ABSTRACT

Airspace fragmentation represents an issue that began to be more frequently mentioned within the Air Traffic Management (ATM) domain in the last two decades. Primarily, it is frequently listed as one of the main causes contributing to inefficiency of the ATM system in Europe. However, even though the issue of the European airspace fragmentation has been recognized back in the 1990s, over the past decades it has neither been frequently studied nor comprehensively addressed. Accordingly, minor progress has been made to describe this issue in more depth. Therefore, this research paper deals with the research of performance-based airspace fragmentation (one of several European airspace fragmentation types). It presents the conceptual and methodological framework of a novel model that can be used to obtain answers to hypothetical questions of where, when, how, and whether it is possible to achieve performance-based airspace defragmentation. Accordingly, it is expected that further studies of the developed model will deliver relevant information that may contribute to a more inclusive, smart, and spatially oriented development of the ATM system in Europe.

KEYWORDS

Air Traffic Management; airspace; fragmentation; performance-based airspace model.

1. INTRODUCTION

National air traffic markets in Europe before 1987 were protected, regulated, and fragmented with a goal to safeguard national interests [1]. Meanwhile, the prerequisites for future economic development have been made (in the sense of air traffic market deregulation and liberalization). However, despite these changes, the ATM system in Europe, i.e., the European airspace, remained fragmented due to national borders. As a result, nowadays, if not otherwise specified, every time an aircraft transits over a national boundary, it is serviced by a different Air Navigation Service Provider (ANSP). Therefore, that often leads to situations in which aircraft are being guided on the base of different operational requirements and restrictions [2]. In such way, fragmentation limits airspace capacity, adversely affects the environment, increases operational costs, and above all, potentially affects safety [3]. Consequently, it threatens further development of the ATM system in Europe, and as such, the development of air transport in Europe.

Airspace fragmentation problems have been officially recognized by the European Commission back in 1996, arguing that the European Union “cannot keep the frontiers in the sky that it has managed to eliminate on the ground” [4]. Although a long time has passed since then, clearly recognizable constraints associated with fragmentation problems are still seriously impeding sustainable growth of the European air traffic market.

Nowadays, the issue of the European airspace fragmentation is considered to be one of the hardest issues to overcome. However, although it is a frequently mentioned issue, there are still many assumptions and unanswered questions regarding this topic. Accordingly, the main goal of this research paper is to provide new insights into the issue of performance-based airspace fragmentation. In principle,
performance-based airspace fragmentation represents a fragmentation type that is a product of partial interactions of several other airspace fragmentation types – such as organizational, operational, technical, and functional airspace fragmentation [5]. This research deals with the establishment of conceptual and methodological model framework that can be used for assessing the performance-based airspace fragmentation level, and above all, for airspace defragmentation.

The research paper consists of three major parts. The first part deals with research background. More precisely, it provides a literature review, regulatory framework review, data processing practice review, and most importantly, identification of their shortcomings (all of which are directly or indirectly contributing to the existence of performance-based airspace fragmentation). The second part presents a conceptual and methodological model framework. It details the methodological assumption, data manipulation framework, and lists several methodological limitations. Therefore, the model is designed so as to test its null hypothesis – that the European airspace is fragmented from the performance-based aspect. The third part provides an overview of model applicability. Finally, the paper is concluded with research overview and further research considerations.

2. RESEARCH BACKGROUND

Airspace is a limited resource and the way it is managed primarily depends on the efficiency of the ATM system. As any other system, the ATM can be more or less efficient. However, the cost of inefficiency is rather high. Since airspace fragmentation represents quite a significant issue, the size of the inefficiency-related cost is even higher. Therefore, it can be noted that there is no consensus on what the real cost is when it comes to the issue of the European airspace fragmentation. For example, Matsoukis and Poulimenakos [6] have estimated that the fragmentation-associated cost is EUR 880 million–1,400 million per year. On the other hand, Grebenšek and Magister [7] consider it to be around EUR 2–3 billion per year. Furthermore, the European Commission mentions two different estimations within its two publications. According to the Aviation Strategy for Europe [8], the costs of fragmented airspace was estimated to be at least EUR 5 billion per year, whereas in the Aviation: Open and Connected Europe document [9], it was approximated to at least EUR 3 billion per year. Therefore, considering the significance and associated relevance of the research topic, this issue should be studied in more detail. However, before proceeding to an assessment model, a brief overview of previous studies in the field of research is followed by the identification of the shortcomings of the regulatory framework and the frequently used data manipulation method.

By reviewing bibliographic sources, it was recognized that significant work has been done so far to study the performances of the ATM system in Europe. However most of the sources do not consider all three correlated features of aeronautical data. Data manipulation is usually based on the analysis of attribute and temporal features of the aeronautical data. As a result, the spatial feature of the data set is frequently underutilized. Since 80% of information requirements stipulated by policy makers are related to spatial location [10, 11], that certainly raises many issues.

Furthermore, summarization of the most recent literature on the topic of European airspace fragmentation [12-15] indicates that there are currently no unambiguous answers to questions on how to define, and more importantly, how to measure airspace fragmentation. Also, the literature review indicates that there are no sources that simultaneously correlate attribute, temporal and spatial features of the performance of the ATM system. Therefore, since performance-based airspace fragmentation has not been comprehensively addressed so far, this research paper complements the existing literature in the ATM domain.

2.1 Regulatory framework review

With the goal to defragment European airspace, in 2004 the European Commission has initiated an ATM regionalization process through the establishment of the Single European Sky (SES) initiative. As a result, nowadays the conceptual assumptions of the development of the ATM system in Europe are primarily based on regulatory-defined performance requirements. More precisely, whether an ANSP is efficient is determined by comparing its performance achievements with those determined by the Performance Scheme. Since the regulatory framework does not consider the casual relationships of the ANSP performance, it may be concluded that it is based on an individualistic approach. Also, the spatial feature of the aeronautical data is
In order to summarize data and to depict their distribution, various data visualization methods are often applied. Within the ATM domain, the most frequently used method to obtain a sense of data distribution is the data grouping method. An example of its application can be found in the publication of the European Organization for the Safety of Air Navigation (EUROCONTROL) [17] shown in Figure 1. In short, the data grouping method breaks down a data set $n$ into a certain number of classes $k$. Thereby, $k$ determination is performed by applying the Sturges’ rule. According to the Sturges’ rule, the number of classes $k$ is approximated by the following equation:

$$n = \sum_{i=0}^{k-1} \binom{k-1}{i}$$

(1)

where the right-hand side of the equation by the binomial theorem equals to:

$$\sum_{i=0}^{k-1} \binom{k-1}{i} = \sum_{i=0}^{k-1} (k-1)^i = (1+1)^{k-1} = 2^{k-1}$$

(2)

Moreover, by taking logarithms with base ten, Sturges' formula equals:

$$(k - 1) \log(2) = \log(n)$$

(3)

$$k = 1 + \frac{1}{\log(2)} \log(n) \approx 1 + 3.3 \log(n)$$

(4)
However, although the application of this method can be frequently found within the ATM related sources, its methodological assumptions do not conveniently reflect the ATM system in Europe. There are three major reasons for this. Firstly, it does not capture performance interactions/compliance between the neighboring ANSPs, as it is based on the assumption of independent observations. That may be seen as a deficiency because the ATM system in Europe is a highly interconnected and interdependent system. Secondly, it may perform poorly in cases when data are not normally distributed. More precisely, it may provide misleading results when extremely high or low values exist. The reason for this is that $k$ depends on the size of the data set $N$, and the difference between the maximum $x_{max}$ and minimum value $x_{min}$. An example of this within the ATM domain can be seen on the example b) shown in Figure 2. Thirdly and most importantly, data visualization is performed by presenting results on a map (and not on a histogram). That is problematic because a map shows “where something is”, whereas a histogram summarizes “how often” measurements occur (regardless of where they occur). In that context, utilization of this data manipulation method within the ATM context is hazardous because at no point are the data georeferenced. For example, when using this method, there is no difference between where any of the analyzed values occur within the ATM network. This is so due to the fact that data distribution is performed according to their attribute distribution and not spatial distribution. Accordingly, it gives no information about the similarity level between the neighboring areas (which should be prioritized within the ATM context). That is also problematic due to the fact that the SES aspirations are to increase the efficiency of the ATM system in Europe by enhancing coordination and collaboration between the neighboring ANSPs. Therefore, utilization of the data grouping method with the goal to display spatial differences between the performances of the ATM should be avoided.

### 3. METHODOLOGICAL FRAMEWORK

The European airspace (de)fragmentation assessment model is based on the quantitative research approach. Therefore, it needs to be emphasized that the model is based on an interdisciplinary research approach. The reason for this is that the applied methodological framework is based on Tobler’s first law of geography (“everything is related to everything else, but near things are more related than distant things”), the origin of which actually derives from Newton’s law of universal gravitation. More precisely, the model is based on the application of the Exploratory Spatial Data Analysis (ESDA) in the ATM domain. Furthermore, the obtained estimations are consolidated and presented in form of a report by using Information Technology – so that they can be used in the strategic planning domain (as one of the management functions). A simplified view of the aforementioned is shown in Figure 3.
ESDA represents a set of techniques aimed at describing and visualizing spatial distribution of data. It identifies “atypical localization” or outliers, and detects patterns of spatial association (e.g. clusters, hot spots, or cold spots) [19, 20]. As a result of the ESDA application, the model determines whether and how the European airspace is fragmented from the performance-based aspect.

3.1 Methodological assumption

A brief review of the ATM system in Europe follows below. Its aim is to argue the methodological assumption applied in the model development. In Europe, if an international commercial flight is subject to Air Navigation Services (ANS) in one state, service provision must also be continued in the next state (into which the aircraft intends to enter), and so on until the aircraft reaches its destination. Hence, it can be inferred that the coordination and cross-border cooperation between the neighboring ANSPs represent a prerequisite that enables the functionality of the ATM system in Europe.

Furthermore, Figure 4 shows ANS distribution per state (according to the International Civil Aviation Organization nomenclature) in 2018. By applying the criterion of whether the service was provided to an international or domestic flight, it can be concluded that an average of 94.05% of the total number of ANS provided in 2018 was conducted in cooperation of at least two neighboring ANSPs. Therefore – as a result of interdependencies between the ANSPs and the differences in their performance levels – performance interdependencies, trade-offs, goal conflicting situations, and spill-over effects occur within the ATM system in Europe. For example, due to the occurrence of capacity congestion in the Area of Responsibility (AoR) of one ANSP, occurrence of such phenomenon will also affect the neighboring areas (ANSPs). Repercussions of such phenomenon will be that the aircraft will go through one of the neighboring areas instead of through the originally planned (congested) area. Thereby, the repercussions of such a phenomenon depend on the significance of the event occurred. In that context, Button and Neiva [22] have also recognized this issue by arguing that since different national ATM systems are not independent of their neighbors, there might be issues of spatial autocorrelation – meaning that the efficiency of one ANSP might be dependent on the efficiency of the neighboring ANSPs. Considering the aforementioned, the model developed is based on the methodological assumption of spatial dependency. As such, the model goes beyond the methods of traditional statistical theory that bases analyses on the assumption of independent observations.

3.2 Data manipulation framework

Before processing input data, they first need to be placed in their spatial context. Since the graph theory can precisely describe the structure of many real-world systems (and thus the ATM system), a network model was designed. It was formed as a directed graph with nodes and arcs. Each node represents one spatial object whereas each spatial object represents AoR, Flight Information Region (FIR), etc. of one ANSP. After forming a network model, a spatial weights matrix \( W \) is derived from it. It represents the \( n \times n \) square matrix that expresses connectivity in a binary form:

\[
W = \begin{cases} 
1 & \text{if } i \text{ and } j \text{ are neighbours} \\
0 & \text{if } i \text{ and } j \text{ aren’t neighbours} 
\end{cases}
\] (5)

![Figure 4 – Air Navigation Services distribution (2018) [21]](image-url)
By convention, the self-neighbor relation \( w_{ii} \) is excluded. Furthermore, since spatial weights are in practice seldom used in binary form, they were row-standardized as follows:

\[
W^* = \sum_j w_{ij}
\]

As a result, a row-standardized weights matrix is built – where each row sum of the row-standardization weights equals 1. After placing input data in their spatial context and after data standardization, data manipulation process can start. In so doing, it includes conducting a few complementary assessments (as shown in Figure 5).

A sense of spatial distribution is firstly obtained by the determination of spatial similarity index \( r \). It is conceptualized so that it quantifies the compliance level between the performances of the neighbors:

\[
r = \sum_{j=1}^{n} \left| \frac{x_i}{x_j} \right|
\]

where the results interpretation is analogous to the Spearman’s correlation coefficient interpretation. After estimation of spatial compliance, further data manipulation process through the study of spatial autocorrelation identifies patterns of spatial association. Spatial autocorrelation represents one of the relatively small sets of methods which deals simultaneously with location and attribute information [23]. Therefore, it is counted globally and locally (whereas both assessments are based on Moran’s I [24, 25]). Global Moran’s I quantifies the spatial autocorrelation across the entire network model. On the other hand, the local indicator of spatial association, global Moran’s I can be obtained. It equals:

\[
I = \frac{\sum_{i=1}^{n} I_i}{n}
\]

Based on the estimated global Moran’s I, it is possible to make conclusions about the spatial autocorrelation. A negative result indicates that spatial objects of similar attribute features are scattered over the network model (and vice versa).

Since spatial autocorrelation is inferential statistics, it enables testing of the null hypothesis. In doing so, global Moran’s I first needs to be standardized as follows:

\[
z - score = \frac{I - E(I)}{\sqrt{Var(I)}}
\]

where the expected value of Moran's I \( E(I) \) is calculated as follows:

\[
E(I) = -\frac{1}{n-1}
\]

while the variance of Moran’s I \( Var(I) \) equals:

\[
Var(I) = E(\hat{I}) - E(I)^2
\]

\[
\hat{E}(I) = A \cdot B
\]

\[
A = n \left[ (n^2 - 3n + 3)S_1 - nS_2 + 3S_0^2 \right]
\]

\[
B = b_2 \left[ (n^2 - n)S_1 - 2nS_2 + 6S_0^2 \right]
\]

\[
C = (n - 1)^3 S_0^2
\]

where \( n \) is the number of spatial objects, \( S_0 \) is the total of the weights matrix \( \sum W_{ij} \), \( n^{(b)} \) denotes the product \( n(n-1)(n-2)(n-3)\ldots(n-b+1) \), while \( S_1 \) and \( S_2 \) are estimated as follows:

\[
S_1 = \frac{\sum_{i=1}^{N} \sum_{j=1}^{n_{ij}} (w_{ij} + w_{ji})^2}{2}
\]

\[
S_2 = \sum_{i=1}^{N} (w_{ii} + w_{ii})^2
\]
where \( w_i \) and \( w_j \) are the row and column totals of the weight matrix, \( \sum w_{ij} \) and \( \sum w_{ji} \) respectively, while \( b_2 \) is the sample kurtosis coefficient that equals:

\[
b_2 = \frac{m_4}{m_2^2}
\]

where \( m_4 \) is the fourth and \( m_2 \) the second sample moment about the mean:

\[
m_4 = \frac{1}{n} \sum (z_i - \bar{z})^4
\]

\[
m_2 = \frac{1}{n} \sum (z_i - \bar{z})^2
\]

After computing \( z \)-score, the null hypothesis may be tested. As defined by the European Commission [26] and advised by EUROCONTROL [27], the confidence level was set at 95%. Hence, in order to reject the null hypothesis, standard deviation should be \(-1.96 < z \)-score > 1.96, while the probability should be \( p \)-value < 0.05. In cases when the null hypothesis cannot be rejected, the spatial distribution is the result of a random spatial process. In other words, the European airspace is fragmented from the performance-based aspect. Otherwise, in cases when the null hypothesis can be rejected, it means that the European airspace is not fragmented, i.e., the high-value and low-value spatial distribution in the data set is spatially clustered.

Further data manipulation framework through utilization of the Moran’s I scatter plot identifies spatial outliers, i.e., identifies local instability in spatial associations. It is conceptualized so that its horizontal axis denotes observed values \( y_i \), while the vertical axis marks the spatial lag \( [Wy]_i \), which equals:

\[
[Wy]_i = \sum_{j=1}^n w_{ij}y_j
\]

Determination of spatial outliers is performed based on four indicators (arising from four quadrants of the scatter plot). Quadrant I (representing high values in a high value neighborhood) and quadrant III (low values in a low value neighborhood) denote spatial objects that are spatially aligned with their neighbors. On the other hand, spatial objects that fall under quadrant II or IV represent spatial outliers. Quadrant II reveals spatial outliers of a low value in a high value neighborhood, while quadrant IV denotes spatial outliers of a high value in a low value neighborhood. Findings obtained by Moran’s I scatter plot are also used in form of a control chart. Its main goal is to determine performance gaps between performance levels of spatial objects, regulatory defined levels and neighbor levels. As a result, it identifies spatial objects with under and over defined performance levels and targets.

Conceptually, the Moran’s I scatter plot is complemented with a linear regression which has global Moran’s I as the slope:

\[ y = a + bx \]

where coefficients \( a \) and \( b \) equal:

\[
a = \frac{\sum y_i x_i^2 - \sum x_i \sum y_i x_i}{n \sum x_i^2 - (\sum x_i)^2}
\]

\[
b = \frac{\sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}
\]

Also, it is supplemented with standard distance assessment and quadrant analysis. Standard distance represents the spatial equivalent of standard deviation applied with a goal to capture absolute dispersion in a point pattern. After determination of the mean center, the standard distance incorporates the straight-line of each point from the mean center as follows:

\[
S_0 = \sqrt{\frac{\sum (x_i - \bar{x})^2 + (y_i - \bar{y})^2}{n}}
\]

Similar to spatial autocorrelation, it is calculated globally and locally. However, unlike spatial autocorrelation, it provides information about spatial compliance by considering performance levels of spatial objects and their neighbors. The shorter the distance between points, the more similar they are (and vice versa). Furthermore, the quadrant analysis is used to determine the frequency of a point pattern distribution. Once a set of 0.5×0.5 cells is superimposed over the Moran’s I scatter plot, the number of points within each cell is calculated as follows:

\[
\lambda = \frac{n}{A}
\]

where \( n \) denotes the number of spatial objects within the cell, while \( A \) marks the cell size.

Since Moran’s I scatter plot gives no information on where significant patterns appear [28,29], data manipulation framework was complemented by two more assessments. Local significance assessment identifies spatial objects whose performance level significantly differs from the neighbor level. On the other hand, local clustering assessment detects locally significant patterns of spatial association. Both assessments are based on standardization of local
indicators of spatial autocorrelation where their significance is tested based on the assumption of a standard normal distribution:

\[ f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} N(0,1) \] (28)

The confidence level is of the order of a two-sigma effect for the local clustering assessment. For the local significance assessment, it is determined based on the empirical rule (also known as the 68-95-99.7 rule). Lastly, to determine spatially optimal performance levels and targets, what-if analysis is included within the data manipulation framework. Its purpose is to identify changes in the network model resulting from the change of one or more input data. As a result, it captures the change ratio in the input and output data. What-if analysis is followed by sensitivity assessment which captures the ratio of the percentage change in the output by the percentage change in the input. The higher the sensitivity figure, the more sensitive the output is to change in the input.

3.3 Methodological limitations

Before the model applicability overview, a few methodological limitations need to be considered. Firstly, it is a particularly undesirable situation when a spatial object does not have neighbors. Hence, each spatial object should have at least one neighbor. In case there is a spatial object with no neighbors, it is referred to as an isolate or island. As a result, all elements in a row in the spatial weights matrix corresponding to such spatial object will equal \( w_{ij} = 0 \), \( \forall j \). As spatial analysis is about interactions, and isolates do not interact, such spatial objects must be excluded from the data set. Example from practice can be found when performing cost-efficiency based assessment. Besides the EUROCONTROL Member States, several affiliate Member States also participate in the ANS charging scheme. One of the affiliate Member States is Uzbekistan. However, Uzbekistan is not spatially connected with the rest of the area where the ANS charging scheme is applicable. Therefore, it represents an island in the context of the model. Consequently, its data, although available, should not be considered. Secondly, no spatial object can be adjacent to all other spatial objects. Also, it should be emphasized that data manipulation is based on first-order adjacency. Nevertheless, it is possible to perform data manipulation so that it considers second-order adjacency. In such case, two spatial objects \( i \) and \( j \) are second-order adjacent if a third area \( k \) exists. Therefore, \( i \) and \( k \) are adjacent, \( k \) and \( j \) are adjacent, but \( i \) and \( j \) are not adjacent. However, due to different sizes of the spatial object, using this approach would result with distance threshold being highly variable. Hence, since empirical studies show that the spill-over effect decreases with spatial distance [30], first-order adjacency if preferred. Also, the issue of input data collection can be outlined as methodological limitation. By obtaining the data capturing performance level of the lower geographical or operational level, i.e., by changing the measurement scale, it would be possible to identify spatial instabilities in spatial associations more accurately. This issue is mainly related to performance data sources (data originators) whereas, for example, the Performance Review Unit (PRU) makes publicly available only the data sets consolidated at national, ANSP, AoR or FIR level. The last methodological limitation is associated with the fact that it is unfortunately often difficult to improve one performance segment without compromising other segment(s) in the ATM system in Europe [31]. Since causal relationship exists within the performances of one ANSP, consequently findings supporting the European airspace defragmentation from one performance segment may also require adjustments within other performance segments.

4. MODEL APPLICABILITY OVERVIEW

With the goal to eliminate the shortcomings identified within the first part of the research paper, the European airspace (de)fragmentation assessment model was developed. From its applicability viewpoint, through determination of spatially optimal performance levels, it contributes to the performance-based airspace defragmentation.

The model applicability arises from several aspects. First of all, the developed model is SES compatible. From the methodological aspect, it enriches the SES postulates of the collaborative and coordinated airspace, and air traffic flow management [32]. Secondly, it meets generally acceptable measurement characteristics such as validity, reliability, sensitivity, repeatability, and objectivity. Its applicability also arises from the fact that air traffic demand could be highly spatially variable in the future [33]. Furthermore, a certain decision can often be deemed good or bad a few years or decades after it was made. However, as the ATM is a safety-critical industry, applying such an approach is not an option. Therefore, models as the one developed are usually
Consequently, the applicability of the developed model stems from the fact that it can facilitate the decision making process and consequently mitigate business risks. This is of a particular importance as the existence of risks can compromise the realization of strategic goals. Therefore, the developed model can be of great value primarily in the domain of strategic planning and development of the ATM system in Europe.

5. CONCLUSION

The problem of fragmented design of the European airspace has been known for a long time. Despite that, the airspace fragmentation issue was not frequently studied. Therefore, this research paper deals with performance-based airspace fragmentation (as one of several European airspace fragmentation types) and its contribution is twofold. Firstly, it complements literature in field of research by providing insights on the causal relationships that are directly or indirectly contributing to the existence of performance-based airspace fragmentation. Secondly, it presents a conceptual and methodological framework of a novel model that can be used for assessing the performance-based airspace fragmentation level. More importantly, as it provides information that contributes to an inclusive, smart, and spatially oriented development of the ATM system in Europe, it contributes to the European airspace de-fragmentation.

Lastly, scarcity of information on the issue of the European airspace fragmentation might be seen as one of the causes for the lack of focus of collaborative initiatives on fragmentation issues. Therefore, future research through the application of the developed model will allow for conducting several case studies. The expected result of model utilization and related studies is to support strategic planning and development function of the ATM system in Europe by delivering information on where, when, how, and whether it is possible to achieve performance-based airspace de-fragmentation.

REFERENCES


