EU-wide target ranges for RP 3

Annex 4. Results of the clustering analysis

12 June 2018

Performance Review Body of the Single European Sky | Rond-Point Schuman 6, 4th Floor, Office 411, B-1040 Brussels
Office Telephone: +32 (0)2 234 7824 | prb-office@prb.eusinglesky.eu | webgate.ec.europa.eu/eusinglesky

steer davies gleave

Memo

To	European Commission
Cc	SDG project team
From	Clémence Routaboul
Date	12 June 2018
Project	Support study on union-wide target setting for RP3
Project No.	23261901

Results of the clustering analysis (for the Ranges Report)

Approach

Introduction

1. The primary goal of the clustering exercise described here is to inform further investigation of the interdependency between capacity and cost efficiency. As a first step, we have sought to control for the impact of exogenous factors that are largely or entirely beyond the control of ANSPs (and outside the control of policy-making/regulatory/advisory bodies such as the Commission, National Supervisory Authorities (NSAs) and the PRB respectively, at least in the short term).

2. This requires identification of factors that can only be changed over the long term, if at all, such as the size and complexity of airspace, allowing to focus on differences in industry performance that are explained by endogenous factors, including investment and innovation. This could mean that different targets could be set for different clusters, recognising that all ANSPs do not face the same environment. This would inform the PRB so that targets can be set based on factors that ANSPs have control upon.

Clustering variables used

3. We reviewed possible exogenous variables with the aim of identifying the most relevant ones. We present here the range of factors we considered, including some that could be classified as endogenous (within ANSP control) or exogenous (beyond ANSP control) to different degrees:

Table 1: Control of ANSPs on airspace factors

Type of variable	Factors	Comment
Endogenous	ATC procedures, technology	
Exogenous	Traffic volumes, seasonality of traffic, weather, number of airports, airspace size, military areas, interface with adjacent units, traffic characteristics (traffic density, traffic in climb or descent, flow structure, traffic mix), cost of living, ATCO-employment costs per ATCO hour	Traffic characteristics is mostly exogenous, as quantified by the Complexity Score, but influenced by the organisation of airspace and air traffic management procedures Within the time-frame of a 5-year reference period, the key drivers of ATCO-employment costs could be considered as exogenous, even if over the long term this variable is within the control of ANSPs.
Mix of both	Airspace organisation, route structure, sectorisation, ATCO profile, level of existing over or under capacity	

Source: Steer Davies Gleave analysis

4. We have sought data representing a range of factors from the PRU, STATFOR, the Network Manager (NM), and other parties, and will refine the analysis as and when further data is received.

5. At this stage, we have not been able to explore the implications of using all these different factors in the analysis due to data availability and the difficulties of isolating purely exogenous factors when variables combine both endogenous and exogenous elements. However, we have been able to repeat the clustering analysis undertaken by the PRU for RP2 using the latest available data (at the time of the analysis this was 2015 data), which enables a clearer comparison of the latest results for RP3 with those from the RP2 analysis.

6. The variables and data sources used for the analysis are presented in the Table below.

Table 2: Summary of variables included in the clustering analysis

Variables included	Source of data set
Traffic volume	ACE data
Traffic complexity	Airspace complexity score from EUROCONTROL
Traffic variability	STATFOR daily traffic data
Cost of living index	GDP (IMF), Eurostat (PPP), ACE data (ANSP exchange rates)
Unit ATCO employment cost	ACE data

Source: Steer Davies Gleave analysis

Clustering technique

7. We consider that the most appropriate clustering technique is multi-dimensional analysis[1], combined with expert review. This balances the need for analytical rigour with the flexibility needed to ensure that the defined clusters appear reasonable to a wider audience, including industry stakeholders.

8. Cluster analysis is based on data covering all the variables for the same unique year. It is important to ensure that the data is as recent as possible, so that the results reflect the current situation rather than the past (although we note that it is also important to avoid allocating ANSPs to clusters based on what might be one-off events in a given year (which could influence certain exogenous as well as endogenous factors). Given the available data, we selected 2015 as the reference year for the analysis.

9. The range of values between variables (e.g. total traffic volumes in the '000s and complexity scores as single digit scores) can be particularly wide, potentially distorting the results (with the largest values having a stronger impact than the smaller ones). We therefore normalised the variables, standardising them relative to the maximum value within the data-set range to ensure that each had an equal weighting within the cluster analysis algorithm.

10. We also applied expert judgment and excluded a number of ANSPs from the analysis:

- Due to the unique nature of its airspace (upper airspace only, across four Member States), we excluded MUAC; and

1 More precisely, we specified a clustering approach based on characteristics as follows: hierarchical (the set of clusters generated is nested and organised as a tree. Each node (cluster) in the tree is the union of its branches (sub-clusters) and the root of the tree is the cluster containing all the objects (in this case ANSPs)), exclusive (individual ANSPs cannot belong to more than one cluster) and complete (every ANSP is assigned to a cluster).

- We excluded BelgoControl and LVNL, as both ANSPs provide lower airspace services only.

Results of the multi-dimensional clustering analysis

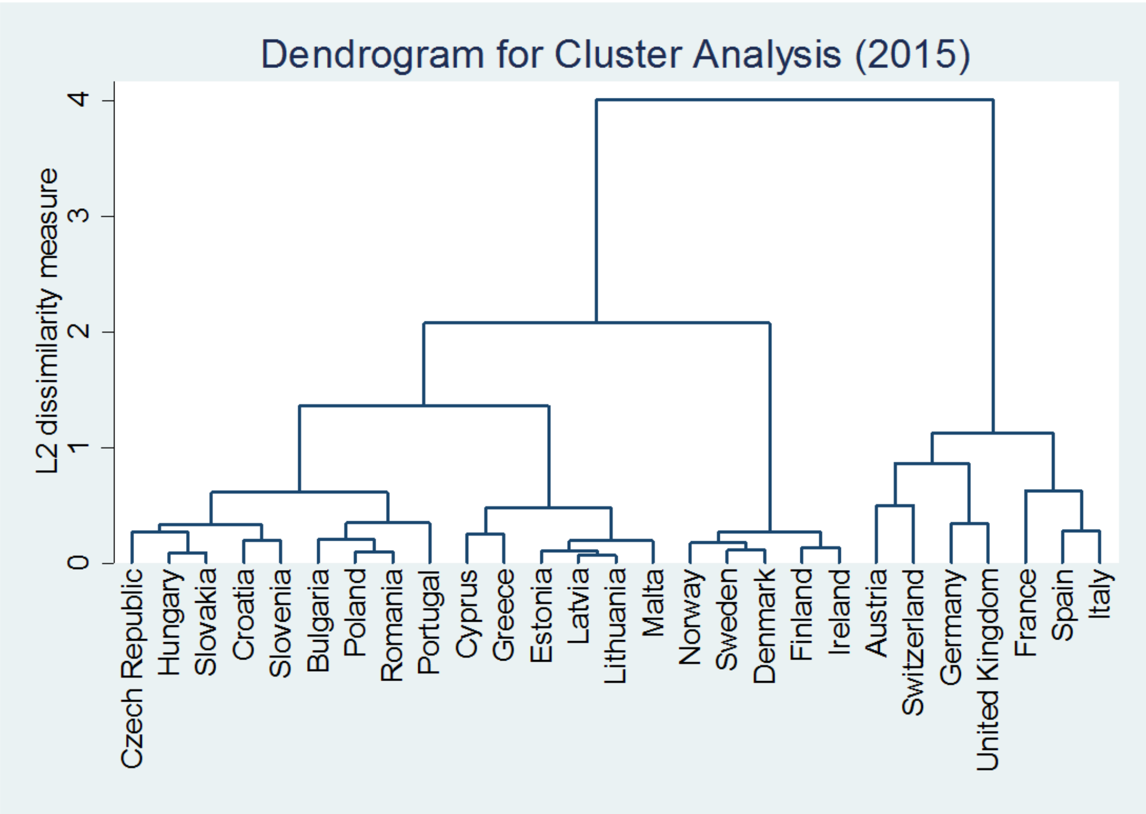
11. For intelligibility, results for RP3 are labelled by Member State, rather than by ANSP.

12. The clusters obtained may be categorised as follows:

- Cluster 1: Austria, Switzerland, Germany, United Kingdom, France, Spain and Italy: a grouping of the largest and usually most complex airspace;
- Cluster 2: Norway, Sweden, Denmark, Finland and Ireland: a grouping with little variability in traffic, limited complexity, and an average level of traffic;
- Cluster 3: Czech Republic, Hungary, Slovakia, Croatia, Slovenia, Bulgaria, Poland, Romania and Portugal: a grouping with a relatively high level of complexity and traffic variability; and
- Cluster 4: Cyprus, Greece, Estonia, Latvia, Lithuania and Malta: a grouping with limited complexity and lower traffic, but with sometimes high variability.

13. The figure below shows a dendrogram, a standard representation of the results obtained from multi-dimensional clustering analysis. The vertical lines provide an indication of how dissimilar two Member States are in terms of the calculated multivariate measure. For example, the vertical height of the connection between Latvia and Lithuania is relatively small, indicating that these two Member States are similar, while the vertical height of the connection between Austria and Switzerland is much greater, indicating less similarity.

Figure 1: Dendogram



Source: Steer Davies Gleave analysis

Conclusions and next steps

14. These results are preliminary results from the clustering analysis performed by Steer Davies Gleave. Additional data could also be usefully considered. The results of the clustering analysis show how Member States could be grouped in clusters so that targets for RP3 performance can be designed with consideration towards the different contexts that ANSPs operate in. The clusters can also be used for benchmarking ANSP performance in a fair manner.

15. A next stage to the target-setting exercise is to now further investigate the interdependencies between capacity and cost efficiency. In principle, the results of the clustering analysis can be used in combination with the results of Data Envelopment Analysis (DEA), to make meaningful comparisons between ANSPs in terms of capacity and cost efficiency. More specifically, the identification of efficient ANSPs handling different levels of traffic and causing different levels of delay would allow us to examine trade-offs between cost-efficiency and capacity along an estimated efficiency frontier.

16. We note that further DEA and Stochastic Frontier Analysis (SFA) has been undertaken in parallel to the clustering work by a group of Academics commissioned by the PRB, and we will consider how this analysis and the results of any further work might be combined with the clustering results to draw conclusions on interdependencies. In addition, we will undertake an analysis of recent trends in cost efficiency, traffic handled and delay across all ANSPs to further investigate the relevant relationships and trade-offs.

steer davies gleave

Memo

To	European Commission		
Cc	SDG project team		
From	Clémence Routaboul		
Date	214 June 2018		
Project	Support study on union-wide target setting for RP3	Project No.	23261901

Technical note T2.D1 - Preliminary results of the clustering analysis

Introduction

1. This note provides an overview of the analysis to identify groupings or clusters of ANSPs as part of the preparatory work for target-setting for Reference Period 3 (RP3) of the Single European Sky (SES) Performance Scheme. These groupings will be used in two ways:

- to inform the determination of EU-wide targets for RP3; and
- to benchmark the projected performance of individual ANSPs, as submitted in their Performance Plans for RP3.

Approach

Introduction

2. The Performance Review Unit (PRU) applied similar analysis to inform target setting for the cost efficiency Key Performance Area (KPA) in RP1 (using 2011 data) and in RP2 (using 2015 data). The purpose of the analysis was to define groups of ANSPs with similar operational characteristics to enable meaningful static benchmarking of performance against a definition of "best in class". A two-step approach was used by the PRB, involving cluster analysis and expert judgement. However, the analysis focused on cost-efficiency, and the interdependency between cost efficiency and other KPAs was not investigated.

3. The primary goal of the clustering exercise described here is to inform further investigation of the interdependency between capacity and cost efficiency. As a first step, we have sought to control for the impact of exogenous factors that are largely or entirely beyond the control of ANSPs (and outside the control of policy-making/regulatory/advisory bodies such as the Commission, National Supervisory Authorities (NSAs) and the PRB respectively, at least in the short term). This requires identification of factors that can only be changed over the long term, if at all, such as the size and complexity of airspace, allowing us to focus on differences in industry performance that are explained by endogenous factors, including investment and innovation. In analytical terms, the aim is to minimise the similarity between the clusters while maximising the similarity between ANSPs within each cluster (with similarity defined by reference to exogenous factors).

Clustering variables used

4. We reviewed possible exogenous variables with the aim of identifying the most relevant ones. The table below identifies a range of factors that we considered, including some that could be classified endogenous or exogenous to different degrees. In practice, differentiating between endogenous and exogenous factors is challenging, and we have not yet obtained data for all of the factors listed. We have nevertheless sought data representing a range of factors from the PRU, STATFOR, the Network Manager (NM), and other parties, and will refine the analysis as and when further data is received.

Table 1: Control of ANSPs on airspace factors

Factors	Endogenous	Exogenous	Mix of both	Comment
ATC procedures	✓			Mostly endogenous, influenced by policy, safety standards, letters of agreement
Technology	✓			Mostly endogenous, but legacy systems and lead-times relevant in the context of potential flexibility during a Reference Period.
Airspace organisation			✓	Arguably mostly endogenous, but influenced by implementation of Free Route Airspace, interfaces with adjacent units, airspace delegation, airports, terrain, noise abatement requirements etc.
Route structure			✓	
Sectorisation			✓	
Traffic characteristics - Traffic density - Traffic in climb or descent - Flow structure - Traffic mix		✓		Mostly exogenous, as quantified by the Complexity Score, but influenced by the organisation of airspace and air traffic management procedures.
Military areas		✓		Exogenous
Interface with adjacent units		✓		Exogenous, transition between standards.
Seasonality		✓		Exogenous traffic variability and predictability.
Airspace size		✓		Exogenous
Number of airports		✓		Exogenous
Traffic volume		✓		Exogenous, particularly where large shifts in traffic patterns are observed
Weather		✓		Exogenous
ATCO profile			✓	Mostly endogenous, but age and skill profile of ATCOs relevant in the context of potential flexibility during a Reference Period.
Level of existing over- or under-capacity			✓	Mostly endogenous, but starting position between available capacity and traffic relevant in the context of potential flexibility during a Reference Period.

Source: Steer Davies Gleave analysis

5. In practice, we have not been able to explore the implications of using all these different factors in the analysis due to data availability and the difficulties of isolating purely exogenous factors when variables combine both endogenous and exogenous elements. Our approach to clustering has therefore been

iterative, allowing for investigation of different clustering options depending partly on the factors treated as exogenous, and it may necessary to revisit the results at a later stage.

6. We also included a cost of living index as an explanatory variable in the clustering analysis, as it controls for different levels of remuneration, reflective of underlying cost levels in different countries. The cost of living index compares GDP measured at current prices and GDP adjusted for Purchasing Power Parity (PPP), following the methodology developed in 2012 for the ATM Cost-Effectiveness (ACE) 2011 Benchmarking Report with 2012-2016 outlook[1].

7. We also considered the clustering analysis undertaken for RP1 and RP2 with a view to updating it for RP3. We observed that in RP2, in addition to the variables available to us (complexity, seasonality, traffic volumes and cost of living), the analysis had also included a data-set on ATCO-employment costs per ATCO hour. Arguably, over the long term, this variable is within the control of ANSPs, can manage employment costs through their recruitment, training or retention policies. However, we concluded that, within the time-frame of a 5-year reference period, the key drivers of ATCO-employment costs could be considered as exogenous. We therefore repeated the RP2 analysis using the latest available data, enabling a clearer comparison of the RP2 and RP3 results.

8. The data sources used for each of the variables is listed in the table below.

Table 2: Summary of variables included in the clustering analysis

Variables included	Source of data set
Traffic volume	ACE data
Traffic complexity	Airspace complexity score from EUROCONTROL
Traffic variability	STATFOR daily traffic data
Cost of living index	GDP (IMF), Eurostat (PPP), ACE data (ANSP exchange rates)
Unit ATCO employment cost	ACE data

Clustering technique

9. We consider that the most appropriate clustering technique is multi-dimensional analysis, combined with expert review. This balances the need for analytical rigour with the flexibility needed to ensure that the defined clusters appear reasonable to a wider audience, including industry stakeholders. Multi-dimensional analysis avoids the need to impose thresholds between clusters, and avoids many of the pitfalls of partitioning based on predetermined threshold values (which may ignore specific characteristics of the ANSPs).

10. There are multiple functional forms for the distance measure used to identify how near factor values for different ANSPs are to each other. There are also multiple criteria against which clusters can be defined, and a range of optimisation strategies that can each lead to different clusters. To ensure that the approach chosen generated meaningful results, we specified a clustering approach with the following characteristics:

[1] ATM Cost-Effectiveness (ACE) 2011 Benchmarking Report with 2012-2016 outlook, DRAFT, December 2012. Sources for the calculation include IMF (GDP data), PPP exchange rate (Eurostat), Market Exchange rates as reported by ANSPs in ACE.

- **Hierarchical:** the set of clusters generated is nested and organised as a tree. Each node (cluster) in the tree is the union of its branches (sub-clusters) and the root of the tree is the cluster containing all the objects (in this case ANSPs).
- **Exclusive:** individual ANSPs cannot belong to more than one cluster.
- **Complete:** every ANSP is assigned to a cluster.

11. At this stage, we have completed the multi-variable analysis and generated a set of results that appear to support meaningful benchmarking of cost efficiency and capacity. However, they will need to be subject to further expert review, in discussion with the Commission and the PRB, to ensure that they are robust.

12. Cluster analysis is based on data covering all the variable for the same unique year. It is therefore important to avoid allocating ANSPs to clusters based on what might be one-off events in a given year (which could influence certain exogenous as well as endogenous factors). In addition, it is important to ensure that the data is as recent as possible, so that the results reflect the current situation rather than the past. Given the available data, we selected 2015 as the reference year for the analysis.

13. For some of the variables, especially traffic volumes, the data range (across different ANSPs) is particularly wide, potentially distorting the results (with the largest values having a stronger impact than the smaller ones). We therefore normalised the variables, standardising them relative to the maximum value within the data-set range and ensuring that each had an equal weighting within the cluster analysis algorithm.

14. We also excluded a number of ANSPs from the analysis:

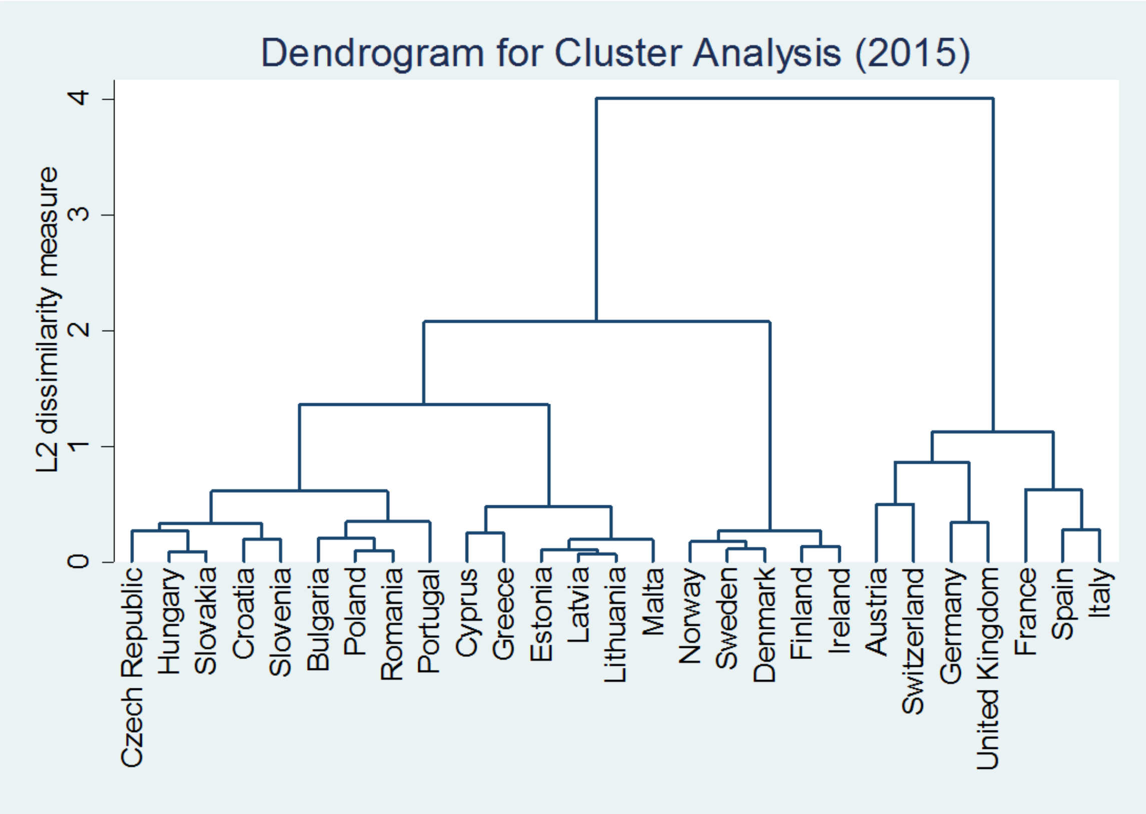
- due to the unique nature of its airspace (upper airspace only, across four Member States), we excluded MUAC; and
- we excluded BelgoControl and LVNL as both ANSPs provide lower airspace services only.

Results of the multi-dimensional clustering analysis

15. Please note that we decided to display the results for RP3 by Member State, and not for individual ANSPs. This is because targets are set at national level.

16. The set of clusters obtained are presented in the figure below. This shows a dendrogram, a standard representation of the results obtained from multi-dimensional clustering analysis. The vertical lines provide an indication of how dissimilar two Member States are in terms of the calculated multivariate measure. For example, the vertical height of the connection between Latvia and Lithuania is relatively small, indicating that these two Member States are similar, while the vertical height of the connection between Austria and Switzerland is much greater, indicating little similarity.

Figure 1: Dendogram



Source: Steer Davies Gleave

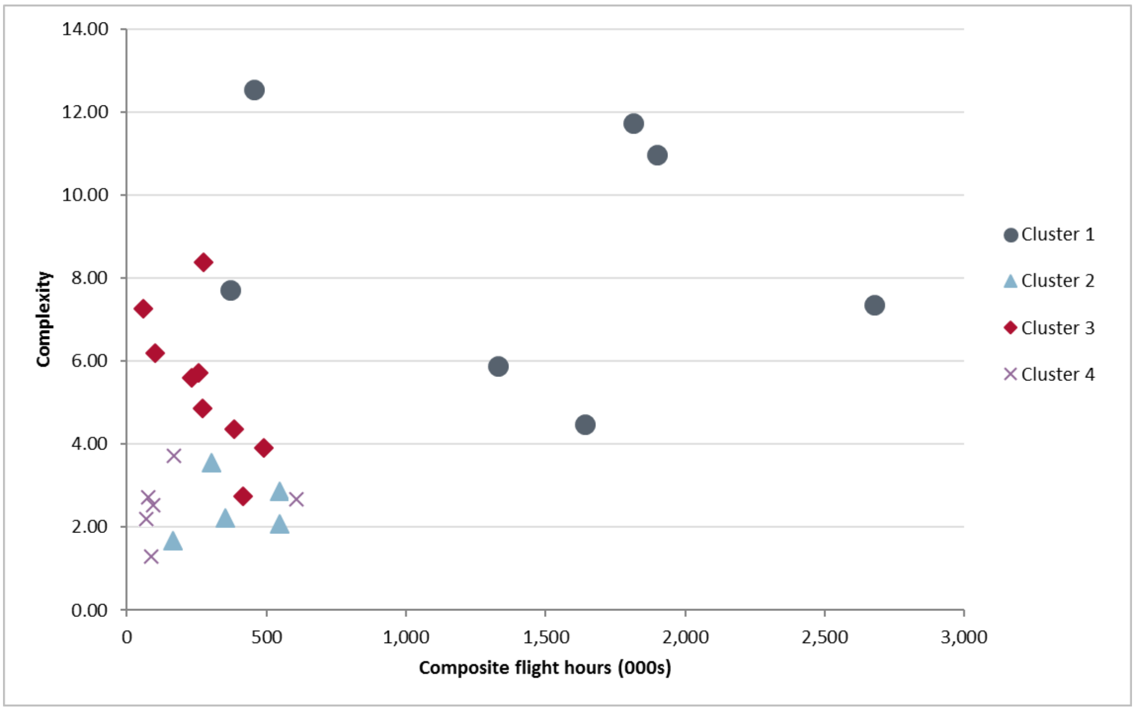
17. The clusters may be categorised as follows:

- Cluster 1: Austria, Switzerland, Germany, United Kingdom, France, Spain and Italy;
- Cluster 2: Norway, Sweden, Denmark, Finland and Ireland;
- Cluster 3: Czech Republic, Hungary, Slovakia, Croatia, Slovenia, Bulgaria, Poland, Romania and Portugal; and
- Cluster 4: Cyprus, Greece, Estonia, Latvia, Lithuania and Malta.

18. In order to investigate the key drivers of these results, we examined the relationship between key pairs of variables, as shown in the following XY scatter charts

19. In Figure 2, we observe that there is a clear grouping of the five largest Member States (Cluster 1) based on traffic volumes. Switzerland highest complexity (12.54) also differentiates it from the other Member States. In contrast the levels of complexity and traffic in Clusters 2 and 4 remain small. Member States in Cluster 3 experience lower levels of traffic than most countries in Cluster 1 but nevertheless exhibit relatively complex airspace.

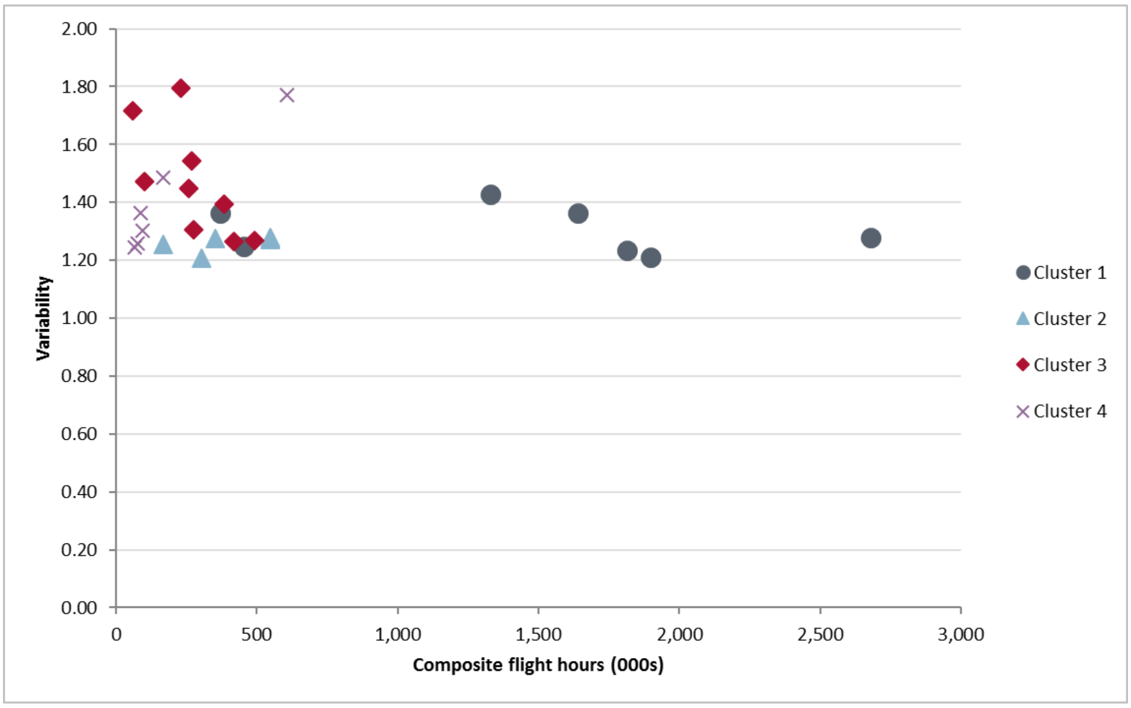
Figure 2: Traffic volume vs complexity



Source: Steer Davies Gleave

20. In Figure 3, the range of traffic variability is highest for Cluster 3and Cluster 4. Both clusters include Member States with significant summer traffic (such as Croatia and Greece). In contrast, traffic in Member States in Clusters 1 and 2 is subject to less seasonal variability.

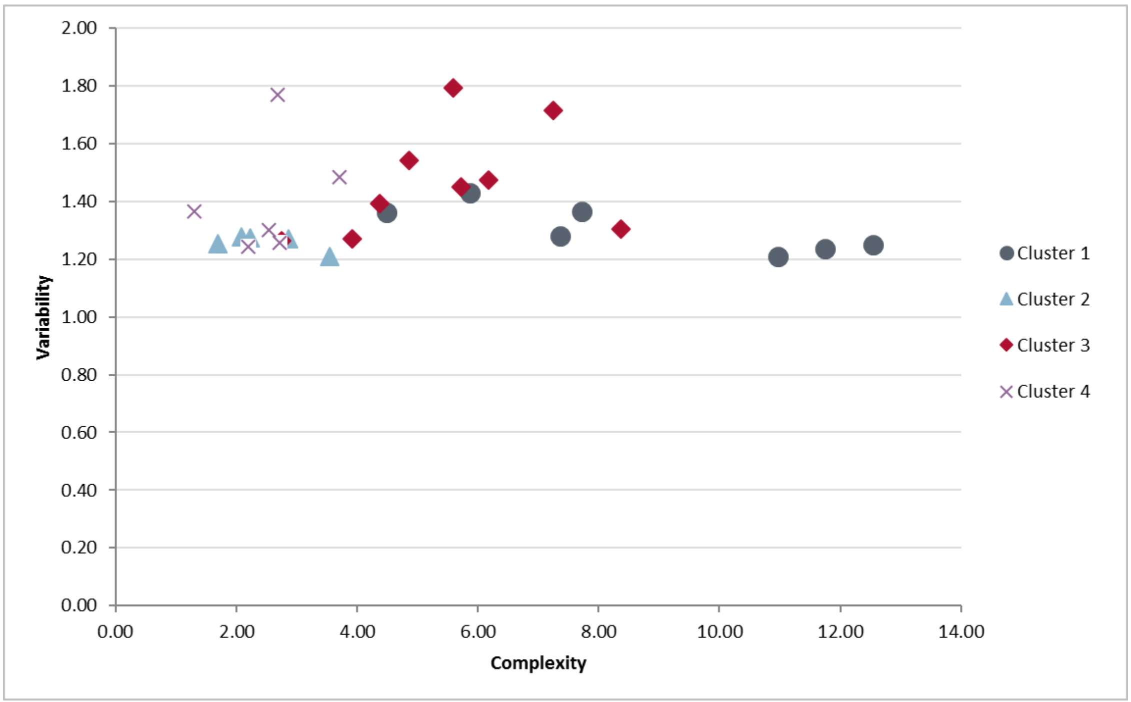
Figure 3: Traffic volume vs variability



Source: Steer Davies Gleave

21. Figure 4 suggests that there is no strong relationship between traffic variability and complexity of airspace. Clusters 3 and 4 exhibit a similar range of variability but airspace in Cluster 3 Member States is significantly more complex. By contrast, the range of variability in Clusters 1 and 2 is more limited but Cluster 1 exhibits greater complexity (and a wider range of complexity across Member States).

Figure 4: Complexity vs variability



Source: Steer Davies Gleave

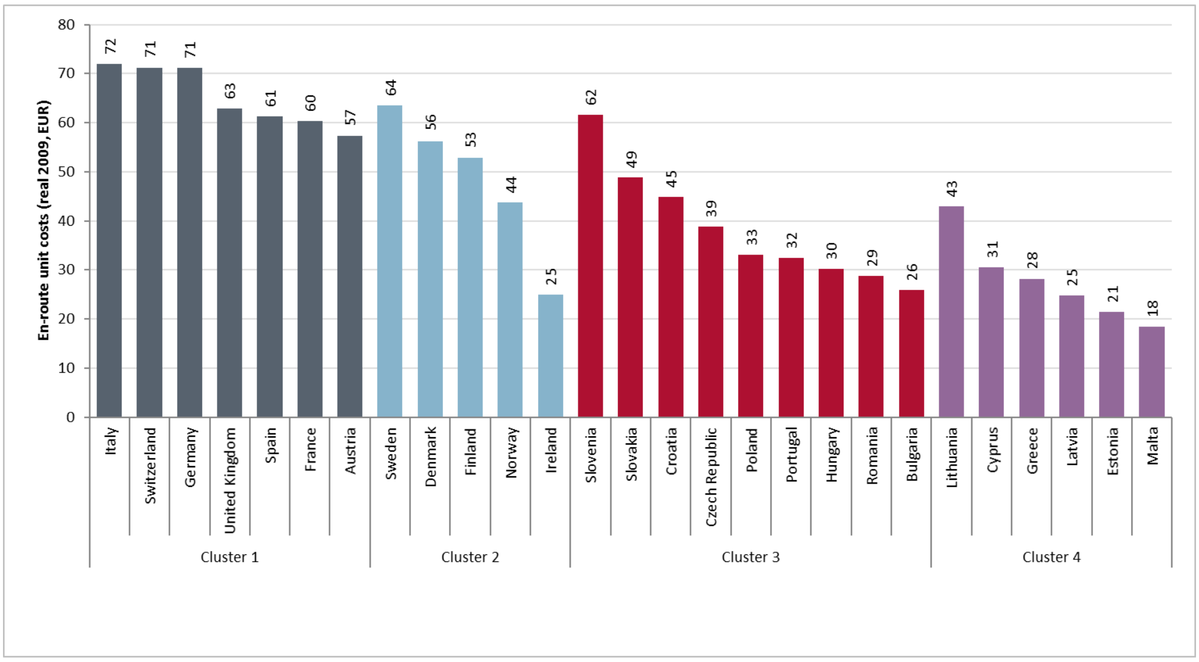
22. With the results above taken into consideration, we can describe more precisely the four clusters identified in the analysis:

- Cluster 1: this is a grouping of the largest and usually most complex States;
- Cluster 2: this is a grouping of States with little variability in traffic, limited complexity, and average level of traffic;
- Cluster 3: this is a grouping of States with relatively high level of complexity and variability.
- Cluster 4: this is a grouping of States with limited complexity and limited traffic, but sometimes high variability.

RP3 benchmarking analysis

23. We have undertaken some initial benchmarking of indicators, including RP3 Key Performance Indicators, within and between clusters. The results are shown in the figures below. Note that the colour-coding of clusters is constant across diagrams, for ease of comparison, but the position of clusters and Member States varies with rank order. The figure below shows en-route unit costs for the 4 clusters. In Clusters 2, 3 and 4, we observe a large variation in unit costs between the best in class and the rest of the cluster, with rations higher than 1:2. It is also difficult to see specific trends in terms of unit cost in these clusters. Within Cluster 1, there is less variance, with overall two groups of unit-costs: four States in the €60 bracket with the other in the €70 bracket.

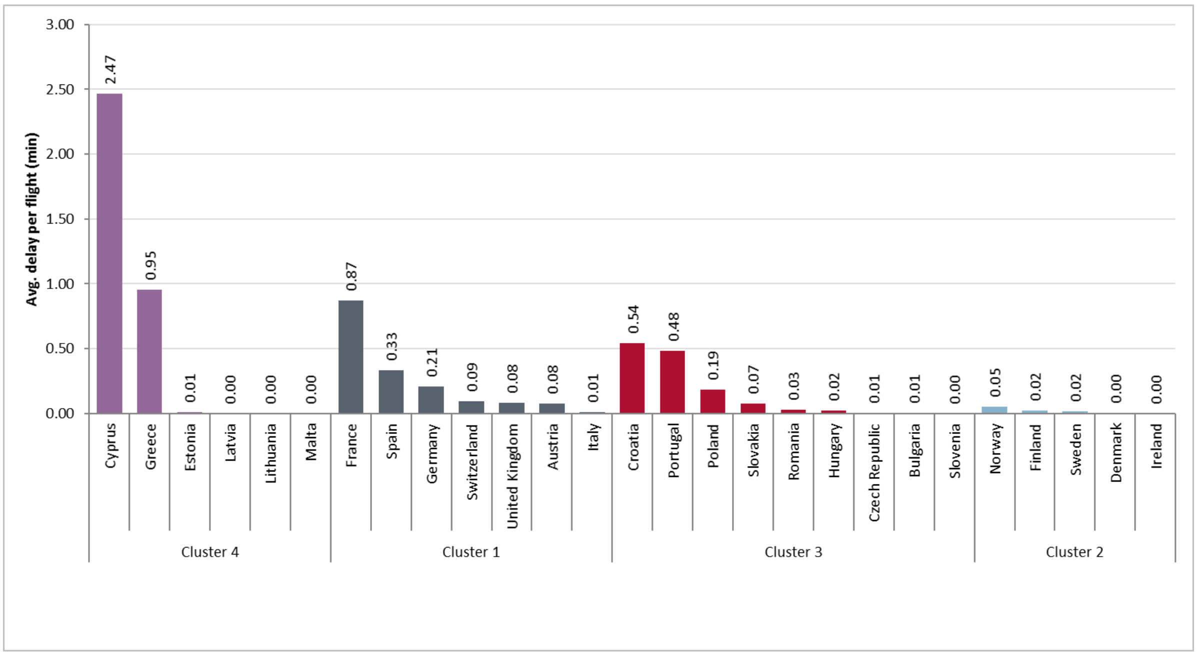
Figure 5: En-route costs (real 2009, converted into EUR) / En-route service units (000s), 2015



Source: Steer Davies Gleave

24. In 2015, we observe in cluster 2 almost no delays. There is more variation in the outturn delay of Clusters 1, 3 and 4, although we observe that delays are usually located in specific Member States and not consistently spread across the European skies.

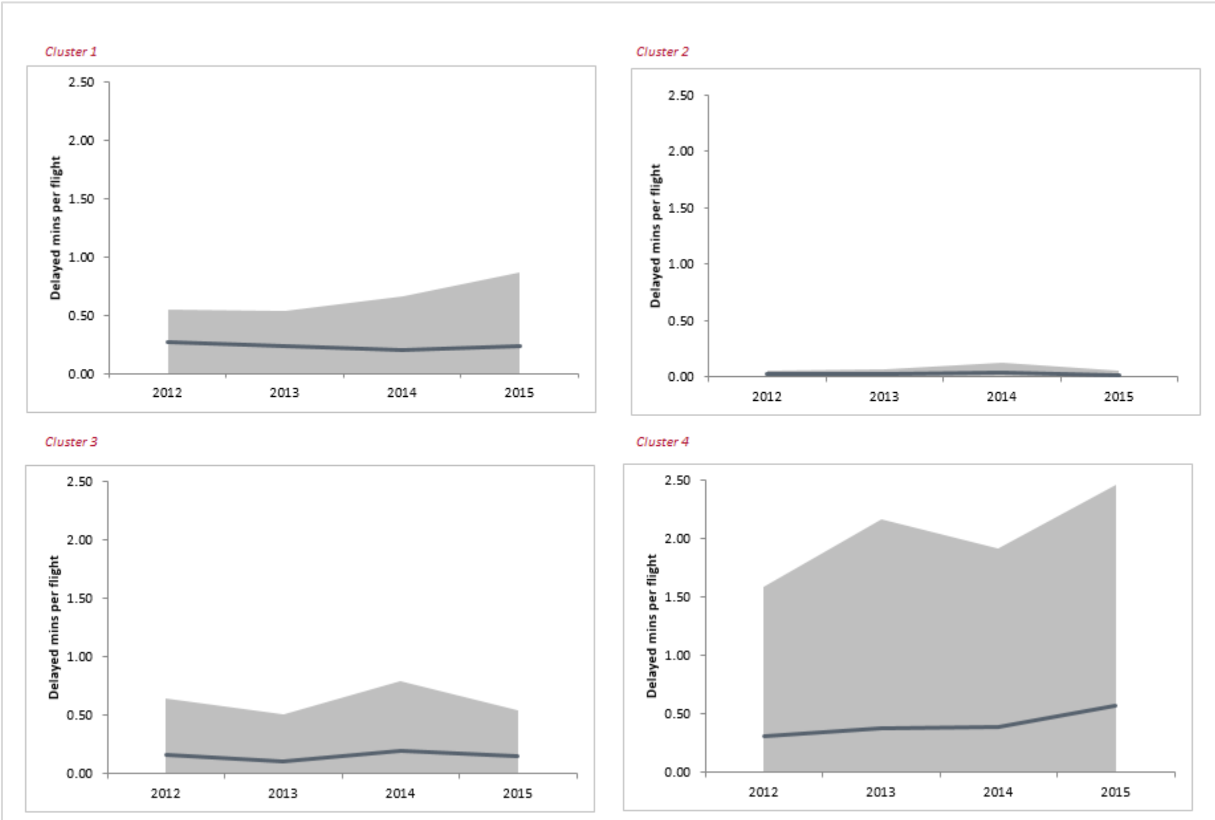
Figure 6: Total en-route delay/Total IFR flights controlled by the ANSP, 2015



Source: Steer Davies Gleave

25. On the graphic below we see the highest, lowest and average values of the mins of delay per flight within clusters. Whilst there are hardly no recorded delays in Cluster 2 and average downward trend in Clusters 1 and 3, average delays are rising in Cluster 4.

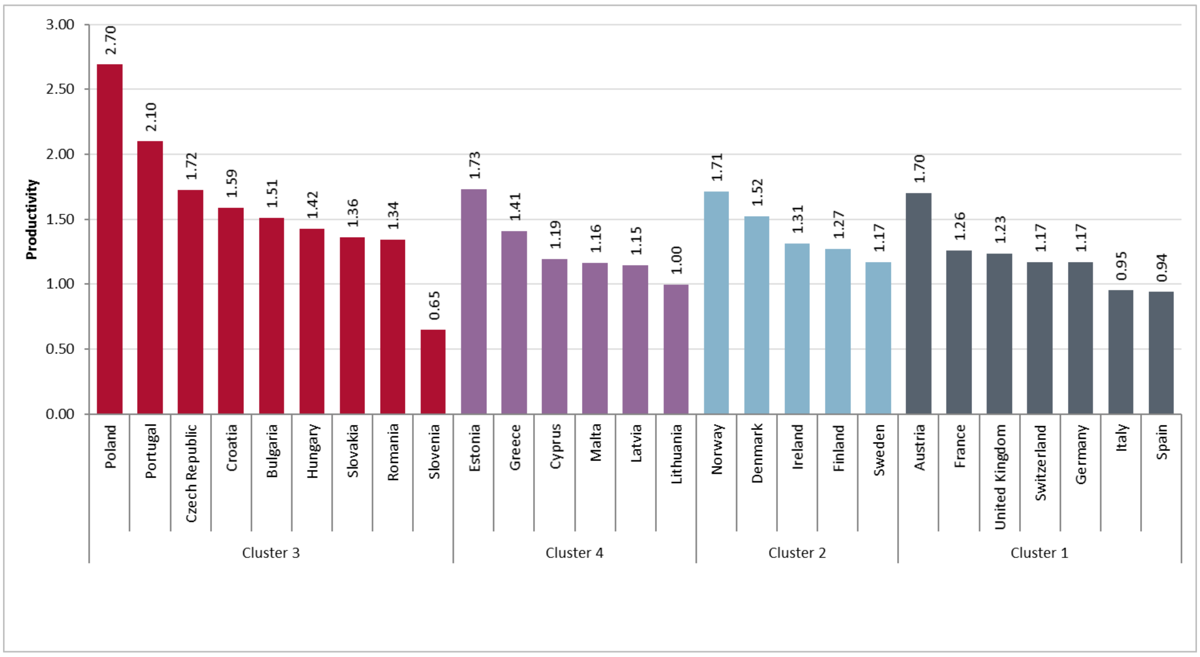
Figure 7: Evolution of delayed mins per flight (Total en-route delay/Total IFR flights controlled by the ANSP) by cluster (2012-2015)



Source: Steer Davies Gleave

26. Although productivity is not an SES KPI, we thought that it might provide some useful context to the PRB for the discussion around the interdependencies between the SES KPIs of cost-efficiency and capacity. The figure below compares the productivity (expressed in terms of IFR flight-hours controlled over ACCs per ATCO OPS hour on duty) in different Member States in 2015. In the case of Clusters 1, 2 and 4, apart from one Member State in each cluster, levels of productivity fall within a relatively limited range, although there is greater variability in the case of Cluster 3.

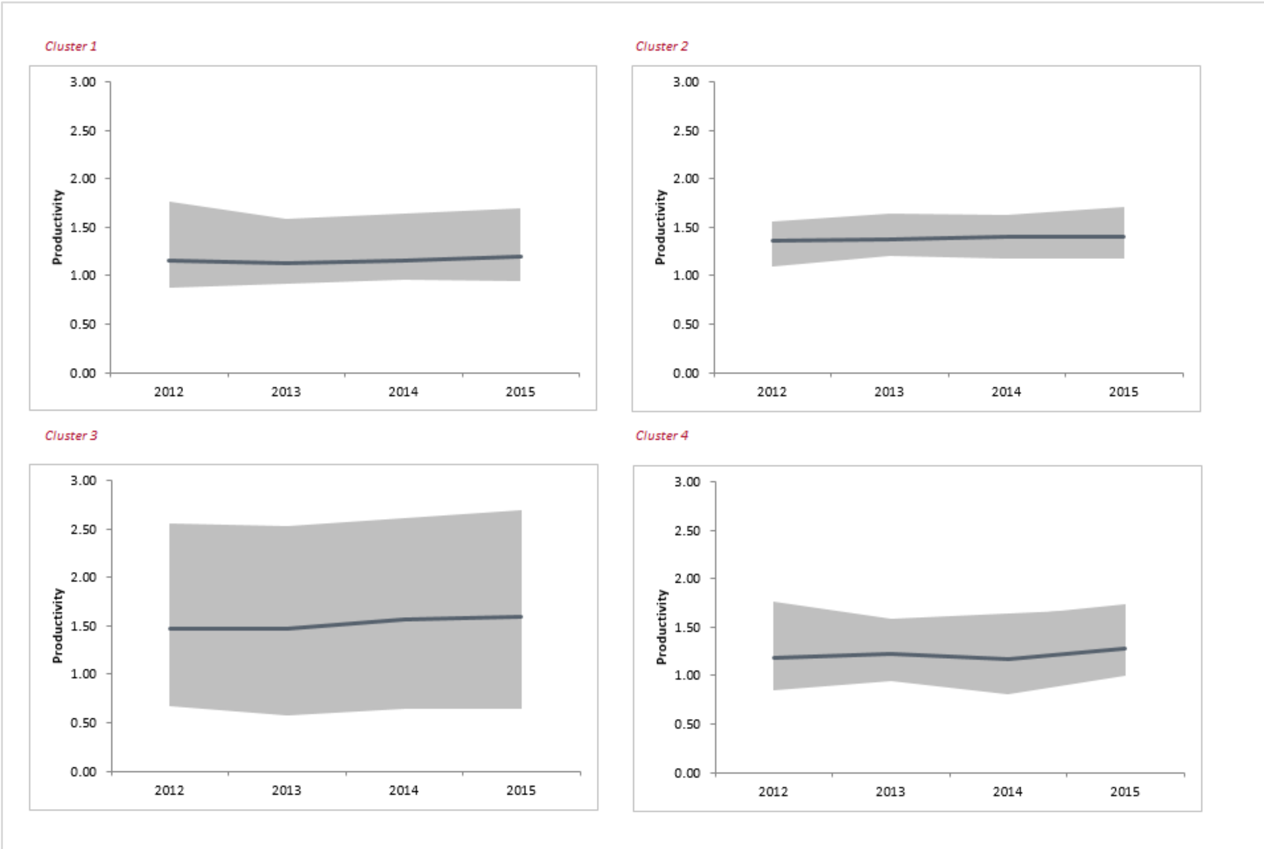
Figure 8: Total IFR flight-hours controlled by the ANSP / ACC ATCOs in OPS hours on duty, 2015



Source: Steer Davies Gleave

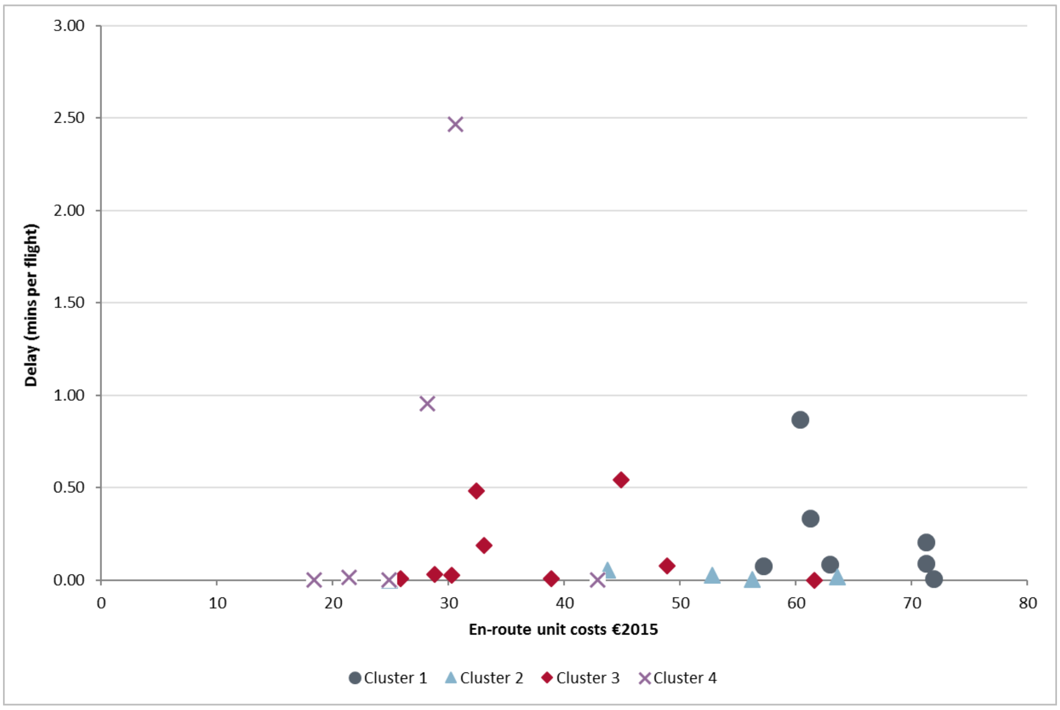
27. We also considered trends in the level of productivity within each of the four clusters (assuming no change in the composition of the clusters) since 2012, as displayed below. We observe similar levels of average productivity for Clusters 1, 2 and 4. In Cluster 3, although there is a much wider range within the level of productivity, in average the productivity is the highest among all clusters.

Figure 9: Evolution of productivity (Total IFR flight-hours controlled by the ANSP / ACC ATCOs in OPS hours on duty) by cluster (2012-2015)



Source: Steer Davies Gleave

Figure 10: Delay / Costs relationship, 2015



Source: Steer Davies Gleave

28. The figure above showing the benchmark between delays and costs Within Cluster 4 it shows wide variation between unit-costs and delays achieved in 2015. There is much less variance between Cluster 1, 2 and 3, although there is no easy pattern based on 2015 only as to the relationship between delays and costs on this figure.

Conclusions

29. These results are preliminary results, from the clustering analysis performed. For instance, we have not applied any expert judgment on them, nor taken into account relevant operating factors. The benchmarking of the clusters realised in this analysis shows that for the cost efficiency KPA there is, in some clusters at least, quite a range of values. This will mean that setting a target in this circumstance will not necessarily be an easy task. Different approaches could be considered: best-in-class, average values, etc.