amplifier); orange for $r \cong 1$ (delay transferor); green for $r \ll 1$ (delay attenuator³⁴). The straight line bisecting the plot shows equal arrival and departure delay whilst the dashed curve shows the fitted quadratic trend (weighted by airport size). The worst five apparent delay amplifiers are: LFPG (Paris Charles de Gaulle), LEMD (Madrid Barajas), EGLL (London Heathrow), EDDF (Frankfurt) and EDDM (Munich), which were also the five busiest in terms of aircraft movements. The (apparent) trend is that larger airports seem more likely to act as delay amplifiers. However, these observations are based on aggregate delay movements without attempting to establish any flight-flight causalities. Whilst the ratios are self-evident, such trends may actually more reflect generalised congestion rather than causal links. Indeed, as we shall see later, *smaller* airports are characterised as delay amplifiers under more discriminatory analytical techniques.

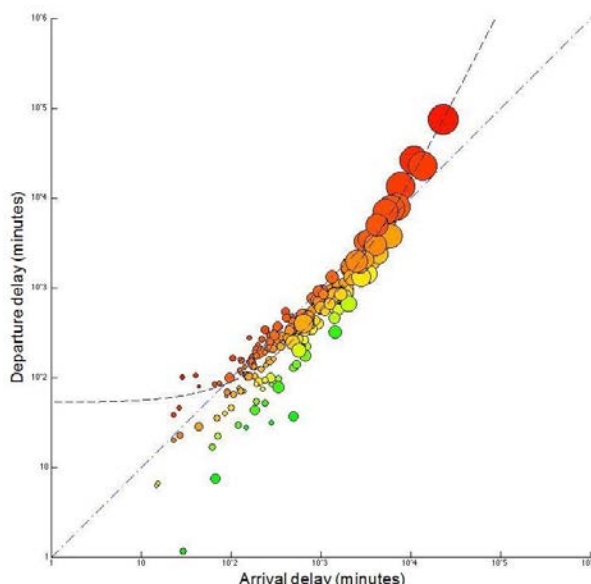


Figure 26. Simple airport delay amplifiers and attenuators

As explained previously, we have selected one combination of simulation day and scenario of special interest from the perspective of the core classical results, to explore in more detail throughout the rest of this section. This combination is scenario A_1 applied to the baseline traffic day, compared with the no-scenario (S_0) baseline traffic day.

³⁴ These terms were introduced in Section 3.1.

3.4.2 Reactionary delay and the propagation of delay

The cause of a reactionary delay should always be identifiable, although this is more difficult in operational practice for longer chains (higher order delay³⁵), as mentioned. In the model, however, we have a (directed) relationship between flights, which can be used to construct a graph that describes how some flights induce reactionary delay and how this propagates through the day. Since the air transportation system is not deterministic, or at least is complex enough to appear non-deterministic (Zanin and Lillo, 2013), we cannot determine *a priori* where or when reactionary delays will appear. Some flights, however, are more likely to be delayed by a reactionary delay than others.

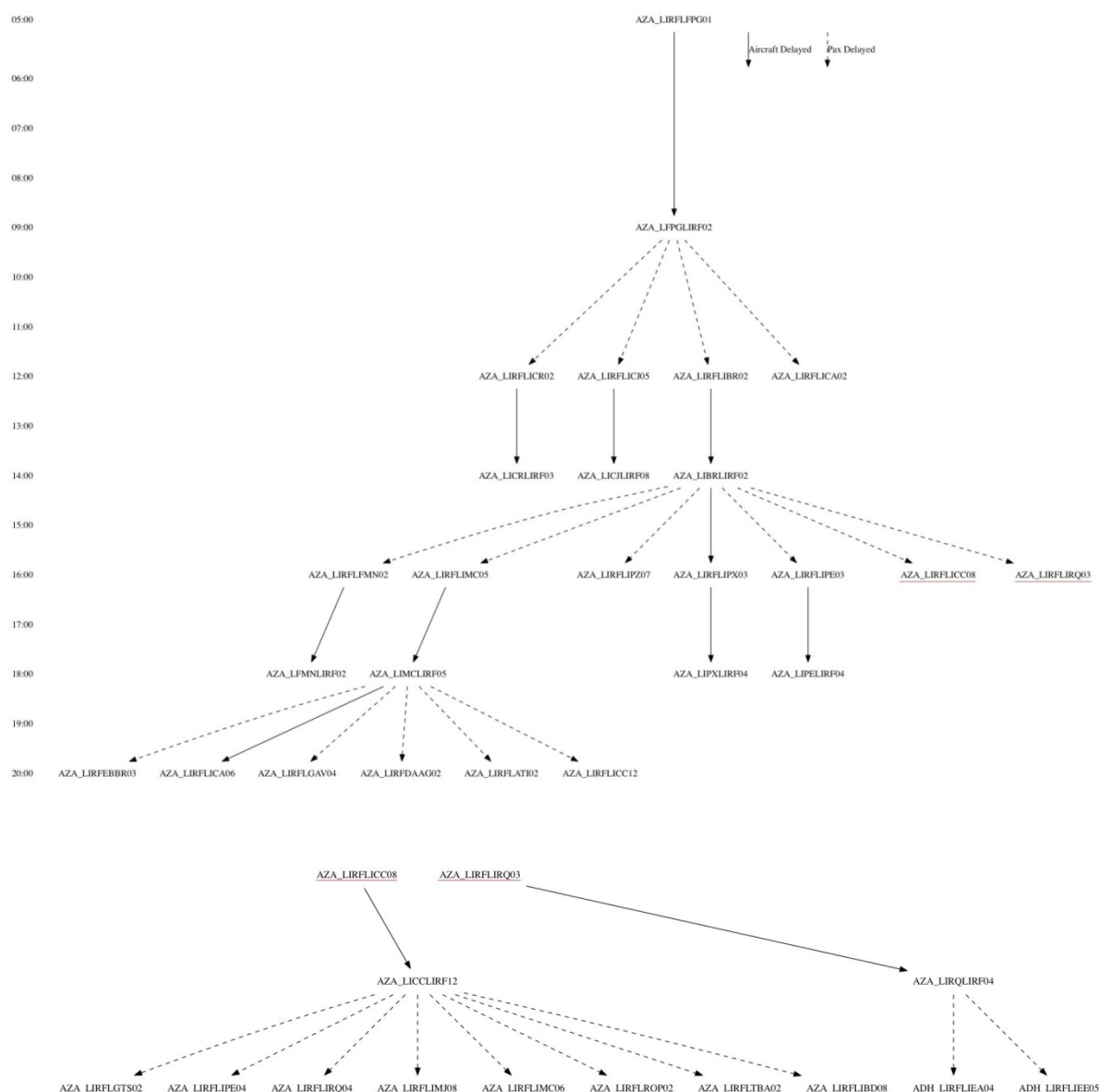


Figure 27. Qualitative reactionary delay propagation tree for AZA_LIRFLFPG01

Note. Lower panel continues upper panel.

³⁵ In these discussions, we include orders higher than one in the term 'reactionary' delay and use the term (unless stated otherwise) to apply to summations of all arrival delay from a given primary delay.

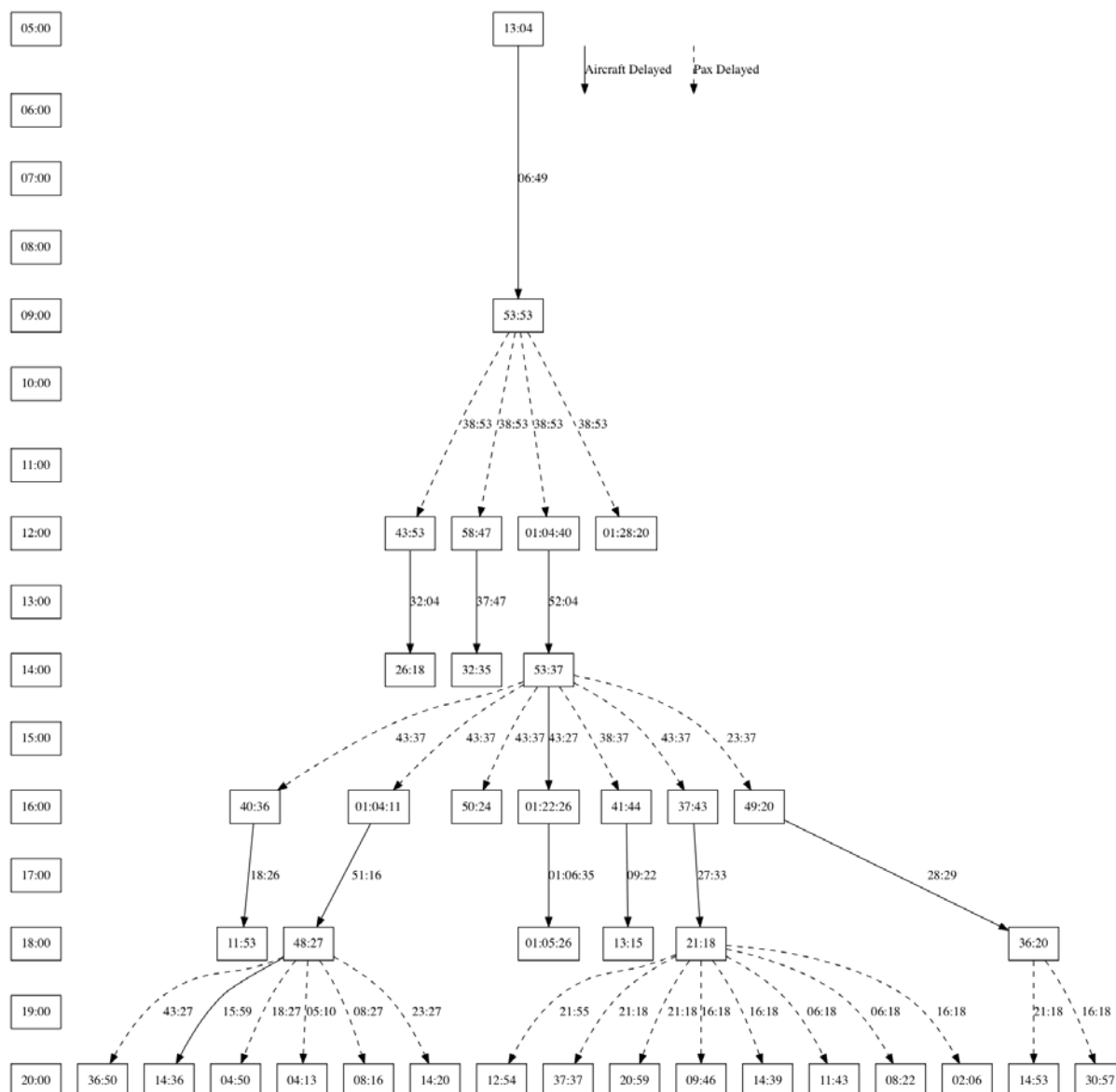


Figure 28. Quantitative reactionary delay propagation tree for AZA_LIRFLPG01

In the POEM model, reactionary delays and their sources can be determined *a posteriori*. In the reactionary delay propagation graph, for a particular model run, nodes are flights and are connected by a directed edge if the first node introduces any reactionary delay in the second. Note that even if several passengers were connecting from different flights and all of them were late, we only consider *the most restrictive*³⁶ connection as the reason for the reactionary delay being induced: therefore, only two flights are linked at a time. In this sense, one flight can delay many others, but any given flight can only be delayed by one previous flight (the most restrictive one). So the graph is in fact a tree. An example is given in Figure 27 for flight AZA_LIRFLPG01.

³⁶ The most restrictive connection is defined temporally. If flight XYZ003 was scheduled to depart at 1000 but actually departed at 1015 because incoming passengers from XYZ001 arrived at 1005 and those from XYZ002 arrived at 1015, we would set XYZ002 as the reason for the reactionary delay. Note that, since the model is run multiple times, it is possible that sometimes XYZ002 would be the most restrictive connection, whilst for other realisations it would be XYZ001.

Since the model is in fact a stochastic model, the structure of the reactionary propagation tree changes from one run to another. Nodes won't change because the flights remain the same between runs, but their connections do. In this sense, we call each tree constructed after one run of the model a 'realisation'. Since the true random tree cannot be determined by analytical methods, we iterate simulations, building up enough realisations such that the error of taking the empirical tree for the true tree is acceptable (i.e. within a 95% confidence interval).

In Figure 28, for example, the (total) reactionary delay of flight AZA_LIRFLPG01, now quantifying the previous example, would be more than 20 hours. By comparing this with the primary arrival delay of the original flight (13 minutes) it might seem that the reactionary delays are routinely very much larger than the primary delays, but the average over the network is of course a lot less than this, as many primary delays are absorbed by buffers or recovered at turnaround, for example. From the tree we can also determine how many flights were delayed (37 in this case) and how long it took to recover from the original delayed flight (here, from 0500 to 2000 UTC).

In order to compare the S_0 and A_1 simulations for the baseline traffic day we present an adaptation of the Kaplan-Meier estimator for the survival function, adapted to give an upper bound of how many flights a delay can persist through ('survive'). The set of all flights is divided into n bins of equal size (C_i), the flights being distributed across the bins by decreasing severity of delay. Then, for each bin, the maximum delay over the subset is computed, $f(i) = \max\{\text{Delay}(j) : j \text{ in } C_i\}$. If the subset consisted of all the flights there would be no possibility for the reactionary and arrival delay to survive, since at the end of the operational day the model resets all delay to zero. Conversely, if single-flight bins were used, the delay corresponding to the first bin would be the maximum possible delay. Note that the values obtained by this method are only upper bounds, such that there may not actually be any specific delay propagating that realises this bound. Note also that the *same* flights are considered under S_0 and A_1 .

We should consider 'survivability' as the gradient of a given curve, whereby a flatter curve (more uniform distribution) represents higher survivability. Hence, from Figure 29, we can conclude that arrival delay shows the greater survivability (albeit with a lower absolute value) and that there is no notable difference between S_0 and A_1 in this respect. Unexpectedly, reactionary delays are distributed quite differently. Nevertheless, under scenario A_1 only a few flights (less than 1% – the area between the dashed and solid blue lines) introduce more reactionary delay (this varies from +10% to +30%) than under scenario S_0 . We remind ourselves that whilst these types of plot offer useful behavioural insights, the model's algorithms under A_1 are geared to cost reduction, such that it is to the cost and passenger-centric metrics that we must turn in order to see the corresponding specific impacts, as concluded in Section 3.2.

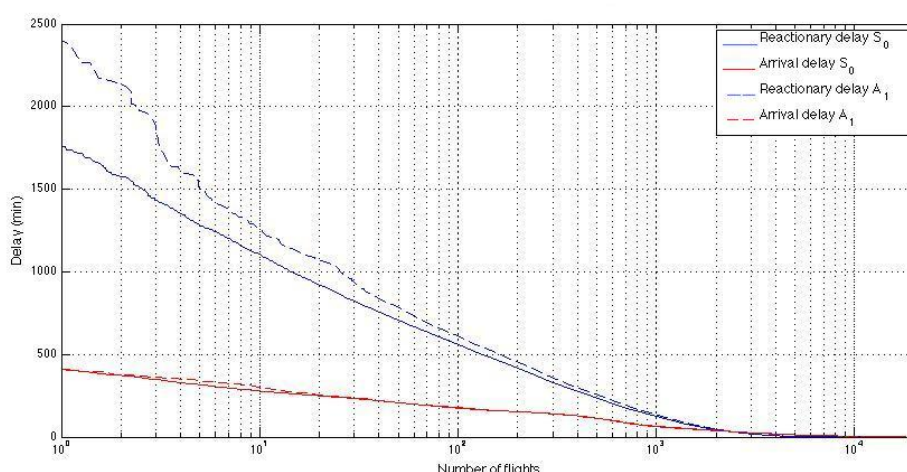


Figure 29. Reactionary and arrival delay distribution

In order to see how these delays are temporally distributed overall in the network³⁷, we first categorise the delays by time of day using time windows of size 5 minutes – see Figure 30. For each time window, we compute the cumulative arrival and reactionary delay for all flights scheduled to be in-block within the time window (normalised by the number of runs). Both scenarios share similar patterns. Again, we observe that reactionary delay generated under A_1 is slightly larger than under S_0 , although the overall shapes of the distributions are similar. On average, reactionary delays affected more flights under scenario A_1 . Also as evidenced before, only a few flights introduce a large increase in reactionary delay. Nevertheless, since this delay is introduced purposefully by the airline through the cost model (i.e. waiting for late passengers) the overall cost to the airline decreases. This trade-off for the airlines was remarked upon at a high level in the findings of Section 3.2.

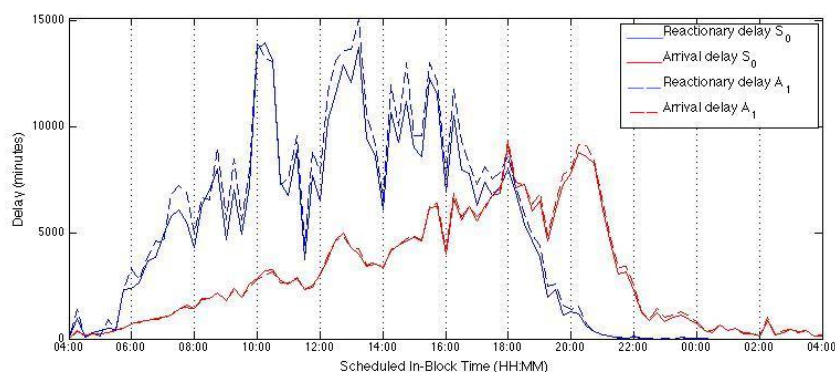


Figure 30. Arrival and reactionary delay distributed by scheduled in-block time

Most of the reactionary delay is generated (but not experienced) earlier in the day; approximately 80% of all reactionary delay is generated before 1615. Arrival delay is experienced later in the day. European departure delay tends to peak towards 2200 UTC³⁸ as the airline focus shifts to connectivity rather than punctuality, although our model does not explicitly model this shift, a similar trend is evident.

By 1615, only 45% of the total arrival delay has been experienced and one would have to wait until 2015 to experience 80% of the total arrival delay – that is, four hours longer to experience the same proportion of arrival delay as reactionary delay. In fact, for proportions of over 50% of the delay, this amount of time remains constant as Figure 31 shows (the dashed line represents the overall trend; graph based on S_0 only).

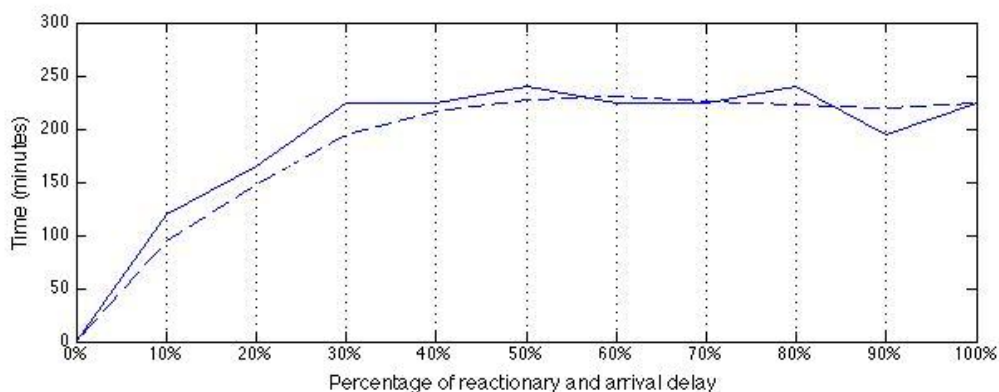


Figure 31. Time required to yield the same proportions of arrival and reactionary delay

³⁷ Further graphics, by selected airports, are to be found in Appendix F.

³⁸ Personal communication, CODA.

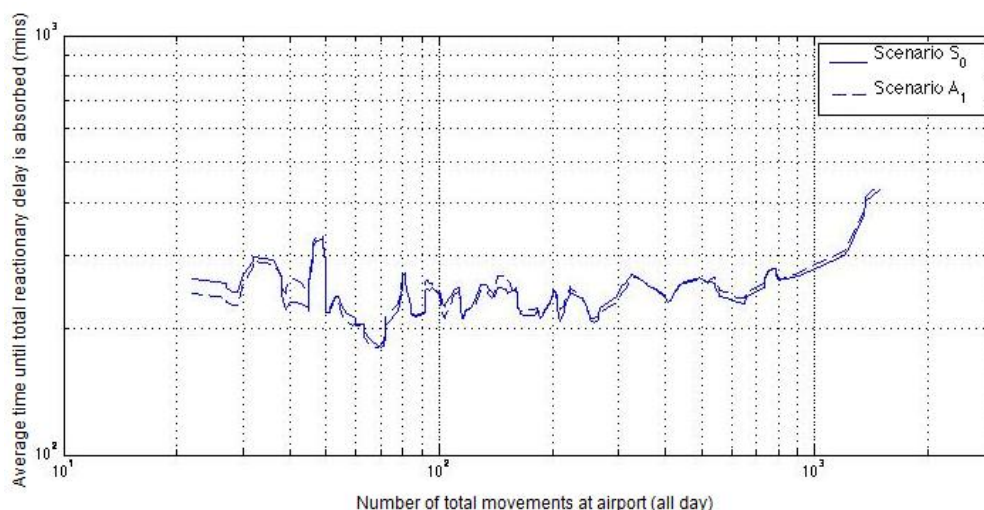


Figure 32. Average time until total reactionary delay is absorbed, by airport size

Some delays (even if they are not severe) cannot be absorbed by the network effectively and it may take some time before they terminate, sometimes first propagating to a large number of flights. For each connected component in the propagation tree, we determine the longest branch by using a recursive deep search. We then take the difference between the arrival times of the flight represented by the terminal node and the origin node. Following this particular metric, the average time for reactionary delays to terminate is 268 minutes (standard deviation: 200 minutes) for S_0 and 272 minutes (standard deviation: 275 minutes) for A_1 – see Figure 32. Both scenarios show very similar patterns. Of note, is that the time it takes until the reactionary delay is absorbed does not increase with respect to the size of the airport where it was generated.

Figure 33 plots total (daily) reactionary and arrival delay as a function of airport movements. Although large airports are associated with more reactionary and arrival delay, there is a considerable relative difference between these delay types at the smaller airports. Some of the forty smaller airports (in terms of movements) behave as delay amplifiers: in some cases arrival delay is doubled (or even tripled) into reactionary delay. This may be caused by late-arriving flights being less able to recover delay at smaller airports, thus propagating it as reactionary delay, echoed to some extent by Jetzki (2009): “The share of reactionary delay for non-hub operations ranges between 34 and 50 percent and is considerably higher than for hub-operations.” Such recoveries may be driven by turnaround and crew resources (see Rule 10, Appendix A) and, of course, onward flight opportunities for passenger reaccommodation.

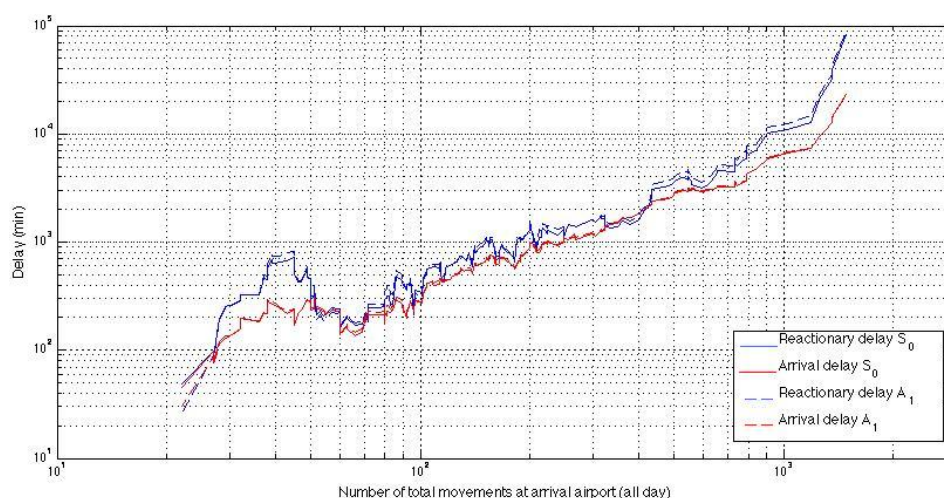


Figure 33. Arrival and reactionary delay, by airport size

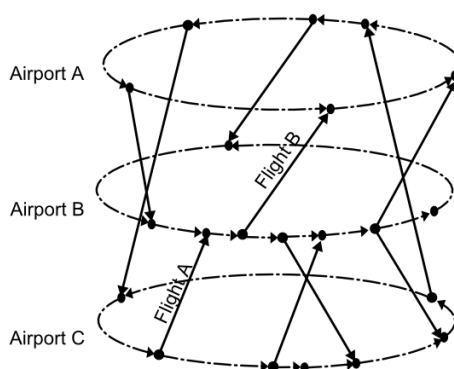


Figure 34. Airport time-line graph

In order to develop the focus on airport metrics, we built an airport reactionary delay propagation graph (as introduced in POEM Deliverable 4.1). In Figure 34, nodes are time-lines at each airport and two airport-time points (Ap_1, t_1) and (Ap_2, t_2) are connected if there is a flight in the reactionary delay propagation tree departing from Ap_1 at t_1 and arriving at Ap_2 at t_2 . This time-line graph also allows the study of how reactionary delay spreads through the network, including through back-propagation. By using exhaustive (path) search algorithms (such as breadth-first searches³⁹) it is possible to determine how reactionary delay travels across the network, from one airport-time point to another airport-time point. By collapsing the whole time-line to a point the problem is simplified and the total amount of interchanged reactionary delay within the simulation time-window can be determined.

Tables 33 and 34 characterise delay propagation from the airport perspective. The highest dozen (as opposed to the highest ten) entries are listed to show that the high-frequency pairing LEMD (Madrid Barajas) – LEBL (Barcelona), identified in Section 3.3 as experiencing high delays during the baseline traffic day, ranks eleventh in the airport pairs sharing the highest levels of propagation delay. Note also that this is the first pairing not to include LFPG (Paris Charles de Gaulle) as one member of the pairing. Since a time-line graph has been employed, none of the pairings include A-B primary-secondary airport pairs in the same city or X-Y proximal pairs, as discussed in Section 3.1.

Table 33. Highest dozen airport propagation pairs

From	To	Propagated delay (hours)
LEMD	LFPG	115
LFPG	EGLL	91
LFPG	LIMC	90
LFPG	EDDF	89
EHAM	LFPG	88
LFPG	LEMD	80
EGLL	LFPG	77
LFPG	EDDL	76
EDDF	LFPG	75
LFPG	EDDM	74
LEMD	LEBL	69
EDDM	LFPG	68

³⁹ BFS is an algorithm used in graph searches that starts at the root node and explores neighbouring nodes until the solution or goal is found. It is an uninformed method in that it is systematic and exhaustive, not using heuristics.

Delays incurred earlier in the day may be expected to back-propagate (often several times per day for a given flight – see Section 3.1) particularly at airports/hubs with higher proportions of short-haul flights and this may also spread tactical demand more evenly than scheduled (Pyrgiotis, 2011). The back-propagation effect at hubs is also investigated by Jetzki (2009), where it is reported that the share of delay back-propagating to the origin airport varied between 20% (Zürich) to 56% (London Gatwick). The average minutes of delay back-propagating ranged from 3 minutes (Zürich) to 12 minutes (Rome Fiumicino). For Frankfurt, London Heathrow and Copenhagen, a notable difference between summer and winter seasons was observed. Primary delays originating from major European hubs were reported as affecting, on average, between 30 and 50 other airports within the ECAC area. The largest impact of primary delays at hubs was through back-propagation. Table 34 quantifies this effect for the POEM model, the list being populated by hub airports and high-frequency pairs as would be expected (e.g. the LEMD – LEBL pair has already been discussed; ENGM – ENBR had 55 flights in total during the baseline traffic day). Most of the delay characterised in these tables is distributed between a relatively limited number of airports. For example, under scenario S_0 , approximately two thirds of the 199 ECAC airports are affected by some level of reactionary delay associated with⁴⁰ LFPG (see Figure 35): also top of the back-propagation ranking. This fraction increases under A_1 to approximately three quarters, whereby the role of LFPG as a (reactionary) delay amplifier increases. More generally under scenario S_0 about 31% of airports are connected⁴¹ by reactionary delay with 90% of all other airports, whilst only 20% of airports are similarly connected with only 50% of all other airports. These values typically increase under scenario A_1 (further figures are in Appendix F).

Table 34. Highest dozen airports in terms of back-propagated delay

Airport		Back-propagated delay (hours)
LFPG	Paris CDG	844
LEMD	Madrid Barajas	424
EDDF	Frankfurt	214
EGLL	London Heathrow	196
LSZH	Zürich	142
EDDM	Munich	115
EDDL	Düsseldorf	91
ENGM	Oslo Gardermoen	73
EKCH	Copenhagen	73
ENBR	Bergen	71
EFHK	Helsinki	67
LEBL	Barcelona	61

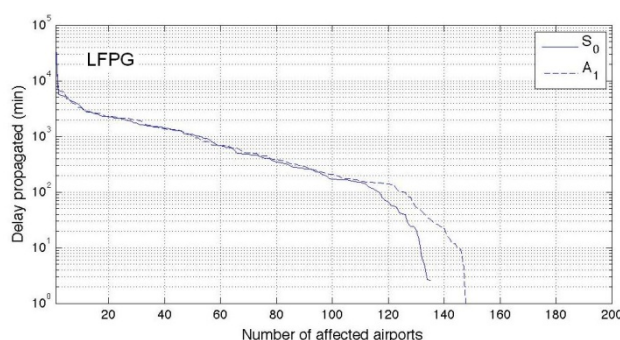


Figure 35. Reactionary delay propagated from LFPG

⁴⁰ Rather than 'originating at', as some flights that propagate reactionary delay from LFPG may have been delayed before arriving there.

⁴¹ 'Connected' in the sense that there is at least one path in the reactionary delay propagation tree connecting both airports.

3.4.3 Granger causality analyses

As introduced in Section 3.1, we next turn to the issue of causality. The Granger causality test is an extremely powerful tool for assessing information exchange between different elements of a system, and understanding whether the dynamics of one of them is led by the other(s).

It was originally developed by Nobel Prize winner Clive Granger (Granger, 1969) and although it was applied largely in the field of economics (Hoover, 2001) it has received a lot of attention in the analysis of biomedical data (Brovelli *et al.*, 2004; Kamiński *et al.*, 2000; Roebroeck *et al.*, 2005).

Classical statistical instruments, like, for instance, correlation analysis, are only able to assess the presence of some common (equivalent) dynamics between two or more systems. However, correlation does not imply causality.

Granger causality, on the other hand, is held to be one of the only tests able to detect the presence of causal relationships between different time series. The two axioms, on which this test are based, are as follows:

Table 35. Granger causality axioms

Axiom	Statement
1	Causes must precede their effects in time
2	Information relating to a cause's past must improve the prediction of the effect above and beyond information contained in the collective past of all other measured variables (including the effect)

Therefore, a time series, q , is considered to Grange-cause another time series, p , if the inclusion of past values of the series q can improve the process of forecasting the values of the time series p . In this case, the future evolution of p also depends on the past values of q . Also, it should be noted that two time series presenting a high correlation, or two time series that are 'forced' by a third system, do not usually pass the Granger causality test: as they have similar values, one of them cannot convey useful information for the forecast of the other.

Nevertheless, claims of causality from (multiple) bivariate time series should always be treated with caution, as true causality can only be assessed if the set of the two time series contains all possible and relevant information and action sources for the problem (Granger, 1980), a condition that real-world experiments can only rarely satisfy (Zanin and Papo, 2013).

A network reconstruction was computed for the flight and passenger layers for both the no-scenario (S_0) and A_1 scenario simulations of the baseline traffic day, i.e. four reconstructions in total. This process comprised three steps.

Firstly, the average hourly arrival delay was calculated for each airport, as a function of the hour of the day of the simulation run, producing data very similar to the matrix $M_{a,t}^{a*}$ used in the factor analyses. Depending on the dataset considered, this delay corresponded to flights' or passengers' delay. The results of different runs were concatenated sequentially, in order to create a time series for each airport corresponding to the delay dynamics of multiple days.

Figure 36 depicts the resulting time series for two example airports (EBCI, Charleroi, blue; EDDB, Berlin Schönefeld, red) for a four-day period (i.e. over 96 hours).

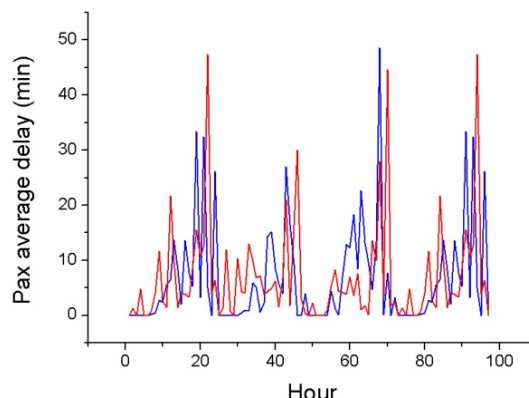


Figure 36. Passenger delay by elapsed time

Secondly, for each pair of nodes (airports), the Granger causality is calculated. This involves, for each time step available, forecasting the next value of the time series by means of a multilinear regression, using the information of the previous 24 hours. Two errors are then compared: one corresponding to the forecast obtained using only information about the first node, and one corresponding to the forecast including information extracted from the time series corresponding to the second node. The result is then expressed as an F statistic's significance level, assessing whether the two forecast errors are significantly different, and thus whether some causality has been detected in the data. Note that the resetting of the delays at the end of each 0400 to 0359 time period (see Section 2.4.1) does not compromise the calculation.

Thirdly, a link between two nodes, A and B, is established when two conditions are simultaneously met: there is a significant causality between A and B ($p < 0.01$) and no causality is detected between B and A ($p > 0.01$). This reduces the effects of confounding factors, e.g. the presence of a third airport forcing both A and B, that would result in bidirectional causalities. The result is an unweighted directed network, where bidirectional links are forbidden. The results are shown in figures 37 through 40, inclusive. Each network reconstruction required approximately 120 hours of computational time.

In the figures, the colour of each node represents its eigenvector centrality, from green (low centrality) to red (most central nodes). The size represents the out-degree, i.e. the number of airports that a given airport 'forces' (in terms of delay). The eigenvector centrality is a metric defined such that this centrality of a node is proportional to the centralities of those to which it is connected. This measure was firstly created for the analysis of social networks, where the 'importance' of a person is proportional to the 'importance' of his or her friends (e.g. one person may be very 'important' even if he/she has only one friend, if that friend is the US President). A node may thus be important, from the point of view of delay propagation, if it is connected to a 'central' airport, in that it can strongly contribute to the propagation dynamics. The position of nodes in the graphics is partly an effect of the representation, although nodes far from the centre tend to be less connected to the core.

In some cases, it is known that the value of the metric alone is insufficient to establish whether the associated topological feature is relevant. For example, the efficiency strongly depends both on the structure of the network and on its number of links. Therefore, in order to have meaningful comparisons of heterogeneous networks, it is necessary to normalise the value obtained against a reference model. This is performed by calculating certain metrics for a large set of random networks, constructed with the same number of nodes and links as the original one. The Z-score, widely used in classical inference testing, is calculated as the difference between the metric in the network under investigation and the average obtained from the random ensemble, divided by the standard deviation of the values obtained in the latter group.

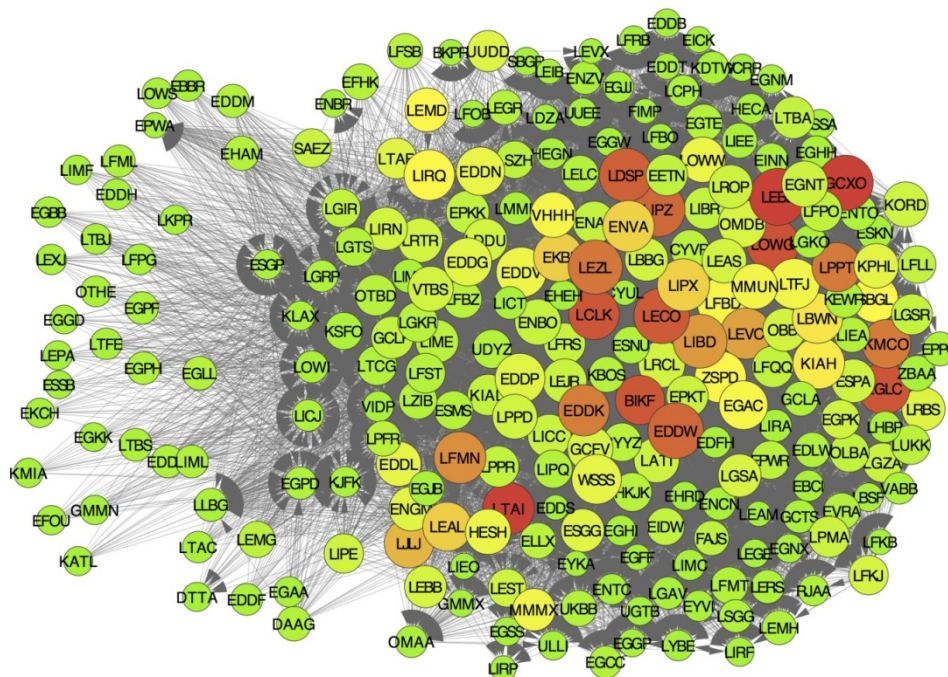


Figure 37. Passenger delay causality network for S_0 simulation

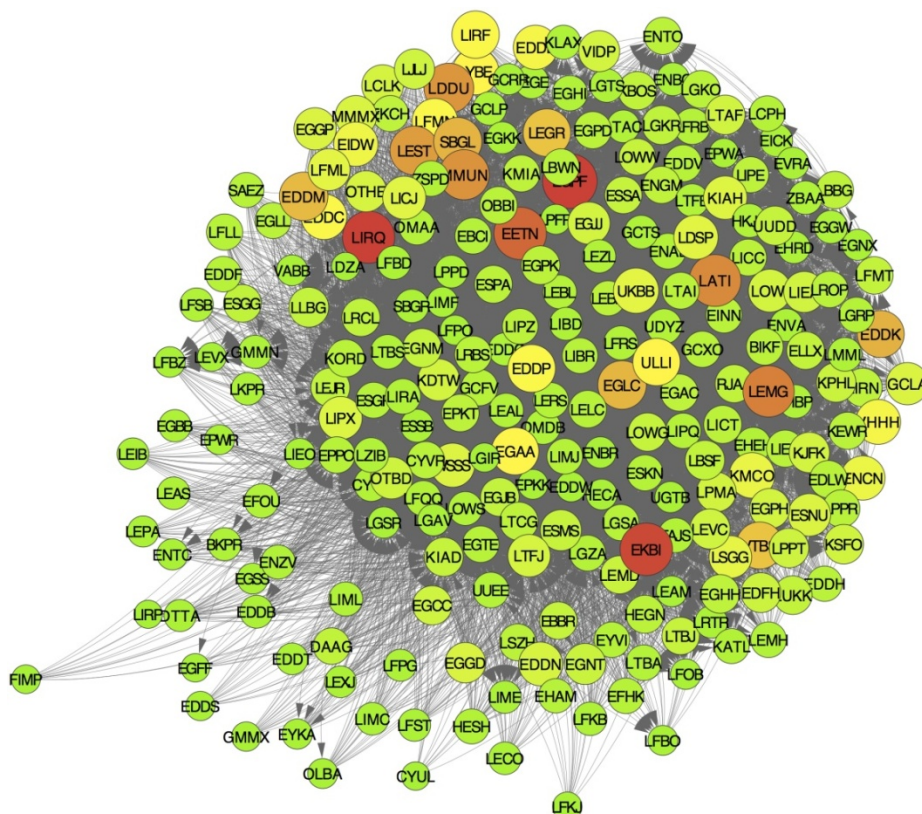


Figure 38. Passenger delay causality network for A_1 simulation

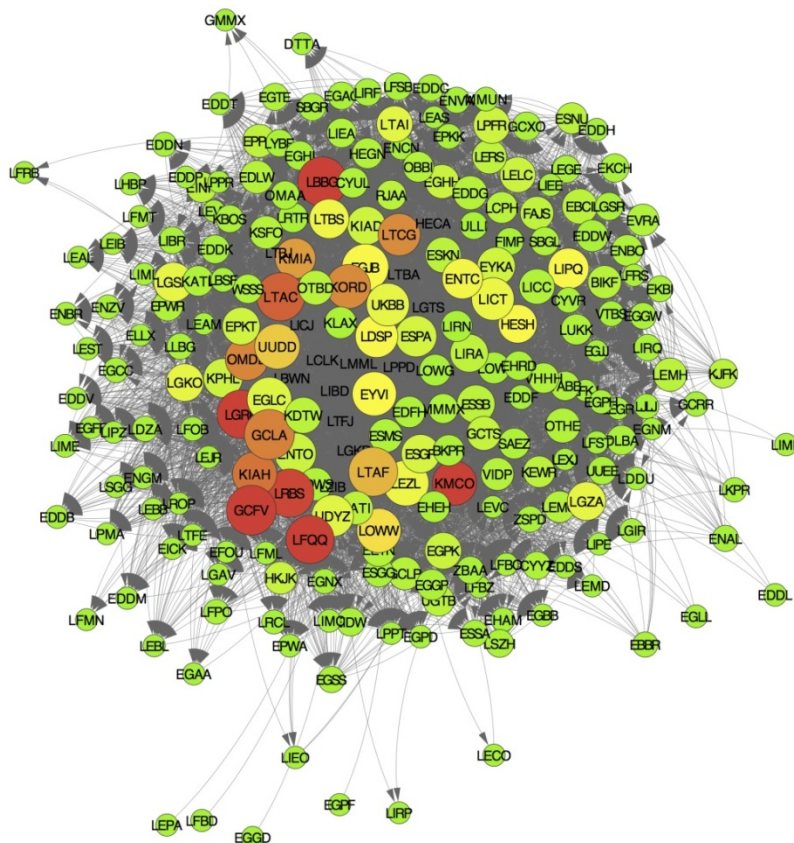


Figure 39. Flight delay causality network for S_0 simulation

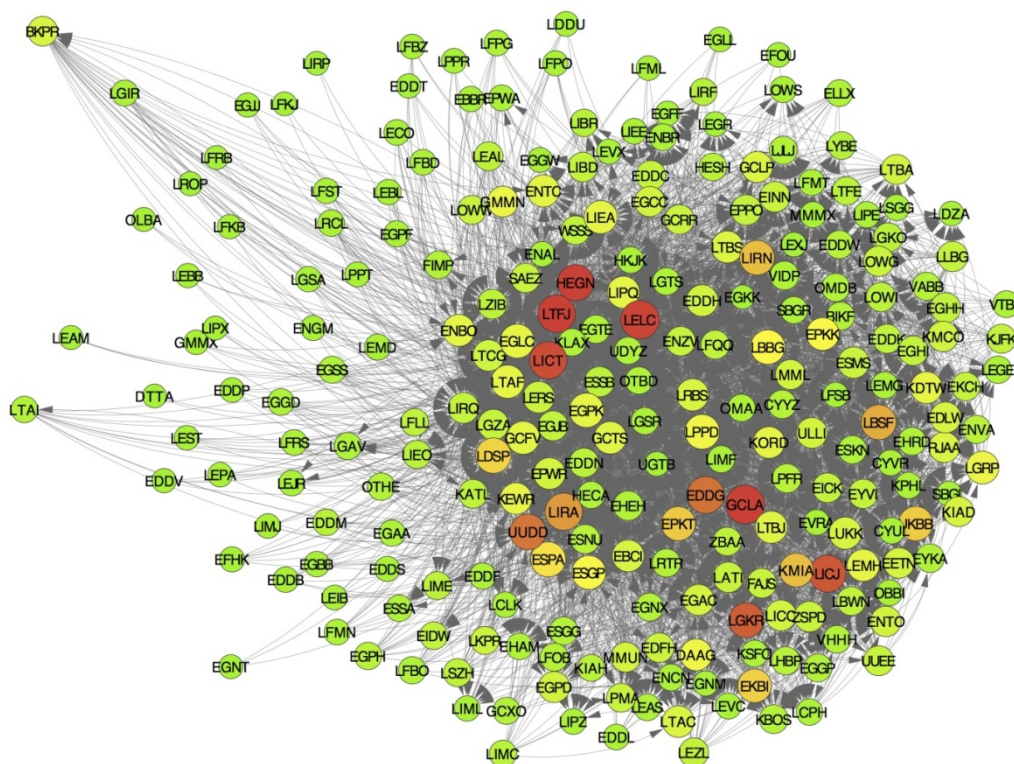


Figure 40. Flight delay causality network for A_1 simulation

Table 36. Summary of key complexity metrics by scenario and layer

Metric	Pax S ₀	Pax A ₁	Flight S ₀	Flight A ₁	Summary implications of effects of scenario A ₁	Metric description
Link density	0.17	0.16	0.13	0.08	Passenger delay propagation occurs more readily than for flights	Proportion of links that are active, with respect to the total number of potential links
Entropy of the degree distribution	5.1	4.8	5.5	4.2	Network more homogeneous; nodes are more similar to each other, with few of them being highly connected	Shannon's entropy of the distribution of node degrees, i.e. of the number of links belonging to each node; provides a measure of the heterogeneity of the network: the maximum value is obtained for a uniform degree distribution, while the minimum (zero) is achieved whenever all vertices have the same degree ⁴² (Wang <i>et al.</i> , 2006).
Degree correlation	-0.00	-0.00	-0.02	-0.01	Practically no effects	Pearson's correlation coefficient of the degrees of nodes at either ends of each link; positive values indicate that highly connected nodes tend to connect with other hubs (Newman, 2002); negative correlations indicate that hubs are mostly connected with nodes with low degree, e.g. with just one or two connections.
Clustering coefficient	0.20	0.26	0.15	0.20	Clustering increases	The clustering coefficient measures the presence of triangles in the network; a high clustering coefficient has been historically associated with social networks; defined as the relationship between the number of triangles in the network and the number of connected triples
Z-score for clustering coefficient	19	73	8.7	60		Defined between minus and plus infinity; values below (above) zero indicate a measurement lower (higher) than expected; the absolute value is the number of standard deviations between the expected and measured values: how divergent the measured value is with respect to what is expected from a random network.
Efficiency	0.46	0.43	0.42	0.34	Delay connectedness reduces, i.e. A ₁ improves the situation	Harmonic mean of the mean path length between pairs of nodes, i.e. the number of links needed to move between two nodes of the network (Latora and Marchiori, 2001); a value of one indicates that it is possible to move from one node to any other in just one step, whereas a value of zero indicates that nodes are not connected; efficiency is an indicator of the connectedness of the network
Z-score for efficiency	-290	-370	-290	-370		See above and main text
Small-worldness	0.90	1.1	0.82	1.7	N/A	Clustering coefficient divided by the mean geodesic distance, the latter being the average number of steps needed to move between two nodes of the network; values between zero and one indicate a random structure (Humphries and Gurney, 2008); the results here are very similar (usually the small-worldness is relevant if above ten, for instance)
Number of (strongly) connected components	57	74	51	84	Delay connectedness reduces, i.e. A ₁ improves the situation; the more connected components, the more the network is divided into different communities	The connected components of a graph are defined as those sets of nodes (or subgraphs) in which any two nodes are connected to each other by at least one path, and which are <i>not connected</i> to any other node of the (original) graph
Size of the giant component	192	175	198	165		Size of the biggest component of the network

(Values, except counts, to 3 s.f.)

⁴² Take, for example, a network with five nodes. The maximum entropy would be obtained for a uniform degree distribution, for instance: node 1 having one link, node 2 having two links, node 3 having three links etc., such that all degrees are equally likely. Thus, if the vector codifying the degrees is (1, 2, 3, 4, 5), its distribution would be flat (uniform), i.e. (1/5, 1/5, 1/5, 1/5, 1/5). In this case, the entropy would be $-5 \ln 0.2 / \ln 2$. In contrast, the minimum entropy is obtained when all nodes have the same degree, e.g. two connections.

Table 37 quantifies the differences between the network layers (i.e. scenario and type of connectivity considered). Table 38 ranks the nodes (for each network layer) according to the number of connections they receive, i.e. the number of airports ‘forcing’ their delays (see preceding explanation). Table 39 does likewise for the out-degree, ranking the nodes according to the number of airports that they force.

Table 37. Percentage of links that are different between network layers

Layer	Passenger S ₀	Passenger A ₁	Flight S ₀
Passenger A ₁	25.3%		
Flight S ₀	23.2%	24.9%	
Flight A ₁	21.3%	19.2%	17.2%

Table 38. Node ranking by in-degree and network layer

Ranking	Passenger S ₀	Passenger A ₁	Flight S ₀	Flight A ₁
1	CYUL (132) Montreal	LGIR (135) Nikos	EGJJ (98) Jersey	LTAF (114) Adana
2	EDDS (114) Stuttgart	ESSB (132) Stockholm Bromma	ESMS (97) Malmo	UDYZ (100) Zvartnots
3	EDFH (106) Frankfurt Hahn	LGSR (128) Santorini	LEXJ (91) Santander	LIRQ (88) Florence
4	EHEH (103) Eindhoven	LGAV (126) Athens	LOWG (90) Graz	LFQQ (86) Lille
5	EHRD (99) Rotterdam	EPKK (115) Krakow	BKPR (83) Pristina	LTCG (83) Trabzon

Table 39. Node ranking by out-degree and network layer

Ranking	Passenger S ₀	Passenger A ₁	Flight S ₀	Flight A ₁
1	LEBL (106) Barcelona	EGPF (99) Glasgow	LTFJ (147) Istanbul Sabiha Gökçen	GCLA (55) La Palma
2	LECO (102) A Coruna	EETN (92) Tallinn	LTBA (142) Istanbul Atatürk	LELC (54) Murcia San Javier
3	EDDW (102) Bremen	EKBI (92) Billund	LTBJ (134) İzmir	LTFJ (51) Istanbul Sabiha Gökçen
4	GCXO (100) Tenerife Norte	LIRQ (91) Florence	LIBD (128) Bari	LIRA (50) Rome Ciampino
5	LIPZ (98) Venice	LATI (90) Tirana	LCLK (124) Larnaca	LICT (50) Trapani Birgi

Table 40 ranks the nodes by eigenvector centrality. This measure includes both in- and out-degree, assessing how important each node is for the propagation of delay. This set of tables concludes with Table 41, wherein we compare these eigenvector centrality rankings through Spearman rank correlation coefficients. A value of +1 indicates that the two rankings are the same (i.e. the nodes are ranked in the same order), whereas a value of -1 indicates the opposite. A value of zero shows no correlation at all between the rankings. The very low values demonstrate the remarkable finding that there is practically no relationship between the importance of airports across the different layers.

Table 40. Node ranking by eigenvector centrality and network layer

Ranking	Passenger S_0	Passenger A_1	Flight S_0	Flight A_1
1	LEBL (0.147) Barcelona	EGPF (0.143) Glasgow	LTBA (0.224) Istanbul Atatürk	GCLA (0.148) La Palma
2	GCXO (0.142) Tenerife Norte	LIRQ (0.133) Florence	LTFJ (0.192) Istanbul Sabiha Gökçen	HEGN (0.137) Hurghada
3	LTAI (0.133) Antalya	EKBI (0.129) Billund	LCLK (0.181) Larnaca	LTFJ (0.136) Istanbul Sabiha Gökçen
4	LOWG (0.127) Graz	EETN (0.123) Tallinn	HECA (0.171) Cairo	LELC (0.130) Murcia San Javier
5	LCLK (0.126) Larnaca	LEMG (0.119) Málaga	LTBJ (0.168) İzmir	LICT (0.128) Trapani

Table 41. Comparison of rankings by network layers for eigenvector centralities

Layer	Passenger S_0	Passenger A_1	Flight S_0
Passenger A_1	-0.007		
Flight S_0	-0.039	0.068	
Flight A_1	-0.037	0.032	0.039

Detailed exploration of the role *specific* airports play with regard to delay propagation is reserved for future research, and we should judge with some caution the nature of delay ‘pushed’ out from the model onto non-ECAC airports (shown italicised), such as Montreal, which are not included in the model itself and hence do not share its delay recovery mechanisms along with the modelled airports. Nevertheless, from these analyses we may draw the following conclusions.

The centrality ranking tables show that different airports have different roles with regard to the type of delay propagated (i.e. flight or passenger delay) and that these are further changed under A_1 . Also apparent is the role of some smaller airports in all the top-ranking lists, a theme to be revisited in Section 3.5.

Notable in the lists is the appearance of numerous Turkish airports. Of the ten European airports with highest absolute growth in 2010, half of them were in Turkey, viz.: Ankara Esenboğa, Antalya, Istanbul Atatürk, Istanbul Sabiha Gökçen and İzmir (EUROCONTROL, 2011a). The growth in Turkey was largely driven by additional international traffic, although also by domestic growth. Ankara Esenboğa is the only one of these five not to feature in the preceding tables, although it did appear in the final airport extraction of the factor analysis carried out on the model input data (see Section 3.3).

High average arrival delay at Istanbul Atatürk in 2010 was commented upon in POEM Deliverable 5.2. With regard specifically to ATFM arrival delays, a high level of capacity-related delays was experienced here, despite a significant reduction compared to the previous year (EUROCONTROL, 2011a). Both Istanbul Atatürk⁴³ and Antalya⁴⁴ served over 10 million passengers in 2010. Both clearly also have significant numbers of connecting passengers. Whilst the presence of Turkish Airlines is more limited at Istanbul Sabiha Gökçen (e.g. operating to Ankara), there is a large low-cost carrier presence here too, notably serving as Pegasus’ main hub, through which connections are sold.

Málaga also served over 10 million passengers in 2010, and although handling fewer transfer passengers, Air Europa sells through tickets for Paris CDG – Malaga – Tenerife, for example. More seasonally peaked, La Palma handled slightly under 1 million passengers in 2010⁴⁵, whilst Hurghada serves Red Sea coastal resorts with passenger volumes of around 8 million per year⁴⁶, now

⁴³ Turkish Airlines’ hub.

⁴⁴ In addition to a large low-cost carrier presence here, Turkish Airlines also operates from Antalya to its hub at Istanbul Atatürk.

⁴⁵ Airports Council International Europe, personal communication.

⁴⁶ <http://www.flightstats.com>

competing with Cairo for direct traffic from Europe and Russia, sharing around 12% growth per year between 1993 and 2006 with the other main airports operated by the Egyptian Airports Company⁴⁷. Data regarding Egyptian traffic are less readily available in the context of European ATF(C)M reporting, since Egypt is a non-ECAC state, although traffic volumes with Europe had resulted in specific initiatives to integrate Egypt into the European ATFCM network in 2010 (EUROCONTROL, 2011b).

Also noted by EUROCONTROL regarding 2010 was that “ninety percent of en-route ATFM delays were concentrated in a comparatively small number of ACCs (17 out of 67), which negatively affected the entire European network [...] the south-east axis (Austria, Croatia, Greece and Cyprus) remains of major concern” (EUROCONTROL, 2011b). Reflecting this, we note the appearance in the preceding lists of: Larnaca; Athens, Nikos and Santorini; Graz; and airports in neighbouring northern Italy.

These contextual comments on some of the airports less commonly featured in delay analyses are to corroborate that their appearance in these lists remains credible, along with those ranging from large-volume, passenger-connecting Barcelona (29 million passengers in total in 2010) to medium-volume Stuttgart (somewhat over 9 million passengers in 2010).

In terms of higher-level findings, the four network layers are clearly substantially different from each other. This may be expected *a priori* in terms of the flight and passenger connectivities, but the differences between the S_0 and A_1 simulations demonstrate that A_1 has had a significant effect on the network dynamics, as indeed evidenced in the previous section.

The flight networks have lower link densities than the passenger networks. This is probably driven by the fact that a given number of causality relationships between airports, in terms of flights, may affect a relatively larger number of passengers connecting through these airports.

A_1 significantly reduces the number of causal delay links, especially for the flight network. Airports no longer form such large (delay) communities. Instead, they form smaller communities, which should promote the containment of delay propagation. However, the Z-scores of the clustering coefficients markedly increase under A_1 , indicating that within these communities, the airports are more strongly connected. From a network point of view, there is thus a trade-off here: the propagation of delay is contained within smaller communities but these communities are more susceptible to such propagation.

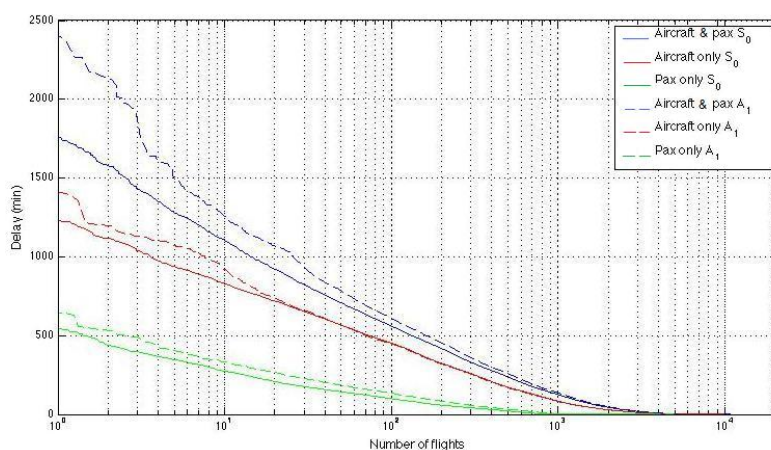


Figure 41. Simultaneous propagation causalities

Finally, in Section 3.4.2, reactionary propagation trees were examined by linking flights through either aircraft rotational or passenger connectivity dependencies. The propagation tree is significantly larger when considering both causes simultaneously (upper curves, “Aircraft & pax S_0 ” and “Aircraft & pax A_1 ”) compared with the simple addition of delays propagated when considering the two causes separately, as shown in Figure 41 (especially for certain flights and under A_1). This effect indeed becomes more relevant under scenario A_1 where the connected components of the propagation tree are significantly larger, as was shown previously. This result further emphasises the difference between the layers and the need to consider joint effects.

⁴⁷ <http://www.eac-airports.com>

3.5 Comparative insights from the analytical methods

In Section 2.2 we presented Figure 42 remarking on the overlap between the metric categories and raising the question of how well non-complexity metrics and methods capture certain features of ATM system dynamics (such as delay propagation) compared with those of complexity science.

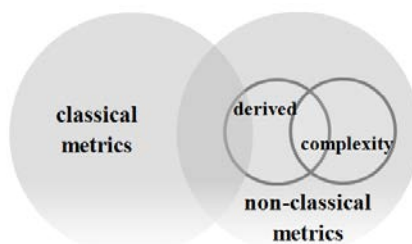


Figure 42. Relationships between metric types (repeat of Figure 2)

Throughout Section 3 we have observed correspondences between the different types of analysis applied to the model data. However, it is difficult to identify one approach on which it would be prudent to wholly rely for extensive insights into the dynamics of delay propagation and prioritisation scenario impacts (nor do we claim, of course, to have concluded such an investigation). One broad conclusion that may be drawn is that the network-level analyses based on graph theory techniques and incorporating causality have produced particularly useful insights at a structural level, allowing us to interpret why and how certain effects (especially the increased reactionary delay under scenario A₁) are manifested. It is unlikely, however, that our attention would have been focused on these characteristics had they not been identified through the classical metric results of Section 3.2. Moreover, the most useful classical metrics have been new and specialist ones, focusing on the passenger, on airline costs and on back-propagation, for example. We conclude that these methods are invaluable in their complementarity to (at least begin to) offer insight into the wider picture of delay propagation and the mitigation thereof.

To some degree, there were also some apparent differences in the extent to which large or small airports were identified as being delay propagators across the analytical methods, as summarised in Table 42. In each case, the input data for the analyses were very similar. The data inputs were all based on arrival delays with the same underlying rules⁴⁸, albeit with some variations in the data structure according to whether a time series was being run or a factor component extraction (as discussed in the corresponding sections).

Despite the selective bias inherent in the filtering process for the propagation node identification in the factor analysis, several airports in the full list of 23 airports (from the model output data) also featured in the Granger causality results, viz.: Barcelona, Athens, Stuttgart and Tenerife Norte. Further analyses would be required before too much interpretative weight should be invested in this correspondence, although the Granger causality methodology is clearly superior in terms of its suitability for the detection of causal relationships. The factor analysis was undertaken to ascertain the extent to which at least one derived (data reduction) technique would compare with a complexity approach.

The Granger causality analysis results are not weighted. This method gives Boolean information in that causality is either established or not. The quantification we expressed was based on the number of causal delay connections at given nodes. It might be expected to find causality between the larger airports, where the delay time series are less noisy and less subject to perturbation from a small number of flights (as may occur for smaller airports). Nevertheless, clear causality relationships have been established for several small airports, in fundamental agreement with the propagation tree analyses.

⁴⁸ Such as recoding early arrivals to zero delay and treating delays from the first minute of delay. See Section 3.1.

Table 42. Comparison of airport type identification across analytical methods

Method	Section	Analytical technique; with critique	Conclusion regarding airport type identification
Simple delay ratios	3.4.1	Analysis of simple departure / arrival delay ratios; based on aggregate delay movements without attempting to establish any flight-flight causalities	Implicates larger airports as delay amplifiers but this may actually more reflect generalised congestion rather than causal links
Factor analysis	3.3.2	Promax (oblique) rotation with identification of factors loading across the (unique) components; gives weighted outputs but the very flexibility of the methods available can sometimes be a disadvantage due to the associated judgement bias	Identifies both smaller and larger airports as propagators of delay, although the processing of components for propagation node identification can be selectively biased towards the detection of larger or smaller airports
Propagation tree	3.4.2	Determination of longest branches through recursive deep-searches, linking flights through either aircraft rotational or passenger connectivity dependencies; only the most restrictive connection is considered (as opposed to a statistically apportioned causality)	Characterises higher volumes of delay at larger airports and identifies certain smaller airports acting in particular as delay amplifiers
Granger causality	3.4.3	Employs a true time series analysis reducing the effect of confounding factors such as high correlations and in which bidirectional links are forbidden; Granger causality analysis is not weighted – it gives Boolean information	Implicates both smaller and larger airports as delay amplifiers but with greater emphasis on the former

Overall, it seems clear that smaller airports are significantly implicated in the propagation of delay through the networks at a level which has hitherto not been commonly recognised. This is probably due to reduced delay recovery potential at such airports (e.g. through flexible or expedited turnarounds, and spare crew and aircraft resources – although the latter are not explicitly modelled).

It is also worthwhile reminding ourselves that the model represents a normative day and that the simulation results thus reflect schedule robustness (e.g. whether a given airport has sufficient connectivity and capacity to reaccommodate disrupted passengers). The precise dynamics and functionalities of the comparative roles in delay propagation of smaller and larger airports would merit further research.

This should be set in the context of back-propagation, which is an important characteristic of the persistence of delay propagation in the network. Paris Charles de Gaulle, Madrid Barajas, Frankfurt, London Heathrow, Zürich and Munich all demonstrated more than one hundred hours of back-propagated delay during the modelled (baseline) day. Most of the delay characterised in the network was indeed distributed between a relatively limited number of airports.

4 Beyond POEM – a look ahead

Although the final project deliverables (Final Technical Report – POEM Deliverable 6.2; Final Strategic Report – POEM Deliverable 6.3) have been completed, continued exploitation will be made of the POEM model along with dissemination of the results.

4.1 Dissemination

For the final deliverables, the following primary dissemination activities are planned:

- to participants attending the POEM workshop (see Section 2.2 and POEM Deliverable 6.1) in London in January 2012 (comprising mostly ANSPs and airlines);
- to members of the University of Westminster's international airline Working Group on delay cost management (with ANSP and SESAR JU members);
- to the SESAR WP-E ComplexWorld community (through its wiki: <http://complexworld.eu/wiki>) and directly with other WP-E researchers engaged in associated research;
- through SESAR WP-E conferences (e.g. a paper to be submitted⁴⁹ to the Third SESAR Innovation Days, Stockholm, 26 – 28 November 2013) and non-SESAR events (e.g. the Second INFORM Airline Forum⁵⁰, London Heathrow, 18 – 20 September 2013);
- through coordination with EU framework projects engaged in associated research (e.g. MetaCDM, FP7);
- through submission of a further journal paper – i.e. in addition to Cook *et al.* (2013); this new paper is expected to cover a specific discussion of the POEM results set in the context of the (limited) published work in this field, with particular comparisons of the cost metrics (including hard and soft cost ratios) and of the flight- and passenger-centric metrics⁵¹.

4.2 Further research

The model is already uniquely able to simulate a normative day with taxi-out and taxi-in functionalities included, and with full passenger itineraries. Furthermore, the functionality of the POEM tool is such that a wide range of additional (flight or passenger prioritisation) scenarios and/or new rules could be implemented. For example, these could be with respect to:

- future (e.g. STATFOR) traffic forecasts;
- changes to EU regulations regarding passenger compensation and care;
- resilience to factors ranging from small changes in delays, cancellations, or load factors through to major network perturbations (such as ash clouds).

These can all be assessed in terms of the flight-centric and passenger-centric performance metrics already incorporated into the model, or new metrics may be built into the platform – these may be aligned specifically with SESAR KPAs and/or SES Performance Scheme objectives.

Future enhancements to the model may include cost recoveries, such as delay savings on crew hours being corrected back to the causal delay or (similar) savings associated with flight cancellation. ANSP cost trade-off tools may also be incorporated.

As concluded in the analyses, the precise dynamics and functionalities of the comparative roles in delay propagation of smaller and larger airports would merit further research. Further work exploring the role of particular (communities of) nodes in network vulnerability is also anticipated to be of value.

⁴⁹ Planned paper and presentation are subject to acceptance.

⁵⁰ POEM presentation slot confirmed.

⁵¹ E.g. the arrival-delayed passenger / arrival-delayed flight ratio: an important comparator metric.

5 References

- Aaker D. A., Kumar V. and Day G.S., 2001. Marketing Research 7th. Ed., John Wiley & Sons.
- Abrantes P.A.L. and Wardman M.R., 2011. Meta-analysis of UK values of travel time: an update, Transportation Research Part A, 45, 1–17.
- Association of European Airlines, 2013. Monthly Monitor – AEA Monthly Traffic Update (historical data available in current file), accessed July 2013.
- Association of European Airlines, 2013. Weekly Monitor – AEA Weekly Traffic Update (historical data available in current file), accessed July 2013.
- Ball M., Barnhart C., Dresner M., Hansen M., Neels K., Odoni A., Peterson E., Sherry L., Trani A., Zou B., 2010. Total delay impact study: A comprehensive assessment of the costs and impacts of flight delay in the United States, Final Report, NEXTOR (October 2010).
- Brovelli A., Ding M., Ledberg A., Chen Y., Nakamura R. and Bressler, S. L., 2004. Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. Proceedings of the National Academy of Science USA, 101 (26), 9849–9854.
- Central Office for Delay Analysis, 2010. CODA Digest: Delays to air transport in Europe, September 2010.
- Cook A. and Tanner G., 2011. European airline delay cost reference values. Commissioned by EUROCONTROL Performance Review Unit, Brussels.
- Cook A., Tanner G., Jovanović R. and Lawes A., 2009. The cost of delay to air transport in Europe – quantification and management, 13th Air Transport Research Society (ATRS) World Conference, Abu Dhabi.
- Cook A., Tanner G. and Lawes A., 2012. The hidden cost of airline unpunctuality, Journal of Transport Economics and Policy, 46 (2), 157-173. ISSN 0022-5258.
- Cook A., Tanner G. and Zanin M., 2013. Towards superior air transport performance metrics – imperatives and methods, Journal of Aerospace Operations, DOI 10.3233/AOP-130032 (in press).
- Department for Transport (UK), 2012. Transport analysis guidance, Unit 3.5.6: Values of time and vehicle operating costs, Transport appraisal and strategic modelling division, Department for Transport (UK), August 2012.
- EUROCONTROL, 2011a. Performance Review Report 2010: an assessment of air traffic management in Europe during the calendar year 2010, EUROCONTROL Performance Review Commission, Brussels.
- EUROCONTROL, 2011b. Network Operations Report for 2010: monitoring & reporting, March 2011.
- EUROCONTROL, 2012. Standard inputs for EUROCONTROL cost benefit analyses (Ed. 5.0), EUROCONTROL Headquarters, Brussels.
- EUROCONTROL, 2013a. Performance Review Report 2012: an assessment of air traffic management in Europe during the calendar year 2012, EUROCONTROL Performance Review Commission, Brussels.
- EUROCONTROL, 2013b. Network Operations Report – Annual 2012, March 2013.
- European Commission, 2004. Regulation (EC) No 261/2004 of the European Parliament and of the Council of 11 February 2004 establishing common rules on compensation and assistance to passengers in the event of denied boarding and of cancellation or long delay of flights, and repealing Regulation (EEC) No 295/91, Official Journal L 046 , 17 February 2004, 1 – 8.

European Commission, 2011a. Flightpath 2050 – Europe's Vision for Aviation (Report of the High Level Group on Aviation Research), ISBN 978-92-79-19724-6, DOI 10.2777/50266, 2011.

European Commission, 2011b. White Paper: Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system, Brussels, 2011.

European Commission, 2011c. Possible revision of Regulation (EC) 261/2004 on denied boarding, long delays and cancellations of flights, Roadmap Version 1, November 2011.

European Commission, 2012. Public consultation on the possible revision of Regulation 261/2004 – results, report prepared by Steer Davies Gleave.

European Commission, 2013. Air passenger rights revision (memo), Brussels, 13 March 2013.

Granger C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37 (3), 424–438.

Granger C.W.J., 1980. Testing for causality: a personal viewpoint. *Journal of Economic Dynamics and Control*, 2, 329–352.

Hess S. and Polak J.W., 2006. Exploring the potential for cross-nesting structures in airport-choice analysis: A case-study of the Greater London area, *Transportation Research Part E* 42, 63–81.

Hoover K.D., 2001. *Causality in macroeconomics*. Cambridge University Press.

Humphries M. D. and Gurney K., 2008. Network 'small-world-ness': a quantitative method for determining canonical network equivalence. *PLoS One* (3), e0002051.

International Air Transport Association, 2002. *Airline Guide to Involuntary Rerouting* (1st Edition, Effective 1 September 2002), Montreal & Geneva.

International Air Transport Association, 2008. *Prorate Manual Passenger*, Part 3, *Prorate Agreements: Section B: Multilateral Proration Agreement* (Effective 01 September 2008), Prorate Agency – Geneva.

International Civil Aviation Organization, 2012. Doc 8643 – Aircraft Type Designators, Edition No. 40, April 2012.

Jetzki M., 2009. The propagation of air transport delays in Europe, Doctoral thesis, Department of Airport and Air Transportation Research, RWTH Aachen University, Germany.

Kamiński M., Ding M., Truccolo W.A. and Bressler S. L., 2000. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological Cybernetics* 85 (2), 145–157.

Kano N., Seraku N., Takahashi F. and Tsuji S., 1984. Attractive quality and must-be quality, *Hinshitsu*, 14(2), 39–48.

Kim J.-O. and Mueller C.W., 1978a. *Factor analysis: what it is and how to do it*, SAGE University Paper series on quantitative applications in the social sciences: 07-013. SAGE Publications.

Kim J.-O. and Mueller C.W., 1978b. *Factor analysis: statistical methods and practical issues*, SAGE University Paper series on quantitative applications in the social sciences: 07-014. SAGE Publications.

Latora V. and Marchiori M., 2001. Efficient behavior of small-world networks. *Physical Review Letters* 87 (19), 198701.

Newman M.E.J., 2002. Assortative mixing in networks. *Physical Review Letters*, 89(20), 208701.

Pels E., Nijkamp P. and Rietveld P., 2003. Access to and competition between airports: a case study for the San Francisco Bay area, *Transportation Research Part A*, 37, 71–83.

Pourtaklo N.V. and Ball M., 2009. Equitable allocation of enroute airspace resources, Eighth USA/Europe air traffic management research and development seminar, Napa, CA.

Pyrgiotis N., 2011. A public policy model of delays in a large network of major airports, Transportation Research Board 90th Annual Meeting, Washington DC.

Roebroek A., Formisano E. and Goebel R., 2005. Mapping directed influence over the brain using Granger causality and fMRI. *NeuroImage* 25 (1), 230–242.

SESAR, 2012. European ATM Master Plan (Ed. 2), The Roadmap for Sustainable Air Traffic Management, October 2012.

Stopher P. and Meyburg A., 1979. Survey sampling and multivariate analysis for social scientists and engineers. Lexington Books.

Tull D.S. and Hawkins D.I., 1993. Marketing research: measurement and method, 6th ed. Macmillan Publishing.

Wang B., Tang H., Guo C. and Xiu Z., 2006. Entropy optimization of scale-free networks robustness to random failures. *Physica A*, 363, 591–596.

Zanin M. and Lillo F., 2013. Modelling the air transport [sic.] with complex networks: a short review, *European Physical Journal Special Topics*, 215 (1), 5 –21.

Zanin M. and Papo D., 2013. Efficient neural codes can lead to spurious synchronization. *Frontiers in Computational Neuroscience* (In press).

Zhang Y. and Hansen M., 2009. Regional GDP — extending ground delay programs to regional airports, Eighth USA/Europe Air Traffic Management Research and Development Seminar, Napa, CA.

Appendix A Key model rules and variables

This appendix summarises the major ('base') rules of the model and is best read in conjunction with Section 2.4.1. The variables referred to are summarised in part (ii), below.

(i) Specification of major base rules

Base Rule 10 (wait for crew & turnaround)

External data sources used:	POEM Deliverable 1.2
Internal rules/variables used:	Variable 51; Rule 61. (Coordinated with Rule 31.)

The turnaround process includes all those actions taking place from on-blocks to off-blocks. Normally, it includes deboarding passengers, unloading cargo, catering, cleaning, loading potable water, changing waste water, refuelling, standard and non-standard maintenance, boarding passengers and loading cargo, among other actions taking place such as changing crews. While some of these processes can happen in parallel, others need to be sequenced.

Usually, the turnaround includes buffer time in the planning phase that can be absorbed in the case of delay, i.e. [scheduled turnaround time] = [minimum turnaround time] + [buffer time]. The buffer time may be designed into the schedule as recovery buffer (schedule padding), or imposed on the airline due to airport slot availability or other scheduling constraints.

We cannot tell *a priori* from the data which factors are in force for any given turnaround time; however, by analysing the data we can classify the operations and observe the distribution of turnaround times. The duration of the turnaround process is modelled on three dependent factors:

- aircraft operator type (Rule 61);
- airport size – with airports classified into three types depending on their number of passengers: 'large' / 'medium' / 'small'. Large includes the top 14 airports by passengers in Appendix B of POEM Deliverable 1.2; the cut-off between medium and small airports is 10 million pax/year;
- wake turbulence category of the aircraft ((H/M/L/J)) (Variable 51).

Analyses were undertaken on cleaned PRISME traffic data (e.g. with cargo flights removed) to determine the overall minimum turnaround times (MTTs) observed in practice. The MTT was approximated as the 2nd percentile of one month (September 2010) of historical turnaround times (provided the sample was large enough to estimate this value within 95% confidence⁵²). This is the MTT that leaves 2% of the sample under that time, as shown in the figure below. By calculating the MTTs we are able to determine if the delay of a flight can (probably) be absorbed by buffer in the turnaround time, in association with Rule 31.

⁵² The p-value of the likelihood of being Normally distributed with an empirical mean and standard deviation, using a simple Kolmogorov-Smirnov test, was established. Assuming the data to be Normally distributed, a confidence interval for proportions was calculated to determine the minimum sample size necessary to produce an error < 2% with a confidence > 95%. If the sample size was not large enough, one or more of the grouping categories was dropped (in the order that the three factors are presented above).

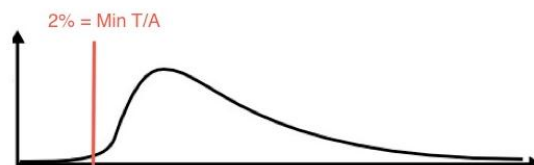


Figure 43. Minimum turnaround time cut-off threshold

The following table and empirical figures show the final values used in the model, with an approximation of the MTT distributions in each category whenever the sample was sufficiently large. Note that applied MTTs are coordinated with Rule 31 on crew densities. Currently, no constraints are placed on the number of MTTs applied in a given model run.

Table 43. Minimum turnaround times by aircraft type, AO type and airport size

Aircraft WTC	AO type	MTTs (minutes and seconds) by airport size		
		large	medium	small
heavy	all	45'1"	48'36"	47'59"
medium	regional	24'2"	15'35"	15'13"
medium	charter	28'38"	21'57"	19'58"
medium	LCC	20'47"	20'11"	19'42"
medium	full-service	29'9"	23'57"	17'57"
light*	all	6'17" (all airports)		

* The MTOW threshold for this category is 7 tonnes
Only 0.02% of passengers with valid metrics are on such aircraft in the POEM model.

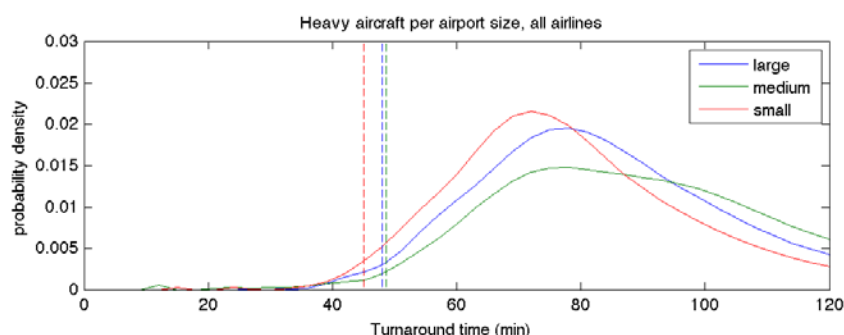


Figure 44. Minimum turnaround times – heavy aircraft, all airlines

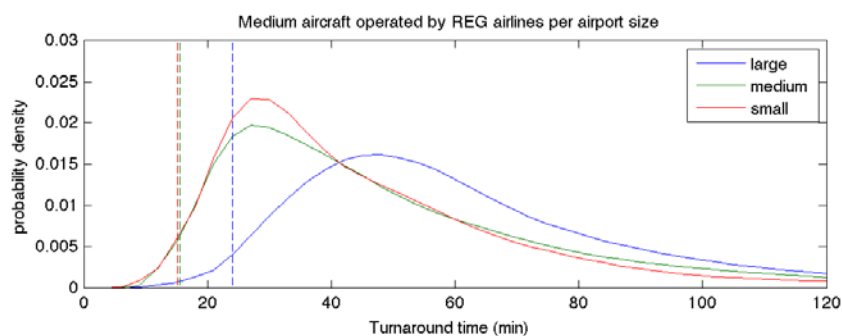


Figure 45. Minimum turnaround times – medium aircraft, regional airlines

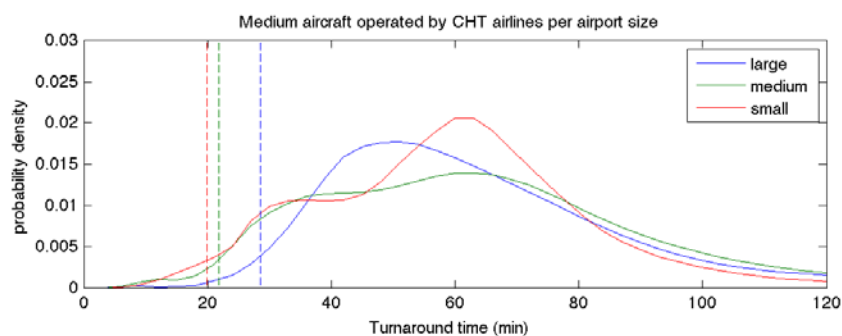


Figure 46. Minimum turnaround times – medium aircraft, charter airlines

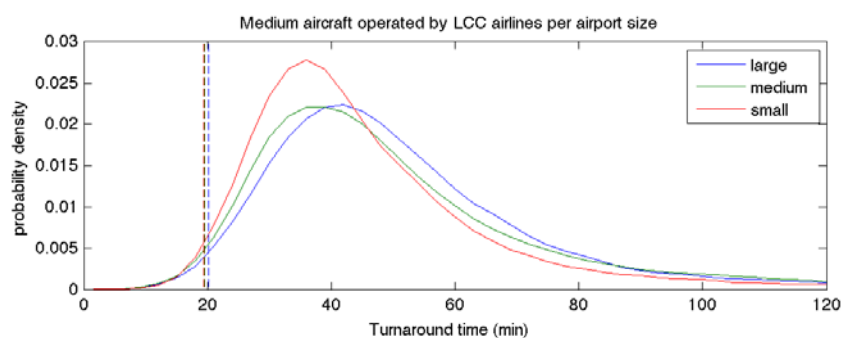


Figure 47. Minimum turnaround times – medium aircraft, LLCs

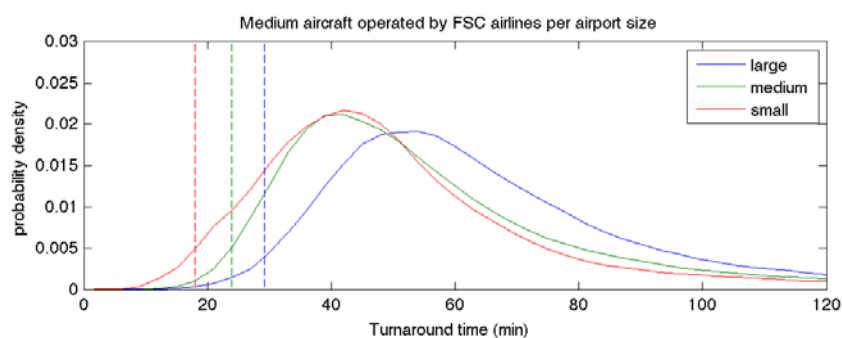


Figure 48. Minimum turnaround times – medium aircraft, full-service airlines

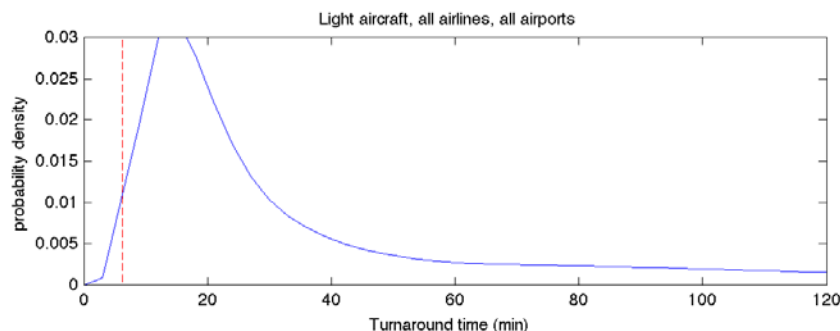


Figure 49. Minimum turnaround times – light aircraft, all airlines

Base Rule 13 (wait for boarding)

External data sources used:	N/A
Internal rules/variables used:	Variables 5, 34, 104; Rules 33, 35, 36, 52, 53, 118

Each flight has awareness of all of its inbound connecting passengers, through Variable 34. We consider that all connecting passengers, i , arrive at the gate at:

$$passenger\ arrival_i^{outbound\ flight} = MIBT_i^{inbound\ flight} + MCT_i^{inbound, outbound\ flights} + u$$

where the modelled (actual) in-block time (MIBT) of the previous flight is given by Variable 104, the MCT is determined by Rule 35, and u is a passenger-independent stochastic component (modelled $\sim N(0, s)$, where s is 5% of the MCT), such that a small number of passengers miss connections even though they are within the MCT (i.e. they become no-shows). However, all passengers making one-leg journeys only are (in the current model) assumed to board their flights. The POEM model allows aircraft to wait for inbound connecting passengers. The time that each aircraft waits for connecting passengers varies according to the conditions specified below:

Table 44. Definition of passenger waiting cases

Case	Condition	Decision
X (‘inflexible’ slots assumed)	MOBT – SOBT* \geq 60 mins (ATFM delay \geq 60 mins)	Do not wait for any pax
Y (‘flexible’ slots assumed)	MOBT – SOBT < 60 mins (ATFM delay < 60 mins, or null)	Wait up to 60 minutes for: (i) any flexibly ticketed [†] connecting pax (ii) any onward long-haul [‡] passenger (iii) until the flight has 80% of expected pax

* MOBT: see Rule 118; SOBT: see Variable 5; [†] See Rule 36; [‡] Beyond ECAC.

Scenario rules

N_1	N/A
N_2	N/A
A_1	Case X: by running Rule 33 passively, compare the costs of (1) departing now and (2) departing another 60 mins later. Wait for 60 minutes if it is cheaper. Case Y: by running Rule 33 passively, calculate costs in increments of 15 mins of delay in addition to current departure time. Take the minimum cost alternative. NB. Delay costs are also determined by rules 52 and 53. See also Section 2.4.1.
A_2	= A_1
P_1	N/A
P_2	N/A

Base Rule 16 (network capacity restrictions)

External data sources used:	N/A
Internal rules/variables used:	Rules 23, 25 and 118.

Network capacity refers to the net total capacity determined by the airport capacities, translated back into a simplified (non-sectorised) en-route environment. There are two kinds of capacity restrictions in the POEM model:

- *Restrictions due to physical limitations of the system under normal operational conditions.*
Capacity at airports is determined by the runway maximum throughput and is estimated from the traffic data through Rule 25. Due to the highly resource-consuming nature of modelling the trajectories of each flight, the airspace capacity is not modelled explicitly, but rather incorporated into the stochastic component of the route length: see Rule 23.
- *Further constraints imposed under the modelled high delay day.*
The increased target departure delay described in Section 2.4.2 is achieved in combination with a stochastic element incorporated into Rule 118. The joint effect of the reduced network capacity and the independently increased at-gate delay can be thought of as a pseudo-ATFM delay, since its delay effect is manifested at-gate.

Base Rule 19 (taxi-out)

External data sources used:	CODA
Internal rules/variables used:	(See also Appendix B.)

Taxi-out times were applied by using means and standard deviations of actual times supplied by CODA for each airport for September 2010, using random sampling in the model from a log-Normal distribution. The 'actual' times in the PRISME data were the same as 'filed' times for non-A-CDM airports (i.e. the vast majority; see also POEM Deliverable 5.2). See also Rule 20.

Base Rule 20 (taxi-out delay)

External data sources used:	N/A
Internal rules/variables used:	Rule 19. (See also Appendix B.)

Since planned taxi times are not available, it is not possible to determine true taxi-out delay. A metric used internally in the model is the average deviation from the mean of the input distribution used in Rule 19, which has a value of zero deviation.

Base Rule 23 (en-route variability, including recovery)

External data sources used:	Case Study 2 (POEM Deliverable 7.2)
Internal rules/variables used:	N/A (Although see also Rules 25, 53 and 118)

Due to the highly resource-consuming nature of modelling the trajectories of each flight, the airspace capacity is not modelled explicitly⁵³, but rather incorporated (in part) into the stochastic component of the route length. This draws on route-based internal analyses of cleaned PRISME data, e.g. to model wind and aircraft performance variabilities.

A delayed flight in the model will try to recover up to a *maximum* of 6% of its *filed* en-route duration (approximately up to the IAF), depending on other simple capacity and stochastic route restrictions, leaving a non-recovered residual of 5 minutes arrival delay in each case (the recovery of which does not normally make economic sense). This setting was established based on feedback from Case Study 2 and in the context of calibrating aggregate arrival and departure delays. Delayed flights are thus more likely to recover time, although non-delayed flights might also do so due to the stochastic en-route component. More advanced delay recovery rules (including an assessment of ATC impacts⁵⁴ and delay recovery costs⁵⁵) may be incorporated into future modelling.

Arrival delay (holding) due to capacity exceedance is actually modelled as part of the total en-route length, using dedicated costs presented in Rule 53.

Base Rule 25 (airport capacity constraints)

External data sources used:	N/A
Internal rules/variables used:	Variables 21 and 27.

Rule 25 establishes both departures and arrivals under normal conditions. The POEM model defines the (hourly) capacity at an airport as the maximum number of possible movements (take-offs and landings) that the airport is able to handle. In addition, the minimum operation time is the inverse of the capacity multiplied by 60 minutes; this measures the average runway occupation times as if the airport had one single runway. To determine the capacity, the time period of the traffic sample is divided into time blocks of 60 minutes length each, overlapping by 45 minutes. For instance, if the

⁵³ See also Rules 25 and 118.

⁵⁴ Such as aircraft bunching; advanced modelling of en-route recovery is being investigated in the WP-E project CASSIOPEIA.

⁵⁵ Not currently modelled; see Rule 53.

traffic sample started at 0400UTC, the time periods would be 0400-0459, 0415-0514, 0430-0529, etc. For each of these time periods, using variables 21 and 27, the number of total movements is calculated.

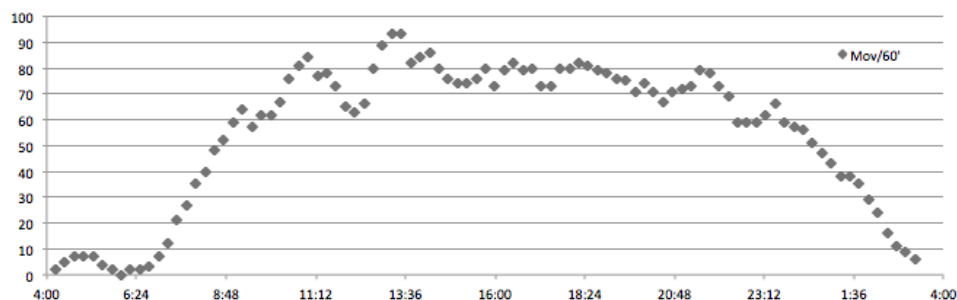


Figure 50. Illustrative movements by time block

Taking the maximum of these values may seem a reasonable choice. However, due to the model's sensitivity to the values chosen, the maximum may not be the best choice. Instead, just over the 98th maximum percentile is taken. That is, the capacity is defined as the minimum integer, which is greater than or equal to the 99%-value of the 60-minute samples taken in overlapping steps of 15 minutes. The transformation is shown in the figure below.

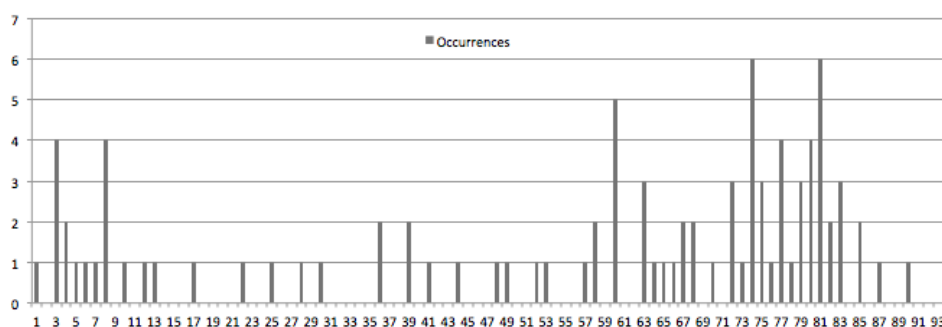


Figure 51. Integral transformation of movements

To avoid assigning an extremely low capacity at some airports due to the fact that they were not operating at 100% capacity at any time during the traffic sample period, a fixed capacity of 45 movements per hour will be assigned as the minimum possible capacity. Arrival capacity is defined as a parameter of the system and does not change as a function of the scenario. It is the way in which this capacity is shared among the users that each scenario will influence.

Base Rule 26 (airborne arrival management)

External data sources used:	N/A
Internal rules/variables used:	Variable 34; Rules 23, 25, 33, 35, 52 and 53

Each airborne aircraft has an Estimated Time Over the IAF. When the aircraft reaches the IAF, it asks for an arrival slot. The airport's arrival manager spaces aircraft at the IAF of the corresponding approach depending on their estimates, and replies with the best available slot after this estimate, provided the aircraft reaches the IAF within a time window [-5, +10] mins.

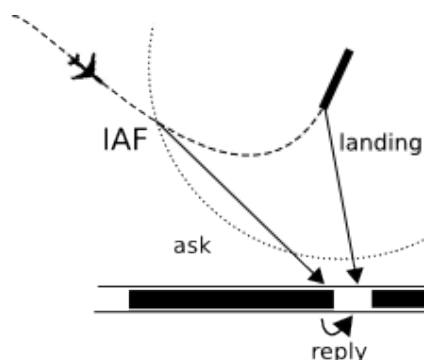


Figure 52. Sequencing over the IAF

Spacing between aircraft is determined by the airport capacity (see Rule 25). Note that in the base rule the sequence is generated following a first-come, first-served basis, such that there is no prioritisation in operation. For holding, see under Rule 23.

Scenario rules

N_1	Respecting the $[-5,+10]$ IAF arrival window estimates, flights are prioritised based on simple passenger numbers: i.e. (passengers * delay minutes) is minimised in the delay sequence.
N_2	Flights arriving more than 15 minutes late are prioritised based on the number of flights that would be delayed by inbound connecting passengers; N_1 resolves any ties.
A_1	N/A
A_2	A_1 is implemented in Rule 13 and flights are (independently) arrival-managed based on delay cost. See also Section 2.4.1.
P_1	N/A
P_2	N/A

Base Rule 29 (taxi-in delay)

External data sources used:	CODA
Internal rules/variables used:	(See also Appendix B.)

Taxi-in times were absent in the PRISME data. These were therefore generated by using actual means and standard deviations of times supplied by CODA for each airport for September 2010, using random sampling in the model from a log-Normal distribution. Since planned taxi times are not available, it is not possible to calibrate taxi-in delay. A metric used internally in the model is the average deviation from the mean of the input distribution, which has a value of zero deviation.

Base Rule 31 (crew hours and aircraft pairings)

External data sources used:	N/A
Internal rules/variables used:	Rule 10, Rule 36

In Rule 10, we described the distribution of aircraft turnaround times used in the model. These distributions are derived for each suitable combination of airport, carrier and aircraft type. Rule 31 is implemented in coordination with Rule 10.

For Rule 31, the extent to which each turnaround time, t_b , can be reduced to a minimum possible turnaround time, $\min(t_b)$, by removing slack time, is determined by the density of crew operations for that airport and airline combination, and subject to the constraints of Rule 10.

For all ECAC airports in the model, a (sorted) list is made based on the following value of crew density, c_d , derived for each airport-airline combination:

$$c_d = \frac{\sum_i (\text{no. of seats on flight})_i}{(\text{time of last flight} - \text{time of first flight})}$$

Since the number of crew is related to the number of passengers onboard a flight, i , the value of c_d gives a value approximately in proportion to the number of crew the airline has operating at that airport per hour. These values are then normalised, c'_d , between 1 and 0 (for the lowest densities), across *all* the airports. Taking a Normal approximation for the distributions shown in Rule 10, this value is then used to determine how many turnaround time standard deviations, s_t , may be subtracted from the mean turnaround time (mean and standard deviation for that *specific* airport-airline-aircraft-type⁵⁶ combination), to give a minimum turnaround time, according to:

$$\min(t_t) = \bar{t}_t - (2c'_d + 1)s_t$$

Thus, for a crew density of one, three standard deviations may be subtracted from the mean turnaround time, for use as a minimum turnaround time. For a crew density of zero, only one standard deviation may be subtracted from the mean, for use as a minimum turnaround time.

Base Rule 33 (passenger care and reaccommodation)

External data sources used:	Regulation (EC) No 261/2004 (European Commission, 2004); IATA Multilateral Proration Agreement (IATA, 2008)
Internal rules/variables used:	Rule 13, Rule 61, Rule 118

This rule was developed and imported from Case Study 2 (POEM Deliverable 7.2) and based in part on Regulation (EC) No 261/2004 (see Section 1.3 for current proposals regarding updates to this Regulation). Passengers are reaccommodated according to the base rule described here, and, alternatively, under the prioritisation scenario rules. The reader is reminded that '(prioritisation) scenario' refers to the various flight priority rules applied to aircraft in the model, to compare the comparative effectiveness of such rules in reducing the costs of delay to the airline, for example. 'Cost scenario' refers to the different assumptions made about the underlying costs of delay to an airline, under either 'low', 'base' or 'high' cost base assumptions.

⁵⁶ Wake turbulence categories (see Variable 51) are used to simplify the categories and ensure suitable sample sizes for each airport-airline combination.

1. Principles and triggering events

The passenger reaccommodation rules are run when one of the three trigger events shown below occurs. (See Section 2.4.1 for further information on the implementation of the cost rules.) These are designed to reasonably reproduce current airline behaviour.

Table 45. Passenger reaccommodation - trigger and control events

Trigger event	Controlled by	When reaccommodation rule is activated	Note
Passenger misses a connection due to the late arrival of an inbound flight	The estimated departure time of the connecting flight (Rule 118; which includes waiting for crew and turnaround through Rule 10) and decisions whether to wait for the connecting passengers (Rule 13; includes MCTs through Rule 35)	When delayed inbound flight is on-blocks (in practice, the airline may do this earlier or later)	The majority of passenger reaccommodations in the model will be driven by missed connections
Passenger arrived too late at the gate and the aircraft left without them	Rule 13 (a randomised event that some passengers no-show at the gate, together with wait rules driven by: pseudo-ATFM slots, aircraft load, length of haul, and passenger ticketing/status)	When flight departs without passenger	-
Flight cancelled	High cancellation day implementation rules	When cancelled	See Section 2.4.2
Passenger denied boarding	Not currently covered in model	-	-

Under some of the prioritisation scenarios, the reaccommodation rules are run **passively** (i.e. without passengers actually being reaccommodated) by another rule, to estimate in advance the cost of particular decisions (see also Section 2.4.1). An example is if an airline is attempting to assess the cost of a slot delay for flights with connecting passengers on-board. This is intended to reflect possible future airline behaviour, as these advance cost estimations are not currently (common) practice. The operational day is deemed to end at 0359 (local time) the following day, as a cut-off time for passenger reaccommodation.

2. Seating / routing allocation

The following sub-rules apply to the passenger reaccommodations:

- (a) passengers are reaccommodated only onto flights where seats are available;
- (b) passengers are reaccommodated according to the minimum number of changes involved in reaching their destination (e.g. they are booked onto direct flights, if possible, before indirect options are considered);
- (c) no originally-booked passenger may be 'bumped' (replaced) by an attempted reaccommodation (although this may occur in practice);
- (d) once a passenger has been reaccommodated, he/she may not be subsequently 'bumped' by a 'higher priority' passenger (i.e. such priorities as determined through sub-rule 3).

Note 1. As a simplifying assumption, the model assumes that all passengers prefer to be rebooked onwards to their final destination, instead of being returned to their point of origin. As shown in Table 47, passengers are entitled to be returned to their point of origin and to receive a ticket reimbursement once their delay exceeds five hours. In practice, some passengers, those travelling on business, for example, may return to their point of origin before incurring this amount of delay, but we do not model this.

Note 2. Passengers are reaccommodated onto aircraft regardless of ticket and cabin class, in that aircraft are simply occupied according to total seat space, following the demands of the sub-rules of Rule 33. No account is taken of the costs associated with upgrading or downgrading.

3. Passenger prioritisation

In order to reflect most closely actual practice, the following prioritisation order would be applied – (a) being the highest priority, (e) the lowest – according to ticket and carrier type:

- (a) passengers with flexible tickets (applies to full-service airline pax only);
- (b) business-class passengers with inflexible tickets (applies to full-service airline pax only);
- (c) other full-service airline passengers;
- (d) regional airline passengers;
- (e) all other passengers (LCC and charter carriers).

However, since the model does not currently differentiate by cabin class, a simplification will be made in this regard, which we will define under sub-rule 4.

Note 1. Two categories of ticket are assigned to passengers flying with full-service carriers (as described in Rule 36):

- **flexible** (first class, plus highest fares of all other fare classes);
- **inflexible** (all others, thus including discounted business).

Note 2. These prioritisation rules are a practical way of allocating at least some form of priority assignments to passengers requiring reaccommodation when whole groups of passengers miss a connection. Clearly, they often cannot all be reaccommodated onto the best alternative flight.

Some passengers will have a degree of autonomy. Some LCCs offer a 'flexible' fare that typically offers the ability to change flights without penalty but *only* onto their own services (i.e. not interlineable) and with no refunds if the trip is cancelled. Passengers on non-LCC flights with fully flexible fares could re-book themselves within an alliance scheme⁵⁷ and those will fully refundable fares could cancel their ticket and independently re-book themselves onto another carrier (subject to seat availability).

Under disruption, however, operations are different. There is more usually a generic reaccommodation policy, such that (a) – (c) (or (d)) *may* be rebooked without priority, or first consideration is given to factors such as: passenger preference and destination; passenger status (such as high-tier frequent flyer programme membership); unaccompanied minors; passengers with reduced mobility (and companions); and, likelihood of baggage being successfully transferred to a new flight (possibly in another terminal).

Under disruption, most non-LCCs have bilateral passenger carriage agreements that do not discriminate by fare type⁵⁸, although there may be one or more non-LCCs with whom a given non-LCC does not have an agreement. There are not usually pre-existing agreements between LCCs and other carriers.

The rules applied in POEM favour those passengers who have paid the most, as a reasonable mechanism for reaccommodating some passengers first. Although somewhat too generic in this sense, they are basically consistent with airline feedback received.

Note 3. Carrier types (full-service, regional, LCC and charter) are allocated by Rule 61.

Note 4. Sub-rule 4 takes precedence, if a conflict arises with sub-rule 3.

⁵⁷ Endorsable fares, i.e. allowing the freedom to transfer onto a competitor's flight, are now rare.

⁵⁸ The International Air Transport Association's 'Airline Guide to Involuntary Rerouting' (IATA, 2002), including Resolution 735d, and other such guidance from IATA, was discussed in some detail in POEM Deliverable 4.2. Most airlines abide by such agreements. Surveys carried out in Germany, Denmark and the UK show that 75% of surveyed passengers facing problems for delays or cancellations were offered re-routing (European Commission, 2013).

4. Hierarchy of interlining

Table 46. Hierarchy of interlining

Carrier type	Ticket type	Rebooking onto next available flight according to departure delay of:		
		up to 2 hours	2 – 5 hours	> 5 hours
full-service	flexible (first/bus.)	any carrier	any carrier	any carrier
full-service	business inflexible	booked/alliance only	any carrier	any carrier
full-service	all other tickets	booked/alliance only	booked/alliance only	any carrier
all other	all tickets	booked carrier only	booked carrier only	booked carrier only

The table shows the best rules for rebooking passengers according to the length of the expected delay and the type of ticket held by the passenger. However, since the model does not currently differentiate by cabin class, the first row is applied to all flexible ticket holders and the third row applies to all inflexible ticket holders. “Booked/alliance only” means that the passenger is only rebooked onto the original carrier, or a co-member of an alliance network (most airlines will try to re-book onto their own flights first). In the POEM model, intra-carrier bookings take precedence over all other rules if any conflict arises. For example, if the sub-rules of Rule 33 dictate that Lufthansa will rebook a passenger onto LH1234, no other airline may have a priority claim to that seat.

On successful reaccommodation, the fare of the remaining legs is transferred to the new carrier, according to principles laid down by IATA (2002)⁵⁹. For further information, see Appendix B of POEM Deliverable 4.2, which discusses IATA interline settlements and involuntary reroutes, proration and endorsement waivers. (The reader is reminded that no compensation for downgrading or cost of upgrading is taken into account in the POEM model). This fare transfer counts as:

- (a) a loss for the transferring carrier (a ‘simple’ hard cost loss);
- (b) a gain for the receiving carrier (a ‘simple’ hard cost gain).

The term ‘simple’ refers to the fact that the value takes no account of the profit ratio. Delay at the final destination is also counted as a passenger value of time cost, see sub-rule (6), below. (See Section 1.3 for current proposals regarding updates to Regulation 261 regarding rerouting.)

5. Ties and unsuccessful reaccommodation

Any ties in passenger reaccommodation are resolved on a first-delayed, first-reaccommodated basis (i.e. passengers with the earlier time of disruption are reaccommodated first) and using a secondary rule of distance to destination (those with the furthest journey are reaccommodated first). Remaining ties are resolved stochastically.

If a passenger has not been reaccommodated by the above process, they are added to a wait list of unaccommodated passengers, one of which is held for each airport. Every hour, on the hour, the reaccommodation rule is re-run for all passengers in the wait list, as new reaccommodation opportunities may arise.

Costs are accumulated for passengers in this list, and assigned to the (original) carrying airline, terminating with an overnight hotel stay if they remain unaccommodated on a flight at the end of the operational day (in practice this decision may be taken much earlier by the airline). The maximum cut-off for this is deemed to be 0359 (local time, next day).

⁵⁹ Only prorated/agreed fares for actually transferred legs should be credited to the new carrier, i.e. this would exclude any subsequent legs where carriage is by the original carrier, although we assume that this applies to all remaining legs, as a simplification that will hold true in most cases anyway.

Disruption costs are also modelled in accordance with Regulation (EC) No 261/2004 as a basic framework (European Commission, 2004). This came into force on 17 February 2005. It relates only to departure delay, denied boarding, and cancellation – the passenger is not granted rights with regard to arrival delay and nothing is currently⁶⁰ due to a passenger as a result of a missed connection *per se*. No additional compensation (i.e. above reimbursement) or rerouting⁶¹ is required by the Regulation in the event of delay. It applies to scheduled⁶² and non-scheduled flights, including those forming part of package tours. The following table shows the rights afforded to passengers when the carrier reasonably expects a flight to be delayed beyond its scheduled time of departure by the delay shown in the first column. The exact terms shown are slightly simplified with respect to Regulation 261⁶³.

Table 47. Regulation 261 rights afforded with respect to delay

Departure delay, t (hours)	Restrictions on applicability (where such apply), in terms of ...		Rights afforded
	...distance, d (km), of flight	... where flight operates	
$t < 2$	-	-	None
$t \geq 2$	$d \leq 1500$	-	
$t \geq 3$	$d > 1500$	intra-Community	(a) meal and refreshments; two telephone calls / telexes / faxes / e-mails
	$1500 < d < 3500$	-	
$t \geq 4$	-	-	
$t \geq 5$	-	-	(b) reimbursement of ticket and return to origin
next day	-	-	(c) hotel accommodation, incl. transportation to/from hotel

The rights specified in the final column, as (a) – (c), are additive, which means, for example, that if a delay of 3 hours caused an unexpected overnight stay to be required, the cost of hotel accommodation would also be afforded to the passenger. The time- and distance-related restrictions remove any obligation, for example, regarding a delay of less than three hours for any long-haul flight. This basic structure is used in the following cost assignment, by carrier type.

⁶⁰ See Section 1.3 for proposals regarding updates to the Regulation regarding this rule. Notwithstanding the current legal situation, the model affords these levels of care to any delayed passengers, more in keeping with understood airline practice.

⁶¹ Although rerouting is commonly effected, according to airlines' internal rules and thus in the POEM model.

⁶² It does not apply to passengers travelling free of charge or at a reduced fare not available directly or indirectly to the public. It does apply to passengers having tickets issued under a frequent flyer programme or other commercial programme by an air carrier or tour operator. None of these are modelled in POEM, however.

⁶³ See Section 1.3 for current proposals regarding updates to the Regulation regarding these rules. Further research would be of value to determine the balance between various factors affecting the application of hard costs. It has been reported (European Commission, 2013) that surveys carried out in Germany, Denmark and the UK show that care was offered in less than 50% of cases. On the other hand, following a judgement by the European Court of Justice (European Commission, 2013), it has now been determined that compensation may be claimed for delays from 3 hours' duration, except in extraordinary circumstances. Where appropriate remedial action is not provided by the airlines, it is likely that the corresponding soft cost of delay will increase with passenger disutility.

Table 48. Costs of provisions made by airlines, by delay duration and cost scenario

Departure delay, t (hours)	Provision (incremental)	Costs (Euro 2010, 2s.f.) by cost scenario		
		High	Base	Low
$1.5 \leq t < 2$	refreshment	2.0	1.7	-
$2 \leq t < 3$	+ tax-free voucher	9.4	7.7	4.7
$3 \leq t < 5$	+ meal voucher, & FFP miles	23	19	12
$t \geq 5$ (no hotel)	+ ticket discount voucher	26	21	13
Overnight	+ hotel accommodation	100	83	51

The costs and categories of provision in Table 48 are slightly adapted and simplified from those reported in Cook *et al.* (2009) for 2005-6, in a study which used (limited) airline data to assess the impact of the introduction of Regulation 261 on the cost of passenger delay to the airline. The costs have been inflated to produce costs in 2010 Euros, with estimates made of the high and low cost scenario values⁶⁴.

As mentioned, the provisions are additive, such that, for example, the meal voucher and frequent-flyer programme (FFP) miles offered at delays exceeding three hours are in addition to the tax-free voucher in the preceding row. Airlines may, of course, apply different rules and make provisions in kind, or otherwise. For example, instead of any of the provisions shown, a passenger may accept a taxi to another airport in the same city to make another connection (although we do not model such level of detail). The actual costs will also depend on the national market in which the airline is purchasing the provision and the level of discount they may negotiate (e.g. with local hotels). There is no other known source of published costs, except that cited in Cook *et al.* (2009). These costs are also applied to business-class passengers, although they may actually: simply spend longer in a business lounge (possibly at lower marginal, additional cost to the airline); impose other, higher hard costs on the airline; and/or, impose higher soft costs (which are explicitly reflected in Rule 52 as a function of fare paid).

Table 49. Final passenger hard cost assignments by carrier type

Departure delay, t (hours)	Full-service	Regional	LCC, charter
$1.5 \leq t < 2$	<i>base, high</i>	0	0
$2 \leq t < 3$	<i>low, high</i>	<i>low, base</i>	<i>low, base</i>
$3 \leq t < 5$	<i>low, high</i>	<i>low, base</i>	<i>low, base</i>
$t \geq 5$ (no hotel)	<i>low, high</i>	<i>low, base</i>	<i>low, base</i>
Overnight	<i>base, high</i>	<i>low, base</i>	<i>low, base</i>

⁶⁴ The values have been inflated according to Euro zone inflation values used in Cook *et al.* (2009) and Cook and Tanner (2011), to produce the base cost scenario values for 2010, and using the ratio of the base/low and base/high cost scenario values reported in the latter reference to produce the corresponding low and high cost scenario values. No low cost scenario value is given for the refreshment, since it would be too low, according to the general method applied, at EUR 1.0.

A cost is determined as shown in Table 49, drawing on a Normal distribution, for the various airline types (as determined by Rule 61). In each case, two cost scenario values are used to determine the mean and standard deviation of the distribution, as shown. The mean of the distribution is the mean of the two cost scenario values. The standard deviation is set at one quarter of the difference between the two cost scenario values, such that, by the known properties of the Normal, 95.5% of the sampled values will lie between the two cost scenario values. The scenarios are chosen judgmentally and aim to at least reflect the different costs expended by full-service carriers, compared with others. In each cell of the table, only the two cost scenario values used are shown, to aid clarity. An explicit example using the base and low cost scenarios (shown as 'low, base') is as follows:

$$\sim N ([base + low]/2, [base - low]/4)$$

For delays ≥ 5 hours, in the absence of supporting data, a judgemental value of 20% of passengers returning to their point of origin is assigned (where applicable), for whom Regulation 261 specifies that a fare reimbursement is also due. This 20% figure is intentionally quite a low proportion bearing in mind that many of the higher priority passengers, who will have particularly needed to continue their journey (such as many of those travelling on premium tickets), will have already been rebooked onwards, and many passengers on return legs of trips will wish to continue home, regardless of the delay.

Passengers who experience a cancelled flight are also entitled to reimbursement and a return to their point of origin (where applicable), or to a rerouting. The same value of 20% is assumed for passengers with cancelled flights requesting a reimbursement. Those who are not already at their origin airport are also returned home, where possible. A number of passengers experiencing a cancelled flight in the morning thus do not start their air transport journey at all, as in practice.

Both types of disrupted passengers may be ultimately obliged to overnight in a hotel. These passengers are assumed to be rebooked either onwards to their original destination, or back to their point of origin, during the course of the following day. This is assumed to be at no additional cost to the carrier, except for those returning to their point of origin. For the latter, rather than assigning a reimbursement cost randomly to a further 20% of overnighing passengers, all passengers who overnight in a hotel have a 20% fare reimbursement cost also allocated.

All reimbursement costs are allocated as a simple hard cost to the original carrier, i.e. taking no account of actual profit ratios. It may be argued that the airline does not actually suffer the full reimbursement as a hard cost, whereas, on the other hand, we do not account for any increased soft cost either. Such cost allocations may be refined in future, if appropriate data are available.

6. Passenger value of time assignment

'Value of time' is a concept widely used in cost-benefit analyses, particularly in transport economics. It is an opportunity cost that corresponds to the monetary value associated with a traveller during a journey. The relationship between the passenger hard cost, soft cost, and value of time is explored in Appendix D. Value of time is separate from the hard and soft costs of delay to the airline, in that it is a cost borne by the passenger.

Relatively few studies have produced values of time for air transport, especially with respect to delay and in a recent European context. We consider there to be insufficient evidence to support a robust waiting time value, such that the POEM model only evaluates value of time as a function of delay at the final (airport) destination. Separate values of time are applied to passengers with flexible tickets, as compared to those with inflexible tickets, as shown in the table below (from Appendix D).

Table 50. Ticket type by value of time applied

Ticket type by value of time	Flexible ticket holders	Inflexible ticket holders
Value of time applied as a function of delay at final (airport) destination	EUR 50/hour	EUR 30/hour
Value of time applied as a function of extra waiting time at airports / hotels	<i>Time counted per passenger, but no cost assigned</i>	

Scenario rules

N₁	N/A (ANSPs not involved)
N₂	N/A (ANSPs not involved)
A₁	N/A (Rule 33 is passive in this respect, in that it does not directly affect slot allocation, it only allocates passengers onto flights with available seats.)
A₂	N/A (Rule 33 is passive in this respect, in that it does not directly affect slot allocation, it only allocates passengers onto flights with available seats.)
P₁	Passengers are reaccommodated based on prioritisation by arrival delay, instead of by ticket type, but preserving interlining hierarchies. Sub-rule 3 is thus ignored, whilst sub-rule 4 is preserved. Passengers are thus reaccommodated in order of delay saving at the final destination, in that if reaccommodation 'A' would reduce a final delay from 3 hours to 1 hour, it is favoured over reaccommodation 'B', reducing a final delay from 4 hours to 2 hours, since 'A' has the smaller residual. This is not a global optimisation, since the priorities are determined locally and sequentially: better solutions arising by jointly considering other airports, or arising later, are not taken into account. Sub-rule 4 is preserved in so far as ticket type still determines which carriers passengers are rebooked onto.
P₂	Passengers are reaccommodated based on prioritisation by arrival delay, regardless of ticket type, and also relaxing all interlining hierarchies. This is thus the same as P ₁ but with the important additional relaxation that sub-rule 4 is now also ignored, such that passengers are reaccommodated regardless of ticket type onto the next available flight operated by any carrier.

Calibration note: The relationship between flexible ticketing and priority reaccommodation has been discussed in detail in the rules above, which include controls on the incurred hard costs through both the sampling from the Normal distributions used and the 'catch-all' provision (for unsuccessful reaccommodation) in sub-rule 5. (Soft costs are addressed by Rule 52.)

Base Rule 35 (airport MCTs)

External data sources used:	IATA (Innovata)
Internal rules/variables used:	These results draw on Case Study 1 (POEM Deliverable 7.1)

The minimum connecting time (MCT) is the shortest time interval required for a passenger/baggage to connect between flights at an airport. Standard MCTs are available for connections between flights that serve: domestic to domestic, domestic to international; international to domestic; and international to international. Standard MCTs and MCT exceptions have been purchased from Innovata who maintain the schedule reference service (SRS) database on behalf of IATA, although the MCT exceptions are not being used. Previously, an average standard MCT was calculated for each of the 199 ECAC airports in scope, ranging from 14 minutes at Oulu Airport to 3 hours 45 minutes at Zvartnots Airport (see Figure 53). This overall average was 45 minutes.

Currently, two MCT values are used by the model, one value for connections onto domestic and non-long-haul international, another for long-haul (flights over 1 500 NM). These two average MCT values are approximately 40 and 75 minutes, respectively.

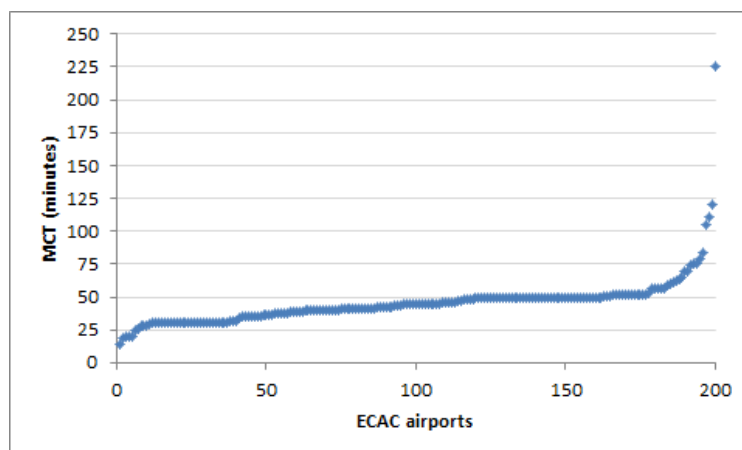


Figure 53. Minimum connecting times (MCTs) for each ECAC airport

Base Rule 36 (passenger assignment, seats & load factors)

External data sources used:	Seats: IATA (Innovata); load factors: AO associations (AEA, BATA, ELFAA, ERA, IATA), ICAO
Internal rules/variables used:	N/A

Aircraft seat numbers and airline load factors determine how many passengers can be reaccommodated from cancelled flights and missed connections by the model. A global aircraft seat database from November 2010 was purchased from Innovata. Although tail numbers were unavailable, total seats were supplied per aircraft per airline fleet, enabling an accurate range of seats to be assigned to most passenger aircraft per airline in the PRISME dataset. Taking the A319 as an example, the maximum total seats available for a selection of airlines include: easyJet: 156 seats; Air Berlin: 150 seats; Adria Airways: 135 seats; Aeroflot: 116 seats (96 economy and 20 business); Iberia: 141 seats (108 economy and 33 business). The model only uses total seats per aircraft, so seat class is unimportant (see Rule 33). For aircraft types not listed in the seat database, a global weighted average (based on other airlines' aircraft seating capacities) was used instead, also per aircraft type.

Generic passenger load factors have been compiled to cover the four aircraft operator types used in the model: full-service, regional, low-cost carrier and charter (refer to Rule 61). Where available from carrier trade associations, summary load factors have been sourced for SEP10 (covering their membership), with additional load factors acquired from individual airlines as required.

Table 51. Generic load factors

AO type	Load factor	Reference values
Charter	90%	Monarch: 89.3%, SEP09; Thomson: 92.8%, SEP09
Full-service	75%	AEA members: 75.3%, SEP10
Low-cost carrier	85%	ELFAA members: 82.3%, JAN10-DEC10; easyJet: 86.2%, SEP09
Regional	70%	ERA members: 69.7%, SEP10

The PaxIS dataset allocates ticket classes according to the following categories:

- first;
- business;
- economy (full);
- economy (discounted);
- 'other'.

For the purposes of the POEM model, two categories of ticket assignment are made:

- **flexible** (all first class, and highest fares of all other classes, combined to form the top decile);
- **inflexible** (all others, thus including discounted business).

Flexible tickets are assigned to first class passengers and others paying the highest fares on a given flight leg such that combined, these two groups make up the top decile of fares. In cases where there are insufficient fare breakdown data (e.g. only one or two ticket categories) on a given leg, no flexible ticket holders are allocated. These assignments are made in the absence of recent (literature) evidence on the actual proportion of flexible fares in the market.

Base Rule 52 (passenger delay soft cost)

External data sources used:	European airline delay cost reference values (Cook and Tanner, 2011)
Internal rules/variables used:	Rule 61

The passenger costs of delay to the airline, central to POEM, manifest themselves as 'hard' and 'soft' costs. 'Hard' costs are due to such factors as passenger rebooking, compensation⁶⁵ and care (see Rule 33). Core cost estimations in the model are with respect to delay costs *to the airline*, since it is these which drive airline behaviour (passenger value of time, on the other hand, is also discussed under Rule 33 and in Appendix D). 'Hard' and 'soft' costs are modelled according to 'high', 'base' and 'low' cost scenarios (see Section 2.3).

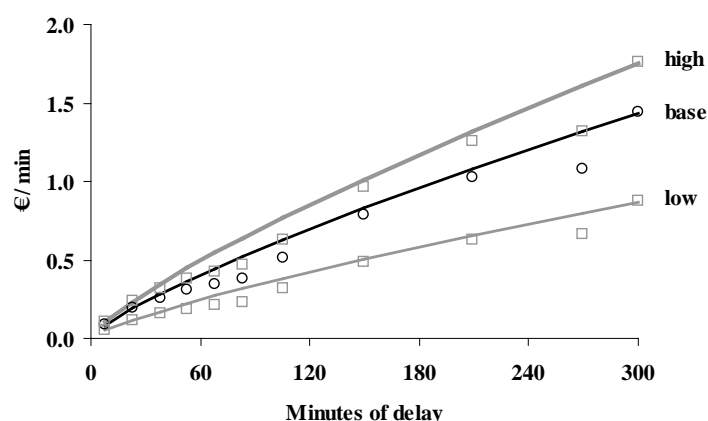


Figure 54. Power curve fit of pax hard costs as a function of delay duration and cost scenario

Source: (Cook and Tanner, 2011)

⁶⁵ Not applicable in POEM.

'Soft' costs manifest themselves in several ways. Due to a delay, a passenger may defect from an unpunctual airline as a result of dissatisfaction (although quite possibly reversing this defection subsequently). Using another airline as an alternative on the day of a booked journey is mostly limited to other members of the same alliance. Endorsable fares, i.e. allowing complete freedom to transfer onto a competitor's flight outside such an alliance, are now rare.

Such a loss may be considered to be largely the gain of another airline, gaining a passenger who has transferred their custom. When such scalable costs (multiplied over a period of time or a network) are assessed, only some *net* loss to the airlines is likely (for example due to trip mode substitution, trip consolidation, trip replacement (such as a teleconference) or trip cancellation). The passenger soft cost of delay thus needs to be appropriately bounded (see later).

For assigning the soft costs of delay, a logit function is used to describe passenger dissatisfaction (δ ; normalised) against various levels of delay. This curve is used to distribute the soft cost as a function of delay duration, and may be thought of as a proxy for the propensity of a passenger to switch from a given airline, to some other choice, after trips with given delay experiences.

$$\delta = \frac{1}{k(1 + e^{a-bt^c})} - k'$$

[5]

This is plotted in Figure 55 (black curve) and has the desirable characteristics of maintaining a low value for some time, then rapidly increasing through a zone of 'intolerance', before levelling off. Quantification of the saturation of delay inconvenience and crossovers in Kano (Kano *et al.*, 1984) customer satisfaction 'requirements' contributed towards this model. Relationships between market share, punctuality and customer satisfaction were also examined.

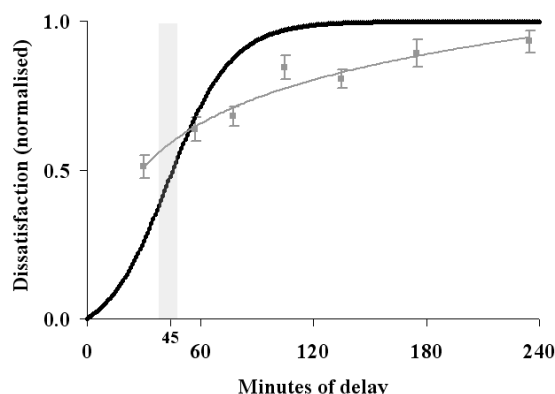


Figure 55. Passenger dissatisfaction as a function of delay duration

Source: Cook and Tanner (2011).

Euro costs (in 2010 Euros) are assigned using δ as a weight, such that when the costs of delay in each delay band of Table 52 are multiplied by the relative proportion of delays in the band⁶⁶, the original, aggregate soft cost value⁶⁷ is obtained, on which the calculations are based. Note that the *per-minute* values are saturated by two hours.

⁶⁶ 2007 delay distributions are used as a base for consistency and comparability across evolving cost models. These delay distributions do not change very much from year to year. The pertinent feature is the weighting of the soft cost distribution appropriately, e.g. by taking into account that most delays occur at relatively low duration (>50% at less than 15 minutes) and that very high delay durations are relatively uncommon.

⁶⁷ EUR 0.183 per passenger, per minute of delay, for the base cost scenario; uses airline data (Cook and Tanner, 2011).

Table 52. Passenger soft costs of delay per minute, by three cost scenarios

Delay (mins)	5	15	30	60	90	120	180	240	300
Low scenario	0.01	0.02	0.07	0.19	0.25	0.27	0.27	0.27	0.27
Base scenario	0.02	0.09	0.25	0.69	0.91	0.96	0.97	0.97	0.97
High scenario	0.03	0.10	0.28	0.77	1.01	1.06	1.08	1.08	1.08

2010 Euros per minute, per passenger.

Note that the values in the table are *per minute* and *per passenger* and that 10% of the values shown in the table are used for scalability reasons⁶⁸. In the POEM model, the soft cost estimate at any given time for a flight, is based on the sum of the individual passenger's delays at their *final destination*. Costs per minute are extrapolated/interpolated (to the nearest minute), from the values in the table, according to the cost scenario being used. The soft cost assignment in POEM is deterministic and based on passenger fares for full-service airline passengers, and stochastic for all other passengers.

For full-service airline passengers: for the final leg⁶⁹ of their journey, all full-service airline leg fares are normalised from 0 (lowest) to 1 (highest), this value being labelled the normalised fare (f_N). This is then used as a weight for the soft cost value in Table 52, with the base cost scenario value at the lower end (i.e. assigned to the lowest weights) and the high cost scenario at the upper end (i.e. assigned to the highest weights). The delay is that (dynamically) estimated at the final destination.

For all other airline types: for each passenger a soft cost is assigned, drawing on a Normal distribution, using the mean of the base and low cost scenario values as the mean of the Normal distribution, with a standard deviation set at one quarter of the difference between the base value and the low value. By the known properties of the Normal, this means that 95.5% of the sampled values will lie between the low and base cost scenario values. The delay is that (dynamically) estimated at the final destination.

Table 53. Summary of soft cost assignments by airline type

AO type	Method type	Summary of method
full-service	deterministic	$f_N \times \text{high} + [1 - f_N] \times \text{base}$
regional, LCC, charter	stochastic	$\sim N ([\text{base} + \text{low}]/2, [\text{base} - \text{low}]/4)$

For further derivation details and critiques of this methodology, please refer to Cook and Tanner (2011) and Cook *et al.* (2012). Although the POEM model estimates soft costs for each passenger dynamically, we can illustrate the magnitudes involved based on a soft cost applied equally to all passengers on a flight (taken from Cook *et al.*, 2012). Although by two hours the *per-minute* soft cost has saturated, the *net* cost continues to increase with total length of delay. If we take a British Airways Boeing 747-400 on a return flight from London to New York, with an average four-class configuration and 80 per cent load factor, this would typically generate revenue of the order of €300k. The soft cost of a 5-hour delay (using the base cost scenario value) is €94k — this equates approximately to the loss in revenue of one in three passengers taking their custom to another airline for one London–New York return trip.

⁶⁸ For single-flight trade-offs, the full soft costs shown may be used, but not in the full network context. For the scalability used in the POEM model, see under calibration, below.

⁶⁹ Which could, of course, be their sole leg.

Calibration note: The soft costs are automatically capped, in terms of their cost per minute, by the nature of the logit distribution described. For regional, LCC and charter passengers, the mean costs will be approximately the mean of the low and base cost scenarios, due to the random Normal sampling. The soft cost for full service passengers is driven by the fare paid, such that higher fares (notionally higher-yield passengers) cost the airline more.

Since soft costs refer to a loss in revenue to one airline as a result of a delay on one occasion, this loss may be considered to be largely the gain of another airline, gaining a passenger who has transferred their custom. When scalable costs are assessed, only some net loss to the airlines is likely. We use 10% of the full soft cost values shown in the table, as discussed and adopted in Cook and Tanner (2011). Soft costs were also applied to passengers overnighing in hotels due to unsuccessful flight re-booking, but an upper cost limit at 5 hours was applied to all passengers.

Base Rule 53 (non-passenger delay cost)

External data sources used:	European airline delay cost reference values (Cook and Tanner, 2011)
Internal rules/variables used:	N/A

The non-passenger costs of delay include the marginal (tactical) costs of additional crew hours, fuel burn and maintenance burden associated with delay incurred by aircraft, separately modelled for the at-gate, taxi, en-route and arrival management phases. A fuel carriage penalty is applied to arrival management. The at-gate phase assumes engines and APUs are off, i.e. no fuel burn. Full details of the methodology are to be found in Cook and Tanner (2011). (The values shown below have been specifically calculated from raw data and are not published in the 2011 report.)

Note that the costs are linear in time, unlike the passenger delay costs discussed in Rule 33 and Rule 52. They are quoted to a relatively high number of significant figures to facilitate the fitting described below; this does not reflect the level of precision of the original cost estimates.

For a given aircraft type, if a match in the tables is not found, an interpolation / extrapolation is made based on the maximum take-off weight (MTOW) of the aircraft required. MTOW values shown are averages, derived from the Central Route Charges Office (EUROCONTROL) data for flights in 2008 and 2009, kindly supplied by EUROCONTROL'S Performance Review Unit.

In many cases the base cost scenario value is very close to the mean of the high and low cost scenario values but these are not perfectly symmetrical about the base value in each case. In order to have absolutely symmetrical distributions about the base cost scenario values, we use a Normal distribution of costs, with the base value as the mean and a standard deviation set at one quarter of the difference between the high cost value and the low cost value.

By the known properties of the Normal, 95.5% of the sampled (assigned) cost values will lie approximately between the high and low cost scenario values, and these will be symmetric about the mean. Note that again, unlike the passenger costs, we treat all carrier types the same with regard to these cost assignments.

Table 54. Summary of non-passenger cost assignment

AO type	Method type	Summary of method
all	stochastic	$\sim N(\text{base}, [\text{high} - \text{low}]/4)$

Table 55. At-gate, non-passenger costs of delay per minute, by three cost scenarios

Aircraft	MTOW (tonnes)	Low cost scenario	Base cost scenario	High cost scenario
B733	60.4	0.2050	8.949	18.26
B734	65.6	0.2384	8.714	18.38
B735	55.2	0.2052	8.418	17.81
B738	72.6	0.1783	9.414	20.07
B752	107.1	0.2826	9.557	18.77
B763	180.7	0.3959	13.585	35.65
B744	392.5	0.8090	17.741	46.15
A319	66.6	0.2202	7.807	15.76
A320	73.6	0.2300	8.259	16.71
A321	86.4	0.2701	8.346	16.73
AT43	16.8	0.1019	5.832	11.77
AT72	22.1	0.1277	6.365	13.30

2010 Euros per minute.

Table 56. Taxi, non-passenger costs of delay per minute, by three cost scenarios

Aircraft	MTOW (tonnes)	Low cost scenario	Base cost scenario	High cost scenario
B733	60.4	6.689	19.35	31.97
B734	65.6	7.388	20.02	33.23
B735	55.2	7.204	19.43	32.29
B738	72.6	6.463	19.36	33.59
B752	107.1	10.326	25.31	39.50
B763	180.7	12.510	32.43	61.22
B744	392.5	23.713	51.23	89.76
A319	66.6	5.939	17.11	28.00
A320	73.6	7.528	19.73	32.24
A321	86.4	7.080	19.19	30.95
AT43	16.8	2.750	10.11	17.35
AT72	22.1	3.278	11.45	19.94

2010 Euros per minute.

Table 57. En-route, non-passenger costs of delay per minute, by three cost scenarios

Aircraft	MTOW (tonnes)	Low cost scenario	Base cost scenario	High cost scenario
B733	60.4	17.83	36.32	54.54
B734	65.6	17.93	36.04	54.51
B735	55.2	16.43	33.44	50.90
B738	72.6	18.52	37.64	58.05
B752	107.1	24.19	46.31	67.43
B763	180.7	34.44	65.55	105.54
B744	392.5	71.38	122.42	184.29
A319	66.6	17.06	34.08	50.57
A320	73.6	17.47	34.81	52.41
A321	86.4	20.66	39.80	58.33
AT43	16.8	3.69	11.56	19.26
AT72	22.1	5.21	14.43	23.88

2010 Euros per minute.

Table 58. Arrival management, non-passenger costs of delay per minute, by three cost scenarios

Aircraft	MTOW (tonnes)	Low cost scenario	Base cost scenario	High cost scenario
B733	60.4	14.07	30.68	47.02
B734	65.6	16.73	34.24	52.11
B735	55.2	11.83	26.54	41.70
B738	72.6	16.40	34.46	53.81
B752	107.1	19.03	38.57	57.11
B763	180.7	31.56	61.23	99.78
B744	392.5	48.74	88.46	139.01
A319	66.6	15.22	31.32	46.89
A320	73.6	17.11	34.27	51.69
A321	86.4	18.66	36.80	54.33
AT43	16.8	3.65	11.50	19.18
AT72	22.1	4.53	13.41	22.52

2010 Euros per minute.

NB. Currently, no cost recoveries are retrospectively corrected back to an earlier phase (e.g. back to taxi-out due to a subsequent en-route recovery) or for any saving associated with cancellations (e.g. due to unused or re-assigned crew hours). See Section 2.4.1 for further information on the implementation of the cost rules.

Calibration note: Through the use of Normal sampling, the cost allocations are symmetric about the base cost scenario and probabilistically bounded by the upper and lower cost scenario values.

Base Rule 61 (AO type assignment)

External data sources used:	AO associations (AEA, ELFAA, ERA, IACA), STATFOR, AO and alliance websites
Internal rules/variables used:	Variable 8, 9, 40

Each airline is assigned one of four categories that broadly describes its primary type of operation: full-service, regional, low-cost carrier (LCC) or charter. Airlines providing services covering more than one category (e.g. charter airlines offering full scheduled services; regional services provided by full-service carriers) are judgmentally assigned the most appropriate type. Other non-commercial IFR passenger AOs such as all-cargo, military transport and private/business aviation are removed from the database.

AO type is partly derived from the PRISME ICAO 3 character Aircraft Operator code (Variable 40), e.g. BAW = British Airways, a full-service operator. In cases where the AO type is not obvious, its general type of operation is determined from its membership of the various airline associations, e.g. Sverigeflyg (Aircraft Operator code = ETS), a member of European Low Fares Airline Association (ELFAA) is hence categorised as an LCC. STATFOR's 'Rules for Classification' of low-cost, cargo and business aviation is another useful reference for classifying operation type. The membership of the three main airline alliances, i.e. Star Alliance, SkyTeam and oneworld, has been compiled (as applicable for 2009/10) as this helps identify full-service operators and also supports likely code shares for connecting passengers (see Rule 33).

In cases where the AO is unknown (i.e. Aircraft Operator code = ZZZ), initial attempts are made to derive the AO code using the aircraft registration (Variable 9, i.e. searching the PRISME dataset for other instances of the same aircraft, which may have a corresponding AO code). Next, the aircraft type (Variable 8) is considered, as wide-bodies are more likely to be operated by full-service carriers and turboprops by regional operators.

The small proportion of remaining uncategorised AOs (excluding non-passenger and business flights), are categorised as regional carriers, since this has a fairly cost-neutral effect on the model.

Base Rule 118 (modelled off-block time)

External data sources used:	N/A
Internal rules/variables used:	Variables 5, 14, 34, 102 and 127; Rules 13, 16

MOBTs (and departure sequences) are assigned such that the generated sequence follows the FPFS principle by default, with reference to the original scheduled off-block time. As discussed in Section 2.4.1, as soon as a flight is flagged as ready for pushback, a departure request is sent to the departure manager of the origin airport so that the departure manager can check, following Rule 16, whether there are any network capacity restrictions. If so, the flight becomes pseudo-ATFM regulated and some departure delay is applied. (Additionally, if the current runway queue exceeds 5 minutes of queuing the flight is further delayed at-gate). Around 10% of all flights are thus pseudo-ATFM delayed, following an exponential distribution with $\lambda = 15$ minutes and truncated at 60 minutes. Furthermore, subject to these constraints, 95% of flights attempt to make an early departure (following the same exponential function).

Calibration note: These settings were assigned empirically in coordination with the target values described in Section 2.5. A value of 13.9 minutes for the average departure delay for a flight (across all flights) is set for the baseline traffic day. This is increased by one minute, to 14.9 minutes (across all flights), for the high delay day.

(ii) Major variables

Table 59 defines the variable types. All variables and rules in POEM belong to one of these three types, with some further division into sub-types.

Table 59. Definition of variable types

Variable code	Variable type	Example	Description	Typically associated with one flight?	Usually derived?	Updated during model run?
I	initialisation	EOBT	initial flight input	✓	✗	✗
D	dynamic	taxi-out delay	core control variable	✓	✓	✓
G	generic	airport MCT	generic data	✗	✓	✗

Table 60 gives the full list of major variables. (The variable numbering is historic through the project.) For initialisation variables, 'I+' indicates that the value is derived, or estimated trivially, from PRISME data. 'I++' indicates that rules need to be applied to compute the variable, which also applies by default to all 'D' and 'G' variables. For variables 1 – 4, all 'estimated' times in ATFM terminology, their dynamic analogues represent modelled 'estimated' times available to the model in real-time. Table 61 maps the corresponding I-type and D-type variables.

Table 60. Master variable list

Var. no.	Variable name / description	Type	Explanation / used to determine
1	EOBT [OOOI]	I	Estimated off-block time (last known planned time)
2	ETOT [OOOI]	I+	Estimated take-off time (last known planned time)
3	ETO [OOOI]	I	Estimated time on (last known planned time)
4	EIBT [OOOI]	I++	Estimated in-block time (last known planned time)
5	SOBT	I	Scheduled off-block time, from Innovata (see Appendix B)
6	SIBT	I	Scheduled in-block time, from Innovata (see Appendix B)
7	CTOT	I+	Establishes whether dep. constraint was imposed (original)
8	Type of aircraft	I	Multiple dependent factors in the model
9	Aircraft registration	I	Tracking of individual aircraft through day
10	Wait for crew & turnaround	D	Time taken for turnaround process
11	Aircraft ready for boarding	D	Dichotomous decision variable
12	ACRT	D	Aircraft ready time (crew available; ready for boarding)
13	Wait for boarding	D	Time taken for passengers to arrive and board
14	Aircraft ready for pushback	D	Dichotomous decision variable
15	MRT	D	Modelled ready time (aircraft ready for pushback)
16	Network capacity restrictions	D	Dichotomous decision variable

Var. no.	Variable name / description	Type	Explanation / used to determine
18	AOBT [OOOI]	I	Actual off-block time (original)
19	Taxi-out	I ⁺⁺	Actual taxi time (original)
20	Taxi-out delay	D	Delay before threshold
21	ATOT [OOOI]	I ⁺	Actual take-off time (original)
22	Planned en-route time	I	Planned en-route time (from schedule)
23	Modelled e/r Δt (incl. recovery)	D	Simulates en-route variability due to wind, etc., and delay recovery
24	PTI	D	Passing time for IAF
25	Airport capacity constraints	D	Non-ATFM, inbound & outbound (previously dichotomous)
26	Airborne arrival prioritisation	D	Driven by arrival capacity constraints
27	ATO [OOOI]	I ⁺	Actual time on (original)
28	Taxi-in	I ⁺⁺	Estimated taxi-in time (previously labelled 'standard' taxi-in)
29	Taxi-in delay	D	Delay before gate
31	Crew hours	D	When flight is departure ready
32	Aircraft-crew pairings	D	
33	Pax care & reaccommodation	D	Passenger care and rebooking onto onward destinations
34	Passengers onboard	D	Stores details of passengers on flights, including any connections
35	MCTs for all airports	G	Determines which pax can make connections for late inbound a/c
36	Seat no.s for all a/c types	G	Determines which pax can be accommodated from cancelled flights and missed connections
37	Load factors	G	
38	Dep. airport	I	Aerodrome of departure (ICAO code)
39	Arr. airport	I	Aerodrome of destination (ICAO code)
40	AO	I	Aircraft operator (ICAO code)
46	Filed NM	I	Filed route length
47	Regulated NM	I	Regulated route length (original)
48	Actual NM	I	Actual route length (original)
49	Filed taxi-out time	I	Filed taxi-out time (also needed to derive ETOT)
50	Regulated taxi-out time	I	Taxi-out time (when flight regulated); needed to derive CTOT
51	WTC	I	Aircraft wake turbulence category (H/M/L/J)
52	Passenger soft delay cost	D	Impact of unpunctuality on AO market share
53	Non-pax delay cost	D	Crew, maintenance and fuel costs, modelled by phase of flight
59	MTOW by a/c type	G	Needed to estimate delay costs
60	TAS	I	True air speed needed for the departure manager
61	AO type	I ⁺⁺	Full-service, regional, low-cost carrier (LCC) or charter
62	Reason for a rerouting	I	To help detect diverted flights
63	Status of reroute	I	To help detect diverted flights
102	MTOT	D	Modelled take-off time
104	MIBT [OOOI]	D	Modelled in-block time
118	MOBT	D	Modelled off-block time
127	MTO	D	Modelled time on
142	Regulation delay mins	D	Modelled ATFM delay (see primarily: rules 13, 16 and 118)

Table 61. Key correspondences between I-type and D-type variables

Var. no.	Variable name / description	I-type	Original / estimated OOOI / taxi / ATFM variable	D-type
1	EOBT [OOOI]	I	Estimated off-block time (last known planned time)	118
2	ETOT [OOOI]	I ⁺	Estimated take-off time (last known planned time)	102
3	ETO [OOOI]	I	Estimated time on (last known planned time)	127
4	EIBT [OOOI]	I ⁺⁺	Estimated in-block time (last known planned time)	104
7	CTOT	I ⁺	Establishes whether departure constraint was imposed (original)	via 142
18	AOBT [OOOI]	I	Actual off-block time (original)	118
21	ATOT [OOOI]	I ⁺	Actual take-off time (original)	102
27	ATO [OOOI]	I ⁺	Actual time on (original)	127
47	Regulated NM	I	Regulated route length (original)	not re-modelled
48	Actual NM	I	Actual route length (original)	via 23

The 'I⁺⁺', 'D' and 'G' variables are carried forward to Table 62, where their derivation and associated rules are summarised. For each variable, either the (base) rule number which handles the variable is shown (in small caps font) or a summary explanation of its derivation is given (in italics). Base rules that are linked to (prioritisation) scenario rules are indicated with an asterisk.

Table 62. Map of variables and their associated rules

Var. no.	Variable name / description	Type	RULE NUMBER - or - summary explanation
4	EIBT [OOOI]	I ⁺⁺	[3] + [28] or [29]
10	Wait for crew & turnaround	D	RULE 10 (PRISME & EUROCONTROL data, considers slack time)
11	Aircraft ready for boarding	D	<i>Dichotomous decision variable, uses [10]</i>
12	ACRT	D	<i>Aircraft ready time, uses [10]</i>
13	Wait for boarding	D	RULE 13* (time taken for passengers to arrive and board, uses [35])
14	Aircraft ready for pushback	D	<i>Dichotomous decision variable, uses [13]</i>
15	MRT	D	<i>Modelled ready time (aircraft ready for pushback), uses [14]</i>
16	Network capacity restrictions	D	RULE 16 (en-route ATFM delay) (previously dichotomous)
19	Taxi-out	I ⁺⁺	RULE 19 (log-Normal sampling from CODA data)
20	Taxi-out delay	D	RULE 20 (log-Normal sampling from CODA data)
23	Modelled e/r Δt (incl. recovery)	D	RULE 23 (route-based, due to wind, etc, uses PRISME)
24	PTI	D	<i>Passing time for IAF, uses [23] and [102]</i>
25	Airport capacity constraints	D	RULE 25 (non-ATFM, inbound & outbound) (previously dichotomous)
26	Airborne arr. management	D	RULE 26* (partly driven by [25])
28	Taxi-in	I ⁺⁺	<i>Estimated taxi-in time, uses [4] and [29]</i>
29	Taxi-in delay	D	RULE 29 (log-Normal sampling from CODA data)
31	Crew hours	D	RULE 31 (artificially estimated, drives departure readiness of flight; crudely modelled only, e.g. with more flexible rules for crew swaps at hub airports or similarly intensified nodes, compared to out-stations)
32	Aircraft-crew pairings	D	

Var. no.	Variable name / description	Type	RULE NUMBER - or - summary explanation
33	Pax care & reaccommodation	D	RULE 33* (passenger care and rebooking onto onward destinations)
34	Passengers onboard [¶]	D	<i>Stores details of passengers on flights, including any connections</i>
35	MCTs for all airports	G	RULE 35 (possible pax connections, uses IATA data)
36	Seat no.s for all a/c types	G	RULE 36 (determines which pax can be accommodated from cancelled flights and missed connections, uses IATA data for seats and multiple external sources for load factors)
37	Load factors	G	
52	Passenger delay soft cost	D	RULE 52 (impact of unpunctuality on AO market share)
53	Non-pax delay cost	D	RULE 53 (based on delay duration & aircraft type; uses [59])
59	MTOW by a/c type	G	<i>PRU, BADA or Innovata (IATA)**</i>
61	AO type	I ⁺⁺	RULE 61 (assigns AO type: full-service / regional / LCC / charter)
102	MTOT	D	<i>Modelled take-off time, uses [118] and [20]</i>
104	MIBT [OOO]	D	<i>Modelled in-block time, uses [102], [23], [26], [28] and [29]</i>
118	MOBT	D	RULE 118 (modelled off-block time, includes ATFM and other delay)
127	MTO	D	<i>Modelled time on, uses [102], [23] and [26]</i>
142	Regulation delay mins	D	<i>Modelled ATFM delay (see primarily: rules 13, 16 and 118*)</i>

* Linked to prioritisation scenarios (see Section 2.3).

Appendix B Final data preparation and cleaning

This appendix summarises the main data preparation and cleaning tasks applied to the PRISME traffic data and PaxIS passenger data. It updates work presented in POEM Deliverable 4.2 and POEM Deliverable 5.2. Model rules referred to are to be found in Appendix A.

(i) PRISME traffic data

Some of the following data preparation and cleaning tasks focused on the selected baseline traffic day (Friday 17SEP10), with other higher level tasks (e.g. related to the allocation of passengers to flights) applied more widely to the whole month. The traffic dataset has been supplemented with additional data fields. For flights in scope for the model these include: operator type, seats, schedule times, flight number and passengers on-board.

Aircraft-specific tasks

As previously reported, logic checks revealed coding inconsistencies between aircraft type, wake turbulence category and aircraft registration. Discrepancies were resolved using ICAO Doc 8643 (International Civil Aviation Organization, 2012) and aircraft registers such as the UK CAA's G-INFO database, the FAA's Civil Aviation Registry and Airframes.org. Unknown aircraft types (coded at source as 'ZZZZ') were resolved wherever possible from other instances of the aircraft in the dataset and from the aircraft registration, if available. Aircraft have been assigned the actual number of seats or the global weighted average (refer to Rule 36) and a pseudo flight number based on the operator combined with (departure/arrival) airport codes, incremented to uniquely identify each flight. For example, KLM_EBBREHAM03 identifies the third KLM flight in the day between Brussels National and Amsterdam Schiphol (a Fokker 70 aircraft with 80 seats).

Airline-related tasks

Each aircraft operator has been categorised by their primary type of operation, i.e. full-service, regional, low-cost carrier or charter (refer to Rule 61). Unknown operators (coded as 'ZZZ') or with indistinguishable operations have been categorised as (medium delay cost) regional carriers. Non-commercial passenger airlines have been identified for demotion in the model and coded accordingly ('XXX').

Flights operated by wet lease / ACMI (aircraft, crew, maintenance and insurance) operators have been categorised by the marketing airline if identifiable (e.g. via the matched schedule), else the charter category has been applied. Instances of the same aircraft flown by different operators on the same day have been categorised to match the operator per individual flight. For example, A320 registration number EID SX (registered in Ireland) was marketed and flown separately by Alitalia ('AZA'), Air One ('ADH') and Volare Airlines ('VLE') on 17SEP10, as illustrated in Table 63. The tracked movements reveal no missing flight legs and the schedule, converted to UTC, matches the actual off-block times. These three airlines are considered 'partner' airlines – Air One and Volare Airlines⁷⁰ are members of the Alitalia Group.

Table 63. Tracked flights of A320 EID SX on 17SEP10

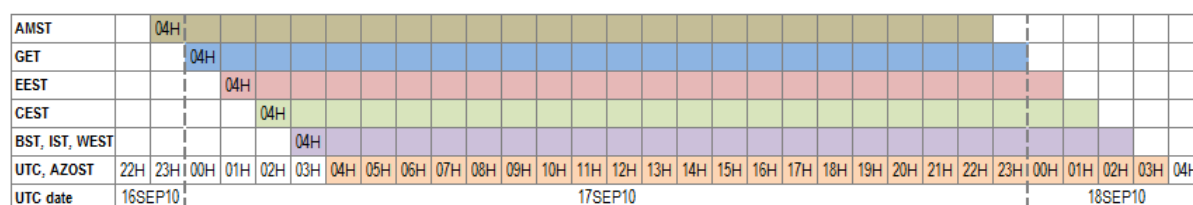
Operator	AO type	ADEP	ADES	Actual off-block time (UTC)	Scheduled departure time (UTC)
AZA	FSC	LIRF	LIML	05:30:00	05:30:00
VLE	LCC	LIML	LIRN	07:24:00	07:30:00
VLE	LCC	LIRN	LIML	10:05:00	10:05:00
ADH	LCC	LIML	LIRN	12:19:00	12:20:00
ADH	LCC	LIRN	LIML	14:33:00	14:30:00
AZA	FSC	LIML	LIRF	16:31:00	16:30:00
AZA	FSC	LIRF	LICJ	19:22:00	19:05:00

⁷⁰ Volare Airlines no longer operates.

Time-related tasks

Departure and arrival times have been converted to local times (in addition to UTC) in order to define local 0400-0359 operational days at ECAC airports and to enable schedule matching (published in local times only). Daylight saving time (DST) adjustments applicable on 17SEP10 have been incorporated per country and per time zone. Although synchronised throughout Europe, there are many different implementations of DST globally (local departure time is required for schedule matching), for example DST in Israel ended on 12SEP10 so only UTC+2H applied on 17SEP10.

The 0400-0359 operational day was based on the conversion of initial estimated off-block times (IOBT, closest PRISME time to schedule) and estimated touchdown times (ARVT_1) at ECAC airports. Flights were in scope if the IOBT at the departure ECAC airport and/or the ARVT_1 time at the destination ECAC airport, when expressed as local times, fell between 0400-0359 on 17SEP10. This led to a *small* over-estimation of the length of the day as estimated touchdown times exclude taxi-in time, thereby flights expected to land shortly before 0359 have been included even if the on-block time may have been after this time threshold. Figure 56 illustrates the offset of local time zones across ECAC. For example, flights departing and/or arriving between 2300Z on 16SEP10 and 2259Z on 17SEP10 at Zvartnots International (within the Armenia Summer Time zone, UTC+5H) were in scope.



AMST - Armenia Summer Time (UTC+5H)

BST - British Summer Time (UTC+1H)

GET - Georgia Standard Time (UTC+4H)

IST - Irish Standard Time (UTC+1H)

EEST - Eastern European Summer Time (UTC+3H)

WEST - Western European Summer Time (UTC+1H)

CEST - Central European Summer Time (UTC+2H)

AZOST - Azores Summer Time (UTC)

Figure 56. The 0400-0359 operational day across ECAC time zones

A series of parses was required to match flights with schedule times (supplied by Innovata), using initial estimated off-block times (IOBT, closest PRISME time to schedule) and estimated touchdown times (ARVT_1) converted to local times. Flights timetabled to operate on Friday 17SEP10 were extracted from the schedule, with care taken to identify the operating carrier from the multitude of code share carriers (PRISME furnishes operator details only).

Flights with IOBTs close to schedule times were initially matched, with greater differences between the times requiring more effort to align. Table 64 illustrates the difficulty of aligning flights that diverge from the timetable. In this example, two Blue Air ('JOR') flights between Cluj-Napoca and Bucharest Baneasa appear in the PRISME data, whereas *three* flights were scheduled. This example raised the following questions: were both flights late or was the first one on-time and the second one late; had the flown schedule been changed after publication of this timetable, such that both flights were actually on-time; had there been a cancellation or was the 'missing' (third) flight even later and outside the 0400-0359 time window? (There were no other flights on this route, so the third flight was not operated by another carrier on behalf of Blue Air.)

Table 64. Example of matching difficult flights with the schedule (local times)

PRISME flights requiring schedule times	Available schedule times for the day
	JOR_LRCLLRBS (dep 08:00)
	JOR_LRCLLRBS (dep 15:30)
JOR_LRCLLRBS01 (dep 19:45)	JOR_LRCLLRBS (dep 19:45)
JOR_LRCLLRBS02 (dep 23:30)	

In the absence of corroborating data, assumptions have been made. In cases where a sequence of flights were offset by a single time difference triggered by one flight – quite common on routes with many flights – the matched times were retained (i.e. flights that were close to the schedule times were assigned these times). This assumed that one flight, or a small number of flights, was/were late rather than a sequence of flights, perhaps held at gate due to a technical issue. Generally charter flights, by definition, are not scheduled so flights operated by charter airlines were assumed to follow their own schedule. Flights with no matchable schedule times were left unmatched and treated differently by the model (see later). A final time conversion was added to the traffic dataset – scheduled departure and arrival times were converted to UTC for use by the model. Overall, 92% of passenger flights on 17SEP10 were matched with the schedule.

Taxi times

Almost all the reported taxi-out times were exactly as filed and taxi-in times are unavailable⁷¹. Instead, monthly taxi-out/in statistics have been sourced from CODA and used to build the model's taxi times. See rules 19, 20 and 29.

Other tasks

Many other *ad hoc* data preparation tasks were carried out. Generic load factors for the four operator types were established (refer to Rule 36) and Pristina Airport coding was standardised as 'BKPR' from the use of both 'BKPR' and 'LYPR'.

Demotion flagging

As introduced in POEM Deliverable 5.2, some flights have been 'demoted' in the dataset in the sense that their metrics are no longer considered valid. Such flights cannot be excluded since they are still required to operate in order to convey their passengers through the network and to retain the correct level of traffic demand. A series of demotion algorithms has been used to flag flights not to be used for the metrics (applicable to SEP10). In summary, these flights have been demoted:

- **by airline type:** in addition to categorising operation by full-service, regional, low-cost carrier and charter, cargo-only operators have been identified (e.g. Air Contractors and AirBridge Cargo);
- **by aircraft type:** non-commercial passenger aircraft have been identified, such as private, military, helicopters and corporate aircraft (e.g. C130 Hercules and PA46 Piper);
- **by aircraft per airline fleet:** although aircraft types may have passenger and cargo variants, some airlines only operate the cargo variant (e.g. all A310 aircraft operated by Emirates and Turkish Airlines are cargo aircraft);
- **by aircraft registration:** tail numbers of non-commercial passenger aircraft have been identified, such as private, military, corporate, business jet, cargo, air ambulance, air taxi, research and state aircraft (e.g. FBUAD, European Space Agency Zero-G A30B research aircraft and N220AU, ORBIS International DC10 flying hospital);
- **by airport discrepancy:** this includes circular flights (i.e. ADEP the same as ADES), unidentified airports (i.e. ADEP or ADES coded as 'ZZZZ') and flights with air-filed flight plans (i.e. ADEP coded as 'AFIL');
- **by flight characteristics:** seven apparent flight characteristic errors (discussed in detail in POEM Deliverable 5.2) –
 - departure delay [AOBT_3]-[IOBT] < -1H or > 24H;
 - actual/filed route length [RTE_LEN_3]/[RTE_LEN_1] < 0.5 or > 2;
 - implied filed speed < 100 knots or > 700 knots;
 - implied actual speed < 100 knots or > 700 knots;
 - actual/filed speed (knots) < 0.5 or > 2;
 - en-route delay recovery > 1H and distance < 1000NM; extra en-route time > 4H;
 - actual speed > 600 knots and actual/filed speed (knots) ≥ 1.5.

⁷¹ The 2012 Network Operations Report highlights the on-going problem of taxi-time data as something that needs to be improved (albeit from a flight planning perspective): "Experiences show that taxi times are rarely correct and often not provided at all – even though EUROCONTROL B2B system can handle it" (EUROCONTROL 2013b, page 35).

- **by missing schedule times:** flights remaining without assigned schedule times (i.e. no schedule times could be matched with 8% of passenger flights);
- **by missing passenger itineraries:** flights populated by 'inert' passengers only (i.e. no matching PaxIS passenger data for these flights during SEP10; such passengers were used to populate flights, e.g. to achieve target load factors, but otherwise had inert characteristics in terms of the model).

5% of passenger aircraft flights were demoted on the basis of 'failing' multiple reasons listed above, for example, a training flight might exhibit unusual flight characteristics whilst also having no published schedule times. In addition to these demotions, duplicate flight records were removed completely. This applies both to 17SEP10 and to the whole month, as monthly seat capacity was integral to the passenger allocation algorithms.

PRISME traffic data cleaning summary

After completion of the data cleaning tasks, a total of 29 555 flights on 17SEP10 were available as model inputs, of which 21 550 passenger flights contributed to the metrics (Table 65).

Table 65. Summary of flight status for 17SEP10

Status	Flights	Flights %
Flights available for the metrics – with passengers	21 550	73%
Demoted flights – with passengers (e.g. missing schedule times)	4 247	14%
Demoted flights – non-pax movements (e.g. cargo, military aircraft)	3 758	13%
Total flights	29 555	100%

(ii) PaxIS passenger data

Compared with the PRISME traffic data, very little cleaning was needed for the PaxIS passenger data – however a considerable amount of recoding, restructuring and remapping (see Section 2.4.3) was necessary.

Recoding tasks

Data in the original PaxIS dataset used IATA codes – airlines and airports were thus recoded to the equivalent ICAO codes. Due to the unavailability of a single dataset valid for SEP10 and suitable for converting IATA to ICAO codes, this was a time-consuming task. For example, IATA airline codes are not unique ('controlled duplicates' exist in different regions) and codes are reissued when airlines delist – hence the timeframe and operating area are required when converting to ICAO codes.

Examples of IATA airline code problems include:

- Pegasus Airlines used IATA code 'H9' until 31JAN10 and 'PC' from 01FEB10 ('H9' was reissued to Helitt Líneas Aéreas from 01NOV11);
- Bahrain Air used IATA code '2B' until 30SEP10 and 'BN' from 01OCT10 ('2B' was reissued to Ak Bars Aero from 01MAY11);
- IATA code 'SQ' applies to both Singapore Airlines and Singapore Airlines Cargo (different ICAO codes are used);
- IATA code 'LH' applies to both Lufthansa and Lufthansa Cargo (different ICAO codes are used);
- IATA code 'BD' applies to both bmi and bmi regional (different ICAO codes are used);

The recoding of IATA airport codes was less problematic as these are generally static, with few unusual cases such as EuroAirport Basel-Mulhouse-Freiburg (two IATA codes: BSL and MLH). As passengers could have travelled to/from anywhere in the world, the recoding task was global rather than constrained to the ECAC area.

Enhancements

To facilitate passenger allocation (see Section 2.4.3), additional data and derived fields were incorporated. Total flight legs per passenger itinerary and flight legs combined into a single string were added, for example the separate origin, connection and destination airport fields DABB > DAAG > DAOO > LEAL became 'DABB/DAAG/DAOO/LEAL'. Minimum connecting times (MCTs) and airline alliance/partnership relationships were also prepared for the passenger allocation process.

Restructuring the raw dataset

POEM Deliverable 4.2 introduced the requirement to restructure the raw passenger dataset. This task has been further refined, with each data record changed from a grouped itinerary covering multiple passengers to an individual passenger itinerary (whilst retaining the flight legs). Figure 57 illustrates the transformation, in this example an itinerary covering four passengers has been duplicated so that each passenger itinerary has become a separate record. The figure also shows some of the additional fields and the result of recoding airlines and airports from IATA to ICAO codes.

Dom_AI	Mar_AI1	Mar_AI2	Mar_AI3	Orig	Connect_2	Connect_3	Dest	Class	Est_Pax	Avg_Fare
KL	KL	KL	AZ	NCL	AMS	FCO	BDS	ECON DISC	4	154

Dom_AI ICAO	Mar_AI1 ICAO	Mar_AI2 ICAO	Mar_AI3 ICAO	Orig ICAO	Connect_2 ICAO	Connect_3 ICAO	Dest ICAO	Class	Est_Pax	Avg_Fare	Est_Pax_One	Leg_Tot	All_Legs
KLM	KLM	KLM	AZA	EGNT	EHAM	LIRF	LIBR	ECON DISC	4	154	1	3	EGNT/EHAM/LIRF/LIBR
KLM	KLM	KLM	AZA	EGNT	EHAM	LIRF	LIBR	ECON DISC	4	154	1	3	EGNT/EHAM/LIRF/LIBR
KLM	KLM	KLM	AZA	EGNT	EHAM	LIRF	LIBR	ECON DISC	4	154	1	3	EGNT/EHAM/LIRF/LIBR
KLM	KLM	KLM	AZA	EGNT	EHAM	LIRF	LIBR	ECON DISC	4	154	1	3	EGNT/EHAM/LIRF/LIBR

Figure 57. The restructured raw passenger dataset

The dataset restructuring was an early task required before any allocation of passengers to flights could be made. Although the four passengers in the example itinerary do not appear complex for the allocation algorithm, some raw itineraries covered many thousands of passengers with the same ticket class and fare paid (e.g. an itinerary with 38 401 passengers paying EUR 88.60 each for discounted economy tickets on Alitalia flights in SEP10, flying one way from Rome Fiumicino to Milan Linate).

PaxIS passenger data cleaning summary

Cleaning iterations were carried out to remove passenger itineraries that were out of scope. Raw passenger itineraries consisted of up to six flight legs (the passenger itinerary in Figure 57 covers three legs – EGNT > EHAM/EHAM > LIRF/LIRF > LIBR). However, passengers with five or six flight legs were excluded due to small numbers (<100 passengers) and the unlikelihood of boarding so many flights within a 24-hour window. Records with flight legs without any corresponding PRISME flights or *via* unknown airports (i.e. the IATA airport code remained unmatched) were also removed. After cleaning, SEP10 itineraries for 58.5 million passengers were available for allocation to individual flights (Table 66). 58.5 million passengers are the equivalent of 71 million passenger movements when multiple flight legs are taken into consideration.

Table 66. Summary of passenger dataset cleaning iterations for SEP10

Cleaning iteration	Total passengers (SEP10)
0. Raw dataset	118 817 314
1. Remove records containing five or six flight legs	118 817 218
2. Remove records containing flight legs without any matched PRISME routes	58 589 053
3. Remove records containing unknown airports ('ZZZZ')	58 562 216

Appendix C Factor analysis methodology

We wish to use factor analysis (principal components analysis, PCA, to be precise) to describe delay propagation by reducing the 199 airports modelled to a much smaller number of groups of airports, sharing common characteristics. As we will explore in these examples, we expect there to be multicollinearities in the data. Since factor analysis takes many forms, it is important that the best approach is taken to give interpretable results. All factor analysis calculations are performed using IBM SPSS Statistics, Version 20.

Table 67. Simple data for three airports – perfect off-set

Time period	LHR	LGW	FRA
01	0	0	0
02	0	0	0
03	0	0	0
04	0	0	0
05	0	0	0
06	100	100	0
07	0	0	200
08	200	200	0
09	0	0	400
10	300	300	0
11	0	0	600
12	500	500	0
13	0	0	1000
14	300	300	0
15	0	0	600
16	100	100	0
17	0	0	200
18	200	200	0
19	0	0	400
20	100	100	0
21	0	0	200
22	0	0	0
23	0	0	0
24	0	0	0

Table 67 shows illustrative values for the total delay by time period at each of three airports. Note that in these first, highly simplified examples, the delays at both the London airports are the same and whenever there is a delay at one of the London airports, Frankfurt has zero delay, and *vice versa*. The delays at Frankfurt are an *off-set linear* function of those in London.

With an unrotated PCA (using the covariance matrix⁷²) on the London airports, a trivial solution is obtained: one component is extracted, which explains 100% of the variance, and with both airports having a factor loading of 1. When Frankfurt is included, still only one component is extracted, with (equal) positive loadings (0.546) on the London airports, a negative loading on Frankfurt (-0.972), with just over 51% of the total variance explained. This result may be considered as if the inclusion of Frankfurt has ‘diluted’ the results by being included in the only component extracted.

⁷² The covariance matrix, as opposed to the correlation matrix, is favoured when we wish to compare results across different samples (in this case different traffic input days) - Kim and Mueller (1978b). Furthermore, our variables are not measured on different scales, which is a typical argument (not applicable here) for using the correlation matrix. Further discussion of the various nuances of this choice are beyond the scope of this Deliverable.

(With (unweighted least squares) factor analysis, loadings of nearly 1 are obtained on the London airports, -0.334 on Frankfurt, and just under 70% of the total variance is explained. However, even these simple results are not reliable, because SPSS reports that one or more communality estimates of greater than unity were encountered during iterations⁷³.)

Factors are produced in descending order of the amount of variance explained by each factor. The extraction stops when some threshold is reached. Results may be improved, in terms of their interpretability, by increasing the number of components extracted (certainly beyond one factor, in this simple case). The number of factors extracted by a factor analysis is driven by several influences, including the structure of the underlying data, whether or not we force a certain number of factors to be extracted, and the eigenvalue setting. As experimental inputs, the data are not under our control. Nor can we determine *a priori* the number of factors to extract for a large number of airports, although we could for very simple examples by inspection. The default value in software is often to use eigenvalues greater than one⁷⁴, such that we only allow factors to be generated which are better than any single variable at describing the delay variance. Unless otherwise stated, the eigenvalue method will be employed, retaining factors with more than 0.5 times the mean eigenvalue⁷⁵. This produces the output of Table 68. Note that component 1 is still the same as when Frankfurt was not included, as described previously (with the same loadings on all three airports and the same variance explained by it). With component 2 included, all the variance is explained, although component 2 is difficult to interpret even in this simple case because it has fairly substantial, positive loadings on all three airports. Whilst it is true that knowledge of component 1 and the underlying data furnishes a reasonable overall interpretation, such a luxury would not be afforded with far more complex data for hundreds of airports and multiple components⁷⁶. We should also note that the same results are obtained if the rows are randomised⁷⁷, i.e. this technique is not a time series analysis, rather it is sensitive to whether delays occur contemporaneously. (An (unweighted least squares) factor analysis failed to extract any factors when either attempting to force the extraction to yield two factors, or by reducing the eigenvalue threshold⁷⁸.)

Table 68. Unrotated extraction for three airports

Component	% of variance explained	
1	51.347	
2	48.653	

Airport	Component (factor) loadings	
	Component 1	Component 2
LHR	0.546	0.838
LGW	0.546	0.838
FRA	-0.972	0.235

⁷³ The communality estimate corresponds to the proportion of the variance accounted for by the extracted *common* factors. A communality value exceeding unity (referred to as a Heywood case) means that the factors account for more than 100% of the variance of the observed variable(s), which clearly cannot be correct. The solution has to be either rejected or replaced by another solution with fewer factors or using a different extraction method. We return to this in the next example.

⁷⁴ This actually applies when using the correlation matrix.

⁷⁵ When using the covariance matrix, the setting relates to components with eigenvalues of higher multiples than the mean eigenvalues, rather than an absolute number.

⁷⁶ Other highly simplified examples with multiple airports the same as LHR and LGW, and multiple airports all with half the delay of Frankfurt (and thus the same as LHR and LGW, but still perfectly off-set), produce a second component with identical loadings for every airport.

⁷⁷ Although for anything but very simple data, this may cause minor variations in the factor loadings.

⁷⁸ The communality exceeding unity is again flagged in SPSS, this time with no factors produced. This type of problem has been reported for factor extractions using MLE, not only in SPSS. The same problem persists even if the data were made far more noisy to remove the perfect correlations and/or perfect off-sets between the variables. Further examination of these problems is beyond the scope of this report, and the (unweighted least squares) factor analyses are not pursued in the subsequent examples.

Holding all the extraction settings the same, but doubling the delay for each of the Frankfurt values produces the extraction shown in Table 69. We note that this has now produced only one component again, this time with just over 42% of the variance explained and a high loading on Frankfurt. Relatively higher delays at some airports may produce higher loadings on the associated components and solutions with fewer factors⁷⁹. Forcing the extraction of two factors produces the results of Table 70, where it is noted that component 1 is unchanged, but we now have a second component with high loadings on the London airports. Notably, all of the variance is again explained.

Table 69. Unrotated extraction for three airports (Frankfurt high delay, one component)

Component	% of variance explained	
1	42.592	
2	-	

Airport	Component (factor) loadings	
	Component 1	Component 2
LHR	-0.374	-
LGW	-0.374	-
FRA	0.999	-

Table 70. Unrotated extraction for three airports (Frankfurt high delay, two components forced)

Component	% of variance explained	
1	42.592	
2	57.408	

Airport	Component (factor) loadings	
	Component 1	Component 2
LHR	-0.374	0.927
LGW	-0.374	0.927
FRA	0.999	0.043

Returning to the original extraction (Table 68), the loadings on the London airports and Frankfurt were not that clear (unlike in the previous example where the delays at Frankfurt were increased, for illustration). Indeed, it is easy to speculate that with slightly more complex data and less clean patterns than the very simple example, it would be hard to deduce that there were two original delay patterns.

Varimax rotation is an orthogonal rotation method that minimises the number of variables with high loadings on each factor, i.e. it would be anticipated that airports would load heavily on one factor or another. This simplifies factor interpretation, as shown in Table 71. (The eigenvalue setting is the same throughout these examples, i.e. retaining factors with more than 0.5 the mean eigenvalue). All

⁷⁹ If the higher delays are high 'enough' – this will vary from context to context but holds as a general principle.

of the variance is explained by the two components, somewhat more by the first component, as compared with the original extraction (Table 68). Furthermore, component 2 now makes intuitive sense based on our knowledge of the simple source data. (The loadings are symmetrical because there are only two, *rotated* components. Other delay patterns, off-set or not, also produce symmetrical loadings under such extraction.)

Table 71. Varimax extraction for three airports

Component	% of variance explained	
1	65.713	
2	34.287	
Airport	Component (factor) loadings	
	Component 1	Component 2
LHR	0.986	-0.169
LGW	0.986	-0.169
FRA	-0.169	0.986

Table 72 shows the results of a promax rotation⁸⁰: this is an oblique rotation method, i.e. we do not wish to force the factors to be orthogonal, as we want to allow residual correlations between the factors. After extraction, we do not expect one group of airports (factors) to have delay characteristics independent of other groups (factors).

As a general principle, unless there is a good theoretical basis to force orthogonality, it is preferable to adopt oblique solutions⁸¹. If the resulting factors are uncorrelated, this orthogonality is not an artefact of the rotation method. Note that under oblique rotation, the sums of squared loadings cannot be added to obtain the total variances because the components are correlated; the percentage of the variance explained by each component is thus not shown.

For each of the preceding extractions, where two components were produced, the components were not correlated (even in the unrotated solutions). With the promax rotation the components are correlated⁸². Evident here is the very clean loadings on the London airports and Frankfurt, differentiated in each component. Indeed, the rationale of the promax rotation is that the orthogonal solutions are usually close to the oblique solutions, such that by reducing the smaller loadings to near-zero values, a simpler result is obtained.

⁸⁰ With Kaiser normalisation, throughout, for all oblique methods.

⁸¹ We therefore do not explore other common methods here, such as quartimax and equimax/equamax, since these are orthogonal.

⁸² Pearson correlation coefficient = -0.332 (with $\kappa = 4$, the SPSS default). The value of κ determines how correlated the factors may be. With $\kappa = 8$, the correlation coefficient is -0.333 and the factor loadings 1.000 and 0.000. With $\kappa = 2$, the correlation coefficient is -0.280 and the factor loadings 0.991 and -0.029. In the absence of very marked effects on results by such changes in κ , or common rubrics for using other values, we retain the default value of $\kappa = 4$ throughout these analyses, also to avoid unnecessary additional layers of interpretation.

Table 72. Promax extraction for three airports

Component	% of variance explained	
1	N/A	
2	N/A	
Airport	Component (factor) loadings*	
	Component 1	Component 2
LHR	1.000	-0.001
LGW	1.000	-0.001
FRA	-0.001	1.000

* Loadings are from the pattern matrix.

Table 73 shows a direct oblimin rotation (using $\delta = 0$, which gives the most oblique results) with the same settings as the promax method. This is also an oblique method⁸³. For ease of comparison, the components have been reversed in order. The only real difference observed is that FRA has a negative loading on component 1, compared with an analogous positive loading under the promax rotation, i.e. the direct oblimin loadings are less intuitive.

Table 73. Direct oblimin extraction for three airports

Component	% of variance explained	
1	N/A	
2	N/A	
Airport	Component (factor) loadings*	
	Component 2	Component 1
LHR	1.000	0.000
LGW	1.000	0.000
FRA	0.000	-1.000

* Loadings are from the pattern matrix.

⁸³ Pearson correlation coefficient = 0.333. Note that this is practically the same as the correlation coefficient under promax, whereas the sign of the promax coefficient makes more sense with respect to the underlying data (the delays at FRA are perfectly off-set to those at the London airports).

Table 74. Simple data for three airports – imperfect off-set

Time period	LHR	LGW	FRA
01	0	0	0
02	0	0	0
03	0	0	0
04	0	0	0
05	0	0	0
06	100	100	200
07	0	0	0
08	200	200	400
09	0	0	0
10	300	300	600
11	0	0	0
12	500	500	1000
13	0	0	0
14	300	300	0
15	0	0	600
16	100	100	0
17	0	0	200
18	200	200	0
19	0	0	400
20	100	100	0
21	0	0	200
22	0	0	0
23	0	0	0
24	0	0	0

Returning to the original data (Table 67), if each of the FRA delays is moved to a window one hour earlier, the same promax components are obtained as the first, second and third FRA delays are moved into alignment with the London delays, although the correlations between the two components progressively weaken (-0.332, -0.308, -0.210, +0.012). (Indeed, these essentially binary component loadings are obtained even if we increase the three FRA delays to 1000 minutes. This increase, for each of the first FRA periods, also makes the inter-component correlations progressively more positive.)

On moving the fourth delay, however, as per Table 74, we obtain only one component – with a higher loading on FRA (0.969, the aligned delays are higher for FRA) than either London airport (0.802). If we moved all the FRA delays to align with the London delays, the single component will have loadings of 1 for all the airports, since the Frankfurt delays are the sum of the London delays. If we increased one of the Frankfurt delays just sufficiently further (or likewise lowered two London delays in a given row), the loadings on the new single component would drop to just below 1, with the FRA loadings being the closest to 1. These key properties are summarised in Table 75, for the oblique case of two components.

Table 75. Properties of rotated, *oblique* components

Rotated, <i>oblique</i> component property	Continuous change observed with:	
	- off-set between delay patterns?	- absolute delays?
loadings on components	x	x
correlations between components	✓	✓

These observations for the rotated, oblique components may be viewed as the rotation ‘absorbing’ the effects of changing the off-set of the delay patterns and/or the absolute delays. These impacts are ‘instead’ manifested in the changing correlations between the components. (Note also that the continuous change is not necessarily simple or linear, as it depends on how much and where the delays change, until it ‘tips’ into another solution, e.g. with fewer components.) The opposite effects are observed for a varimax rotation, for example, whereby changing the off-set of the delay pattern and/or the absolute delays will change the component loadings, but not the component correlations – since these are, by definition, zero.

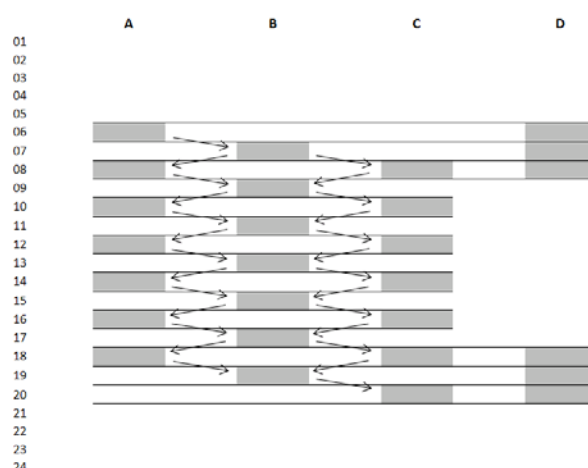


Figure 58. Four-airport example of delay patterns

We may now extend these concepts just a little further to complete the theoretical grounding. Consider the 24-hour period now extended to four airports, A through D. Hours during which heavy delays occur are shaded as grey blocks (all modelled as having the same value, for simplicity). These off-sets are not as perfect as the previous examples. The arrows represent flights, which may propagate delay from one airport (group) to another, establishing a causal relationship between the observed delays (although we noted in Section 3.1 that many observed relationships between airports are not causal). Three relationships represent the basic types of correlation we can expect between delays at (groups of) airports, whereby any of the correlations present may be causal or non-causal:

- positively correlated (e.g. A and C);
- negatively correlated (e.g. A and B);
- weakly/not correlated (e.g. A and D).

With all the same settings as for the previous promax rotation, the loadings obtained are shown in Table 76 (with D excluded for now). Here we can see that the correlation between A and C is reflected in their common loading on component 1, whilst the negative correlation between A (or C) with B, is reflected in B loading separately and highly on component 2. Even in this simple example, the components emphasise the correlations. Both A and C have exactly equal and high loadings on component 1 (0.946), whereas their Pearson correlation coefficient is (only) $r = 0.795$. Furthermore, the negative correlation between the two components (-0.456) is also slightly higher than that between A or C, and B (-0.438).

Table 76. Promax extraction for A, B and C airports

Component	% of variance explained	
1	N/A	
2	N/A	
Airport	Component (factor) loadings*	
	Component 1	Component 2
A	0.946	-0.004
C	0.946	-0.004
B	-0.003	0.998

* Loadings are from the pattern matrix.

Including D, an airport with blocks of continuous delay in the morning and evening, with three periods of overlap with A and C, and two with B, it is remarkable that the same promax rotation extracts three clean components (with convergence in only five iterations), as shown in Table 77.

Table 77. Promax extraction for A, B, C and D airports

Component	% of variance explained		
1	N/A		
2	N/A		
Airport	Component (factor) loadings*		
	Component 1	Component 2	Component 3
A	0.946	-0.003	0.000
C	0.946	-0.003	0.000
B	-0.005	0.998	0.000
D	0.001	0.000	1.000

* Loadings are from the pattern matrix.

It will be recalled that each block in the figure represents an equal amount of delay. If we quadruple the amount of delay for the six blocks of airport D, this causes the solution to collapse to one factor with a loading of 1 on D and of less than 0.3 for the other airports. Extending the dominance of airport D further, if it were to have these same delays in every time period from 06 to 20 (inclusive), Table 78 is obtained: a collapse to two components and a substantial loading of D on both of them.

Table 78. Promax extraction for A, B, C and D (extended) airports

Component	% of variance explained	
1	N/A	
2	N/A	

Airport	Component (factor) loadings*	
	Component 1	Component 2
A	0.946	-0.038
C	0.946	-0.038
B	-0.461	0.893
D	0.528	0.826

* Loadings are from the pattern matrix.

Since D is correlated identically with A, C and B, the higher loading of D on component 2 (the 'B' component) than on component 1 (the 'A and C' component) is potentially misleading. In actual data, where we are not afforded the luxury of the insight of simple original structures, this type of effect needs monitoring. It can be expected that airports demonstrating high levels of delay throughout the day will load highly on a number of components, with potentially less interpretability with respect to their *relative* loadings across such components. Forcing the extraction of three factors, or reducing the eigenvalue threshold, both have exactly the same effect. This causes a collapse of the factor structure, essentially retaining component 2 (now as the first component) and splitting A and C across components 2 and 3. Such problems may be offset to a certain extent by different types of rotation, but these will bring other associated interpretation issues, in that there is no perfect extraction method which fits all situations.

The main conclusions from Appendix C are summarised in Section 3.3.

Appendix D Passenger value of time

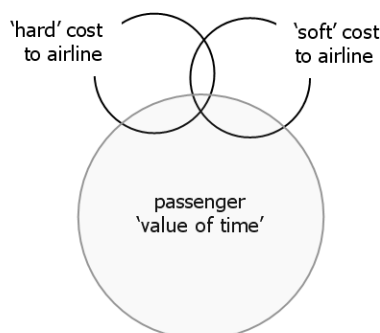


Figure 59. Value of time in context

'Value of time' is a concept widely used in cost-benefit analyses, particularly in transport economics. It is an opportunity cost that corresponds to the monetary value associated with a traveller/passenger during a journey. It is, usually, how much a traveller would be willing to pay in order to save time during such a journey (for example by travelling on a quicker service or a faster mode), or how much 'compensation' they would accept, directly or indirectly, for 'lost' time. The relationship between the passenger hard cost, soft cost, and value of time is simplified in the figure above. These values may be mutually dependent in a complex way. If an airline pays more compensation to a delayed passenger (a hard cost), the passenger may be less likely to defect to another airline (thus reducing the associated soft cost) and may make an adjustment to their in-trip or post-trip value of time. The hard and soft costs combined should cover every aspect of delayed passenger costs *that impact the airline*.

As discussed in Cook and Tanner (2011), the value of time is widely used in cost-benefit analyses for planning improvements in transport services and infrastructure, which deliver improved (usually faster) services to the traveller. It corresponds to an estimate of the 'non-monetary' cost associated with a journey. When this is added to the monetary cost (actual 'out-of-pocket' costs; such as fares paid) this produces the generalised cost (equivalent to the price of a good). The value of time is itself often split into work-related and non-work-related time. The former relates to trips undertaken as part of the traveller's work. It is often simply assessed as non-productive ('wasted') time, i.e. an opportunity cost to the employer (generally equivalent to the worker's rate of pay, with on-costs). It is more difficult to assign a value to non-working time. This is associated with the concept of 'utility theory' (the desirability of consuming a good or service, or the satisfaction derived therefrom – a relative measure). Values of time are expressed variously in terms of:

- access time (e.g. to an airport, and mode thereof) / search time (e.g. for parking);
- wait time / interchange time / headway / service frequency;
- in-vehicle time (congested / free-flow);
- departure/arrival time shift (planned, e.g. schedule delay);
- delay (unplanned arrival time shift).

Accurate evaluations should take into account the amount of productive time during the trip (e.g. working on-board an aircraft, or in an airport lounge during a delay). The value of time varies as a function of many variables: country, person, season, mode(s) used (air normally producing the highest values) and by journey distance and purpose (business-purpose usually being the highest). It may also vary by the *stage* of the journey and according to the method used for estimation (e.g. type of stated preference model). For a comprehensive meta-analysis, see Abrantes and Wardman (2011), which covers 226 studies carried out between 1960 and 2008, yielding a total of 1 749 valuations for the UK alone.

In a generic cost function (Pourtaklo and Ball, 2009) based on direct operating costs and passenger value of time (both sourced from ATA), the cost of delayed passengers *to airlines* was considered linear and approximated as one sixth of passengers' value of time, capped after 15 hours; the first 15 minutes of delay of the cost function (airline and passenger costs) had a zero cost. Cost parameters have also been developed (Zhang and Hansen, 2009) that include the passenger value of time at

USD 28.60 per hour and an aggregate flight delay cost of USD 2 000 per hour (based on earlier cost of delay reporting from the University of Westminster).

Ball *et al.* (2010) have calculated costs to passengers due to: delay and disruption (through simulations using detailed passenger itinerary and flight delay data); passenger inconvenience resulting from (airport capacity-induced) schedule delay (through statistical analysis and optimisation to find differences between capacity-constrained and ideal schedules); and 'schedule adjustment' costs (adapting personal travel schedules to mitigate delay impacts). Through econometric modelling these authors also calculate 'lost demand' costs such as welfare losses due to suppression of demand in delay-impacted markets and modal shift. These calculations thus cover both value of time *and* soft costs. Table 79 summarises three sources of value of time values presented in the EUROCONTROL 'Standard Inputs' document (EUROCONTROL, 2012). These are adjusted to either 2010 or 2011 Euro values, although we do not need to adjust the 2011 to 2010 values, considering the relative degree of uncertainty associated with the original estimates. Source (2) shows an asymmetry between travel time and delay time, which is not reflected in source (3).

Table 79. Value of time per passenger, per hour

Source	Basis	Generic value	Travel time	Wait time	Delay	Source/comment
1	Generic	46 - 59	-	-	-	EUROCONTROL 'Standard Inputs' 'preferred' value - although unclear whether marginal or average values; European source.
2	Private travel	-	27	3	41	EUROCONTROL 'Standard Inputs', 'alternative' value; based on income; Norwegian study from 1997; 'delay' values (not given in text) cited as being 50% higher than the travel time value (calculated to nearest Euro).
	Business travel	-	34	12	51	
3	Generic	-	26*	-	26	EUROCONTROL 'Standard Inputs', 'alternative' value; based on FAA report; values cited as "value of passenger time saved or lost" (marginal values), hence entered in 'travel time' and 'delay' columns; values based on income.
	Private travel	-	18*	-	18	
	Business travel	-	33*	-	33	
4	Private travel	-	-	-	30	Adopted values in this report, based on a crude evaluation of the values above and limited additional evidence, as discussed in the main text.
	Business travel	-	-	-	50	

* Travel time saved

A 2010 value of working time (i.e. to be used as a generic value of time for journeys made in the course of work) of GBP 39.65/hour (EUR 46/hour⁸⁴) for rail passengers is published by the UK Department for Transport in its economic assessment guidance documentation (Department for Transport, 2012); no value for air travel is cited. A non-working time, 'resource cost' (i.e. based on gross wages and on-costs) value of GBP 4.80/hour is cited, with no variation by mode (although the evaluation does not include air). Abrantes and Wardman (2011) state that air transport users have values around 4.25 times larger than car users, although recognising that a number of confounding effects make such comparisons difficult. This would give an extremely crude estimate for leisure purpose air travel of GBP 20.40/hour, although the 4.25 factor itself already includes business and leisure purpose air travel. Using the 'generic' / 'private travel' ratio of source (3) in Table 79, gives a further crude adjustment to GBP 14.10/hour (EUR 16/hour). This fairly close to the value cited in source (3), in fact, i.e. EUR 18/hour. Pels *et al.* (2003) use a nested logit model that estimates combined access mode and airport choice, in a San Francisco Bay area case study. For August, business travellers had a value of time of USD 2.90/min, whilst for leisure travellers in the same month, the value was USD 1.57/min. These estimates are based on 1995 data, and adjusting them to

⁸⁴ GBP/EUR exchange rates sourced from the European Central Bank's Statistical Data Warehouse (<http://sdw.ecb.europa.eu>); values calculated to the nearest whole Euro.

current Euro values would be rather dubious, especially as value of time does not follow simple inflationary patterns (Department for Transport, 2012). We note that this ratio, approximately 1.8, is almost the same as source (3). Hess and Polak (2006) use a cross-nested logit model on data collected in the Greater London area to jointly analyse the choice of airport, airline and access-mode. The result revealed significant influences on passenger behaviour by access-time, access-cost, flight-frequency and flight-time. A mean value of GBP 15.96 /hour is cited for the trade-off ratio of in-vehicle access-time over access cost. The data relate to business travellers in 1996. The authors comment on the potential value of further analyses using more recent data. The extent to which this value can be used as a proxy for a (current) delay cost estimate, through use of the flight-time coefficient of nearly unity⁸⁵, is open to further exploration. We cite this 1996 value of GBP 15.96/hour more as a relatively rare example of a European value of time set in the air transport context, also remarking that it falls in the same order of magnitude as several values cited above.

It is difficult to set the values given in source (2) for wait time in the comparative context of other results. As mentioned above, they depend on the productivity of such waiting time, compared with extended in-vehicle time and the disutility of delay at the final destination. Abrantes and Wardman (2011), in their meta-analysis, comment that wait time showed less than its usual premium of twice in-vehicle time (the overall time multiplier cited is 1.70). Although it is difficult to translate these findings into the air transport context, for comparison, the values presented in EUROCONTROL (2012) (source (2), Table 79) appear to be quite low.

In conclusion, relatively few studies have produced values of time for air transport, especially with respect to delay and in a recent European context. More work has been carried out on airport choice factors, and far more research has been published for road and rail modes. The extent to which these various evaluations can be compared is limited, and there is certainly scope for much more research in the context of European value of time estimations for air transport delay. Nevertheless, there is fair agreement between existing values for the air transport mode, which also broadly correspond to very crude estimates made from other modes. This is not completely surprising, since many evaluations are based to a greater or lesser extent on income. We have adopted a value of EUR 50/hour for business travel (see Table 80), which agrees closely with the European source (2), and is only somewhat less than the upper range of the generic derivation for source (1). Applying the 1.8 ratio of business/leisure to this value, suggests a value of EUR 28/hour for private travel, which we have rounded to EUR 30/hour. This is still less than the values in sources (1) and (2), but rather greater than that suggested by source (3). These values adopted are based on crude arguments only, and more work needs to be done in this area. The values adopted are, however, close to two European (2011 Euros) values for business travel, and preserve a ratio of 1.8 for business/leisure travel. In particular, we note that the values of time discussed are not cited as a function of delay *duration*, which is an important consideration currently absent from these calculations, as only flat rates are available to be applied. Furthermore, values of time may be adjusted by travellers as a direct result of previously experienced delay, and delays incurred during separate phases of a journey (e.g. waiting and at final destination) may often not be independent. Neither of these effects is taken into account. We consider there to be insufficient evidence to support a robust waiting time value, such that the POEM model only evaluates value of time as a function of delay at the final (airport) destination. For each passenger, additional waiting time and delay at the final (airport) destination are recorded, but only the latter is assigned a cost. (See Rule 36 (Appendix A) regarding flexible and inflexible ticket holders.)

Table 80. Ticket type by value of time applied

Ticket type by value of time	Flexible ticket holders	Inflexible ticket holders
Value of time applied as a function of delay at final (airport) destination	EUR 50/hour	EUR 30/hour
Value of time applied as a function of extra waiting time at airports / hotels	<i>Time counted per passenger, but no cost assigned</i>	

⁸⁵ Although researchers are mostly concerned with variations in monetary valuations, it is common practice to express valuations of non-in-vehicle time in equivalent units of in-vehicle time.

Appendix E Baseline traffic day disruption

In total, 19 days in September 2010 were affected by ATC industrial action by varying degrees of severity, for example, almost 4 000 flights were delayed by 67 minutes on average on the most severely affected strike day (23SEP10).

In selecting nominal days as candidates to be modelled (as fully discussed in POEM Deliverable 4.2), badly disrupted days were excluded if their total airport and en-route ATFM strike (or weather) minutes were greater than three standard deviations above the mean for their respective categories – 07SEP10 and 23SEP10 were thus identified for exclusion.

In selecting ‘typical’ quiet and busy ATFM-delay days, Tuesdays and Fridays were preferred, respectively. However, no Tuesday was unaffected by strikes and two Fridays (the busy days) had ATFM delays attributable to strike. Friday 17SEP10 had approximately half the ATFM delay of the other strike-affected Friday and was selected as the baseline traffic day.

Approximately 5% of the (cleaned) flights for the baseline traffic day were affected by ATC industrial action, which was restricted (in source) to France. The average departure delay of flights with such a regulation reason assigned was 23 minutes. For all airports where more than 20 flights were affected, the number of such flights and the average departure delay (to the nearest whole minute), specifically due to this regulation reason, are shown in Table 81.

Table 81. Flights affected by strike action in France on baseline traffic day

Airport		Number of flights	Average dep. delay
LFPO	Paris Orly	96	21
LEMD	Madrid Barajas	75	28
LFLL	Lyon	60	15
LFPG	Paris CDG	54	19
LFBO	Toulouse	43	14
LFML	Marseille	40	11
EGKK	London Gatwick	39	32
LEBL	Barcelona	36	20
LFRS	Nantes	32	23
EBBR	Brussels	26	13
LEPA	Palma de Mallorca	26	14
LIRF	Rome Fiumicino	26	8
LFMN	Nice	24	8
EGSS	London Stansted	23	28
EHAM	Amsterdam	23	34
LFBD	Bordeaux	23	12
EGCC	Manchester	22	37

The average departure delay for each of these airports in the PRISME data for the baseline traffic day was in any case less than their CODA annual averages for the same year, with the exception of Barcelona. For Paris Orly, Paris Charles de Gaulle, Barcelona, Brussels, Palma de Mallorca and Rome Fiumicino, the strike-related average delays were no more than five minutes greater than their CODA annual average. For Madrid Barajas the strike-related average delays were 8 minutes higher than its CODA annual average, for London Gatwick 11 minutes higher, London Stansted 16 minutes, Amsterdam 20 minutes and Manchester 17 minutes. These larger values, above 15 minutes, related to fewer than 25 flights at each of London Stansted, Amsterdam and Manchester.

Above all, these delays are reset in the model such that the simulations represent a normative day to the greatest extent possible (with the exception of residual cancellations, as discussed in Section 2.4.2).

Appendix F Further analyses of airport reactionary delay

Results in this appendix complement those of Section 3.4.2 by offering further insights into higher-level principles already illustrated. The airports selected are the top five with respect to traffic movements for the baseline traffic day.

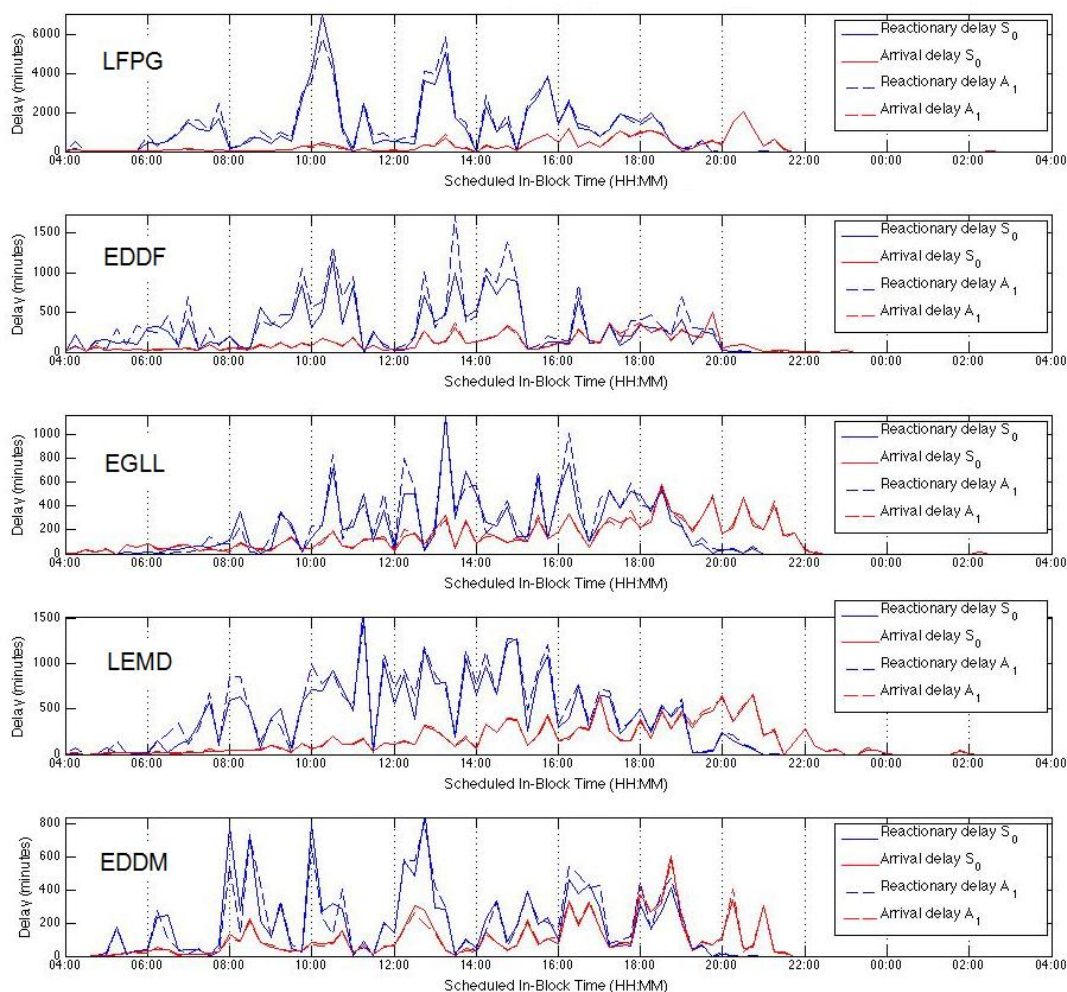


Figure 60. Arrival and reactionary delay distributed by scheduled in-block time for busiest airports

Figure 60 shows the distribution by time of day for arrival and reactionary delay for these airports. Arrival delay often peaks towards the end of the operational day, which is not the case for reactionary delay *generation* (as would be expected). Note that the scales differ: LFPG (Paris Charles de Gaulle) is a large generator of reactionary delay, as explored in Section 3.4.2. Figure 61 shows the proportion of (other) airports affected by some level of reactionary delay associated with these five busiest airports, such fractions typically increasing under A_1 .

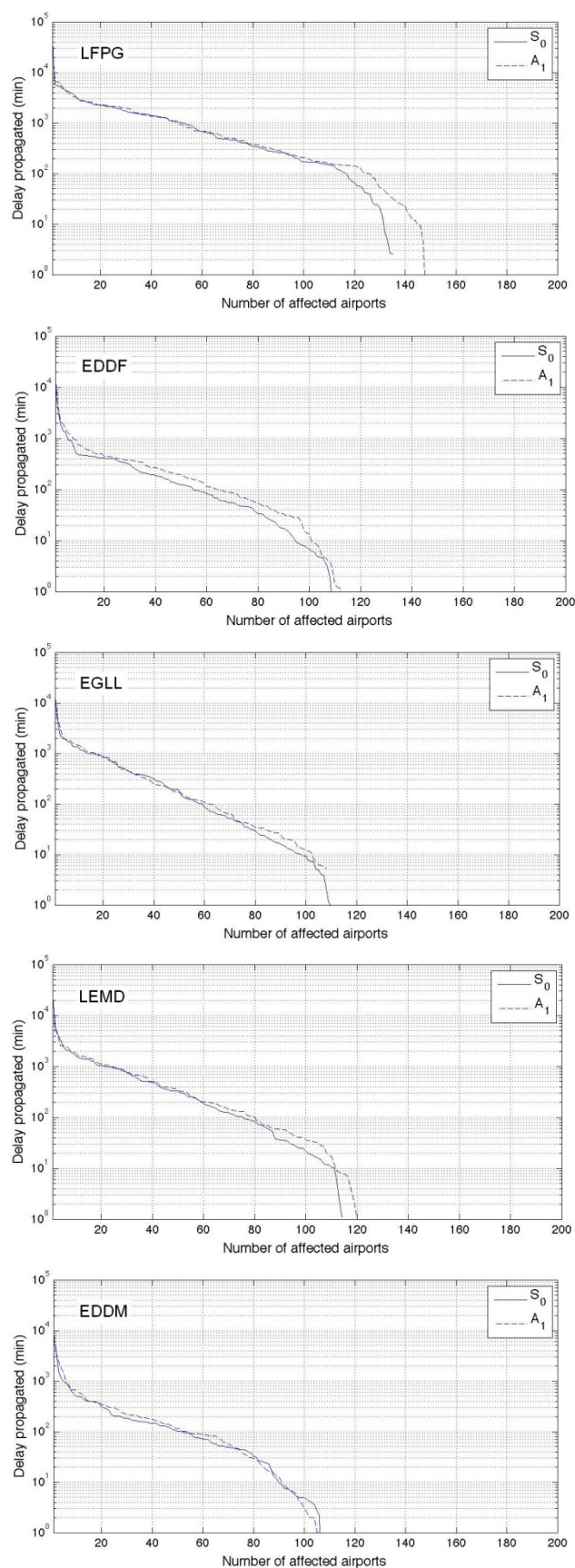


Figure 61. Reactionary delay propagated from busiest airports

-END OF DOCUMENT-