6-2009

Towards a single European sky

Yael GRUKSHA-COCKAYNE
University of Virginia

Bert DE REYCK
Singapore Management University, bdreyck@smu.edu.sg

Citation

GRUKSHA-COCKAYNE, Yael and DE REYCK, Bert. Towards a single European sky. (2009). *Interfaces*. 39, (5), 385-501.
Available at: https://ink.library.smu.edu.sg/lkcsb_research/6764

Interfaces
Vol. 39, No. 5, September–October 2009, pp. 400–414
ISSN 0092-2102 | EISSN 1526-551X | 09 | 3905 | 0400

DOI 10.1287/inte.1090.0436

Towards a Single European Sky

Yael Grushka-Cockayne
Darden Graduate School of Business Administration, University of Virginia, Charlottesville, Virginia 22906-6550,
ygrushka.phd2003@london.edu

Bert De Reyck
Department of Management Science and Innovation, University College London,
London, United Kingdom WC1E 7JE; and Department of Management Science and Operations,
London Business School, London, United Kingdom NW1 4SA, bdereyck@london.edu

We describe an integrated decision-making framework and model that we developed to aid EUROCONTROL, the European air traffic management organization, in its vital role of constructing a single unified European sky. Combining multicriteria decision analysis with large-scale optimization methods, such as integer programming and column generation using branch and price, our model facilitates the process by which the numerous European aviation stakeholders evaluate and select technological enhancements to the European air traffic management system. We consider multiple objectives and potential disagreements by stakeholders regarding the impact of proposed system enhancements and allow for different priorities for each key performance area. In an earlier paper, we described the mathematical programming model in detail. In this paper, we elaborate on the broader decision framework and supporting methodologies to help EUROCONTROL in its facilitation role. Using our model and decision framework, EUROCONTROL is currently selecting a set of enhancements to the European aviation system upon which all stakeholders have agreed.

Key words: decision analysis: multiple criteria; transportation: air; programming: integer, applications; research and development: project selection.
History: Published online in *Articles in Advance* June 10, 2009.

Eurocontrol, the European air traffic management (ATM) organization headquartered in Brussels, Belgium, has as its primary objective harmonizing and integrating air navigation services in Europe. Its ultimate goal is the creation of a uniform pan-European ATM system. In 1963, the International Civil Aviation Organization (ICAO) founded the EUROCONTROL agency to create a single body with responsibility for the entire European airspace. At that time, however, most European states were not prepared to give up sovereignty over their own airspace. Nevertheless, recent increases in air traffic have made an integrated European sky a necessity. According to the latest forecasts, air traffic in Europe is expected to reach up to 16 million flights annually by 2020 (Grushka-Cockayne et al. 2008). Such a traffic increase will cause severe congestion in the sky and around airports. Beginning then, 3.7 million flights per year might not be accommodated, causing a potential yearly loss of 50 billion euros.

Currently, European air traffic management cannot fully exploit available capacity, mainly because of segregated systems, lack of standardization, and restrictive regulations. Airport congestion, already a problem at many major airports, will become more widespread, especially at the international hub airports that serve major European cities. In addition, growing environmental concerns will also restrict any increase in capacity unless changes to the ATM systems are made.

A Single European Sky

According to Victor M. Aguado, the former Director General of EUROCONTROL, "In order to handle the levels of traffic we will face in 2020, we need to begin working now to build a pan-European network of air navigation services, airports, airlines, and airspace users whose evolution is planned and designed to meet traffic loads" (EUROCONTROL 2006). With this goal in mind, the European Commission launched the Single European Sky (SES) initiative in March 2004. The SES program aims to eliminate the current system fragmentation by restructuring the European airspace to create additional capacity, increase the overall system efficiency, improve safety, and reduce the environmental impact.

In the past, EUROCONTROL and other stakeholders have undertaken many initiatives to improve the European ATM systems. However, these initiatives never achieved their full potential, mainly because of a lack of commitment by the stakeholders. Now, for the first time in European ATM history, all aviation players are united in defining, committing to, and implementing a European ATM master plan. A consortium of representatives from airspace users, airports, air navigation service providers, the supply industry, safety regulators, the military, pilots, research centers, and EUROCONTROL is implementing this project, which the European Commission and EUROCONTROL has funded.

The ATM Master Plan

The ATM master plan (Figure 1) consists of a series of operational improvement (OI) programs that include the operational, technical, and institutional changes that are required to meet future performance requirements and to improve key performance areas, such as capacity, safety, cost-effectiveness, flexibility, and the environment.

The estimated cost of the master-plan development phase is 2.1 billion euros.

In constructing the ATM master plan, the consortium of stakeholders must identify the OI programs to be included. Because each OI program has advantages and disadvantages with respect to the system's key performance areas, trade-offs must be made. These trade-offs are complicated by the interactions that exist among the OI programs; the inclusion of one OI pro-

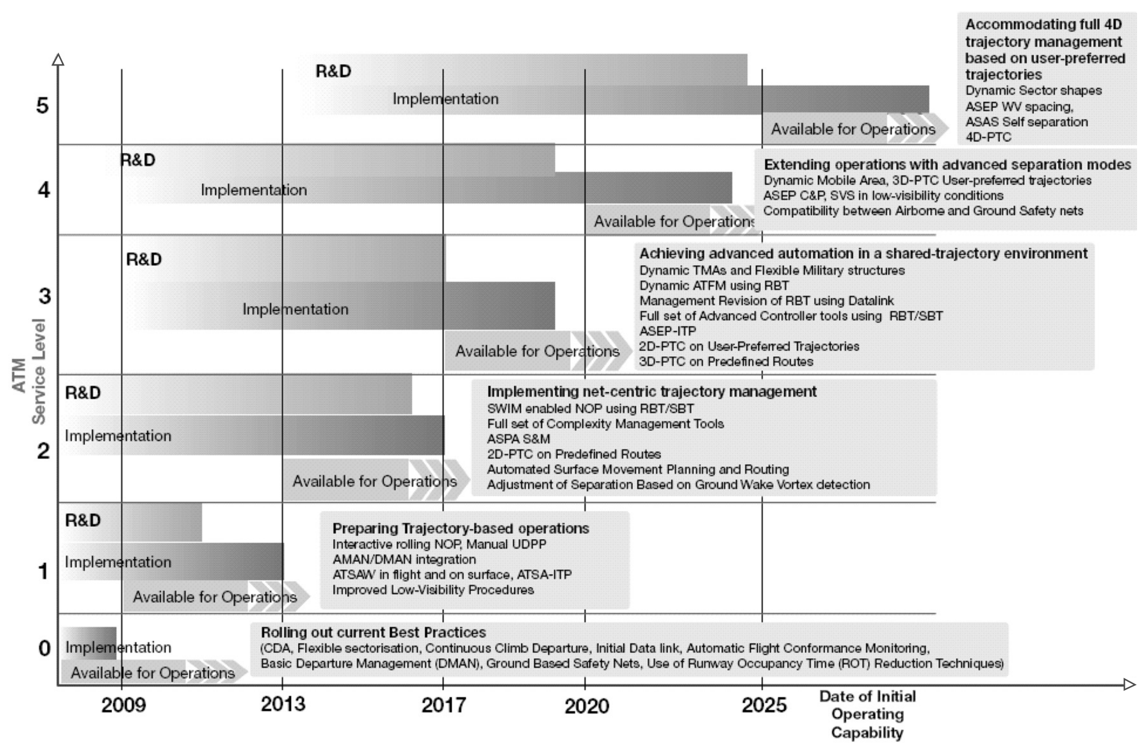


Figure 1: The European ATM master plan contains a series of operational improvement programs foreseen for the period 2009–2025.

gram might affect the impact delivered by others. In addition, a stakeholder can implement each program in one of several different ways, each of which is characterized by a different cost and expected performance impact. In addition, the consortium must consider numerous performance targets, including safety, capacity, predictability, and environmental impact. The most challenging issue, however, is that the various stakeholders might have different priorities and opinions concerning the expected cost and performance benefits of each program.

Decision-Making Framework

Up to now, both the European ATM systems and the decision-making processes have been fragmented. The stakeholders typically carried out their assessments of improvement programs individually and the evaluations of each objective separately. Experts typically did not discuss their assessments jointly, and they used different frameworks. Some emphasized quantitative assessments, such as cost-benefit analyses; others used qualitative frameworks, such as assessments of environmental benefits.

Following an in-depth review of existing trade-off methods (De Reyck et al. 2005), we developed an integrated decision-making framework and a set of models for OI evaluation and selection. Unlike other methods, our framework incorporates the views of multiple stakeholders, multiple objectives, and qualitative as well as quantitative information in a mathematically rigorous model that can simultaneously consider multiple interrelated programs.

Figure 2 presents a general overview of our decision-making framework.

It consists of nine iterative stages. We revise the model as deemed necessary until we obtain a requisite representation of the decision problem (Phillips and Stock 2003). The first three stages identify the OI programs requiring further analysis, the relevant stakeholders and their objectives, and any risks or constraints that might exist. Stage 4 uses both qualitative and quantitative ratings to evaluate the performance of different combinations of OI programs in terms of the various objectives. In Stage 5, we determine the trade-offs among the objectives by assigning weights to performance criteria to reflect the relative importance of each. Because the stakeholders have different weights and ratings, each stakeholder performs Stages 4 and 5 separately.

In Stage 6, we determine the preferred combination of OI programs using an integer mathematical programming model in which we combine the different stakeholder perspectives and multiple objectives. Because the number of combinations of OI programs could be huge, we use large-scale optimization methods, including column generation and branch and price. In Stages 7 and 8, we examine the results and conduct advanced sensitivity analyses to ensure robust recommendations. The final stage of the decision-making process, Stage 9, is the actual decision. A successful application of the framework will result in a shared understanding of the problem and achieve a joint commitment to action by all involved parties. Table 1 summarizes the desired properties of the OI evaluation and selection methodology, as determined by EUROCONTROL, and states how we addressed these in our proposed framework.

The contributions of our work are fourfold. First, we demonstrate how decision analysis can be applied

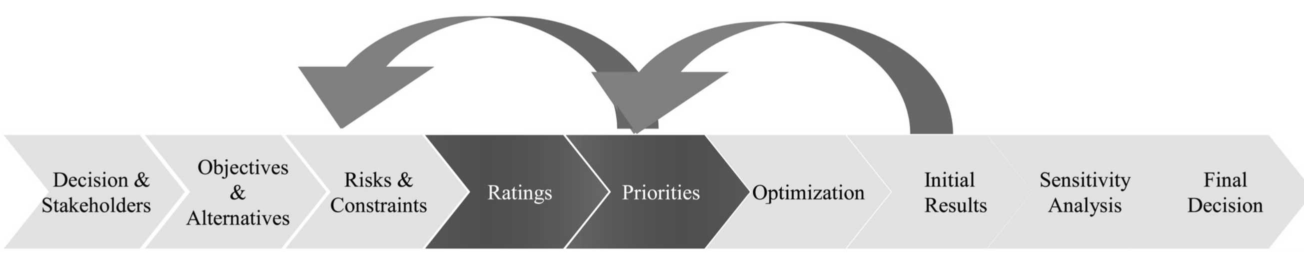


Figure 2: In our decision-making framework, the stages in light shade of grey are carried out by all stakeholders; the darker stages are performed by each stakeholder separately.

Requirement	Reason	Decision-making framework
Capability to represent multiple objectives.	Because of the nature of the ATM system, the expectation is to pursue multiple key performance areas, i.e., safety, capacity, efficiency, and environmental considerations.	We establish the list of relevant objectives in Stage 2 of the process. In Stage 4, we assess the alternatives' performance on each objective.
Support of qualitative and quantitative assessments.	The ATM key performance areas include areas such as safety and environment that are typically rated on a qualitative scale. Quantitative data might not always be available or disclosed by the stakeholders.	The model can be adapted to use a variety of assessments, including qualitative or quantitative, and point estimates or ranges.
Sufficient robustness to value ATM systems across a wide range of possible views.	Because many stakeholders play an active part in the decisions, all their views must be part of the discussion and support the selected alternatives.	We develop a separate model for each stakeholder, allowing for different assessments of the alternatives and a different hierarchy of objectives and associated weights. We combine the different views and contrast them while looking for a compromise solution.
Objectivity and lack of biases.	Because of the involvement of multiple stakeholders with different agendas, a bias-free decision-making platform is necessary.	An unassociated party proposed the methodology; to support the objectivity of the process, a neutral facilitator will provide facilitation.
Academically sound and rigorous process.	This is required to gain buy-in and the trust of the parties involved because they view an academic perspective as neutral and objective.	We structure the framework as a variant of multiple-criteria decision analysis models, which are well established and widely used today.
Understandable to all participants and transparent to all stakeholders and decision makers.	The methodology must allow for traceability of all input data and support the challenging of all assumptions to avoid "black-box" characteristics.	Workshops and interviews, with much interaction with the stakeholders, serve to validate assumptions and gain support for the model input.

Table 1: EUROCONTROL put forward a list of characteristics that any proposed methodology should follow, all of which we incorporated into our decision-making framework.

when making high-impact, strategic, and global decisions to promote the recognition of issues, such as environmental impact and safety. Second, we extend the use of multicriteria decision analysis models to situations in which multiple stakeholders with conflicting objectives and different perspectives must agree upon a single course of action. Third, we combine existing large-scale optimization methods such as integer programming with multicriteria decision analysis methods such as pairwise comparisons. Thus, we advocate the use of an integrated multicriteria decision analysis model; the model also incorporates the use of problem-structuring methods for framing a complex problem within a large organization, combining qualitative and quantitative assessment methods, and allowing for interactions among the alternatives—a challenge that, to our knowledge, has not received much attention in the literature. In this way, we contribute by suggesting a novel approach for advancing the application of decision analysis techniques to project portfolio problems. Finally, our fourth contribution is in providing an elicitation technique for solving multicriteria problems in cases in which there are many alternatives to consider, i.e., extrapolating information for a combination of a few discrete options to infer about a large set of alternatives.

Evaluating Alternative OI Programs

We will now discuss in detail the model we developed to support the decision framework that we presented in the previous section. Table 2 summarizes the decision frame.

As part of the framing process, the stakeholders generate their lists of objectives and criteria against which they will assess the alternatives. We classify these objectives either as *filtering criteria*, which we use to screen and eliminate alternatives if any one of these criteria is not met, or *comparison criteria*, which we use for assessing and comparing the alternatives. Next, through a series of interviews and workshops, we obtain ratings measuring each alternative's attractiveness with respect to each performance criterion. We assess the alternatives on a variety of scales,

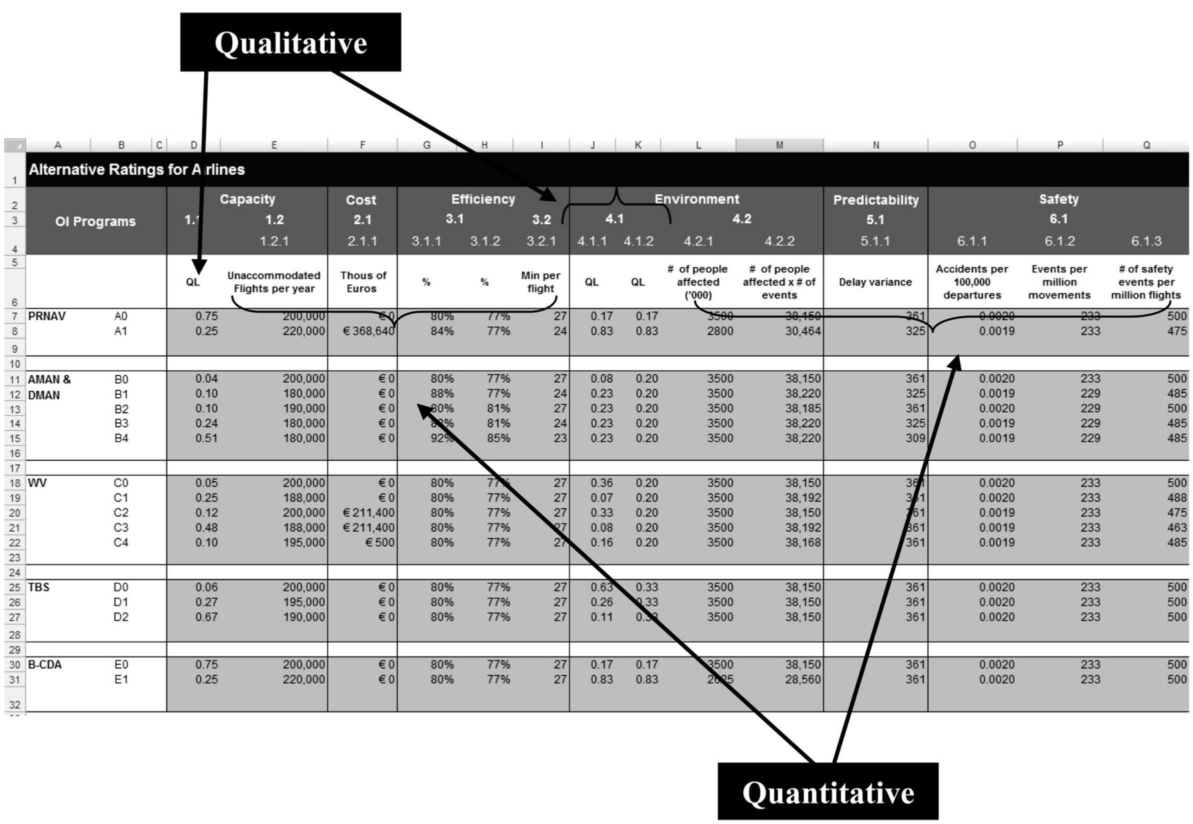
Decision frame	Elements
Decision context	• OI programs in the ATM master plan
Stakeholders	• Airspace users (airlines and private aircraft owners), air-navigation service providers, airports, aircraft manufacturers and the aeronautic industry, the military, society, member states, European commission
Objectives (key performance areas)	• Access and equity, capacity, cost-effectiveness, efficiency, environmental sustainability, flexibility, predictability, safety, security, participation, interoperability
Alternatives	• Different combinations of OI programs performed in a specific way
Constraints	• Performance targets
Risks	• Uncertainty regarding the performance of the alternatives on the different objectives • Uncertainty concerning the priorities of stakeholders • Availability and disclosure of data by the stakeholders

Table 2: The first step in our OI program evaluation and selection is framing the problem.

including monetary values, percentages, kilometers, and qualitative comparisons, and convert these into a numerical value function (Keeney and Raiffa 1976).

To support the evaluation process, we developed an Excel-based system using Visual Basic for Applications (VBA). Figure 3 illustrates how one stakeholder, the airlines, can rate each alternative OI program for one specific element, "Arrival and Departure with P-RNAV Support," which examines changes to the aircraft arrival and departure processes at airports to improve the safety, capacity, and efficiency of the airspace operations in the terminal area.

We list the alternatives on the left in column A and possible options of implementing these alternatives in column B (also referred to as *modes*). The relevant criteria appear on the top in rows 2, 3, and 4. Stakeholders might require different types of assessments. For example, columns D, J, and K require qualitative mea-



Alternative Ratings for Airlines

OI Programs		Capacity 1.1 QL	Capacity 1.2 / 1.2.1 Unaccommodated Flights per year	Cost 2.1 / 2.1.1 Thous of Euros	Efficiency 3.1 / 3.1.1 %	Efficiency 3.1 / 3.1.2 %	Efficiency 3.2 / 3.2.1 Min per flight	Environment 4.1 / 4.1.1 QL	Environment 4.1 / 4.1.2 QL	Environment 4.2 / 4.2.1 # of people affected ('000)	Environment 4.2 / 4.2.2 # of people affected x # of events	Predictability 5.1 / 5.1.1 Delay variance	Safety 6.1 / 6.1.1 Accidents per 100,000 departures	Safety 6.1 / 6.1.2 Events per million movements	Safety 6.1 / 6.1.3 # of safety events per million flights
PRNAV	A0	0.75	200,000	€ 0	80%	77%	27	0.17	0.17	3500	38,150	361	0.0020	233	500
	A1	0.25	220,000	€ 368,640	84%	77%	24	0.83	0.83	2800	30,464	325	0.0019	233	475
AMAN & DMAN	B0	0.04	200,000	€ 0	80%	77%	27	0.08	0.20	3500	38,150	361	0.0020	233	500
	B1	0.10	180,000	€ 0	88%	77%	24	0.23	0.20	3500	38,220	325	0.0019	229	485
	B2	0.10	190,000	€ 0	80%	81%	27	0.23	0.20	3500	38,185	361	0.0020	229	500
	B3	0.24	180,000	€ 0	[illegible]	81%	24	0.23	0.20	3500	38,220	325	0.0019	229	485
	B4	0.51	180,000	€ 0	92%	85%	23	0.23	0.20	3500	38,220	309	0.0019	229	485
WV	C0	0.05	200,000	€ 0	80%	77%	27	0.36	0.20	3500	38,150	361	0.0020	233	500
	C1	0.25	188,000	€ 0	80%	77%	27	0.07	0.20	3500	38,192	[illegible]	0.0020	233	488
	C2	0.12	200,000	€ 211,400	80%	77%	27	0.33	0.20	3500	38,150	361	0.0019	233	475
	C3	0.48	188,000	€ 211,400	80%	77%	27	0.08	0.20	3500	38,192	361	0.0019	233	463
	C4	0.10	195,000	€ 500	80%	77%	27	0.16	0.20	3500	38,168	361	0.0019	233	485
TBS	D0	0.06	200,000	€ 0	80%	77%	27	0.63	0.33	3500	38,150	361	0.0020	233	500
	D1	0.27	195,000	€ 0	80%	77%	27	0.26	0.33	3500	38,150	361	0.0020	233	500
	D2	0.67	190,000	€ 0	80%	77%	27	0.11	[illegible]	3500	38,150	361	0.0020	233	500
B-CDA	E0	0.75	200,000	€ 0	80%	77%	27	0.17	0.17	3500	38,150	361	0.0020	233	500
	E1	0.25	220,000	€ 0	80%	77%	27	0.83	0.83	[illegible]	28,560	361	0.0020	233	500

Figure 3: Each stakeholder can assess each of the alternative OI programs on each of the criteria using qualitative and quantitative ratings.

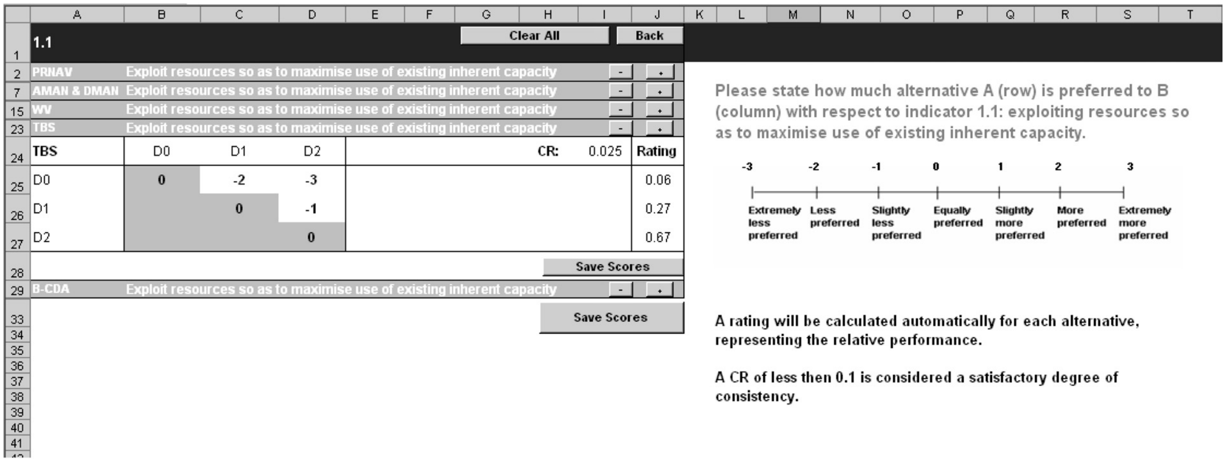


Figure 4: We use pairwise comparisons for qualitative ratings.

surements, whereas columns E, F, and G require quantitative measurements—a number of flights, thousands of euros, and percentages, respectively.

We obtain qualitative ratings using pairwise comparisons (Figure 4) and provide a scale to assist the stakeholders with expressing their preferences. When developing this scale with EUROCONTROL experts, we found that a numerical scale of −3 to 3, representing *Extremely less preferred* to *Extremely more preferred*, was more intuitive than the conventional 1–9 scale (Saaty 2005). The stakeholders enter the data in the upper triangle of the matrix; the model then automatically calculates the alternatives' ratings for each criterion.

We found that the process of rating the alternatives on the various criteria was the most time-consuming task in the project. Although we expected that the experts would be least comfortable using the pairwise comparisons, we found that several experts felt uncomfortable rating the alternatives on quantitative scales because of the lack of data available; we obtained qualitative assessments more easily. The general preference by users for using qualitative pairwise comparisons has long been accepted as the strength of qualitative techniques, such as the analytic hierarchy process (AHP) (Saaty 2005). On the other hand, some experts have raised serious doubts about the theoretical foundations of the AHP and about some of its properties, such as rank reversal (Dyer 1990). Thus, we did not use the AHP framework throughout but limited our use to pairwise comparisons when we solicited preferences among the combinations on only a subset of the criteria.

Assessing Interactions

The next step is to assess the performance of a combination of OI programs, rather than each program separately, on the different criteria. Although we can assess combinations of OI programs in a fashion similar to that used for the assessment of each program individually, this is likely to be very time consuming because the number of combinations could be huge. For example, for "Arrival and Departure with P-RNAV Support," there were 300 combinations obtained by combining the two options (or modes) for precision area navigation (A0 and A1 in Figure 3) with the five options for arrival and departure manager (B0, B1, B2, B3, B4), five options for wake vortex (C0, C1, C2, C3, C4), three options for time-based separation (D0, D1, D2), and two options for basic continuous descent (E0, E1). For other elements consisting of more OI programs and options, the number of combinations was substantially higher. Qualitative pairwise comparisons of program combinations would be an even more cumbersome task; the number of pairwise comparisons required would be $n * (n-1)/2$, with n being the number of combinations considered for each qualitative criterion.

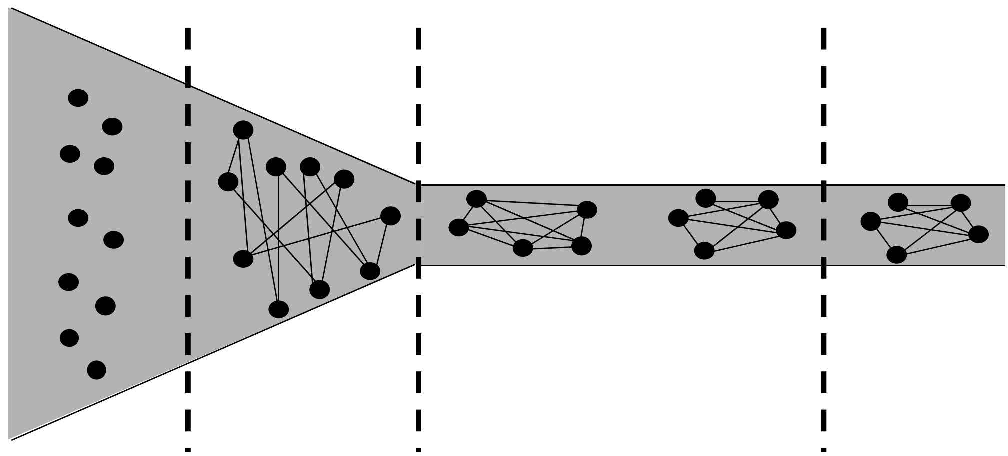
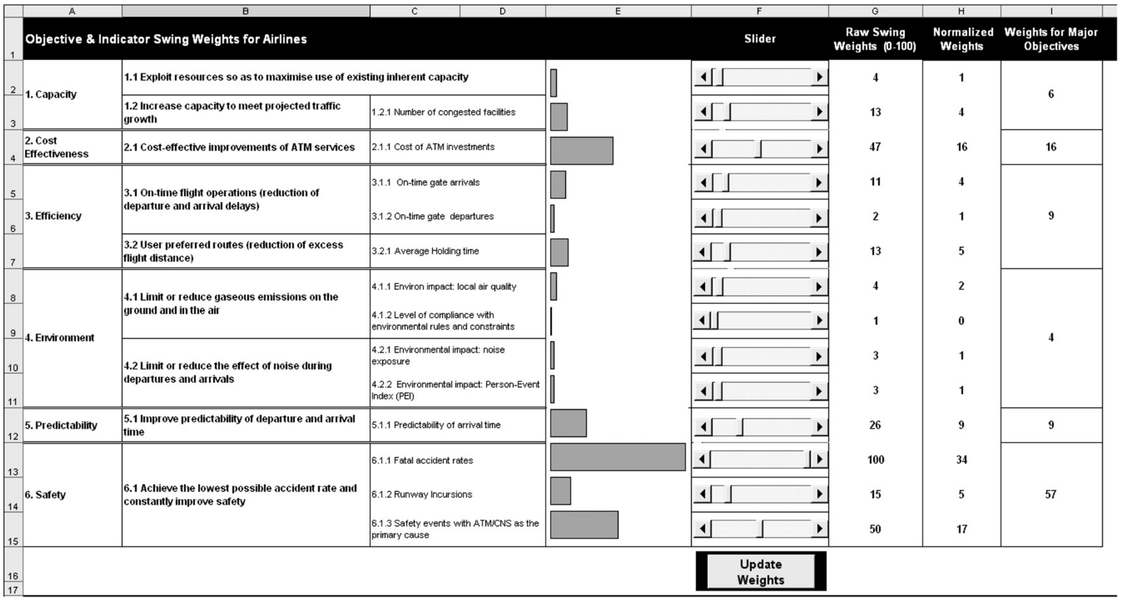


Figure 5: We use a phased approach for assessing the performance of combinations of OI programs. First, we use the separate impacts to estimate the joint impact using a set of heuristics, followed by a direct assessment on a subset of preferred combinations.

Although technically possible, it is unlikely that even experts can do this in a consistent manner. Hence, we use a phased approach to assess the performance of OI program combinations. First, we use the individual performance assessments of each OI program to automatically generate an estimate of the performance impact of each combination. Then, we revisit the assessments and amend the combinations' ratings as necessary (Figure 5).

To estimate the joint impact of a combination of OI programs, we analyze the nature of the interactions among the different OI programs. Because independence and additivity do not necessarily apply, we developed several heuristics to best describe these interactions.

- **Additive:** Experts deemed some OI programs to be independent with regard to some performance criteria; the result was a total combined effect equal to the sum of the expected performance of each individual OI program. For example, the cost of multiple OI programs was considered to be additive. Although this probably results in a slight overestimation of development costs, EUROCONTROL preferred this conservative approach to estimating costs.
- **Synergy:** Experts found some programs to be complementary; their combined effect was expected to be larger than the sum of the individual impacts. For example, the combined effect of implementing A1, B1, and D2 on on-time arrivals was expected to be synergistic. In one case, A1 was expected to deliver a 10 percent improvement, B1 a 7 percent improvement, and D2 a 3 percent improvement; the estimated joint effect would be 21.2 percent, a value that exceeds the sum of the individual effects.
- **Antagonism:** Some OI programs acted as substitutes; their joint impact was lower than the sum of the different effects. For example, a 10 percent improvement in capacity because of implementing B2, a 10 percent improvement as a result of implementing C2, and a 5 percent improvement as a result of implementing D1 result in an estimated total effect of only 21.5 percent.
- **Maximum:** In some cases, such as OI programs that are perfect substitutes, a combination of multiple programs will perform no better than the strongest



Objective & Indicator Swing Weights for Airlines			Raw Swing Weights (0-100)	Normalized Weights	Weights for Major Objectives
1. Capacity	1.1 Exploit resources so as to maximise use of existing inherent capacity		4	1	6
	1.2 Increase capacity to meet projected traffic growth	1.2.1 Number of congested facilities	13	4	
2. Cost Effectiveness	2.1 Cost-effective improvements of ATM services	2.1.1 Cost of ATM investments	47	16	16
3. Efficiency	3.1 On-time flight operations (reduction of departure and arrival delays)	3.1.1 On-time gate arrivals	11	4	9
		3.1.2 On-time gate departures	2	1	
	3.2 User preferred routes (reduction of excess flight distance)	3.2.1 Average Holding time	13	5	
4. Environment	4.1 Limit or reduce gaseous emissions on the ground and in the air	4.1.1 Environ impact: local air quality	4	2	4
		4.1.2 Level of compliance with environmental rules and constraints	1	0	
	4.2 Limit or reduce the effect of noise during departures and arrivals	4.2.1 Environmental impact: noise exposure	3	1	
		4.2.2 Environmental impact: Person-Event Index (PEI)	3	1	
5. Predictability	5.1 Improve predictability of departure and arrival time	5.1.1 Predictability of arrival time	26	9	9
6. Safety	6.1 Achieve the lowest possible accident rate and constantly improve safety	6.1.1 Fatal accident rates	100	34	57
		6.1.2 Runway Incursions	15	5	
		6.1.3 Safety events with ATM/CNS as the primary cause	50	17	

Figure 6: Stakeholders can modify the criteria weights by manipulating the assigned swing weights. Columns A through D list the objectives and criteria. Column F contains sliders to allow stakeholders to amend the weights.

improvement among them. For example, we expect that implementing A1, expected to reduce noise pollution by 7 percent, together with E1, expected to reduce noise pollution by 15 percent, will result in a total reduction of only 15 percent in noise pollution.

To ensure that the performance impact is comparable across the various criteria, we use a value function (Keeney and Raiffa 1976) to represent the stakeholders' preferences on a common scale. Once we assign a rating to each combination, we normalize these ratings using linear value functions (Dyer et al. 1998), although the model can also support nonlinear value functions.

Setting Priorities

In the fifth stage of the process, we determine a weight for each criterion to represent the subjective importance of the objectives for each stakeholder. Because the relative importance associated with a specific criterion is sensitive to the variability in values that a criterion can take (Keeney 2002), the model requires complete ratings of the alternatives on the different criteria prior to assigning weights.

The model uses two techniques to solicit the weights: pairwise comparisons and swing weights (Goodwin and Wright 2004). We first assign weights using pairwise comparisons. We then validate, and if necessary modify, the weights in the "Objective & Indicator Swing Weights for Airlines" section of the model (Figure 6).

This validation aims to overcome some concerns regarding the use of qualitative pairwise comparisons by providing a graphical presentation of the importance of each criterion; this stimulates additional discussion about the accuracy of these assessments. If desired, a stakeholder can also modify the weights by manipulating the swing weights directly.

The boxplot in Figure 7 illustrates the range of weights assigned to a set of performance criteria by several stakeholders for one specific element. Typically, some ranges are wider than others, allowing us to identify main sources of disagreement and commonalities among the stakeholders.

Determining a Recommended Combination

When recommending a particular combination of OI programs for implementation, we need to find a compromise that is acceptable to all stakeholders. The rec-

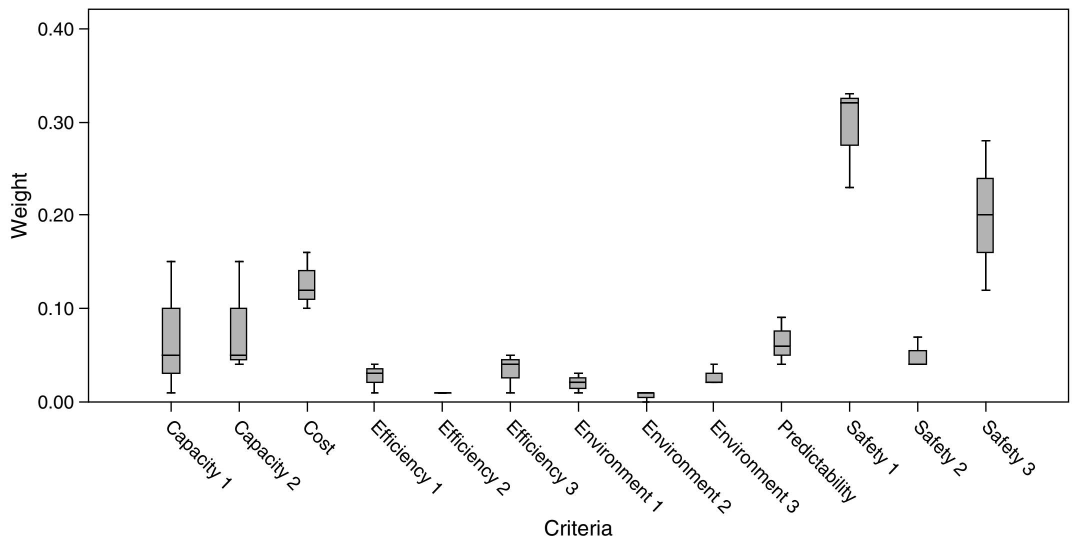


Figure 7: Boxplots show the range of weights assigned to criteria by stakeholders.

ommendation must take into account that stakeholders might have different opinions about the expected performance impact of each OI program and different priorities about the importance of each objective.

First, we determine the preferred—or *optimal*—combination of OI programs for each stakeholder separately by using separate impact assessments, value functions, and priority weights. We then derive an overall performance score and rank for each combination of programs according to each stakeholder's perspective (Keeney and Raiffa 1976).

Because of the different perspectives, a combination that is optimal for one stakeholder might not be optimal for another. Thus, to ensure that our recommendation is acceptable to each stakeholder, and in support of the SES master plan harmonization activities, we define the maximization of the anticipated performance score over all the stakeholders as the overall objective. We can achieve this by maximizing a weighted average of the anticipated performance scores, where the weights, which a steering committee headed by EUROCONTROL determines, are used to represent trade-offs among stakeholders. To ensure that no stakeholders feel that they could improve their situation if a different combination of OI programs was selected, the objective of minimizing the maximum regret (French 1986) could also be used.

Performance targets for each stakeholder are determined either exogenously by European regulators or endogenously by the stakeholders themselves. For example, the current level of accident risk was set as a performance target for the future level of safety. Targets are typically set as a minimum or maximum percentage of change from the status quo. We model the OI selection problem as an integer programming model (Figure 8). Grushka-Cockayne et al. (2008) provide model details.

The downside of this model is that it can grow enormously, depending on the number of available combinations of OI programs. When the problem size becomes intractable, a column-generation approach can be used: a small subset of OI program combinations, which serve as the columns, is collected in a master program whose dual prices are used by subproblems to evaluate new combinations. We continue this process until we find a solution to the LP relaxation of the master, upon which we use a branch-and-price procedure (Barnhart et al. 1998) to find the optimal integer solution.

We decompose the problem by constructing subproblems for each stakeholder separately. The selected combination of OI programs in each subproblem provides insight into potential disagreements among the stakeholders concerning the preferred course of

Maximize the overall performance score of the selected combinations of OI programs, weighted by the stakeholders' importance weights, or minimize the maximum regret.
Subject to:
(1) Ensuring that the performance targets are achieved;
(2) A limited available budget of each stakeholder;
(3) Enforcing the selection of a single combination;
(4) The selection (or not) of each combination (binary choice).

Figure 8: Using an integer programming model, we determine the combination of OI programs that best meets the stakeholders' preferences.

action. The master program's objective is to maximize the overall performance score of the selected combination across the stakeholders, subject to ensuring that all stakeholders achieve the global performance targets and that an overall compromise is reached. We enforce the latter in the master problem by means of coordination constraints. In the subproblems, we consider meeting the stakeholders' performance targets and ensuring that the stakeholders do not exceed their individual budgets with each subproblem minimizing the reduced cost. The constraints in the subproblem ensure that the stakeholders (1) meet the performance targets, (2) do not exceed their budget, and (3) select only a single OI program combination.

Degraeve (1992) suggests that adding a fixed number of potentially "good" columns initially will result in excellent integer solutions. Specifically for our case, the initial columns could consist of a fixed number of low-cost OI program combinations for each stakeholder that achieves the desired performance targets.

Sensitivity Analysis

To ensure the robustness of our recommendation, we included advanced sensitivity analyses methods in the model. First, *intrastakeholder* analyses examine the sensitivity of the recommendations to the criteria weights assigned by each stakeholder and to the performance impact assessments made by the stakeholders. Second, *interstakeholder* analyses investigate the robustness of the recommendations when the weights assigned to the stakeholders, reflecting their importance or bargaining power, are varied.

Many established multicriteria decision analysis methods focus on a series of one-dimensional sensitivity analyses to establish the robustness of recommendations. Although these analyses do provide insights, they might also be misleading because they ignore the potential interaction that could result from simultaneously manipulating multiple uncertainties. Therefore, we use a random weights simulation technique in which we randomly generate the criteria weights assigned by the stakeholders according to a specific distribution (Butler et al. 1997). The boxplot in Figure 9 shows the results of such an analysis and highlights the resulting variability in the rankings of considered OI combinations.

We represent high rankings toward the bottom of the boxplot, i.e., closer to the x-axis, with rank 1 as the highest. We recommend that combinations that might be ranked first under certain combinations of criteria weights should not be excluded from the discussions because a stakeholder might prefer them under certain conditions.

A simulation analysis also allows us to provide recommendations in the absence of complete information about the expected performance impact of some OI programs. Therefore, we use ranges rather than single values to capture the anticipated impact. In addition, we found that significant uncertainty exists because of lack of availability and disclosure of data by the stakeholders. Past experience has shown that when assessing the performance of system enhancements, some stakeholders prefer to provide ranges of expected costs and benefits to avoid the disclosure of strategic information. Figures 10 and 11 illustrate the running of a simulation while randomly drawing the performance impact ratings from the ranges provided.

Figure 10 identifies the top-ranking clusters according to all stakeholders. Figure 11 shows the frequency of a specific OI program combination scoring the highest from a single stakeholder's perspective. For example, according to the airlines' ratings and weights, combination 140 is preferred in over half of the simulation runs. Combination 290 is preferred in approximately 15 percent of the cases. Thus, simulation analyses provide additional information regarding the value of the alternatives that might affect the recommendations.

Because of political concerns, it is often difficult to assign weights to the stakeholders as the calculation of the objective function value in our model requires. Using robust portfolio modeling methodology (RPM) (Liesiö et al. 2007), we can model this as a case of incomplete information regarding the

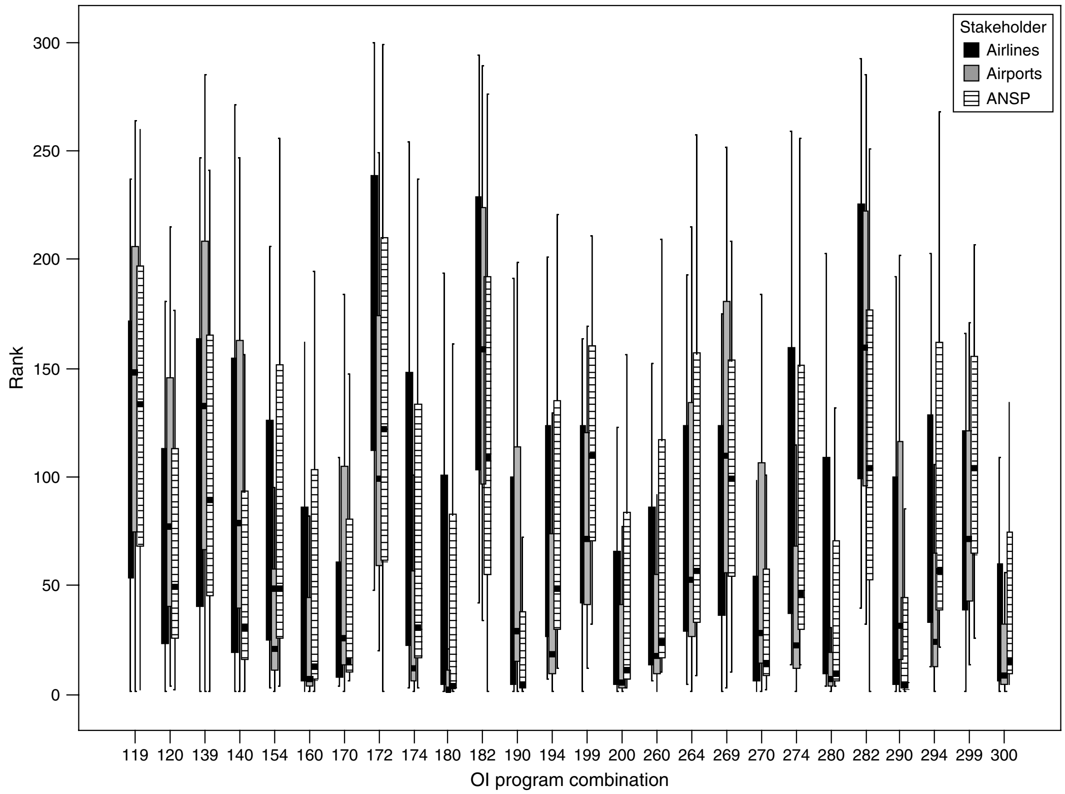


Figure 9: Boxplots indicate the robustness of scores assigned by airlines, airports, and air navigation service providers (ANSP) to each OI program combination when criteria weights are varied using a simulation analysis.

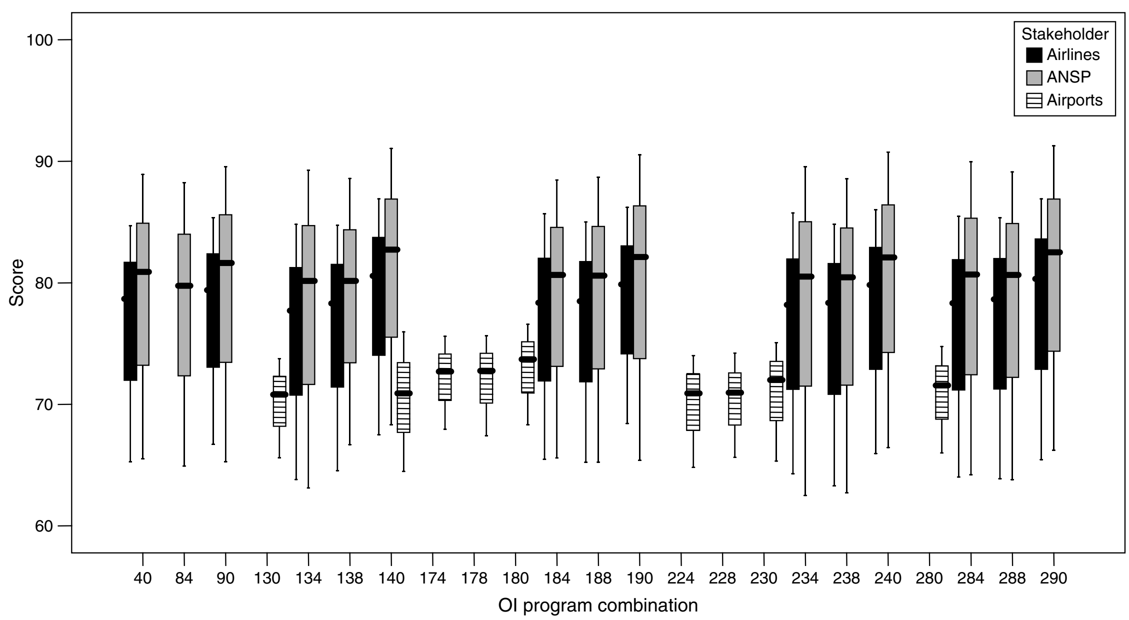


Figure 10: Boxplots visualize the variability in OI program combinations' scores when the expected performance impact is varied using a simulation analysis.

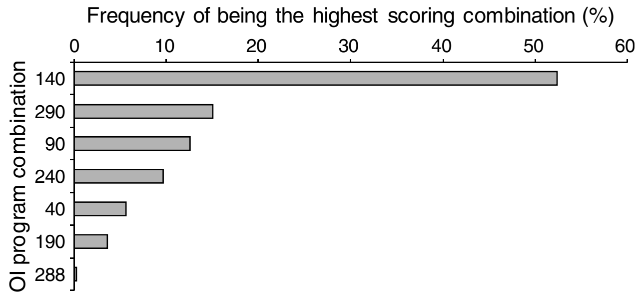


Figure 11: Graphs show how frequently a specific OI program combination obtained the highest score using a simulation analysis.

OI program	*CI*	Result	OI program	*CI*	Result
A0	0.00	Exterior	C2	0.40	Borderline
A1	1.00	Core	C3	0.60	Borderline
B0	0.00	Exterior	C4	0.00	Exterior
B1	0.30	Borderline	D0	0.33	Borderline
B2	0.00	Exterior	D1	0.30	Borderline
B3	0.30	Borderline	D2	0.37	Borderline
B4	0.40	Borderline	E0	0.43	Borderline
C0	0.00	Exterior	E1	0.57	Borderline
C1	0.00	Exterior			

Table 4: Programs with a core index of 1, e.g., implementing P-RNAV (option A1), appear in all nondominated combinations; programs with a core index of 0 do not appear in any nondominated combinations.

stakeholder weights. We can use incomplete or partial information about these weights to exclude certain combinations of OI programs that will never be preferred even if more information becomes available at some time. Such partial information could, for example, specify that one stakeholder is more important than another without the need for specifying how much more important, a value that might be difficult to determine. Grushka-Cockayne et al. (2008) provide more details on the application of this methodology.

Without any assumptions on stakeholder importance whatsoever, we are typically able to exclude many alternatives. For example, for the program on "Arrival and Departure with P-RNAV Support," we were able to reduce the set of alternatives to 30 nondominated combinations (Table 3).

Borrowing the terminology provided by Liesiö et al. (2007), we define the *core index* (*CI*) of an OI program as the proportion of nondominated combinations that contains that program. Programs with $CI = 1$ are robust choices in the sense that they would definitely be recommended, even if additional information were to become available. *Exterior* programs, where $CI = 0$, can safely be rejected because they will never be included in the set of nondominated combinations, despite potentially new information. *Borderline* programs in which $0 < CI < 1$ will require further analysis. The *CI* results for "Arrival and Departure with P-RNAV Support" are shown in Table 4.

We can use a three-dimensional scatter plot (Figure 12) to combine the perspectives of the different stakeholders to highlight efficient solutions, i.e., combinations that are not dominated by any other. Typically, only a small subset of combinations dominates, thus allowing the discussion to focus on those options.

No.	Combination	OI programs	No.	Combination	OI programs
1	30	A1 B4 C2 D0 E0	16	180	A1 B4 C2 D0 E1
2	34	A1 B1 C3 D0 E0	17	184	A1 B1 C3 D0 E1
3	38	A1 B3 C3 D0 E0	18	188	A1 B3 C3 D0 E1
4	40	A1 B4 C3 D0 E0	19	190	A1 B4 C3 D0 E1
5	80	A1 B4 C2 D1 E0	20	228	A1 B3 C2 D1 E1
6	84	A1 B1 C3 D1 E0	21	230	A1 B4 C2 D1 E1
7	88	A1 B3 C3 D1 E0	22	234	A1 B1 C3 D1 E1
8	90	A1 B4 C3 D1 E0	23	238	A1 B3 C3 D1 E1
9	124	A1 B1 C2 D2 E0	24	240	A1 B4 C3 D1 E1
10	130	A1 B4 C2 D2 E0	25	274	A1 B1 C2 D2 E1
11	134	A1 B1 C3 D2 E0	26	278	A1 B3 C2 D2 E1
12	138	A1 B3 C3 D2 E0	27	280	A1 B4 C2 D2 E1
13	140	A1 B4 C3 D2 E0	28	284	A1 B1 C3 D2 E1
14	174	A1 B1 C2 D0 E1	29	288	A1 B3 C3 D2 E1
15	178	A1 B3 C2 D0 E1	30	290	A1 B4 C3 D2 E1

Table 3: Using robust portfolio modeling, we can typically reduce the set of available options by more than 90 percent by focusing on nondominated combinations.

Implementation

The definition phase of the SES ATM Research Program (SESAR) has now been completed and the development phase (2008–2013) will soon commence. The SESAR consortium has been using the decision-making framework we developed throughout the definition phase, i.e., during the initial construction of the master plan, based on guidelines described in De Reyck et al. (2006). The stakeholders will continue

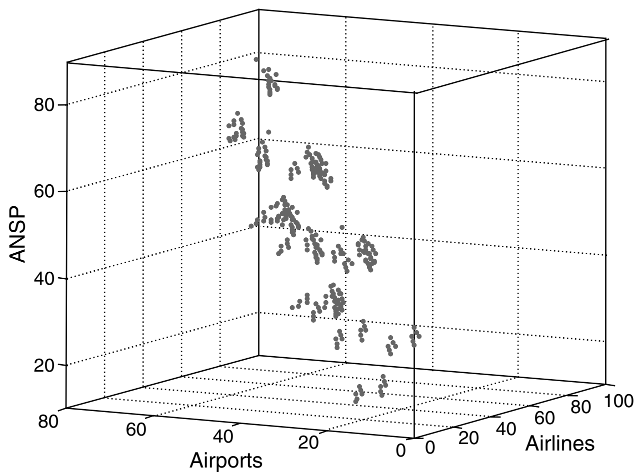


Figure 12: A three-dimensional perspective of the performance of the alternative OI program combinations allows us to focus on efficient combinations.

to use the framework during the development phase to support the decision making throughout the life cycle of the OI programs, from research and development to production and implementation efforts. During the SESAR definition phase, the framework was also used in cases with only a single option to be assessed, whereby the output from the framework served as a checklist to ensure that the option met all the stakeholder criteria.

Figure 13 shows the activities involved in the facilitation process and the nine stages of the decision-making framework.

A strategy team carries out the initial framing of each OI program evaluation and selection problem; the team includes members of EUROCONTROL serving as facilitators—and stakeholder representatives—taking part in interviews, workshops, and data collection activities. The framing task, which is initiated during a kickoff meeting, begins with identifying the decision context and the stakeholders involved. During this meeting, participants also identify the alternatives; they highlight the OI programs to be considered and examine the feasibility of alternative combinations of these programs.

The facilitator brings the kickoff meeting to a close by identifying the relevant experts and data sources that will serve for aggregating the OI programs' ratings and the criteria weights. Following this, we carry

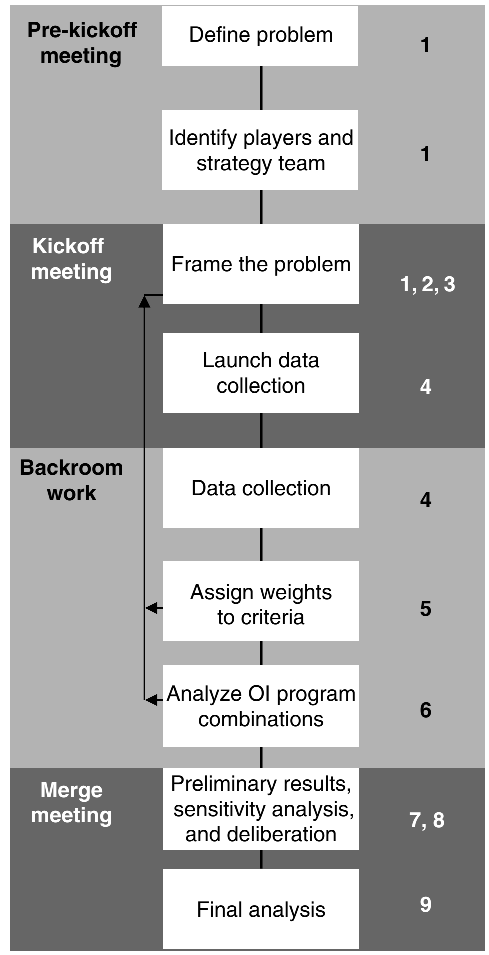


Figure 13: The facilitation process consists of a series of meetings combined with backroom work. We indicate the corresponding decision-making framework stage on the right.

out backroom work (Belton and Stewart 2002), collect all necessary data, validate the model assumptions, and identify heuristics for characterizing the OI program interactions. The majority of the backroom work is done via one-on-one interviews or workshops with a small team of experts. We found that it is important to maintain transparency by monitoring and documenting every source of data used in the model, because this will aid in justifying the model inputs, which the stakeholders might challenge.

A merge meeting (Phillips and Bana e Costa 2007) is typically necessary to reach a stable consensus and a joint commitment about the way forward by all stakeholders. The involved parties will view the results of the analysis as either confirming their intuition or surprising. In each case, they can gain much value from the discussions that happen in the merge meeting in terms of developing a shared understanding.

Integration with Existing Systems

The SESAR development phase will include the integration of our decision-making framework with EMOSIA (the European Model for Strategic ATM Investment Analysis), the European approach for ATM cost-benefit analysis (Purves et al. 2008). With EUROCONTROL, we have tailored the integrated approach to support decisions on the OI programs throughout their life cycles.

EMOSIA follows a six-stage process that is similar to our decision-making framework. However, it focuses exclusively on quantitative assessments for cost-benefit analyses. As such, our OI program evaluation and selection model is complementary to EMOSIA. Adding a multistakeholder dimension and allowing for qualitative assessments, our method will help overcome some issues such as the reluctance of stakeholders to quantify expected performance; thus, it delays the entire assessment of the OI programs to later stages. In the early phases of the OI programs' life cycle, our decision-making framework assumes a leading role, benefiting from its ability to integrate qualitative and quantitative data on a high level and to filter alternatives that do not meet certain criteria. EMOSIA is then used to assess the costs and benefits of the OI programs as they progress, and as uncertainty decreases and the data become more quantitative (Figure 14).

Frequent exchanges of information and recommendations between the two methods are fundamental because the information required for the selection of the OI programs is different throughout the program life cycle. For example, there is no reason to reduce uncertainty if the OI program evaluation and selection model indicates that additional information is unlikely to change the recommendation. Thus, the decision-making framework we have developed provides EMOSIA with a clear definition of the criteria and filtered alternatives for further analysis; it also identifies the key benefits of the filtered alternatives.

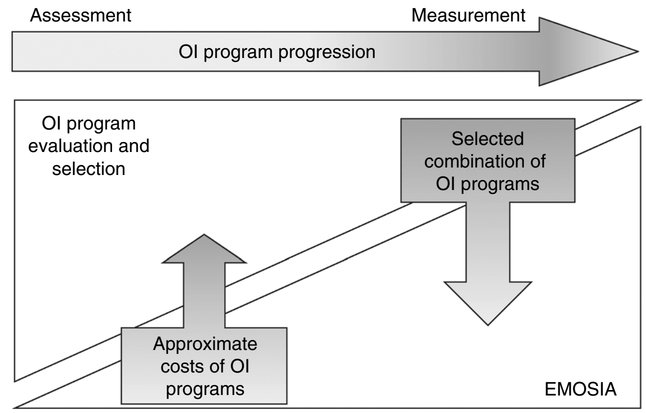


Figure 14: Our OI program evaluation and selection model is integrated with the EMOSIA cost-benefit analysis during the OI program life cycle.

Impact Beyond Europe

Whereas European ATM systems are undergoing modernization efforts as part of the SES initiative, the US Joint Planning and Development Office (JPDO) program is also proposing ambitious modernization goals for its domestic ATM systems to cope with the limited system capacity in the face of continuously increasing demand for travel. Mozdzanowska et al. (2006) identify and describe the dynamics of the air transportation system transition and processes for reviewing and implementing new system capabilities. To implement the significant changes currently envisioned for the US ATM systems, JPDO has recognized that it will be critical to structure system changes in such a way as to anticipate and overcome stakeholder disagreements and improve the efficiency of the approval and implementation processes. Our framework complements that presented by Mozdzanowska et al. (2006) by providing a decision-making framework for supporting the negotiation loop, structuring the stakeholders' preferences, and supporting the decision making (Figure 15).

Thus, the decision-making framework we have presented in this paper has made a positive contribution to EUROCONTROL and to the European aviation community by allowing EUROCONTROL to resume its visionary role as facilitators in European aviation decision making. In addition, it can benefit global players such as the US Federal Aviation Administration.

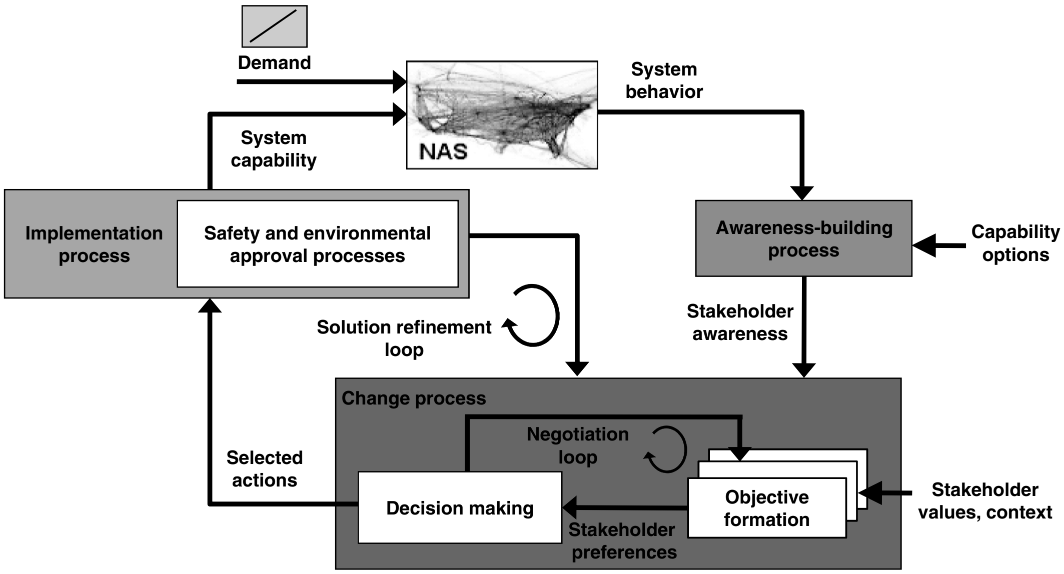


Figure 15: Our framework can support the negotiation loop in the US transition in the air transportation system model.

Acknowledgments

We gratefully acknowledge Zeger Degraeve from the London Business School, and Patricia Cauwenbergh, Robert Graham, Henk Hof, Paula Leal De Matos, and Peter Eriksen from EUROCONTROL for their cooperation and assistance throughout this project.

References

Barnhart, C., E. L. Johnson, G. L. Nemhauser, M. W. P. Savelsbergh, P. H. Vance. 1998. Branch-and-price: Column generation for solving huge integer programs. *Oper. Res.* **46**(3) 316–329.

Belton, V., T. J. Stewart. 2002. *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishers, Boston.

Butler, J., J. Jia, J. Dyer. 1997. Simulation techniques for sensitivity analysis of multi-criteria decision models. *Eur. J. Oper. Res.* **103**(3) 531–546.

Degraeve, Z. 1992. Scheduling joint product operations with proposal generation methods. Doctoral dissertation, Graduate School of Business, The University of Chicago, Chicago.

De Reyck, B., Z. Degraeve, Y. Grushka-Cockayne. 2005. System performance trade-offs: Survey of state-of-the-art methods and tools. Technical report, EUROCONTROL, Brétigny-sur-Orge, France.

De Reyck, B., Z. Degraeve, Y. Grushka-Cockayne. 2006. Decision support using performance driven trade-off analysis. Technical report, EUROCONTROL, Brétigny-sur-Orge, France.

Dyer, J. S. 1990. Remarks on the analytic hierarchy process. *Management Sci.* **36**(3) 249–258.

Dyer, J. S., T. Edmunds, J. C. Butler, J. Jia. 1998. A multiattribute utility analysis of alternatives for the disposition of surplus weapons-grade plutonium. *Oper. Res.* **46**(6) 749–762.

EUROCONTROL. 2006. Europe addressing capacity crunch. Press release (January 18).

French, S. 1986. *Decision Theory: An Introduction to the Mathematics of Rationality*. John Wiley & Sons Australia, Brisbane, Australia.

Goodwin, P., G. Wright. 2004. *Decision Analysis for Management Judgment*. John Wiley & Sons, Chichester, UK.

Grushka-Cockayne, Y., B. De Reyck, Z. Degraeve. 2008. An integrated decision-making approach for improving European air traffic management. *Management Sci.* **54**(8) 1395–1409.

Keeney, R. L. 2002. Common mistakes in making value trade-offs. *Oper. Res.* **50**(6) 935–945.

Keeney, R. L., H. Raiffa. 1976. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. John Wiley & Sons, New York.

Liesiö, J., P. Mild, A. Salo. 2007. Preference programming for robust portfolio modeling and project selection. *Eur. J. Oper. Res.* **181**(3) 1488–1505.

Mozdzanowska, A., R. Weibel, E. Lester, R. J. Hansman. 2006. The dynamics of air transportation system transition. Massachusetts Institute of Technology, Cambridge.

Phillips, L., C. A. Bana e Costa. 2007. Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. *Ann. Oper. Res.* **154**(1) 51–68.

Phillips, L., A. Stock. 2003. Use of multi-criteria analysis in air policy. Report, Department for Environment, Food and Rural Affairs, London.

Purves, K., M. Gibellini, P. L. de Matos, B. Lacroix. 2008. An integrated approach to decision making in air traffic management. Working paper, EUROCONTROL, Brétigny-sur-Orge, France.

Saaty, T. L. 2005. The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision making. J. Figueira, S. Greco, M. Ehrgott, eds. *Multiple Criteria Decision Analysis, State of the Art Surveys*. Springer Science + Business Media, New York, 345–408.