PRB | Performance review body of the single european sky

Annex IV
Clustering Analysis

Revised narrative

22 February 2019

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 | regula.dettling-ott@prb.eusinglesky.eu / prb-office@prb.eusinglesky.eu | webgate.ec.europa.eu/eusinglesky

To	European Commission	**Memo**	
Cc	Steer project team		
From	Clémence Routaboul		
Date	13 February 2019		
Project	Support study on union-wide target setting for RP3	Project No.	23261901

Clustering analysis – revised narrative

Introduction

Purpose of clustering

1. EU-wide targets are determined before the start of each new SES Reference Period. These targets set the level of ambition for each of four Key Performance Areas (KPAs), and are a key input used by charging zones as they draft their Performance Plans. Setting EU-wide targets for en-route services is a complex exercise, requiring an understanding of the interdependencies between different KPAs, notably between capacity and cost-efficiency.

2. Air navigation service providers (ANSPs) face operating environments that differ significantly in terms of ownership and governance, the configuration and size of the airspace that they manage, the traffic that they handle and their operational practices and staffing levels. Whilst some of these variables result from decisions made by individual ANSPs and/or can be managed by them to some degree, others are exogenous and cannot be influenced by management action, at least in the short term.

3. Effective economic regulation involves setting performance targets reflecting a regulated entity's ability to manage risk. In the case of the SES, targets are set at EU-wide level, but performance plans are developed at the national level or at the level of functional airspace blocks and must therefore take account of the factors that individual ANSPs cannot easily influence. Against this background, we have undertaken clustering analysis to identify groups of en-route charging zones where ANSPs face similar exogenous[1] environments (operational and economic) and hence similar constraints affecting their ability to deliver improvements in capacity and cost efficiency.

4. By taking these groups into account in the preparation and assessment of individual performance plans, it is possible to ensure that plans, while stretching in terms of their ambition and consistent with the EU-wide targets, are also realistic.

Clustering variables

5. We examined a wide range of variables, including some that could be classified as either endogenous or exogenous to some degree. Because of data limitations for some variables, we have not been able to explore the implications of using all the possible factors identified at the start of the analysis. Nevertheless, our approach to clustering has been iterative, allowing for investigation of different combinations of exogenous factors.

[1] Exogenous factors are external circumstances that are faced by ANSPs and that are largely or entirely beyond their control.

6. After reviewing the options, we selected traffic demand (i.e. volumes of traffic), traffic complexity and traffic variability (i.e. seasonality) as exogenous variables. We also included a cost of living index[2] as an exogenous explanatory variable as it controls for different levels of remuneration, reflecting underlying cost levels in different countries, and an ATCO-employment costs per ATCO hour variable. Arguably, over the long term, employment costs are within the control of ANSPs, who can manage them through their recruitment, training or retention policies. However, we concluded that, within the 5-year time-frame of the third reference period (RP3), the key drivers of ATCO-employment costs could be considered as exogenous.

7. The data sources used for each of the exogenous explanatory variables are listed in the table below.

Table 1: Summary of variables included in the clustering analysis

Exogenous variables	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

8. Ireland and Portugal handle a greater number of overflights than the other ANSPs in their respective clusters because of their geographical location. This tends to depress unit costs, since overflights typically require fewer resources to manage and often generate a higher number of service units (since the relevant aircraft are often larger, long haul aircraft). While the proportion of overflights in the total handled is an exogenous factor, we chose not to include it (and other potential factors) in the clustering analysis to avoid both identifying too many clusters and reducing the numbers of ANSPs within any single cluster to three or less. The corollary is that comparisons of cost efficiency levels within clusters must be informed by careful consideration of operational factors that are specific to an individual ANSP.

9. In addition, due to the unique nature of its airspace (upper airspace only, across four Member States and three charging zones), we initially excluded MUAC from the clustering analysis. We also initially excluded BelgoControl and LVNL as both these ANSPs provide lower airspace services only (with upper airspace control being provided by MUAC). However, the appropriate treatment of these organisations in the definition of clusters is considered further below.

Initial results of the clustering analysis

Overview

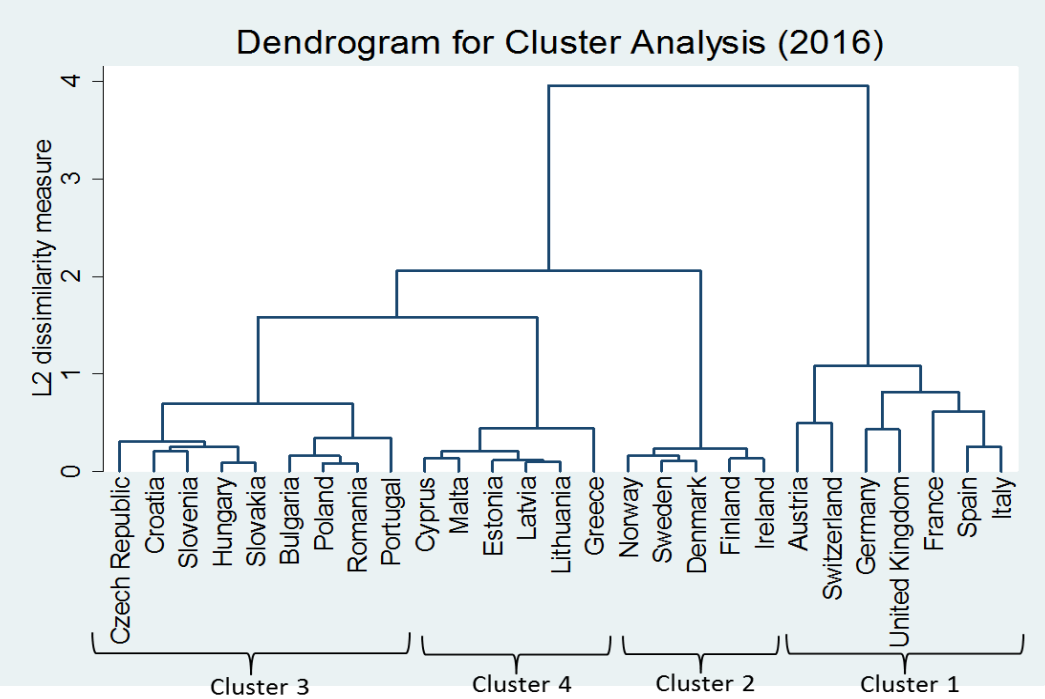
10. The clusters[3] obtained from our initial analysis performed in May and June 2018 are presented as a dendrogram in Figure 1[4].

[2] 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.

[3] The clusters are presented as groups of Charging Zones since targets are set at this level. Note that Portugal, Norway and United Kingdom refer to their respective continental Charging Zone only. For Spain, both Charging Zones have been consolidated as one.

[4] A dendrogram is a standard representation of the results obtained from multi-dimensional clustering analysis. The vertical lines provide an indication of how dissimilar two Charging Zones 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 Charging Zones are

Figure 1: Dendrogram



Source: Steer

11. From examination of the dendrogram, four clusters emerge and can 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, Croatia, Slovenia, Hungary, Slovakia, Bulgaria, Poland, Romania and Portugal; and
- **Cluster 4** - Cyprus, Malta, Estonia, Latvia, Lithuania and Greece.

12. These clusters are illustrated in Figure 2.

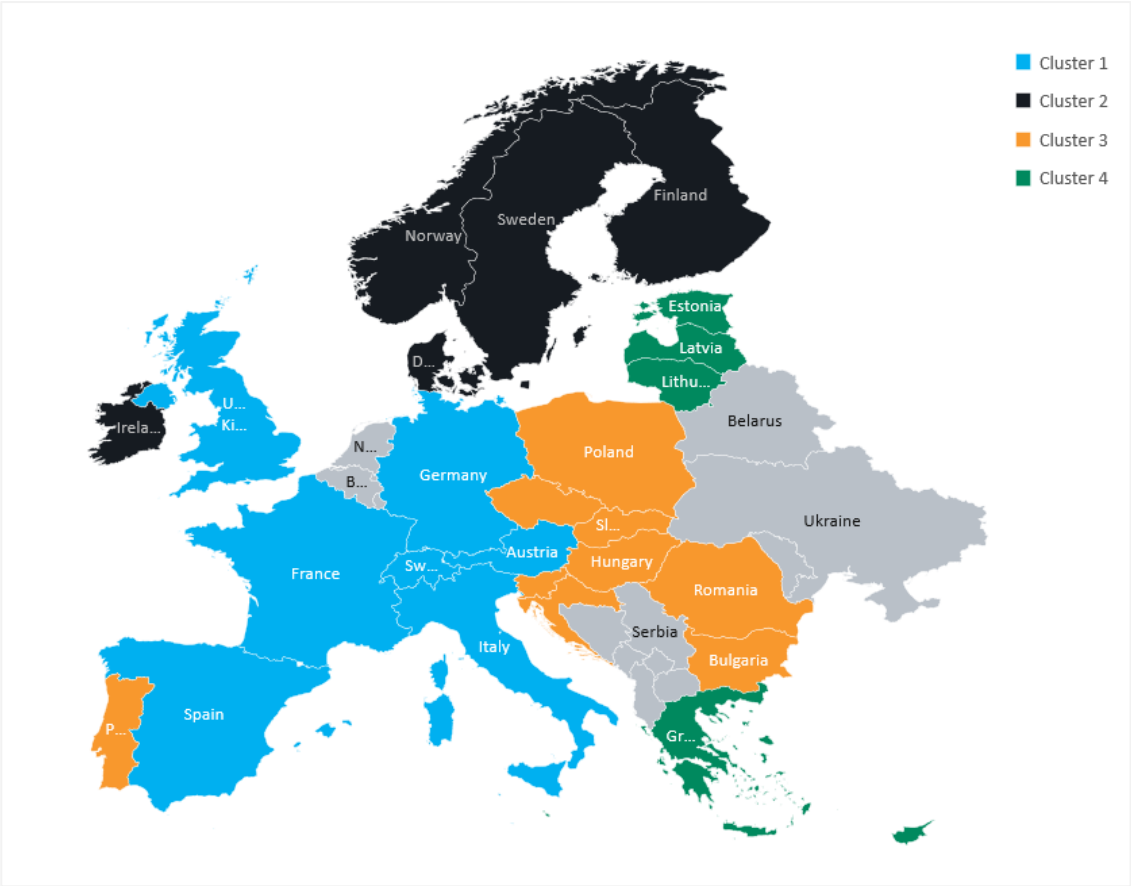
13. All Charging Zones can be readily allocated to clusters; there are no Charging Zones that are substantially different from all their potential peers. Further, there is a high degree of similarity between Charging Zones in Clusters 2, 3 and 4. The similarity is highest in Cluster 2, as shown in the dendrogram by the horizontal line joining Norway/Sweden/Denmark and Finland/Ireland, which is closer to the horizontal axis than any of the corresponding lines in the other clusters. The lines joining Charging Zones in Clusters 3 and 4 are also relatively close to the horizontal axis.

14. We note that there is a case for creating two sub-clusters from Cluster 3 (consisting of, in one case, Czech Republic, Croatia, Slovenia, Hungary and Slovakia and, in the other, Bulgaria, Poland, Romania and Portugal). However, we grouped all these Charging Zones together as there is more similarity between them than between the individual Charging Zones in Cluster 1 (for example, there is less similarity between Austria and Switzerland or between Germany and the UK than there is between all the Charging Zones in Cluster 3). Cluster 1 also includes France, which is the most distinct of all the Charging Zones

similar, while the vertical height of the connection between Germany and United Kingdom is much greater, indicating less similarity.

included in the analysis (as indicated by the height of the vertical line connecting it to the rest of its peers)[5].

Figure 2: Map of the Clusters generated without expert judgment



Source: Steer

Further development of cluster definitions

15. The purpose of the clustering analysis is to support Union-wide target-settings as well as the review and approval of performance plans. However, as we excluded BelgoControl and LVNL from the initial analysis (see paragraph 9) a question arises as to how to assess the performance plans of Belgium-Luxembourg and the Netherlands. More specifically, as there are no other ANSPs that are similarly concerned solely with the management of lower airspace, it is not clear that they have any appropriate comparators. We have examined two possible responses to this issue.

16. In the first approach we establish a fifth cluster alongside those identified in the analysis described above. Cluster 5 includes Belgium-Luxembourg and the Netherlands as well as MUAC (since MUAC provides the upper airspace service provision for Belgium-Luxembourg and the Netherlands)[6]. This grouping can be supported on the grounds that these ANSPs face particularly different operational circumstances to others

[5] France was identified as the most distinct Charging Zone among all those considered in this analysis, based on two separate runs of the analysis (i.e. using 2015 and 2016 data).

[6] MUAC also provides upper airspace services above Germany but MUAC costs represents a much lower share of the German Performance Plan than of the respective Belgium-Luxembourg and Dutch plans.

in the SES while having similar characteristics to one another. However, such a small cluster does not offer sufficient scope for benchmarking or comparator analysis.

17. The second approach, which is the one adopted for the analysis presented in the remainder of this paper, is based on the observation that, notwithstanding the unique operational characteristics of Belgium-Luxembourg and the Netherlands, both are subject to similar exogenous factors as Switzerland and Austria in Cluster 1 (which themselves handle far less traffic than the largest five Charging Zones in Cluster 1). This suggests the creation of a fifth cluster including Belgium-Luxembourg, the Netherlands, Switzerland, Austria and MUAC (for the reasons given in paragraph 16), and redefinition of Cluster 1 to include only the five largest Charging Zones (Germany, United Kingdom, France, Spain and Italy). All other Clusters remain unchanged.

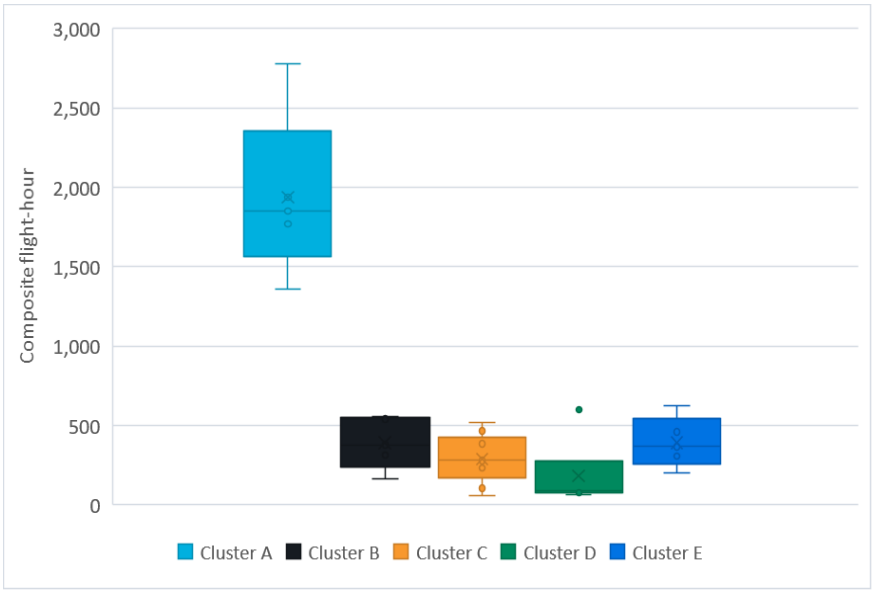
18. For clarity we have labelled the revised in the second, preferred, approach A to E to distinguish them from the clusters derived using the analysis described in paragraphs 10 - 14.

Analysis of exogenous factors

19. In the following paragraphs, we discuss how the clusters compare in terms of the five exogenous factors used in the analysis (traffic volumes, traffic variability, complexity, Unit ATCO employment cost and cost of living). The discussion is based on a series of 'box and whisker' diagrams - in the charts shown, the top and bottom 'whiskers' show the maximum and minimum values (excluding outliers) for a given cluster, whilst the 'o's show the other individual Charging Zone values. The 'x' represents the mean of the cluster, whilst the horizontal line within a box shows the median. The bottom of the box shows the first quartile, whilst the top of the box shows the third quartile.

20. As shown in Figure 3,Cluster A which includes the five largest Charging Zones in terms of traffic (Italy, Spain, United Kingdom, Germany and France) records on average 4 times more traffic than the Charging Zones within each of Clusters B, C, D and E which have broadly comparable levels of traffic, except in the case of Greece[7], which is an outlier within Cluster D.

Figure 3: Traffic volume comparison for the five clusters

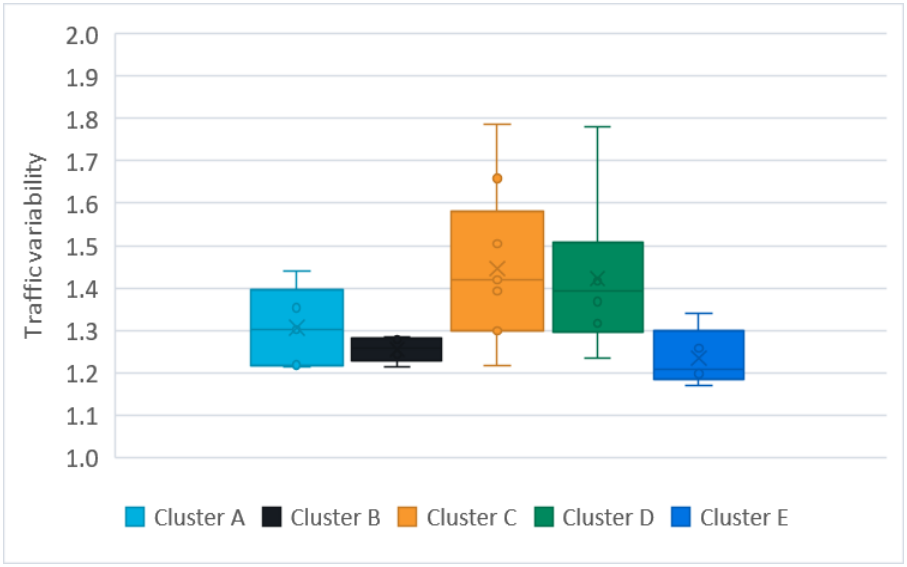


Source: Steer. Note: No maximum whisker shows for Cluster 4 as it is covered by the box.

[7] Greece is defined as an outlier here because its traffic volume value is more than 1.5 times the interquartile range above the third quartile of Cluster D.

21. Figure 4 shows that that there are important differences between clusters in terms of traffic variability. Charging Zones in Cluster C and Cluster D (especially those with significant summer traffic, such as Croatia, Slovenia and Greece) are subject to relatively high levels of variability, while variability within Clusters A, B and E is generally lower. In Cluster B, all Charging Zones exhibit similar variability, whilst the range is larger in Cluster A (with, for example, Italy's variability equalling the mean for Cluster C).

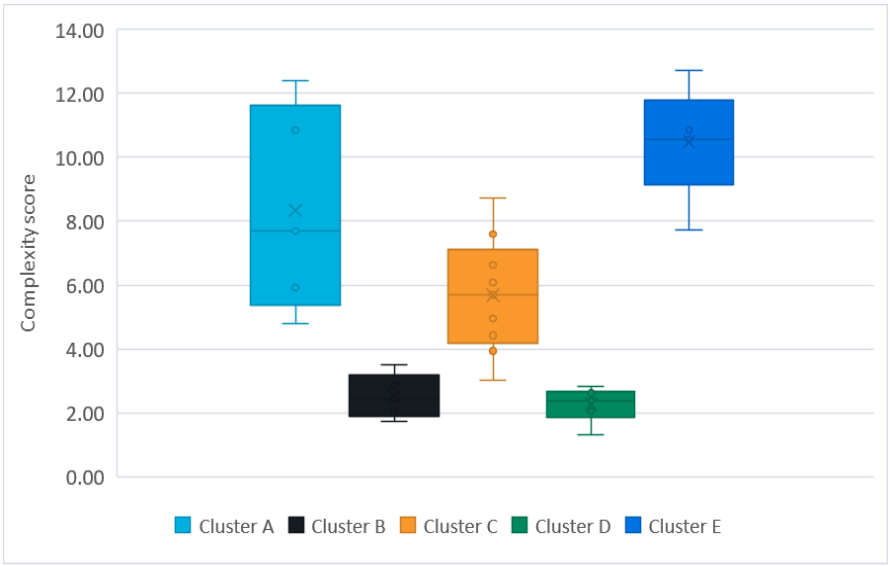
Figure 4: Traffic variability comparison for the five clusters



Source: Steer

22. The complexity of airspace within the clusters is presented in Figure 5. Clusters B and D have relatively low complexity, whilst Clusters A, C and E are subject to both a higher average level and a greater range of complexity. In Cluster A, UK, Germany and France have particularly complex airspace, as do all Charging Zones in Cluster E apart from Austria. Note that some clusters also include Charging Zones with oceanic traffic, with a greater proportion of overflights and less complex traffic, such as Ireland and Norway in Cluster B and Portugal in Cluster C.

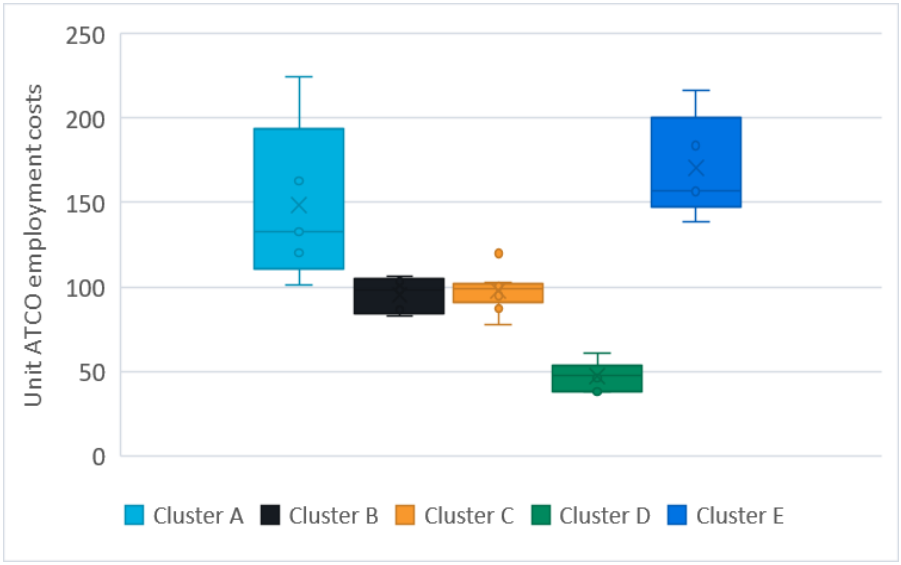
Figure 5: Airspace complexity comparison for the five clusters



Source: Steer

23. Turning to unit ATCO employment costs, Figure 6 shows that Cluster D has the lowest unit costs of all clusters by a considerable margin (the unit costs of Clusters B and C are approximately twice, and those of Clusters A and E three times as high as those of Cluster D). At the same time, the Charging Zones within Clusters B and C exhibit broadly similar unit costs. The range of unit costs within Cluster A and to a lesser extent Cluster E is greater than in any of the other clusters.

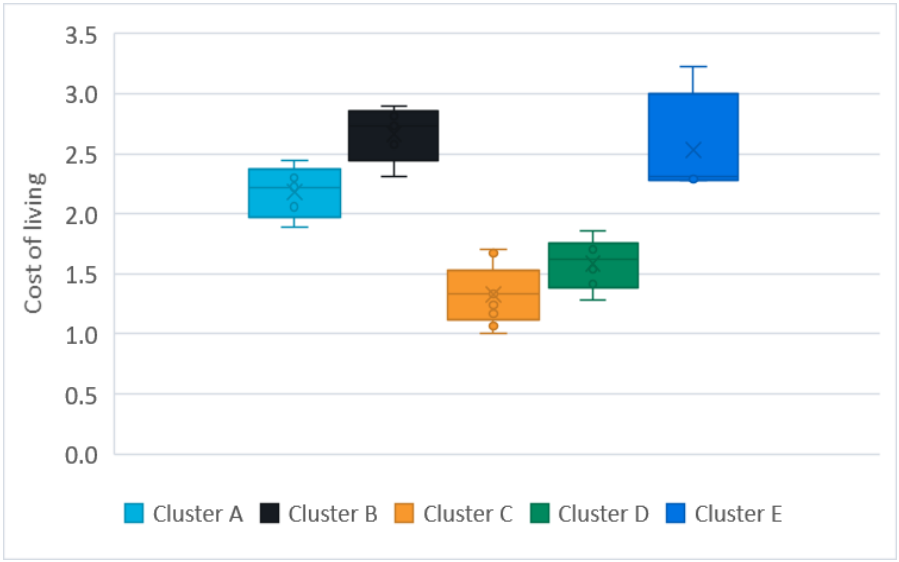
Figure 6: Unit ATCO employment costs comparison for the five clusters



Source: Steer

24. The cost of living within each Charging Zone also differs markedly between clusters. The cost of living in most Charging Zones in Cluster E is comparable with those in Cluster B as well as those in Cluster A. However, living costs in Switzerland are considerably higher than in all other Charging Zones, increasing both the range and average level of costs in Cluster E. The cost of living in Clusters C and D is substantially lower.

Figure 7: Cost of living comparison for the five clusters



Source: Steer

Conclusions

25. The results of the clustering analysis suggest that the clusters of Charging Zones can be characterised as follows:

- Cluster A is a grouping of the five largest Charging Zones in terms of traffic. They are responsible for some of the most complex airspace, although most are not exposed to particularly variable traffic. Living costs are high and unit ANSP employment costs are among the highest of all Charging Zones included in the analysis.
- Cluster B consists of Charging Zones in Northern Europe as well as Ireland, exhibiting relatively low levels of traffic, traffic variability and airspace complexity. Whilst the cost of living in these countries is the highest of all Charging Zones considered, unit ATCO employment costs are in the middle of the range.
- Cluster C is a grouping of Charging Zones that do not experience particularly high traffic levels but which are subject to relatively high traffic variability and moderate airspace complexity. They are in the centre of Europe, between the largest European markets to the west and countries with similarly variable traffic levels to the north and south. The cost of living in these countries is among the lowest of all those covered, although unit ATCO employment costs are comparable with those in Cluster B.
- Cluster D includes Charging Zones with relatively variable traffic but low airspace complexity. They are located on the periphery of the SES, either in the far north of Europe or at the eastern end of the Mediterranean. The cost of living in these countries is low and unit ATCO employment costs are significantly lower than in all other Charging Zones included in the analysis.
- Cluster E includes Charging Zones with relatively low levels of traffic and traffic variability. Airspace complexity differs significantly across the countries within the cluster, some of which are responsible for the most complex airspace in the SES. Costs of living and employment costs are among the highest in Europe.

26. The table below summarises the characteristics of the clusters in terms of the exogenous factors used in the analysis. Note that the terms 'high', 'moderate' and 'low' indicate general tendencies within the cluster; as indicated above, there may be exceptions, reflecting the balancing of different exogenous factors within the analysis.

Table 2: Summary characteristics of the clusters

	Cluster A	Cluster B	Cluster C	Cluster D	Cluster E
Traffic volume	High	Low	Low	Low	Low
Traffic variability	Low	Low	High	High	Low
Airspace complexity	High	Low	Moderate	Low	High
Unit ATCO employment costs	High	Moderate	Moderate	Low	High
Cost of living	High	High	Low	Low	High

27. Clustering analysis provides an interesting tool to group Charging Zones together. However, relevant operational factors (such as the proportion of oceanic or lower airspace traffic) must be taken into consideration when making comparisons between entities within clusters. This will help to ensure that the level of ambition within individual performance plans can be properly assessed and that the expected contribution of individual ANSPs to the achievement of Union-wide targets is neither unduly onerous nor unduly easy to deliver.

Appendix

Stability of the clustering analysis

28. At first sight, it may seem that the two economic factors (unit ATCO employment costs and the general cost of living) had a particularly strong influence on the results. To test the sensitivity of the results to these factors, we also performed the clustering analysis using the following exogenous variables:

    i. Traffic volumes, complexity, variability and unit ATCO employment costs: this assumes that there are no differences in cost of living across Europe.
    ii. Traffic volumes, complexity, variability and cost of living: this assumes that ANSPs can manage unit ATCO employment costs in the short term.

29. Please note that we did this on the outputs of the initial clustering analysis (the four clusters 1 to 4), undertaken in May and June 2018 and presented in paragraph 10. We did not test the stability of the five A to E clusters as the choice of clusters A to E involves some expert judgment.

30. The results of these tests were similar to those described above, with 22 of 27 Charging Zones remaining within broadly equivalent groupings across all three of the analytical scenarios. More specifically, only the following countries were allocated to different clusters under one or both sensitivity tests:

- Austria was allocated to a cluster with Czech Republic, Croatia, Slovenia, Hungary and Slovakia under test i) and ii).
- The sub-group of Bulgaria, Poland, Romania and Portugal moved to a cluster with Norway, Sweden, Denmark, Finland and Ireland under test i) and to a cluster with Cyprus, Malta, Greece, Estonia, Lithuania and Latvia under test ii).

31. The following table provides an explanation of these movements.

Table 3: Explanations for allocation of Charging Zones moving between clusters under the stability tests

Charging Zone	Explanation
Austria	Austria shares many of the traffic patterns that the Czech Republic faces. Both Charging Zones are grouped together when unit ATCO employment costs or when the cost of living index is considered separately. When both economic variables are taken into consideration, Austria is grouped with the largest six countries.
Portugal (continental)	As already noted, Portugal shares similar operational characteristics (large oceanic sectors) with Ireland, with overflights by larger aircraft tending to increase cost-efficiency. However, Portugal and Ireland are only in the same Cluster under test b) (i.e. assuming no difference in living-costs across Europe), when they are grouped with the Nordic States (Norway, Sweden, Denmark, Finland). In the main results reported above, Portugal is grouped with Czech Republic, Croatia, Slovenia, Hungary and Slovakia.
Bulgaria, Poland and Romania	Bulgaria, Romania and Poland share many traffic similarities, although Poland has slightly more traffic. When both cost of living and unit ATCO employment costs are considered as exogenous variables, all three charging zones are grouped with Czech Republic, Croatia, Slovenia, Hungary and Slovakia. Under test a), which assumes that there is no difference in the cost of living, these three States are grouped with Norway, Sweden, Denmark, Finland and Ireland, whilst under b), they are grouped with Cyprus, Malta, Greece, Estonia, Lithuania, Latvia.