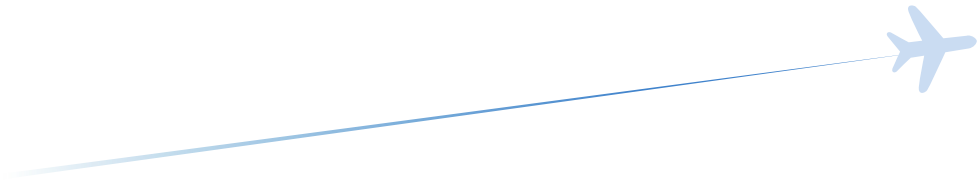


D5.1 Methodology for the empirical investigation of aircraft trajectories

Deliverable 5.1

ADAPT

Grant:	**783264**
Call:	**H2020-SESAR-2016-2**
Topic:	**SESAR-ER3-03-2016 Optimised ATM Network Services: TBO**
Consortium coordinator:	**Università degli Studi di Trieste**
Edition date:	**28 February 2019**
Edition:	**01.01.00**

Founding Members

EUROPEAN UNION EUROCONTROL

SESAR JOINT UNDERTAKING

Authoring & Approval

Authors of the document

Name/Beneficiary	Position/Title	Date
Salvatore Miccichè / Università degli Studi di Palermo	WP5 Leader / Associate Professor	26 February 2019
Giuseppe Pappalardo / Università degli Studi di Palermo	WP5 Participant / Postdoc	26 February 2019
Salvatore Miccichè / Università degli Studi di Palermo	WP5 Leader / Associate Professor	26 February 2019

Reviewers internal to the project

Name/Beneficiary	Position/Title	Date
Rosario N Mantegna/ Università degli Studi di Palermo	WP5 Participant / Full Professor	27 February 2019
Tatjana Bolic / Università degli Studi di Trieste	WP 3 Leader / Senior researcher	27 February 2019

Approved for submission to the SJU By — Representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date
Lorenzo Castelli / Università degli Studi di Trieste	Project coordinator / Assistant professor	28 February 2019

Rejected By - Representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date
N/A		

Document History

Edition	Date	Status	Author	Justification
01.00.00	21 December 2018	Release	ADAPT Consortium	New document for review by SJU
01.01.00	28 February 2019	Release	ADAPT Consortium	Review comments addressed

ADAPT

ADVANCED PREDICTION MODELS FOR FLEXIBLE TRAJECTORY-BASED OPERATIONS

This deliverable is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement No 783264 under European Union's Horizon 2020 research and innovation programme.

Abstract

This deliverable presents the results from the empirical investigation performed in Task 5.1.1 to identify trajectory preferentialities and define new metrics to measure sector congestion.

Table of Contents

Executive summary

The scope of ADAPT is to propose a set of methods and tools (a solution) at the strategic and/or pre-tactical level of network management that is conducive to the trajectory-based operations, which clearly demonstrates the flexibility, information exchange responsibilities, and benefits for all the stakeholders. The aim of the ADAPT project is to adapt, create and test models and metrics that enable strategic planning (early information sharing), by providing the information on flight flexibility and network hotspots, which can eventually be integrated into the Network Operations Plan (NOP) and serve as a basis for stakeholder collaboration.

The main objective of this deliverable is to provide and illustrate a set of metrics that will be used to detect network hotspots and that can also be used as a quantitative instrument to check whether the strategic solution developed in WP3 are effective in the goal of providing strategic solutions in the flight planning operations that clearly show flexibility

As indicated in the proposal, the main tool we propose as a metric for measuring hotspots is the di-fork metric that will be dealt with in chapter 3.

However, we will try to reinterpret the concept of hotspots by enlarging its operational meaning as to include three different facets:

- Hotspots as places where controllers' actions are taken.
- Hotspots as places where congestion is present.
- Hotspots as places where disruptions are present.

The message we want to convey with this document is that: when considering controllers' actions the di-fork metric discussed in chapter 3 is the right metric to consider; when considering congestion, the complexity metrics discussed in chapter 4 are the right metrics to consider; when considering disruption the right way to address this issue is the percolation approach discussed in chapter 5.

The approaches discussed in chapter 4 and 5 are less mature than the di-fork approach discussed in chapter 3. Next steps will include further refinements of such metrics that will be presented in deliverable 5.2 (August 2018).

1 Introduction

The scope of ADAPT is to propose a set of methods and tools at the strategic and/or pre-tactical level of network management that is conducive to the trajectory-based operations, which clearly demonstrates the flexibility, information exchange responsibilities, and benefits for all the stakeholders. The aim of the ADAPT project is to adapt, create and test models and metrics that enable **strategic planning** (early information sharing), by providing the information on **flight flexibility** and **network hotspots**, which can eventually be integrated into the Network Operations Plan (NOP) and serve as a basis for stakeholder collaboration.

The ADAPT project consists of three main activities:

1. Development of the ADAPT strategic solution.
2. Tactical assessment.
3. Visualisation.

This deliverable essentially is part of the tactical assessment activities. In fact, it reports on the activities done within subtask 5.1.1 that deals with an investigation of the micro (within the sector) impact, the aircraft trajectories generated in the strategic phase (WP3). Our main scope in this document is to describe the metrics that will be used to assess the existence of hotspots in the air traffic system.

One of the peculiarities of the work presented here is the approach we will follow to investigate the ATM system. In fact, our analyses will be based on (i) the use of techniques typical of Complex Systems theory and Statistical Physics and (ii) a description of the ATM system in terms of complex networks. In fact, we will base our research activities on the fact that the ATM system can be described as a network, i.e. a graph were nodes are airports or navigation points or sectors and links are set between any pair of nodes whenever there is a physical connection, i.e. a flight, between them.

The main aim is that of characterising preferential patterns in the considered trajectories. In particular, we will perform an empirical detection of the preferentialities in the navigation points network constructed from the trajectory files obtained as an output of the strategic phase (WP3). These preferentialities will be initially identified by detecting deviations from carefully selected null hypothesis [1, 2]. The intention is to use these analyses to create new network based metrics that can measure, in a quantitative way, aspects of the traffic congestion. These will be based on the complexity indicators already studied in [3,4].

Driven by some input taken during the Advisory Board meeting held on 17/05/2018 in London, the first activity we have done was to redraw the network of navigation points that will be the basis of our microscopic analyses of the ATM system. In fact, the investigation of empirical data relative to day 12/09/2014 shows that in that day all trajectories recorded in the M1 DDR2 files containing the last files flight plans -- and then actually flown according to what is recorded in the M3 files -- crossed over *1169158* navigation points (NVPs). This is a huge number that nevertheless can be substantially

decreased when considering that most NVPs are very close to each other. This reduction will be surely beneficial from a computational point of view, however the most important aspect is that from an operational point of view air traffic controllers (ATCs) do not work at the level of single navigation points (unless for some specific and well known specific navigation points), but rather on larger areas around them. We therefore devised a brute-force procedure such that when two navigation points have a 2-D distance smaller that d_{thresh}=5 km, then they are considered to be the same. We then consider such aggregated navigation points, rather than single navigation points, as the nodes of the network. In chapter 2 we show the details of the aggregation procedure as well as a quantitative estimate of the distortion that such procedure brings into the system.

The analyses of the di-forks in chapter 3 and of the complexity metrics in chapter 4 will be done at the level of these aggregated navigation points that we call navigation balls. Starting from the navigation balls network created from DDR2 M1 and M3 files, in chapter 3 we compute the statistically validated di-forks that are over-(under-) expressed according to a null hypothesis that takes into account the number of connections of each navigation ball in the system. We then compare the geographical localization of these statistically validated di-forks with the critical sectors that are detected in the optimization procedure that brings to the strategic solution proposed in WP3. The comparison shows the mixed nature of the di-fork metric. In fact we may have over-expressed di-forks that are outside of the critical sectors. We take this as a suggestion of the fact that during night the over-expressed deviations detected by the di-forks probably are essentially directs given to facilitate the air traffic flow. However, we may also have localized just at the borders of the critical sectors. We take this as a suggestion of the fact that (especially during daytime) deviations might occur not only to facilitate the air traffic flow but also as a way to solve congestion issues. In chapter 4 we compute the complexity metrics and we show how by considering their global behaviour, rather than considering each metric separately, it is possible (i) to observe clusters of metrics that behave similarly in ta given sector, and (ii) to detect air traffic sectors that become critical for the sectors network, with respect to the specific metric considered.

In chapter 5 we present a novel analysis that will allow us at investigating the presence of hotspots from another point of view. While di-forks can be usefully employed to detect trajectory segments where controllers' actions occur – either for solving congestion problems or for giving directs – the complexity metrics mainly look at the hotspots from the point of view of congestion declined in different ways (ascending, descending, head-on crashes, …). However, we believe that there is another point of view worth of further investigations: hotspots might also be seen as places where potential network disruptions occur. In other words hotspots might be trajectory segments where there is an high risk that, if any problem occur, the networks gets fragmented meaning that it will not be possible to go from any point to any other point of the network. We might investigate these issues by using a percolation approach that has already been used to detect potential bottlenecks in the urban traffic network and other model complex systems [5]. When applied to our system, the percolation approach might be useful to identify those network links (trajectory segments) potentially responsible for massive network disruptions.

In order to provide metrics whose interpretation can be easily done in terms of operational standards, we will try to move from the microscopic level of the navigation balls network to a macro-level closer to the operational scenario. Such macro-level might be that of the air traffic sectors where capacity issues might show up thus limiting the efficiency of the ATM system.

The output of this task is a set of metrics to be used in the next sub-task (5.1.2) that will deal with the tactical assessment itself of the strategic solution developed in WP3. The results of next sub-task (5.1.2) will be described in Deliverable D5.2 (August 2018).

2 Data

In this section we will describe the data used in the deliverable. We will consider both the original data downloaded from EUROCONTROL and those produced by the UNITS unit as the result of the optimization activities performed in WP3 at the strategic phase level.

Moreover, we show in some detail the procedure the follow to process these data before using them for computing the relevant metrics. In fact, rather the working at the level of single navigation points, we work at the level of so-called navigation balls. Essentially, we aggregate navigation points within a radius of d_{thresh}=5 km. We show that this aggregation procedure introduces a negligible distortion in the system, when we represent it as a network.

2.1 Real trajectories

Our database contains information on all the flights that, even partly, cross the ECAC airspace. Data are owned by EUROCONTROL (http://www.eurocontrol.int), the European public institution that coordinates and plans air traffic control for all of Europe. Data were downloaded from the DDr2 historical data section https://ext.eurocontrol.int.

The data come from two different sources. First, we have access to the Demand Data Repository (DDR) from which we have all the trajectories followed by any aircraft in the ECAC airspace. In this deliverable we primarily consider one single day of data, specifically the 12/09/2014. A trajectory, called indifferently flight plan here, is made by a sequence of navigation points crossed by the aircraft, together with altitudes and timestamps. The typical time between two navpoints lies between 1 and 10 minutes, giving a sufficient time resolution for trajectories. In this deliverable we only use the “last filed flight plans” (M1 SO6 files) and the “realized trajetories” (M3 SO6 files), which are the planned trajectories – filed from 6 months to one or two hours before departure and the real trajectories updated with radar tracks, respectively.

The other source of information are the NEVAC files that contain all the elements allowing the definition (borders, altitude, relationships, time of opening and closing) of the elements of airspaces, namely airblocks, sectors, airspaces (including FIR, National Airspace, ACC, etc.). The active elements at a given time constitute the configuration of the airspace at that time. These files allow to determine the configuration of the airspaces for an entire AIRAC cycle. Here we only use the information on sectors, airspaces and configurations to rebuild the European airspace. Specifically, at each time we have the full three dimensional boundaries of each individual sector and airspace in Europe.

2.2 WP3 strategic solution trajectories

Another source of data consists in the set of trajectories produced in WP3 [6]. We considered data generated for the day 12/09/2014, as a result of the WP3 activities . The trajectories coming from WP3 are of the same format as the actual M1 data for the chosen test day. These trajectories are the result of the EFPS model optimisations. The EFPS model assigns the trajectory, departure time and flexibility measure (TW) for all the flights of the ECAC network for the entire day of traffic. This enables to identify constrained flights, and the sector-hours that are imposing the constraints. As the EFPS models use as input a set of trajectories (larger than what was flown on any day), and minimizes the total flight costs, subject to capacity constraints, the resulting M1 trajectories are different from the actual M1 trajectories submitted on the chosen test day. The details on the data input preparation can be found in D2.1 Support to modelling activities [11], and the description of EFPS model and the results are given in deliverable D3.1 Flight flexibility and hotspots in the ADAPT solution [12].

For the selected day 12/09/2014, the WP3 has provided a file containing all the OD pairs of that day, their departure time, the list of main navigation points crossed during the flight and the time at which they entered each air traffic sector traversed by the considered flight. A set of additional information was also provided that allowed us to retrieve the M1 trajectories that represent the ADAPT solution results for that day.

2.3 Aggregated navigation points

The investigation of empirical data relative to day 12/09/2014 shows that in that day all trajectories recorded in the M1 files -- and actually flown according to what is recorded in the M3 files -- crossed over *1169158* navigation points. This is is a huge number that nevertheless can be substantially decreased when considering that most of them are very close to each other.

In fact we devised a brute-force procedure such that when two navigation points have a 2-D distance smaller that d_{thresh}=5 km, then they are considered to be the same. The specific procedure used to aggregate the navigation points closer than d_{thresh} is as follows:

We first compute the distance between all navigation point pairs. In order to speed up the process and reduce the computational effort needed to process all possible pairs, we do not consider pairs were latitude and longitude differ already at the level of the second decimal digit.

Then we select the first navigation point N_1 in a randomly formed list L_{day} of all navigation points in that day and start selecting all other navigation points N_{1i} that are far away from N_1 for less than d_{thresh}. We thus form a list L_1. At the second step we consider the second navigation point N_1 belonging- to L_{day} and not to L_1. We therefore start selecting all other navigation points N_{2i} that are far away from N_2 for less than d_{thresh}. We thus form a list L_2. We thus iterate such procedure until all navigation points from L_{day} are considered.

Starting from a list L_{day} of 1169158 navigation points we ended up with 148743 such lists L_i, i=1, …, 148743. We therefore aggregated the original 1169158 navigation points into a smaller set of 148743 so-called navigation balls.

Then we select the first navigation points in the list L_i to represent the i-th navigation ball. Afterwards we substitute any navigation points N in the M1 and M3 files with the representative navigation point associated to the navigation ball whom N belongs to. As an example, in Figure 1

we show a pictorial representation of the severe filtering introduced by the aggregation process in the case of the Sicilian airspace.

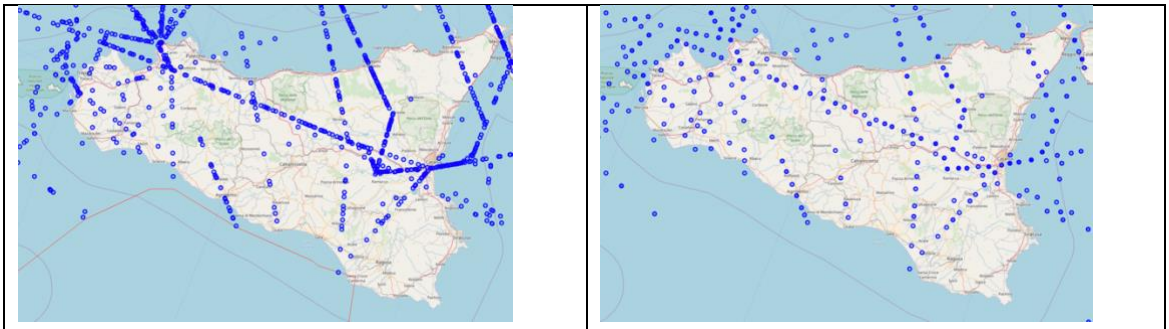


Figure 1: *Pictorial representation of the severe filtering introduced by the aggregation process in the case of the Sicilian airspace.*

In order to check the effectiveness of the aggregation procedure, we constructed the navigation points network and the navigation balls network in order to highlight possible differences in their statistical properties as measured by some standard network metrics. In the navigation points network, the nodes are given by all considered navigation points, and a link is set between two navigation points whenever there is a flight connecting them, the total number of such flights being the link weight. Analogously, In the navigation balls network, the nodes are given by all considered navigation balls, and a link is set between two navigation balls whenever there is a flight connecting any two navigation points inside them, the total number of such flights being the link weight. When construction these networks we considered the flights recorded in the M1 files only.

The navigation points network of day 12/09/2014 has 283471 nodes and 470393 links. The corresponding navigation balls network has 82895 and 219668 links. In Figure 2 we show the degree distribution (top-left) and the strength distribution (top-right) of the two networks. As expected the degree distribution spans a larger set of values. In fact, points and many more than balls and therefore there are more connections between points than between balls. Analogously, the strength distribution of the navigation balls network shows higher values than the one associated to the navigation points network, given that the balls are aggregation of pints and therefore for each ball there are more flights than for single points. A similar argument holds for the betweenness distribution shown in the bottom-left panel. However, in all cases the two distributions show tails that decay in a very similar way, thus confirming that the distortion introduced by the aggregation process is negligible.

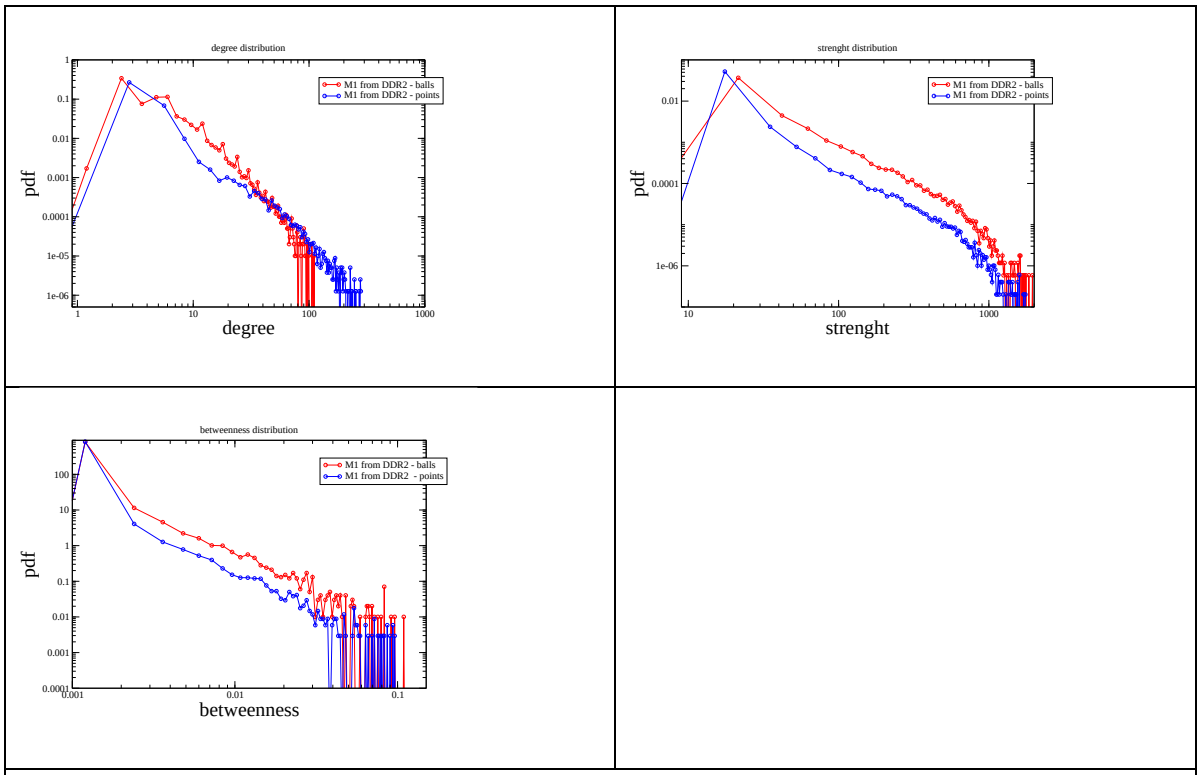


Figure 2: *Comparison of the navigtion points network and navigation balls network constructed from the M1 files related to day 12/09/2014. The top-left panel shows the degree distribution. The top-right panel shows the strenght distribution. The bottom-left panel shows the betweenness distribution, while the bottom-right panel shows the average-path length distribution.*

The above result suggests that the severe filtering introduced by the aggregation procedure induces negligible distortions on the system when we represent it through a network approach.

2.4 Comparison of real and WP3 strategic solution trajectories

In this section we again want to use network theory to compare the real trajectories with those obtained with the WP3 strategic solution.

In particular, we compared the navigation balls network constructed starting from the DDR2 M1 files of day 12/09/2014 and the navigation balls network constructed starting from the synthetic M1 file reconstructed from the data produced in WP3 by the UNITS unit. The corresponding navigation balls network has 72538 and 185039 links. In Figure 3 we show the degree distribution (top-left), the strength distribution (top-right), the betweenness distribution (bottom-left) and the average-path-length distribution (bottom-right) of the two networks. The network properties of the two navigation balls networks are essentially the same, thus indicating that the models used in WP3 gives a reasonable replication of the real trajectory planning, at least when network metrics are considered.

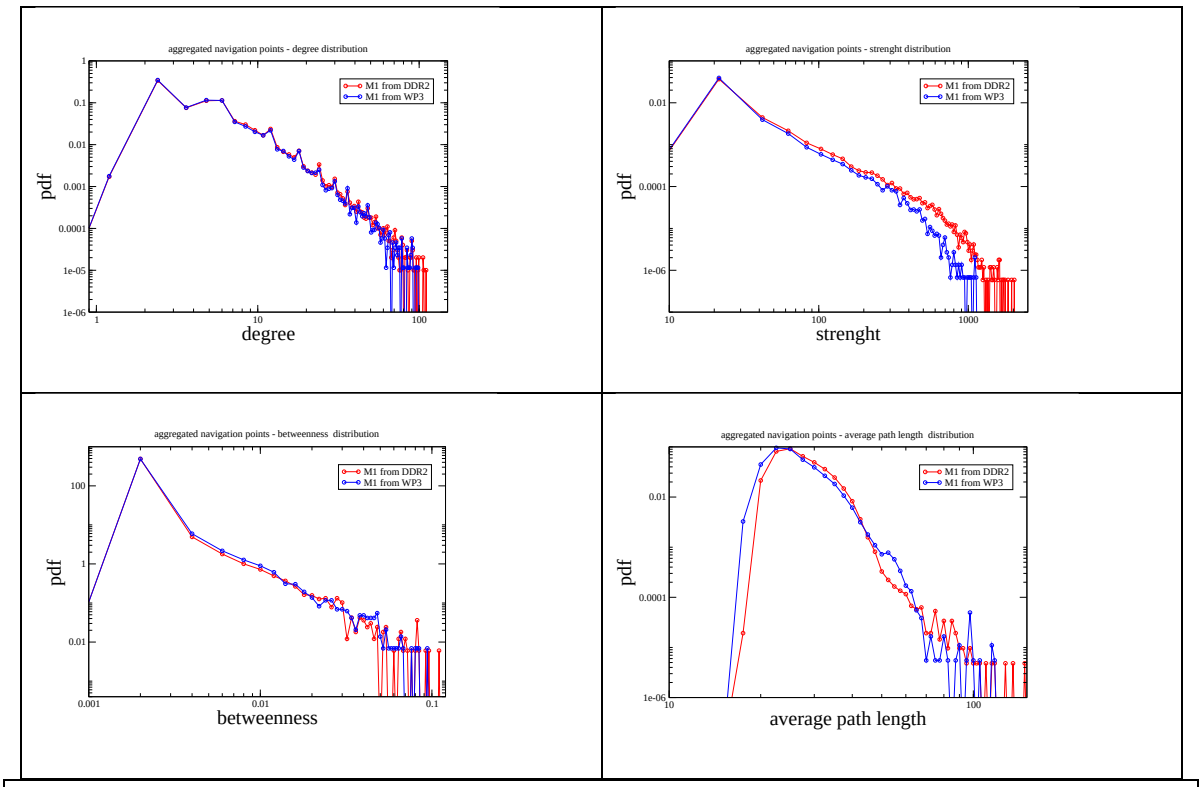


Figure 3: *Comparison of standard network metrics obtained from the M1 files downloaded from DDR2 and from the M1 files reconstructed starting from the WP3 strategic solution trajectories. The top-left panel shows the degree distribution. The top-right panel shows the strenght distribution. The bottom-left panel shows the betweenness distribution, while the bottom-right panel shows the average-path length distribution.*

As a further test, we decided to investigate how our networks can be partitioned in smaller clusters of elements by applying a community detection algorithm.

In a first investigation we considered the Infomap community detection algorithm. In this case the network generated starting from the original M1 files is partitioned into 3438 communities of sizes ranging from 3 to 221. The partitioned network retains 178645 out of the 219668 original links. The network generated starting from the synthetic M1 files obtained from WP3 trajectories is partitioned into 3119 communities of sizes ranging from 3 to 195. The partitioned network retains 150564 out of the 185039 original links. The structure of the communities found in the two networks is essentially the same, as shown in the left panel of Figure 4. In order to quantitatively check whether the two partitions are similar, we computed the ARI (Adjusted Rand Index) metrics. This is defined as the ratio between the number of node pairs that are classified either in the same cluster or in different clusters in the two partitions, divided by the total number of node pairs. Since we have 219668 pairs in the M1 network and 185039 pairs in the one reconstructed from the WP3 data, we performed the analysis on the 149328 pairs that the two networks have in common. We find that 109481 pairs are classified in the same cluster in both partitions, while 17112 pairs are classified in different clusters in both partitions. Therefore we have an ARI value of ARI=0.848, which is quite high.

In a second investigation we considered the Radatool community detection algorithm. In this case the network generated starting from the original M1 files is partitioned into 89 large communities of sizes ranging from 2 to 27286. The partitioned network retains 213527 out of the 219668 original links. The network generated starting from the synthetic M1 files obtained from WP3 trajectories is partitioned into 88 communities of sizes ranging from 2 to 24671. The partitioned network retains 180409 out of the 185039 original links. The structure of the communities found in the two networks is essentially the same, as shown in the right panel of Figure 4. In order to quantitatively check whether the two partitions are similar, we computed the ARI (Adjusted Rand Index) metrics: we have an ARI value of ARI=0.951, which even higher than the previous value.

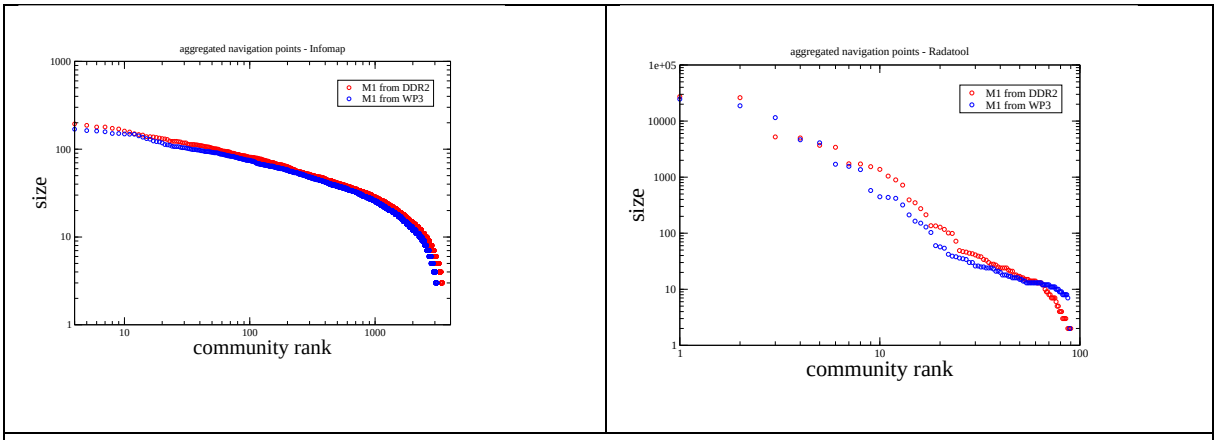


Figure 4: *Comparison of the partition sized obtaiend with two different community detection algorithms for the real trajectories (red points) and WP3 trajectories (blu points). Data on the left refer to the INFOMAP algorithm and data on the right refer to the RADATOOL algorithm.*

The partitions thus obtained can be displayed at the following links:
INFOMAP real M1: https://www.dropbox.com/s/cv12rxjuyzefx5y/clusters_INFOMAP.html?dl=0
INFOMAP WP3 M1: https://www.dropbox.com/s/qg5n9nqklthfj8n/clusters_SAIPE_INFOMAP.html?dl=0
RADATOOL real M1: https://www.dropbox.com/s/v1t1r4wpp960fmi/clusters_RADATOOL.html?dl=0
RADATOOL WP3 M1: https://www.dropbox.com/s/zy11n021a9nuwiy/clusters_SAIPE_RADATOOL.html?dl=0
Our results show that the model developed by the UNITS unit reproduces the real data quite well, as the statistical properties investigated by using a network based approach are essentially the same in the two cases.

3 The di-fork metric

The main aim in this chapter is that of characterising preferential patterns in the considered trajectories. In particular, we will perform an empirical detection of the preferentialities in the flown trajectories. These preferentialities will be identified by detecting deviations from carefully selected null hypothesis [1,2].

The analyses of the di-forks in this chapter will be done at the level of the aggregated navigation points, i.e. the navigation balls. Starting from the navigation balls network created from DDR2 M1 and M3 files, we compute the statistically validated di-forks that are over-(under-) expressed according to a null hypothesis that takes into account the number of connections of each navigation ball in the system.

While the di-fork metric has already been introduced in Ref. [2] and Ref. [1], in this deliverable we show a simple generalization of this metric that also takes into account the flight level at which the considered trajectory segments are flown.

The aim of the work presented in this chapter is to illustrate the di-fork metric and its capabilities in detecting preferentialities in the flown trajectories. To this end, we will consider here the M1 and M3 trajectories that come from the DDR database, see section 2.1 and not the WP3 trajectories described in section 2.2. In fact, as we will see, the di-fork metric is based on a comparison between planned and realized trajectories. According to the project plan, the WP3 trajectories play the role of M1 planned trajectories and will be used to feed an agent-based model that will generate simulated trajectories that will play the role of M3 realized trajectories. These activities will be described in deliverable D5.2.

We compare the geographical localization of these statistically validated di-forks with the critical sectors that are detected in the optimization procedure that brings to the strategic solution proposed in WP3. Although the di-forks are obtained with real M1 and M3 trajectories while critical sectors are relative to the WP3 trajectories, we nevertheless deem useful to present this analysis because our results in the previous chapter show that the statistical properties of the real M1 and WP3 trajectories are similar and therefore we believe that the comparison might still give useful insights about the air traffic system. Indeed, the comparison shows the mixed nature of the di-fork metric. In fact we may have over-expressed di-forks that are outside of the critical sectors. We take this as a suggestion of the fact that during night the over-expressed deviations detected by the di-forks probably are essentially directs given to facilitate the air traffic flow. However, we may also have localized just at the borders of the critical sectors. We take this as a suggestion of the fact that (especially during daytime) deviations might occur not only to facilitate the air traffic flow but also as a way to solve congestion issues.

Such findings confirm our view, according to which the di-fork metric is the right metric to measure the air traffic controllers' actions when managing the aircraft trajectories. We believe this is one of relevant aspect to be considered when one is interested in the detection of the hotspots in the ATM system.

3.1 Definition of di-fork

In this section we first discuss a metric, called ***fork***, that is used for characterizing the differences between planned and realized flight trajectories at the level of each single navigation point. The fork metric was introduced in Ref. [2]. Let us first provide a qualitative description of it. For each flight, we consider the last navigation point which is common to the planned and the realized flight trajectory. At this point, we consider that a "fork" happens when the flight trajectory is deviated from the planned one. By counting the number of flights which are deviated from the considered navigation point and by dividing it by the total number of flights flying through the navigation point in the selected time interval, we obtain a quantitative indicator of how much the navigation point is a "source" of deviations for the planned flight trajectories. This quantity varies between 0 and 1 and is computed for each navigation point.

Hereafter we are providing a more formal definition. Let us consider a certain time window Δt. Let us consider a generic navigation point P appearing in at least one of the realized flight trajectories. Let us call $pF_{\Delta t}(P)$ the number of flights passing through P as observed in the planned flight trajectories. Let us call $dF_{\Delta t}(P)$ the number of flights passing through P, as observed in the realized flight trajectories, and missing the next navigation point Q as indicated in the corresponding planned flight trajectory, see figure 5. The fork $F_{\Delta t}(P)$ is defined as the ratio $F_{\Delta t}(P) = dF_{\Delta t}(P)/pF_{\Delta t}(P)$. By construction, this metric aggregates the information on the different flight trajectories that in a certain time window Δt are passing through P.

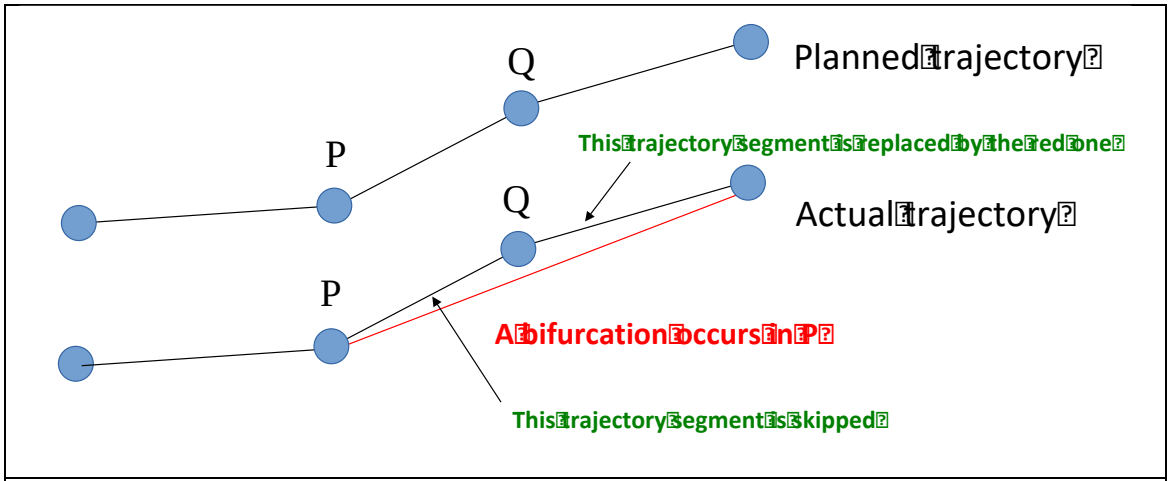


Figure 5: *Illustrative example of the fork metric.*

This metric already produced some interesting results presented in [2]. However, its weakness relies in the heterogeneity of trajectories which can cross a single navigation point. Indeed, controllers are managing flows, i.e. ensemble of trajectories, and for them the navigation point is a support to the flow. As a consequence, different flows crossing at a given navigation point can be managed differently.

Therefore, in Ref. [1] we introduced a slightly different metric, where we take into account the direction of the flow as well as the navigation point itself.

Let us consider pairs $C(P_j,P_k) = (P_j,P_k)$ of navigation points that are consecutively crossed according to a certain flight plan. The navigation point pairs we consider are ordered and therefore (P_j,P_k) and (P_k,P_j) are different pairs describing flights passing through the same pair of navigation points but moving in the opposite direction.

Similarly to the previously mentioned fork metric, the directional fork (or di-fork) $\Phi_{\Delta t}(C)$ associated with an ordered navigation point pair C is defined as the ratio $\Phi_{\Delta t}(C) = DF_{\Delta t}(C)/PF\Delta t(C)$ where $PF_{\Delta t}(C)$ is the number of flights planned to flow through P_j and P_k in the direction from j to k and $DF_{\Delta t}(C)$ is the number of flights actually crossing P_j and then deviated to a navigation point different from P_k in the considered time window Δt. In other words, the first navigation point is the one crossed by the aircraft and the second one is the navigation point present in the planned flight trajectory but not present in the realized flight trajectory. This definition allows us to investigate the deviations as a function of the different directions, and to have a more flow-based metric. It is worth emphasizing again that the di-fork metric refers to navigation point pairs, while the fork metric of Ref. [2] refers to single navigation points.

In this deliverable for the first time we introduce a further enhancement of the metric, by taking into account the flight level. In fact, the definition in itself remains the same. The only difference is that navigation points are now identified not only by a 2-uple containing their 2-D coordinates but we instead consider a 3-uple that in addition contains the information about the flight level at which it is crossed.

Below we investigate the capabilities of the di-fork metric in providing a statistical characterization of the deviations occurring in the flight trajectories. More specifically, we are interested in seeing how the statistical facts we found in section IV are present at the microscopic level, i.e. at the navigation point pair level.

3.2 Statistical Validation of di-fork

Here we investigate whether the flight trajectory deviations are randomly distributed over the day or rather if they are over-expressed or under-expressed for specific navigation point pairs. This type of investigation cannot be done only in terms of the occurrence of the flight trajectory deviations because the number of flights passing through a specific navigation point pair in a given time interval is a quite heterogeneous variable. We therefore estimate the over-expression and under-expression of flight trajectory deviations by considering navigation point pairs and trying to compare the occurrences of flight trajectory deviations observed in this pair with an appropriate null model.

In this section we investigate the navigation point pairs $C(P_j,P_k)$ for which the air traffic flow is from P_j to P_k. Suppose that during a specific day we have PF_{day} flights with planned flight trajectories connecting P_j to P_k in a step. Suppose also that DF_{day} is the number of flights passing through the first navigation point P_j and deviating from the successive navigation point P_k in the same day. Let us now define $PF_{\Delta t}$ the flights that are planned to fly through P_j and P_k during an intraday time

interval Δt. We can estimate what is the probability of observing a number $DF_{\Delta t}$ of flights flying through P_j and then deviating from P_k during the same time interval. By assuming that for each navigation point pair, the flight trajectory deviation events are independent the one from the other, a good approximation of the probability of detecting $DF_{\Delta t}$ is given by the hypergeometric distribution:

$$H\left(DF_{\Delta t}\middle|PF_{day}, DF_{day}, PF_{\Delta t}\right) = \frac{\binom{DF_{day}}{DF_{\Delta t}}\binom{PF_{day}-DF_{day}}{PF_{\Delta t}-DF_{\Delta t}}}{\binom{PF_{day}}{PF_{\Delta t}}}$$

By using this value of the probability of observing $DF_{\Delta t}$ deviated flight trajectories we can obtain for each navigation point pair $C(P_j,P_k)$ a p-value for the over-expression or the under-expression of $DF_{\Delta t}$. The probability of Eq. (3) allows us to associate a p-value $p(DF_{\Delta t})$ with the actual number $DF_{\Delta t}$ of detected deviation. Specifically, for over-expression (OE) we have

$$p_{OE} = 1 - \sum_{X=0}^{DF_{\Delta t}-1} H\left(X\middle|PF_{day}, DF_{day}, PF_{\Delta t}\right)$$

whereas for under-expression (UE) we have

$$p_{UE} = \sum_{X=0}^{DF_{\Delta t}} H\left(X\middle|PF_{day}, DF_{day}, PF_{\Delta t}\right)$$

Since we are performing this test for all possible navigation point pairs $C(P_j,P_k)$, we have to use a correction for multiple hypothesis test comparison. Specifically we use the FDR multiple hypothesis test correction therefore, after sorting the p-values in increasing order, in order to reveal over-expression or under-expression we select those navigation point pairs with p-values which are below the straight line with null intercept and slope equal to $0.01/(2\ N_{pair}\ N_t)$ where N_t = 24 is the number of used time bins and N_{pair} is the number of possible pairs we tested. The factor 2 is taken into account because we want to consider both over- and under-expressions.

Finally, we want to emphasize that over-(under-)expressed di-forks do not in general correspond to having more (less) deviated aircraft. Not in absolute terms. As also explained in Ref. [1], over-expressed di-forks measure deviations that occur more frequently than expected on the basis of a realistic null hypothesis of randomness that takes into account the heterogeneity present in the system, i.e. specifically the number of aircraft crossing the navigation-points in the planned and realized trajectories. For a detailed discussion of this point, we refer the reader to section V.B of Ref. [1].

3.3 Definition of di-fork with altitudes

The di-fork definition reported above does not take into account the altitude at which aircraft cross each of the two navigation points in a trajectory segment.

In fact, this aspect can be taken into account by simply considering that each navigation point pair $C(P_j,P_k)$ is determined by the 3-D coordinated of P_j and P_k rather than its 2-D coordinates as in the previous section. In other words, each navigation point pair $C(P_j,P_k)$ is identified by a six-tuple of numbers rather than a four-tuple as previously done.

In the present case, we detect a deviation even if an aircraft flying from P to Q is still crossing Q in the flown trajectory, although at a flight level different from the planned one.

3.4 Results

3.4.1 Results for the real trajectories

In Table 1 for the whole ECAC airspace, we show the result of the detection of over-(under-) expressed di-forks. In particular, we show the number of tested segments (second column) the number of Over-(under-) expressed trajectory segments (third and fourth column) for each of time-intervals shown in the first column. The univariate Bonferroni threshold used in this analysis is 0.01. The results shown in the table and in the subsequent figure are obtained implementing the FDR multiple comparison correction. One can notice that there are very few under-expressed segments. Data in the fifth and sixth column refer to the percentage of detected over-expressions (OEs) and under-expressions (UEs) with respect to the number of considered trajectory segments shown in the second column. One can see that although in absolute numbers the detected OEs are larger during daytime, in percentage, we have many more OEs during night-time than during day-time.

Day and time interval	Trajectory segments	OE	UE	perc OE	perc UE
20140912_010000	5584	158	0	0,028	0,000
20140912_020000	7461	176	0	0,024	0,000
20140912_030000	11307	233	1	0,021	0,000
20140912_040000	19420	271	1	0,014	0,000
20140912_050000	30140	378	3	0,013	0,000
20140912_060000	30491	397	1	0,013	0,000
20140912_070000	34330	419	3	0,012	0,000
20140912_080000	36716	536	6	0,015	0,000
20140912_090000	36524	508	0	0,014	0,000
20140912_100000	36036	602	0	0,017	0,000
20140912_110000	37187	539	0	0,014	0,000
20140912_120000	34839	512	0	0,015	0,000
20140912_130000	36159	564	0	0,016	0,000
20140912_140000	36543	479	0	0,013	0,000
20140912_150000	36153	537	0	0,015	0,000
20140912_160000	36287	470	0	0,013	0,000
20140912_170000	33966	423	0	0,012	0,000
20140912_180000	32096	447	0	0,014	0,000
20140912_190000	29547	465	0	0,016	0,000
20140912_200000	25286	332	0	0,013	0,000
20140912_210000	17001	237	0	0,014	0,000

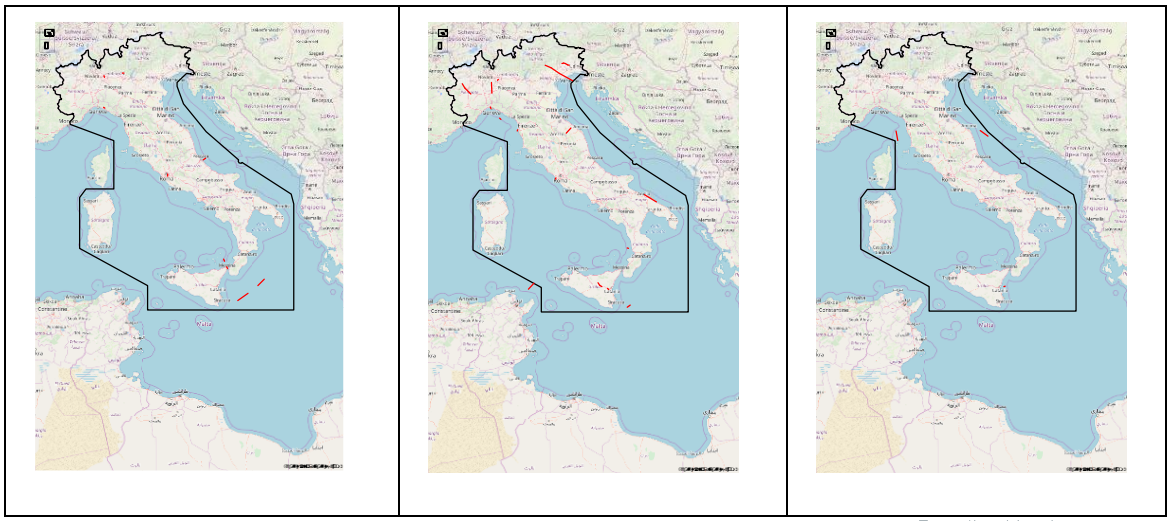
20140912_220000	10855	224	0	0,021	0,000
20140912_230000	7139	174	0	0,024	0,000

Table 1: *For the whole ECAC airspace, we show the number of tested segments (second column) the number of over-(under-) expressed trajectory segments (third and fourth column), the percentage of over-(under-) expressed trajectory segments (fifth and sixth column) for each of time-intervals shown in the first column. The univariate Bonferroni threshold used in this analysis is 0.01. These results are obtained implementing the FDR multiple comparison correction.*

It is worth mentioning that the statistical validation of di-fork is quite sensible to the choice of the airspace that one wants to investigate. Let us explain it by an example: the number of trajectory segments that are active in the ECAC airspace between 13:00 and 14:00 is 36159. Therefore the Bonferroni threshold is set to: $0.01/36159=2.59\ 10^{-7}$. If we want to limit our investigation to Germany only, we have 4821 active trajectory segments, which leads to a Bonferroni threshold of $2.07\ 10^{-6}$. If we want to limit our investigation to Italy only, we have 4821 active trajectory segments, which leads to a Bonferroni threshold of $2.77\ 10^{-6}$. Depending on the considered airspace, we might have different Bonferroni thresholds. Therefore, when performing the analysis over the whole ECAC, the over- (under-) expressed trajectory segments in one country might be a subset of those observed if we perform the analysis at a country level. This limitation is intrinsic in the methodology and it is unavoidable. In fact, this is the price we have to pay in order to avoid false positive due to the large number of performed tests. The Bonferroni correction increases the statistical precision (by minimizing the number of false positive of the test) but decreases the accuracy of the estimation because the severe reduction of the statistical threshold can be associated with the presence of a large number of false negative.

The FDR correction usually is a good compromise between these two opposite requests. As such, we expect that the issues mentioned above would be less relevant with the choice of the FDR correction for multiple comparison. In fact, as an example, when starting from the ECAC airspace, the number of OEs in the time interval between 13:00 and 14:00 in Germany and Italy is 8 and 22, respectively. When performing a country-based analysis we observe 8 OEs in Germany and 23 OEs in Italy. As one can see the differences are quite limited.

For illustrative purposes in Figure 5 we show the geographical localization of over-expressed and under-expressed di-forks in Italy (top panels) and Germany (bottom panels) in the different time intervals of the day: from 04:00 to 05:00 am in the left panels, from 13:00 to 14:00 in the central panels and from 22:00 to 23:00 in the right panels. The tested segment pairs in these three time windows were 1633, 3611 and 746 respectively for the Italian case and 3043, 4821 and 1118 respectively for the German case. The number of over-expressed pairs was 2, 23, 3 respectively for the Italian case and 6, 8, 3 for the German case.



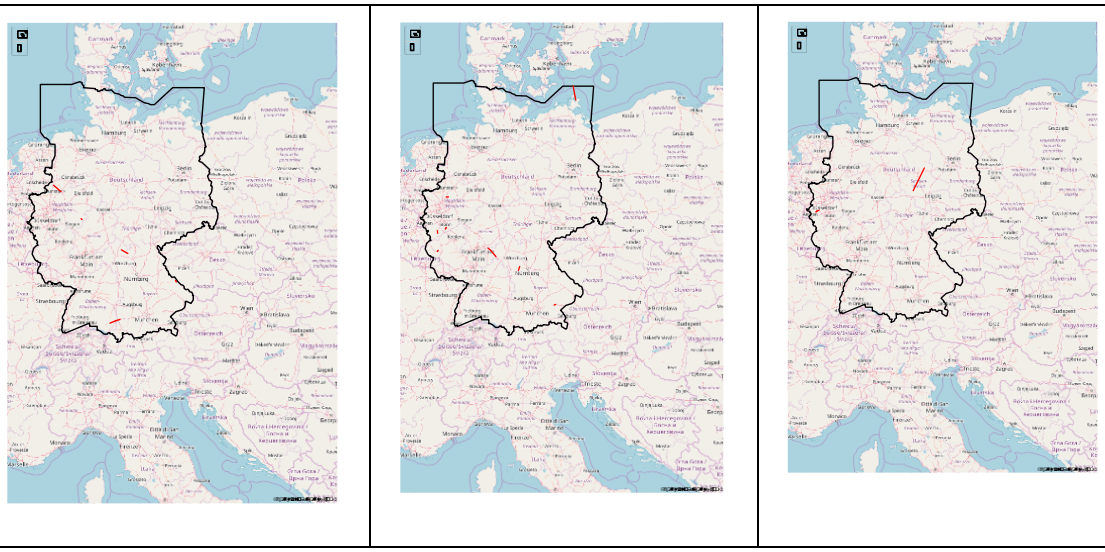


Figure 5: *Geographical localization of over-expressed and under-expressed di-forks in Italy (top panels) and Germany (bottom panels) in the different time intervals of the day: from 04:00 to 05:00 am in the left panels, from 13:00 to 14:00 in the central panels and from 22:00 to 23:00 in the right panels. The results shown in these panels are the result of a country-by-country analysis.*

We have also tested whether the observed di-forks were somehow in connection with the critical sectors that are detected by the UNITS unit when computing the maximal time-window (TW) as part of their strategic optimization. Such sectors are those that limit the length of the TW to a minimum value because extending the TW length would cause capacity infringements that cannot be managed within the model. In Figure 6 we show the geographical localization of the critical sectors and of the di-forks for the German airspace in the time-intervals considered above. The blue lines mark the borders of the critical sectors. The figure shows the mixed nature of the di-fork metric. In fact in the left panel we have over-expressed di-forks that are outside of the critical sectors. We take this as a suggestion of the fact that during night the over-expressed deviations detected by the di-forks probably are essentially directs given to facilitate the air traffic flow. In the central panel we have again di-fork localized outside the sectors as well as a few di-forks that occur close to the borders of the critical sectors. We take this as a suggestion of the fact that during day deviations might occur not only to facilitate the air traffic flow but also as a way to solve congestion issues.

By construction, di-forks involve a comparison between planned and realized trajectories. Specifically, di-forks measure in a quantitative way deviations from planned trajectories and over-(under-)expressed di-forks measure those deviations that are not expected from a null hypothesis of randomness that however takes into account the natural heterogeneity present in the system for that regards the number of flights crossing a certain navigation point.

Our results show that deviations and the associated controllers'actions may arise for two main reasons:

1. during night-time the over-expressed deviations detected by the di-forks probably are essentially directs given to facilitate the air traffic flow. This is in line with the fact that the localization of di-forks and critical sectors is not the same. In this case, di-forks measure points where controllers'actions occur but this are not related to hotspots.
2. During day-time, traffic conditions are obviously different from night-time and we do not expect deviations due to directs. Rather we expect that deviations from planned trajectories might be due to congestion problems. In fact, we observe over-expressed di-forks during day-time and in some case these are close to the WP3 critical sectors. This corroborates the fact that di-forks observed during day-time are a good metric to detect hotspots.

The reasoning we have done so far considers hotspots as synonymous of congestions.

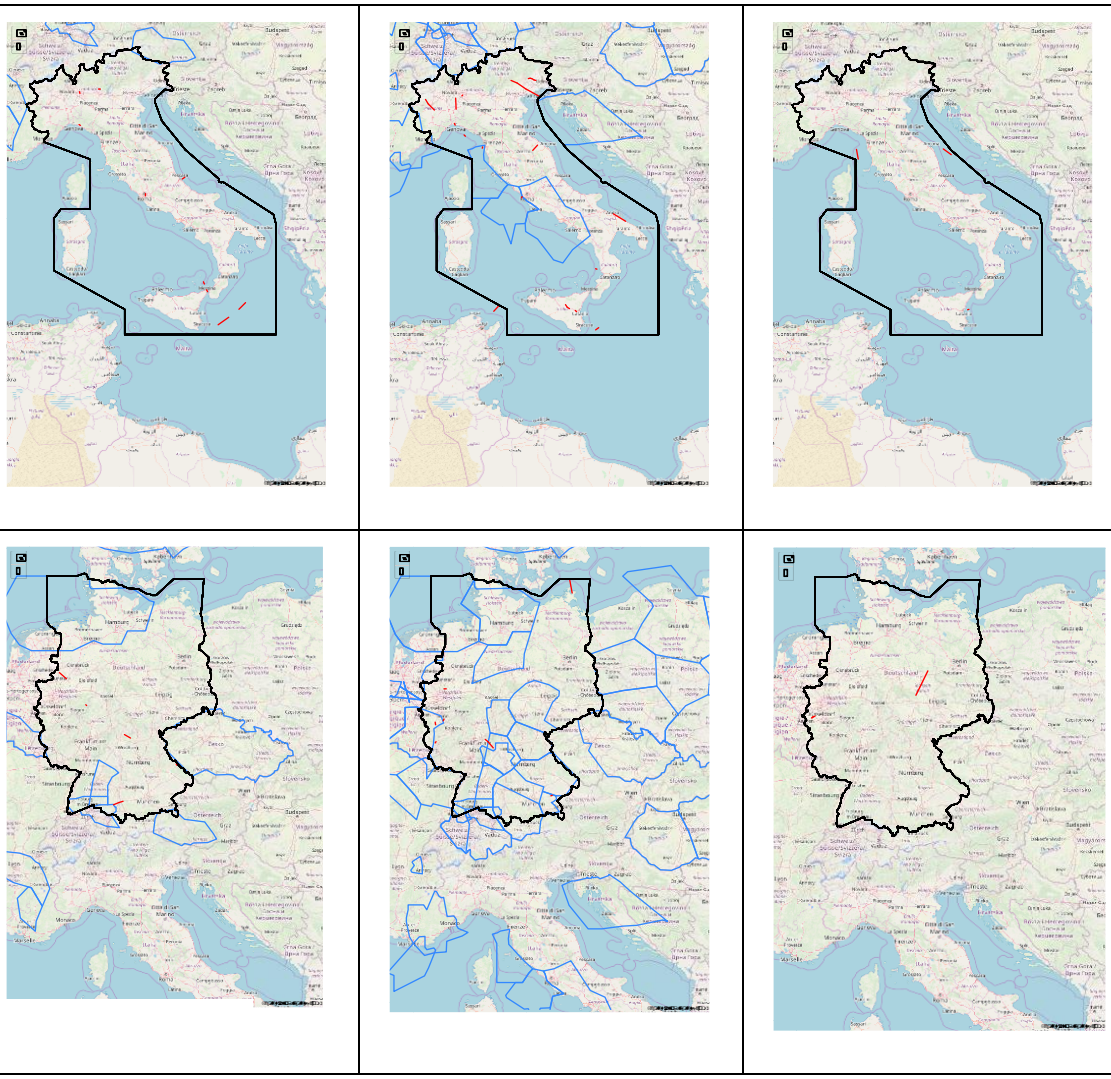


Figure 6: *Geographical localization of the critical sectors and of the overexpressed di-forks for the Italian (top panels) and German (bottom panels) airspaces in the time-intervals considered above. The blue lines mark the borders of the critical sectors. The results shown in these panels are the result of a country-by-country analysis.*

3.4.2 Results for the real trajectories: altitudes

In Table 2 for the whole ECAC airspace, we show the result of the detection of over-(under-) expressed di-forks with altitudes. In particular, we show the number of tested segments (second column) the number of Over-(under-) expressed trajectory segments (third and fourth column) for each of time-intervals shown in the first column. The univariate Bonferroni threshold used in this analysis is 0.01. The results shown in the table and in the subsequent figure are obtained implementing the FDR multiple comparison correction.

Day and time interval	Trajectory segments	OE	UE	perc OE	perc UE
20140912_000000	6640	181	0	0,027	0,000
20140912_010000	6919	150	0	0,022	0,000
20140912_020000	9544	205	0	0,021	0,000
20140912_030000	14273	264	0	0,018	0,000
20140912_040000	25131	294	0	0,012	0,000
20140912_050000	41647	452	0	0,011	0,000
20140912_060000	42977	433	0	0,010	0,000
20140912_070000	48291	439	0	0,009	0,000
20140912_080000	52323	600	0	0,011	0,000
20140912_090000	52214	543	0	0,010	0,000
20140912_100000	51288	636	0	0,012	0,000
20140912_110000	53681	565	0	0,011	0,000
20140912_120000	49558	533	0	0,011	0,000
20140912_130000	51614	590	0	0,011	0,000
20140912_140000	51961	569	0	0,011	0,000
20140912_150000	51312	559	0	0,011	0,000
20140912_160000	51396	496	0	0,010	0,000
20140912_170000	47792	492	0	0,010	0,000
20140912_180000	44546	420	0	0,009	0,000
20140912_190000	40996	492	0	0,012	0,000
20140912_200000	34282	341	0	0,010	0,000
20140912_210000	22543	272	0	0,012	0,000
20140912_220000	13909	247	0	0,018	0,000
20140912_230000	8729	146	0	0,017	0,000

Table 2: *For the whole ECAC airspace, we show the number of tested segments (second column) the number of over-(under-) expressed trajectory segments (third and fourth column), the percentage of over-(under-) expressed trajectory segments (fifth and sixth column) for each of time-intervals shown in the first column. The univariate Bonferroni threshold used in this analysis is 0.01. These results are obtained implementing the FDR multiple comparison correction. In this analysis trajectory segments have been considered with their altitudes.*

The results obtained in this new investigation are qualitatively similar to those described in the previous section. Also in this case, as shown in Figure 7, we have over-expressed di-forks that are outside of the critical sectors as well as a few di-forks that occur close to the borders of the critical sectors, especially during day-time. As already mentioned, we take this as a suggestion of the fact that during night the over-expressed deviations detected by the di-forks probably are essentially directs given to facilitate the air traffic flow, while during day deviations might occur not only to facilitate the air traffic flow but also as a way to solve congestion issues.

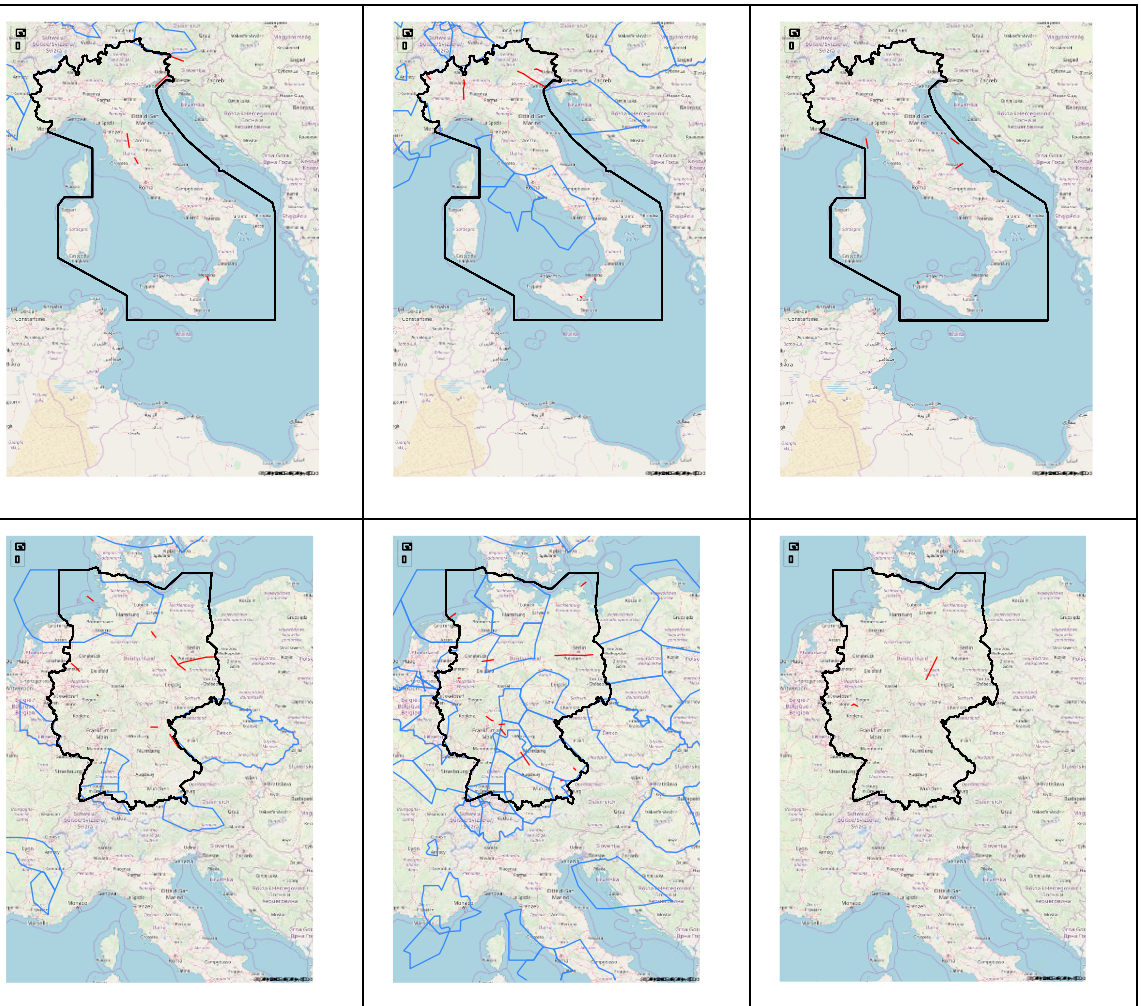


Figure 7: *Geographical localization of the critical sectors and of the over-expressed di-forks with altitude for the Italian (top panels) and German (bottom panels) airspaces in the time-intervals considered above. The blue lines mark the borders of the critical sectors. The results shown in these panels are the result of a country-by-country analysis.*

3.5 Conclusions

As already observed in Ref. [1] our analysis confirms that statistically validated trajectory deviations in percentage occur more frequently during night rather than during day, although in absolute terms the reverse is true. As already, pointed out in Ref. [1] this might be due to the fact that trajectory deviations during night are essentially directs that the ATC controllers issue in order to facilitate the traffic flow in presence of low traffic levels. On the contrary, during daytime, deviations might also be due to the necessity of solving congestions problems. This is somehow reinforced by the comparison between the geographical localization of di-forks and the critical sectors that were detected by the UNITS unit as a result of their strategic optimization. In fact, during day-time we detect di-forks that occur close to the borders of the critical sectors. We take this as a suggestion of the fact that during day deviations might occur not only to facilitate the air traffic flow but also as a way to solve congestion issues.

Next steps will involve repeating these analyses directly starting from the M1 file reconstructed from the output of the strategic optimization performed in WP3. Since the computation of the di-forks involve a comparison between planned and flown trajectories, we preliminary need to perform numerical simulations that would reproduce what would have been the flown s9olutions starting from such reconstructed M1 file. To do so, we will take advantage of the capabilities offered by the tactical layer of the ELSA agent based model. However, we are also planning to make a preliminary improvement of the model, by adding a few features that would make it more realistic and apt for the purposes of the present analysis.

When this will be done, we will generate surrogate M3 trajectories and we will then compute the di-forks directly starting from the M1 file reconstructed from the output of the strategic optimization performed in WP3. Questions we want to investigate are:

- Is the total number of observed di-forks smaller than those observed in the present investigation? Are the number of di-forks presumably due to congestion smaller than those observed in the present investigation?
- Is the geographical localization of di-forks different than in the present investigation, specifically with respect to the borders of the critical sectors?

These are a few issues that we will investigate in the analysis that will be presented in the next deliverable D5.2 (August 2018).

4 The complexity metrics

The main aim in this chapter is that of detecting congested areas in the European airspace by using well-known complexity metrics already existing in the literature [3]. The analysis will be done at the level of the collapsed sectors, and starting from the DDR2 M1 trajectories.

The considered metrics are essentially useful to measure the congestion of the airspace and its level of complexity as seen from the point of view of the controllers. Such metrics have already been studied in Ref. [4]. Here we enlarge the analysis to the whole ECAC airspace and we use a reduced time resolution to compute these metrics, down to the minute.

As a novelty, rather than focusing on the results of a single metric, we will consider the global behaviour of the 15 considered metrics at the level of each sector and compare the obtained results with the critical sectors that are detected in the optimization procedure that brings to the strategic solution proposed in WP3. The comparison shows that in many cases the critical sectors devised in the WP3 procedure are also highlighted when globally considering the complexity metrics.

Such findings confirm our view, according to which these metrics are a useful tool to measure occupancy and congestion in the airspace, when their global behaviour is considered.

4.1 Definition of the complexity metrics

We will use a set of 15 complexity metrics defined in the 2001 paper by Chatterji and Sridhar of Ref. [3]. The full definition of the metrics is given in that paper where it is also reported that "… The flow complexity metrics C1 to C16 are computed with every track update. At each instant of time, the complexity measures characterize the spatial distribution of traffic, Thus, their time history characterizes both the spatial and the temporal patterns of the traffic ….". According to this paragraph we decided to compute the metrics every minute. This would ensure that at the typical velocity of an aircraft the safety threshold of 10 NM is not traveled. The meaning of the 15 considered metrics is reported in Table 3.

Metric	Description
C1	Density
C2	Number of climbing aircraft
C3	Number of level flying aircraft
C4	Number of descending aircraft
C5	Inverse of the mean weighted horizontal distance between aircraft pairs
C6	Inverse of the mean weighted vertical distance between aircraft pairs
C7	Mean inverse horizontal distance of aircraft below vertical separation distance
C8	Mean inverse vertical distance of aircraft below horizontal separation distance
C9	Inverse minimal horizontal distance among pairs of aircraft
C10	Inverse minimal vertical distance among pairs of aircraft
C11	Number of aircraft pairs with positive time-to-conflict less than or equal to Δt
C12	Number of aircraft with at least one positive time-to-conflict divided by mean time-to-conflict

C13	Inverse of minimal time-to-conflict
C14	Variance in the groundspeed
C15	Contrast ratio (ratio of standard deviation to mean groudspeed)

Table 3: *Description of the complexity metrics.*

As mentioned above we will compute these metrics at each minute in day 12/09/2014 and for each collapsed sector in that day. However, rather than considering the value of the single metrics, we will consider their global behaviour following two complementary correlation–based approaches:

- On one side we will focus on each collapsed sector and starting from the empirical values of the 15 metrics we will compute the correlation matrix of the 15 metrics. This will be used to understand whether there are metrics with similar behaviour at the level of sectors.
- On the other side, for each metric, we will compute the correlation matrix of the sectors. This will be used to understand whether there are sectors that show similar behaviours with respect to a given metrics and therefore to highlight critical sectors.

The correlation-based approach allows us to empirically investigate similarities between metrics in the first case or amongst sectors in the second case. Such similarities must be understood not in terms of physical relationships but rather in terms of logical dependencies. For example, if we fix a metric and consider the correlation matrix computed over the different sectors, we can investigate the similarities between pairs of sectors with respect to that metric. High values of correlation between sector A and B might tell us that if we have an high value of the metric C_i in sector A it is likely that the same will occur in sector B. Therefore, if we have a critical situation with respect to a certain metric C_i in sector A, we expect that it is likely that a critical situation would also occur in sector B and, in general, in all other sectors that have high correlation with A. On the other side, low values of correlations between two sectors would indicate that a critical situation in sector A very likely will not propagate to sector B.

In the first approach, for each of the thousand sectors we will have a correlation matrix of the 15 metrics. In the second approach, for each metric we will have a correlation matrix of the thousand sectors active in the considered time-interval. In both cases we have a large amount of information from which we need to filter the relevant one, in order to have meaningful insights about the ATM system. This goal can be accomplished in a simple way by using the Single Linkage clustering algorithms that allows extracting from the correlation matrix a dendrogram, i.e. the Hierarchical Tree (HT), and the Minimum Spanning Tree (MST). The dendrogram reveals the hierarchical dependencies of its elements, while the Minimum Spanning Tree reveals their topological organization. These are classical, consolidated and nowadays standard tools for the analysis of large complex systems whose implementation is particularly simple and requires relatively low computational efforts. A full guide to these methodologies can be found in Ref. [7].

4.2 Results

We show here the results for day 12/09/2018 in the time-interval [13:00, 14:00]. The metrics are computed at each minute, so we have 60 snapshots of the system, as measured by the 15 considered metrics. We prefer to show the results relative to a specific time-interval in order to avoid issues related to the intraday behaviour of the metrics. However, the presented methodologies as well as the qualitative results thus obtained are easily transferable to other time intervals and scalable at the level of the whole day. In the considered time-interval we have 1178 active collapsed sectors.

In the first investigation we focus on each collapsed sector and starting from the empirical values of the 15 metrics we compute the correlation matrix of the 15 metrics, by using the Pearson correlation coefficient. As a result, for each of the 1178 sectors, we have a 15x15 correlation matrix. In Figure 8 we show the hierarchical trees of the 15 metrics for two different sectors: LIRROVD (left panel) in Italy and EDMMFRKU (right panel). LIRROVD is located in Central Italy, in the Rome ACC. EDMMFRKU is located in Southern Germany, in the Munich ACC. The two HTs show that the hierarchical organization of such two different sectors shows similarities. In fact, one can notice that metrics C1 and C11, metrics C7 and C9, metrics C14 and C15, metrics C9 and C10 appear in the same cluster in both cases. Indeed, these findings are also observed in other HTs relative to other sectors. In fact, metrics C1 and C11 appear in the same cluster in the 98% of cases, metrics C7 and C9 appear in the same cluster in the 73% of cases, metrics C14 and C15 appear in the same cluster in the 98% of cases, metrics C8 and C10 appear in the same cluster in the 83% of cases. Our approach is therefore able to detect regularities in the metrics that hold across sectors.

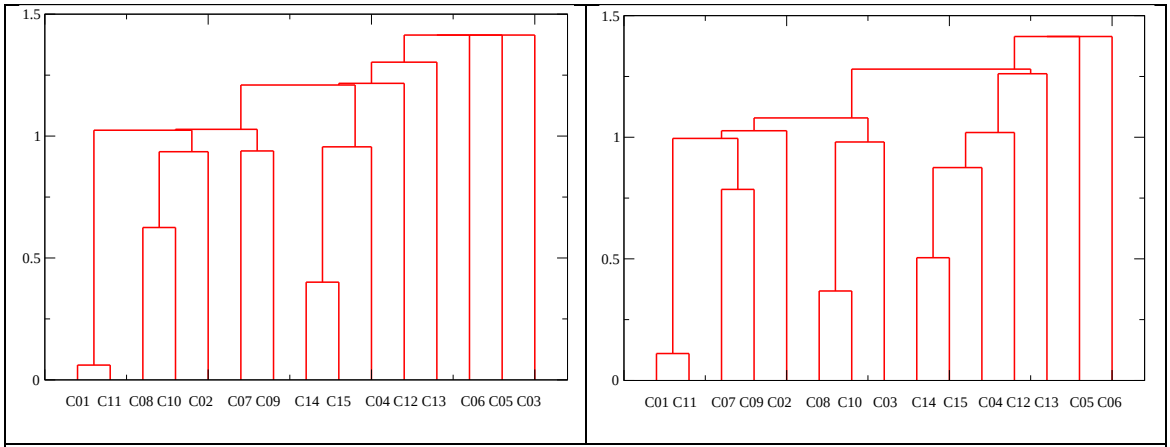


Figure 8: *Hierarchical Trees of the 15 metrics for two different sectors: LIRROVD (left panel) in Italy and EDMMFRKU (right panel).*

In the second investigation we focus on each metric and starting from the empirical values of the 15 metrics we compute the correlation matrix of the 1178 collapsed sectors, by using the Pearson correlation coefficient. In Figure 9, we show the Minimum Spanning Tree (MST) of the sectors starting from the correlation matrix that measures how similar are two sectors on the basis of the complexity metrics C9. The red sectors are the critical sectors that are detected in the optimization procedure that brings to the strategic solution proposed in WP3. One can see that there is one sector, i.e. the 730th sector LFFFTML, whose removal would cut the MST in two parts. Similar situations occur for other sectors in the same tree as well as in the other MSTs associated to the other metrics. Examples are: the and 493th sector LDZOULA for the metric C1, the 1123th sector

LYBAUWES for the metric C2, the 174th sector EDYYBOLN for the metric C3, the 987th sector LOVVEAL1 for the metric C4, etc. This shows that also in this case this filtering methodology highlights the presence of sectors that are crucial with respect to the complexity metrics.

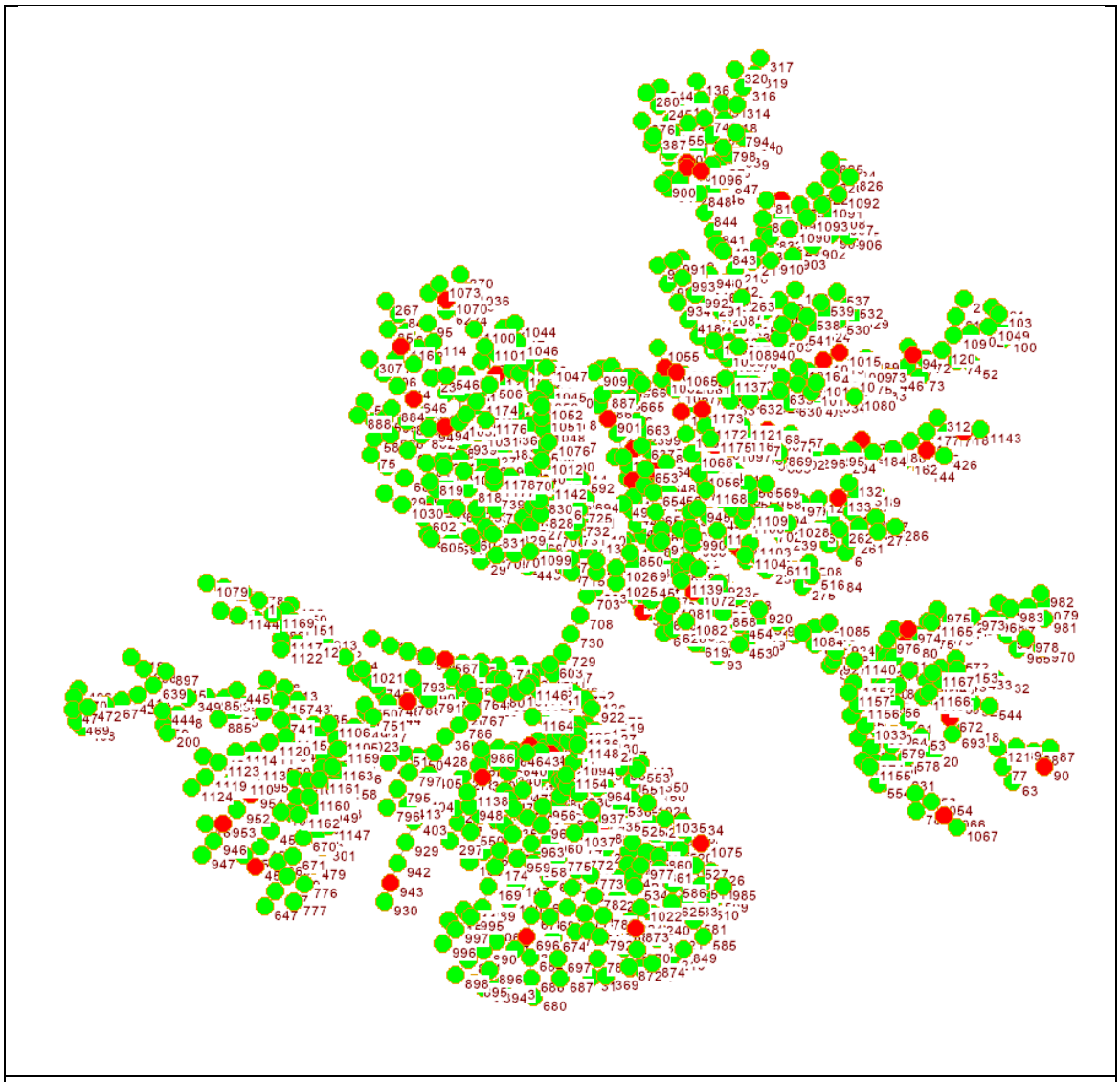


Figure 9: *Minimum Spanning Tree (MST) of the sectors starting from the correlation matrix that measures how similar are two sectors on the basis of the complexity metrics C9. The red sectors are the critical sectors that critical sectors that are detected in the optimization procedure that brings to the strategic solution proposed in WP3.*

In general, these links can be identified by looking at the nodes of the MST that have degree equal to 2. In the third column of Table 4 we give the number of nodes with degree 2 in the MSTs relative to the 15 complexity metrics. Among those, we then consider the ones that are also critical according to the optimization procedure developed in WP3. This is reported in the fifth column of

Table 4. One can see that a percentage of the MST nodes around 1% is critical both because they have degree 2 and because they were identified as critical in the optimization procedure developed in WP3. It is worth mentioning that the optimization procedure developed in WP3 selects 72 distinct critical sectors open at least from 13:00 to 14:00 on 12/09/2014. With our procedure we validate 43 of them, as shown in Figure 10, where we show the critical sectors detected according to the WP3 optimization procedure (blue borders). The red borders indicate the sectors that are also have degree 2 in the MSTs obtained starting from the 15 complexity metrics.

Metric	MST nodes (not isolated)	MST nodes with degree 2	MST nodes with degree 2 and not critical	MST nodes with degree 2 and critical
C1	916	249	238	11
C2	915	258	243	15
C3	922	264	251	13
C4	907	254	237	17
C5	919	279	264	15
C6	927	240	229	11
C7	910	245	235	10
C8	919	223	213	10
C9	910	222	213	9
C10	910	224	213	11
C11	924	263	252	11
C12	904	267	252	15
C14	912	272	259	13
C15	916	257	249	8

Table 4: *The second column reports the number of nodes in the MSTs. The third column reports the number of nodes in the MSTs with degree 2. The third column reports the difference between second and fourth column., The fourth column report the number of nodes in the MSTs with degree 2 and that are also critical according to the optimization procedure of WP3. Data refer to 12/09/2014 in the time interval between 13:00 and 14:00.*

As mentioned above, if we fix the complexity metric and study how the metric behaves in different sectors, we can get information about sectors that behave in a similar way with respect to the considered metric. This is the first step for a more accurate study of the logical dependencies amongst sectors, with respect to a specific complexity metric.

The MST is one simple instrument that allows studying such dependencies. Other instruments might be the hierarchical trees already considered in figure 8. Here we devoted our attention to MSTs because they allow us to study the topological dependencies between sectors.

In general the MST is a tool that allows us to filter the information present in the whole correlation matrix by selecting the most relevant links. We emphasize here that

1) the relevance of the links must always be understood with respect to the considered complexity metric and
2) the links we consider here are not physical links, given the fact that we start from a correlation-based network and not a physical one.

As such, the fact that the MST selects the most relevant links is relevant for us because it allows us to understand that if we have a critical situation in a sector, then we can follow the full path through which the criticality might propagate over the other sectors.

In this context we use the fact that certain MST nodes have degree 2 as a measure of the centrality of these nodes, i.e. as a measure of their relative importance in the MST. In fact, if we are interested in detecting the paths through which any criticality might propagate over the network, having a node with degree 2 is of the utmost importance because this means that the path must necessarily involve that node. With this proviso, the fact that 43 out of the 72 critical sectors of WP3 have a degree 2 in the MST confirms the relevance of these sectors in the topological organization of sectors, with respect to the specific complexity metrics considered.

4.3 Conclusions

The considered metrics have been investigated by following an approach whose peculiarity consists in the fact that one considers their global behaviour rather than looking at them separately.

The key point is to follow a correlation–based approach that involves the computation of a correlation matrix that can then be investigated by using filtering tools such as the Single Linkage algorithm. The previous results shows that by following this approach one is able to detect metrics whose behaviour is similar within a sector and sectors that are instrumental in ensuring the connectedness of the system.

The first aspect can be relevant from an operational point of view because it tells us that a certain situation of congestion may be seen from different aspects and/or might have different sources. For example, while metric C1 signals a potential risk due to high density of aircraft, metrics C11, which turns out to be highly related to C1 in 98% of cases, signals that such high density is also associated to the similar time-to-conflict, which is a complementary and extremely useful information. Similarly, the fact that C8 and C10 as well as C7 and C9 are usually in the same clusters of metrics is another indication of the fact that these two pairs of metrics put forward different aspects of the possible conflicts arising in the horizontal and vertical dimensions, respectively. On one side our findings provide a new way at looking at these complexity metrics. However, on another side, our findings can be seen as an internal check that confirms their goodness and the appropriateness in measuring congestion. One further aspect worth of further investigation is that of looking for metrics that somehow play the role of predictors for complexity. This can be easily investigated by looking at the lagged correlations between metrics and to the Granger causality tests. These are a few issues that we will investigate in the analysis that will be presented in the next deliverable D5.2 (August 2018).

The second aspect mentioned above is instead more related to the structure of the airspace. We are able to show that the critical sectors detected in the WP3 analysis in many cases are also crucial for having a fully connected MST, i.e. a fully connected topological structure of the system. Our procedure is in fact able to show that 43 out of the 72 critical sectors detected in the WP3 analysis have degree 2 in the MSTs of the 15 correlation-based matrices computed for each

complexity metrics and therefore removing such sector from the MST would cause a partition of the MST into two distinct sets of nodes. This issue will be again addressed in the next chapter.

Such findings confirm our view, according to which these metrics are a useful tool to measure occupancy and congestion in the airspace, when their global behaviour is considered.

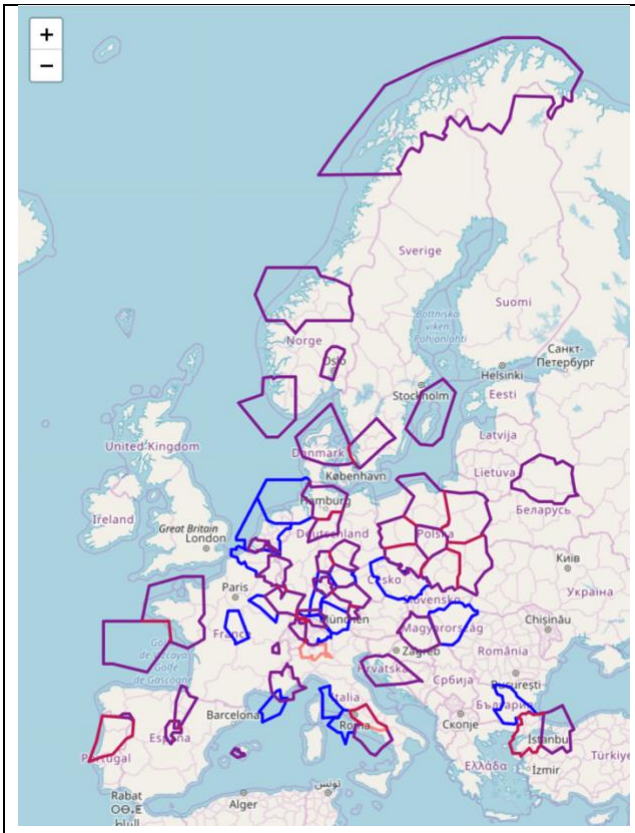


Figure 10: *Critical Sectors according to the WP3 optimization procedure (blue borders). The red borders indicate the sectors that are also have degree 2 in the MSTs obtained starting from the 15 complexity metrics. Data refer to the time-interval [13:00-14:00] of day 12/09/2014.*

5 The Percolation approach

The main aim in this chapter is that of detecting critical links in the ATM network by using a percolation approach. The analysis will be done at the level of the aggregated navigation points, i.e. the navigation balls, and starting from the navigation balls network created from DDR2 M1 and M3 files.

While the percolation approach has already been considered in other model complex systems, to our knowledge it is the first time that an empirical study using this methodology is performed at the level of the network that originates from the navigation points. In our investigation we will follow the approach of Ref. [5], where a urban road network is investigated.

We compare the geographical localization of the critical links detected with percolation with the critical sectors that are detected in the optimization procedure that brings to the strategic solution proposed in WP3. The comparison shows that there is not a total superposition between critical sectors and critical links. In fact, the percolation approach identifies critical links that are located outside the critical sectors, although in many cases around the airports. This is not necessarily surprising as the percolation approach provides a more “microscopic” description of the system and therefore captures aspects that a network-wide, macroscopic approach cannot detect.

Our findings confirm our view, according to which the percolation approach is the right way to microscopically detect the links that have the higher probability to cause disruptions in the network. In this context by disruption we mean the fact that several network connections are removed from the network so that it is not possible to reach any given node of the network starting from any other node. This is usually the case when the network gets fragmented in many small parts that are internally connected but not connected with the outside.

The analyses presented in this chapter are still at a preliminary level. In fact, this is a new type of metrics that we believe can complement the information provided by the other two types of metrics presented in the previous chapters.

Further efforts will have to be done in order to link the microscopic level addressed by the percolation approach to the macroscopic level at which, in particular, air traffic sectors and their capacity are the relevant objects.

5.1 Definition of the main concepts about percolation

In recent years there has been a growing awareness that the air transportation system cannot be planned, optimized, monitored and investigated by separately focusing on the different modules building up the system. In fact, although the majority of the interactions among air traffic actors

are interactions relatively localized in space and time, the complex interconnections present among all the actors of this socio-technical system make the air transportation system a "system-of-systems" structured as a layered collection of interacting networks [8]. In each of these network layers the interconnection of a plurality of heterogeneous socio-technical actors produces emergent phenomena impacting multiple spatial regions on multiple time scales.

A key concept in the description of emergent properties of complex systems is the concept of phase transition. Phase transitions occur in physical, biological and social systems both in the presence and in the absence of tuning parameters. A phase transition refers to many locally interacting elements causing a collective phase change (returning to the example of water, a physical analogue is the melting of ice, i.e. a transition from the solid to liquid phase). Typically, there exists a critical point that marks the passage from one phase to another [9].

The original concept of “phase transition” refers to the fact that a thermodynamic system can change its state of matter. In the complex systems context, “phase transition” refers to the fact that a complex systems made of many elementary elements locally interacting, under certain conditions may present a collective state (called phase of the system). The properties of this state are controlled by a variable named “order parameter”. It also said that the system shows **criticality**, meaning that there exists a “critical point” that marks the passage from a disordered state in which the local interactions are predominant to the collective state.

A paradigmatic example of phase transition is provided by the phenomenon of percolation that was originally introduced in simple geometrical lattices and therefore used in the modelling of many real systems. The original concept of “percolation” concerns the movement and filtering of fluids through porous materials. In the complex systems context, “percolation” refers to the fact that given two neighboring sites, a link between them is present with probability p and its absence is observed with probability (1-p). Percolation is relevant in the study of the spreading of information over a network.

In general, percolation is a random process exhibiting a phase transition. In the simplest setting percolation is investigated in simple geometrical systems such as regular lattices covering a 2D surface. Even in the simplest setting there are different variants of the percolation problem. Specifically, one speaks about *bond percolation* when a link between two neighbouring sites is present with probability *p* and its absence is observed with probability (*1-p*). In this variant all sites are present in the system and links between any pair of them may or may not be present. In the other variant of *site percolation* the links of the lattice are always present between two occupied sites, but each site is occupied with probability *p* and empty with probability (*1-p*).

In our context, percolation refers to probabilistic, network-wide emergent behaviour, between sites or sub-systems, across links in the network. In air transportation, there are several networks where percolation may happen.

Road traffic in urban areas, as a large-scale and complex dynamical system, has attracted much attention, especially on its dynamical transition between free flow and congestion. The dynamics of traffic have been studied using many types of models. However, there is still a gap between the microscopic behavior of individual vehicles and the emergence of macroscopic city traffic. Indeed, a fundamental question has rarely been addressed: how the local flows in roads interact and organize collectively into global flow throughout the city network. This knowledge is not only necessary to bridge the gap between local traffic and global traffic, but also essential for developing efficient traffic control strategies.

In this context the paper by Havlin [5] is of the utmost importance. As illustrated in Figure 11, reproduced from Ref. [5], they showed that by using a percolation approach it is possible to identify critical links that carry a high probability of causing disruptions in the network. It is our aim that of translating the investigation performed in that paper to the Air Traffic Management system.

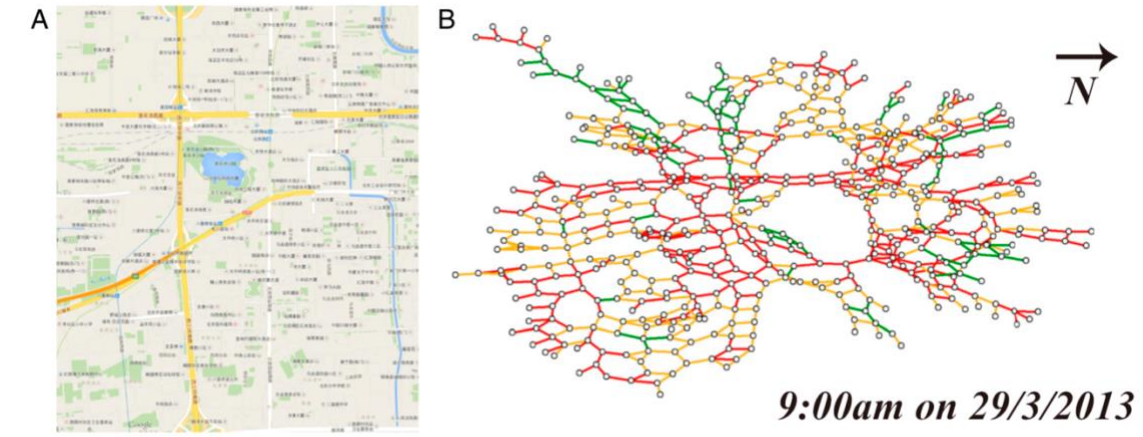


Fig. 1. Road network of the observed district. (*A*) Map of the investigated district. (*B*) Road network of the investigated district. Road network at 9:00 AM on March 29, 2013 is shown, where links are classified into three categories according to their velocity ratio r_{ij}: velocity ratio below 0.4 (red), between 0.4 and 0.7 (yellow), and above 0.7 (green). Note the clustering of each color.

Figure 11: *Illustrative example of the percolation approach used in Ref. [5] to study a road traffic network. Reproduced from Ref. [5].*

5.2 Percolation in the ATM networks

Percolation theory was successfully applied to model complex systems in statistical physics, economy and recently in social networks studies. We can compute the possible relevant critical thresholds and study the evolution of the system related to the variation of characteristic quantities such as the average size of the clusters and its distribution using the mathematical methods developed in Percolation Theory. To our knowledge, the use of this approach in the ATM context is very limited [10].

The basic and general setting we will consider is that of *bond percolation* and can be described as follows:

. 1) Suppose we have a graph G with a vertex set V and edge set E.

. 2) We make all the edges independently *"open"* with probability p_b and *"closed"* with probability $(1 - p_b)$.

We want to investigate whether there exists an abrupt change in the state of the network structure and whether this change is obtained around a critical value of a certain key parameter.

We will consider the navigation balls networks *G(d)* of each day *d* from 01/09/2014 to 14/09/2014 and for each link V_{ij} between ball *i* and ball *j* of the networks we will compute the number of aircraft $N_{ij}(d,t)$ flying over that segment in a hourly time interval *t*. For each segment we will compute the maximal number of aircraft $Q_{ij}(t)$ in a hourly time interval *t*. We will therefore associate to each link the ratio

$$r_{ij}(d,t)=N_{ij}(d,t)/Q_{ij}(t)$$

We will assume that this ratio play the role of the bond probability p_b mentioned above.

For the specific day d*=12/09/2014, and for each hourly time interval t, starting from the original network $G(d^*)$, we will compute the filtered networks $G_q(d^*,t)$ given by those links in $G(d^*,t)$ such that $r_{ij}(d^*,t)>q$, where q is a number in the interval [0,1]. For all these filtered network we then compute:

1. The number $L_q(d^*,t)$ of the links in the filtered networks $G_q(d^*,t)$.
2. The number $n_q(d^*,t)$ of connected components present in the network $G_q(d^*,t)$.
3. The size $LC_q(d^*,t)$ of the largest connected component in the network $G_q(d^*,t)$.
4. The size $SLC_q(d^*.t)$ of the second largest connected component in the network $G_q(d^*,t)$.

5.3 Results

5.3.1 A link percolation approach

In Table 5 we show the number $L_q(d^*,t)$ of the links in the filtered networks $G_q(d^*,t)$. As expected the number of links in day-time are larger than those during night-time and analogously the number of filtered links follows the same pattern. The reduction in the number of filtered links is severe, as it amount to 30-35% for each time interval.

day 12/09/2014	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5	0,55	0,6	0,65	0,7	0,75	0,8	0,85	0,9	0,95	
00:00-01:00	11920	11914	11889	11842	11740	11471	11016	10995	10836	10826	9101	9044	8950	8597	8551	8442	8347	8313	8292	829
01:00-02:00	11826	11819	11787	11748	11630	11407	10948	10930	10753	10747	9188	9132	9033	8666	8627	8498	8414	8375	8361	836
02:00-03:00	13641	13631	13562	13506	13355	13082	12490	12447	12271	12256	10411	10357	10200	9734	9683	9474	9362	9328	9316	931
03:00-04:00	18332	18317	18263	18190	18047	17632	16828	16805	16589	16573	14020	13950	13784	13180	13127	12874	12706	12650	12623	1262
04:00-05:00	29200	29182	29096	28994	28787	28330	26942	26878	26546	26529	22141	22013	21686	20605	20507	20135	19896	19795	19765	1976
05:00-06:00	44662	44635	44504	44368	44054	43193	41059	40968	40431	40399	32814	32578	31981	30129	29909	29123	28584	28353	28269	2826
06:00-07:00	46529	46503	46382	46243	45893	45000	42637	42531	41925	41886	34280	34029	33447	31561	31342	30577	30057	29822	29718	2971
07:00-08:00	51699	51677	51536	51367	50994	50026	47476	47368	46700	46658	38186	37914	37258	35407	35154	34383	33844	33625	33541	3352
09:00-10:00	54856	54830	54659	54431	54017	52886	49969	49834	49072	49001	40116	39855	39179	37126	36886	36088	35525	35321	35221	3521
10:00-11:00	54939	54926	54757	54580	54158	53040	50123	49967	49188	49113	40454	40149	39512	37498	37282	36484	35958	35751	35639	3563
11:00-12:00	53567	53536	53344	53165	52661	51462	48799	48658	47891	47807	39499	39188	38538	36727	36483	35771	35271	35036	34948	3494
12:00-13:00	55527	55499	55336	55136	54677	53575	50727	50588	49849	49760	41309	40966	40251	38246	37942	37109	36502	36288	36194	3619
01:00-02:11	52257	52235	52064	51893	51493	50401	47764	47623	46860	46786	38771	38488	37848	36014	35787	35053	34587	34403	34310	3430
13:00-14:00	54464	54431	54265	54070	53598	52488	49642	49506	48792	48728	40361	40076	39413	37528	37245	36460	35924	35714	35625	3561
14:00-15:00	55815	55782	55646	55459	55024	53894	51184	51066	50343	50262	41639	41337	40621	38714	38465	37687	37158	36959	36882	3687
15:00-16:00	55518	55478	55309	55135	54707	53680	50901	50782	50098	50053	41484	41186	40538	38546	38265	37476	36916	36703	36609	3660
16:00-17:00	55950	55915	55739	55564	55114	53999	51268	51131	50391	50332	41418	41157	40589	38660	38393	37561	37019	36837	36752	3674
17:00-18:00	52793	52750	52608	52432	52030	50929	48283	48161	47475	47413	39011	38751	38106	36269	36015	35191	34623	34395	34320	3431
18:00-19:00	49537	49512	49380	49190	48811	47781	45360	45220	44543	44488	36761	36494	35914	34119	33886	33182	32715	32531	32470	3246
19:00-20:00	45661	45635	45475	45305	44952	43957	41700	41558	40923	40868	33448	33214	32654	30900	30698	30015	29542	29357	29281	2927
20:00-21:00	38776	38751	38605	38445	38126	37287	35245	35145	34577	34538	28383	28187	27708	26324	26148	25597	25198	25037	24974	2497
21:00-22:00	27808	27776	27656	27510	27221	26541	25162	25066	24649	24587	20529	20393	20127	19216	19107	18769	18518	18417	18386	1838
22:00-23:00	20024	20019	19930	19835	19647	19151	18232	18163	17866	17821	15134	15016	14834	14206	14132	13879	13708	13624	13606	1360
23:00-24:00	14377	14370	14314	14253	14101	13717	13000	12952	12718	12702	10606	10550	10426	9990	9939	9755	9638	9591	9576	957

Table 5: *Number $L_q(d^*,t)$ of the links in the filtered networks $G_q(d^*,t)$.*

A similar pattern can be traced in the number $n_q(d^*,t)$ of connected components present in the network $G_q(d^*,t)$, as shown in Table 6.

12/09/14	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5	0,55	0,6	0,65	0,7	0,75	0,8	0,85	0,9	0,95
00:00-01:00	3764	3766	3766	3776	3786	3808	3837	3841	3861	3862	3730	3735	3749	3764	3767	3770	3783	3788	3792
01:00-02:00	3707	3709	3715	3715	3724	3739	3751	3758	3762	3764	3688	3709	3728	3727	3734	3735	3747	3753	3758
02:00-03:00	3924	3924	3938	3931	3964	3995	4077	4089	4111	4114	4089	4097	4121	4126	4130	4131	4122	4124	4127
03:00-04:00	5210	5213	5221	5245	5266	5334	5429	5441	5493	5501	5532	5539	5573	5578	5574	5602	5619	5631	5627
04:00-05:00	7001	7003	7029	7057	7112	7217	7443	7454	7523	7527	7800	7833	7890	8010	8040	8098	8125	8139	8144
05:00-06:00	7844	7854	7872	7909	7975	8190	8566	8594	8754	8769	9442	9504	9669	9935	10003	10139	10214	10259	10279
06:00-07:00	7909	7915	7939	7967	8035	8231	8677	8704	8865	8885	9711	9785	9962	10293	10363	10498	10636	10710	10757
07:00-08:00	8663	8669	8710	8753	8832	9039	9479	9517	9715	9736	10608	10707	10916	11295	11396	11577	11725	11770	11792
09:00-10:00	8498	8502	8534	8574	8653	8872	9425	9464	9693	9719	10782	10865	11072	11428	11520	11734	11854	11931	11966
10:00-11:00	8576	8577	8609	8636	8727	8956	9519	9574	9813	9847	10892	11007	11213	11606	11695	11891	12039	12090	12126
11:00-12:00	8476	8481	8518	8556	8669	8926	9440	9495	9724	9766	10727	10852	11062	11388	11491	11659	11792	11866	11888
12:00-13:00	8433	8440	8492	8537	8637	8869	9382	9420	9619	9665	10590	10721	10941	11337	11450	11677	11868	11926	11957
01:00-02:11	8144	8152	8194	8226	8321	8543	9043	9095	9331	9367	10425	10547	10761	11106	11194	11374	11528	11577	11607
13:00-14:00	8255	8261	8298	8355	8457	8702	9208	9257	9474	9492	10616	10734	10930	11289	11396	11606	11742	11781	11812
14:00-15:00	8907	8911	8939	8977	9080	9313	9798	9842	10050	10080	11095	11216	11438	11864	11957	12148	12301	12368	12395
15:00-16:00	8636	8642	8689	8741	8838	9083	9584	9629	9838	9863	10939	11044	11240	11651	11750	11940	12113	12170	12198
16:00-17:00	8866	8876	8929	8965	9050	9271	9783	9837	10052	10077	11152	11241	11416	11801	11917	12097	12231	12280	12309
17:00-18:00	8843	8853	8875	8912	8994	9238	9664	9703	9903	9922	10867	10975	11166	11498	11586	11789	11940	11990	12010
18:00-19:00	8774	8787	8822	8869	8963	9170	9623	9684	9904	9933	10728	10821	11015	11329	11402	11559	11632	11665	11690
19:00-20:00	8362	8365	8395	8438	8536	8753	9194	9248	9411	9435	10266	10352	10528	10763	10832	10932	11008	11056	11074
20:00-21:00	8067	8076	8116	8182	8268	8434	8754	8792	8930	8938	9427	9485	9597	9748	9802	9872	9943	9994	10004
21:00-22:00	7107	7110	7147	7189	7268	7410	7627	7655	7776	7796	7922	7953	7996	8062	8085	8123	8146	8159	8159
22:00-23:00	5541	5542	5559	5582	5620	5702	5842	5867	5920	5929	5986	6004	6019	6016	6024	5993	6002	6004	6006
23:00-24:00	4504	4507	4516	4519	4547	4577	4601	4609	4635	4638	4438	4433	4434	4413	4422	4419	4425	4432	4432

Table 6: *Number $n_q(d^*,t)$ of connected components present in the network $G_q(d^*,t)$.*

In Figure 12 we show the size of the largest and second largest connected component $LC_q(d^*,t)$ and $SLC_q(d^*,t)$ as a function of the thresholding parameter q for the 12 time-intervals from t_1=[07:00, 08:00] to t_2=[18:00, 19:00] of day d^*=12/09/2018. The considered time-intervals are all during day-time, i.e. when ATM operations are at their full level. One can see that as long as q increases, at a certain critical level q_c the difference between $LC_q(d^*,t)$ and $SLC_q(d^*,t)$ gets small. This is a clear indication of the fact that while originally all navigation balls are included in a large connected component, as soon as the thresholding process -- consisting in the removal of links with $r_{ij}(d^*,t)<q$ -- takes place, such large connected component becomes smaller and comparable with other connected components. As such while originally it was always possible to reach any point in the network starting from any other point, as soon as q increase the network gets fragmented into smaller pieces, which prevents a potential user to reach any point in the network starting from any other point.

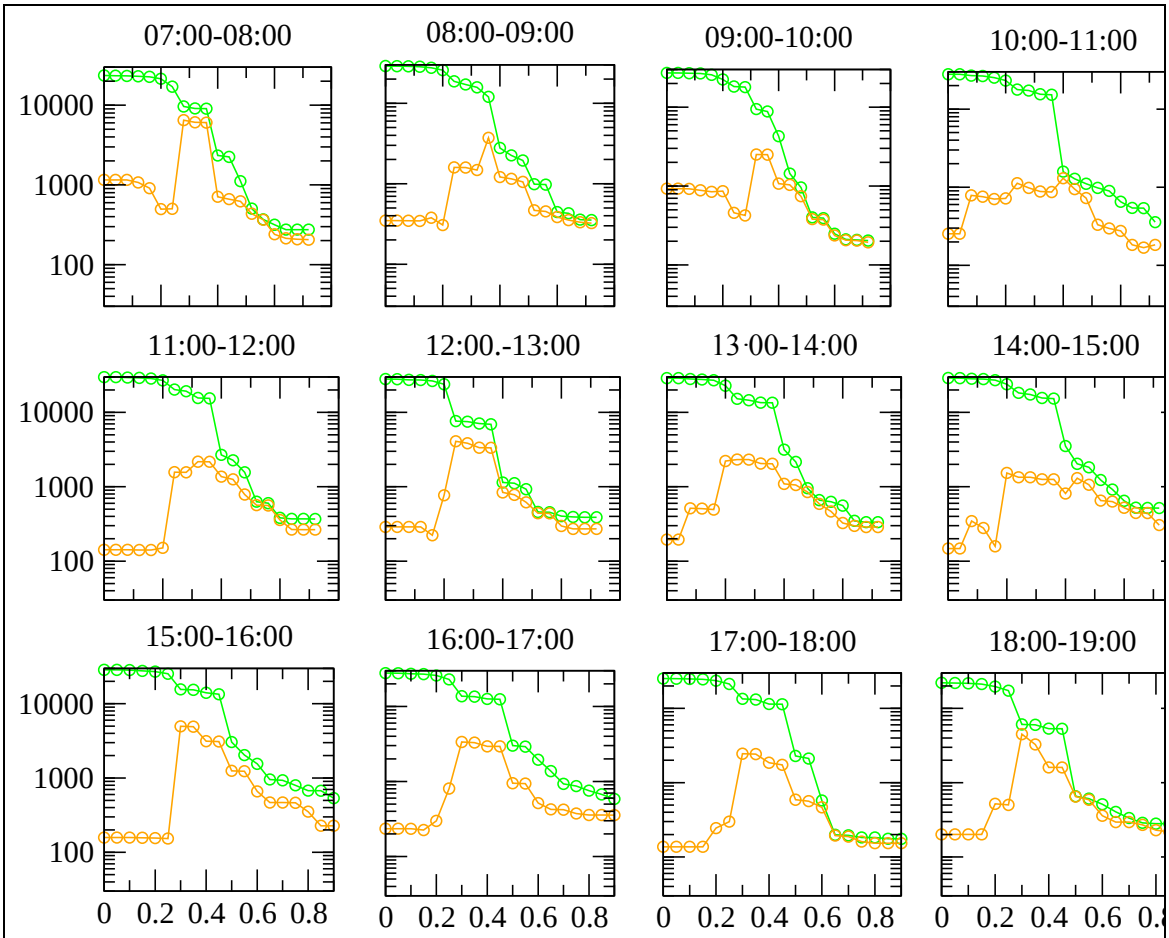


Figure 12: *Size of the largest and second largest connected component $LC_q(d^*,t)$ and $SLC_q(d^*,t)$ as a function of the thresholding parameter q for the 12 time-intervals from t_1=[07:00, 08:00] to t_2=[18:00, 19:00] of day d^*=12/09/2018. The vertical axes are logarithmic for the sake of visibility.*

In Figure 13 we show the size of the largest and second largest connected component $LC_q(d^*,t)$ and $SLC_q(d^*,t)$ as a function of the thresholding parameter q for the 12 time-intervals from t_1=[00:00, 01:00] to t_2=[06:00, 07:00] and from t_1=[19:00, 20:00] to t_2=[23:00, 24:00] of day d^*=12/09/2018. The considered time-intervals are all during night-time, i.e. when ATM operations are at a reduced level. One can see that in such set of time-intervals he previous effect is not present, apart from two time-intervals at [04:00, 05:00] and [05:00, 06:00].

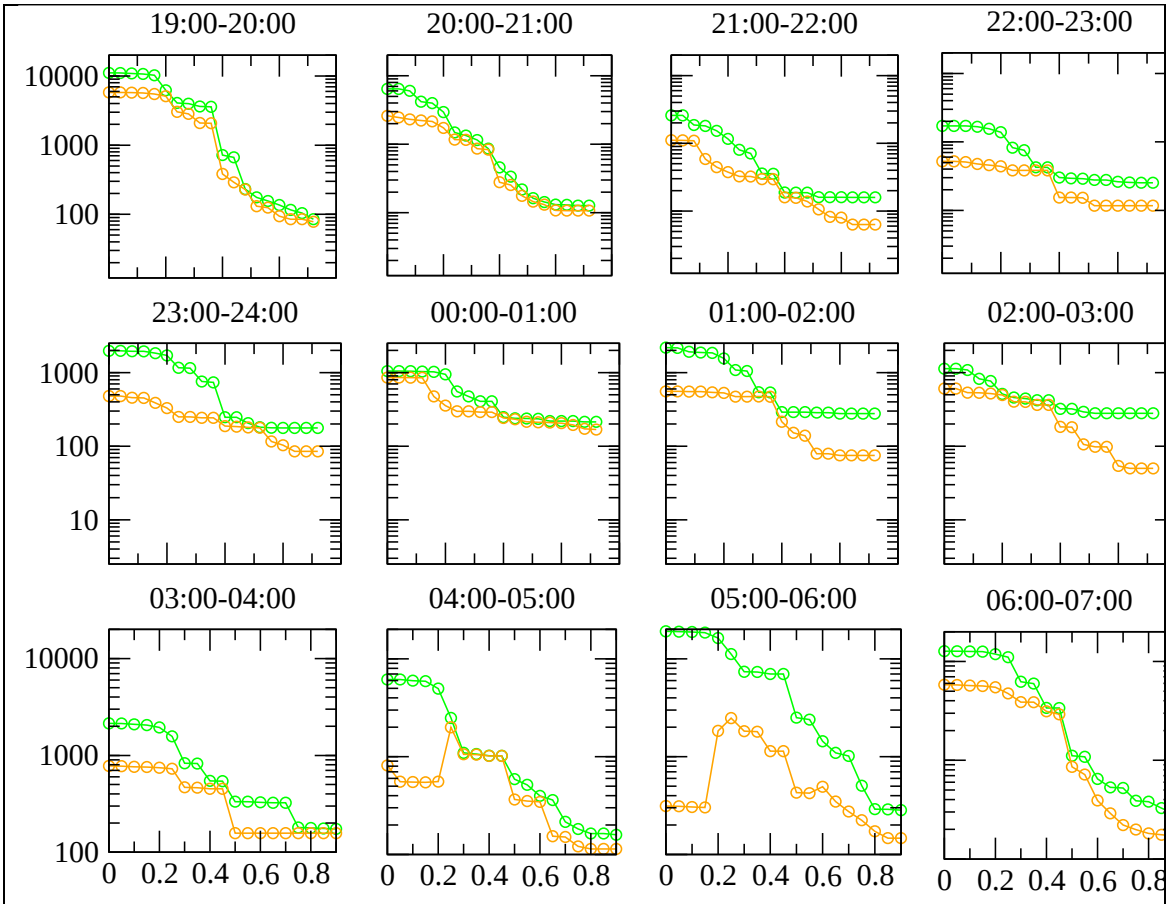


Figure 13: *Size of the largest and second largest connected component $LC_q(d^*,t)$ and $SLC_q(d^*,t)$ as a function of the thresholding parameter q for the 12 time-intervals from t_1=[00:00, 01:00] to t_2=[06:00, 07:00] and from t_1=[19:00, 20:00] to t_2=[23:00, 24:00] of day d^*=12/09/2018. The vertical axes are logarithmic for the sake of visibility.*

The investigation so far performed, although preliminary, yet clearly shows that the percolation approach is able to detect links, characterized by certain q values, which carry a potential risk of disruption in the network. The idea, here, is that the q parameter can be considered as an order parameter that would allow to track the system changing from the case when a large connected component is present to the case when the network gets fragmented. A precise procedure to identify such critical threshold is not yet available. However, the above results show that it is already possible to identify a range of q values, which q_c presumably belongs to.

In Figure 14 we show the segments V_{ij} between ball *i* and ball *j* whose weight $r_{ij}(d^*,t)$ is in the range $r_{ij}(d^*,t) \in [0.35,0.55]$ in the time interval t=[12:00, 13:00] of day d^*=12/09/2018. The considered weight values are those around the expected critical threshold. The blue lines represent the 2-D borders of the critical sectors that are detected by the UNITS unit as part of their strategic optimization. One can see that many segments are inside the critical sectors. However, a

consistent fraction of them is located outside the borders and a closer look reveals that in many cases these segments are located around the main European airports.

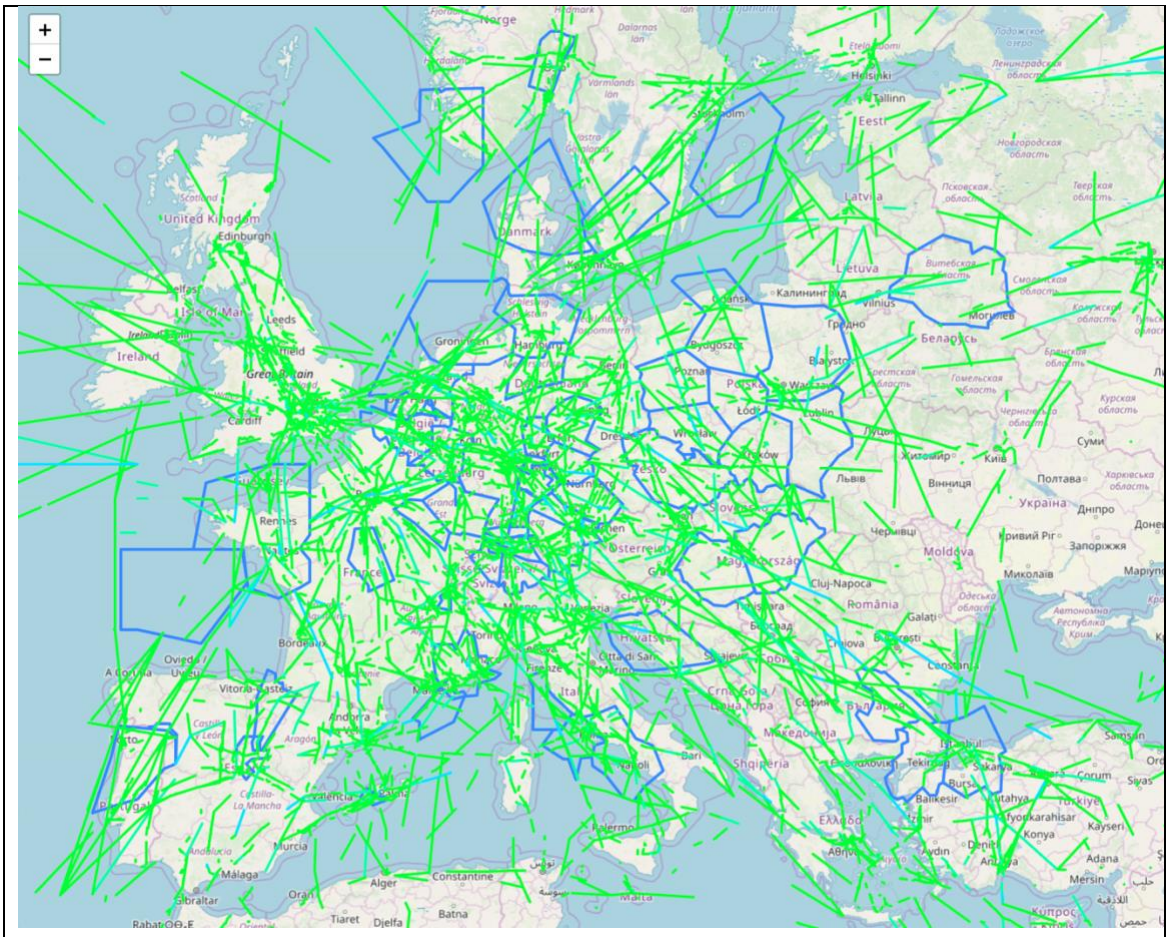


Figure 14: Segments V_{ij} between ball *i* and ball *j* whose weight $r_{ij}(d^*,t)$ is in the range $r_{ij}(d^*,t) \in [0.35,0.55]$ in the time interval t=[12:00, 13:00] of day d^*=12/09/2018. The blue lines represent the 2-D borders of the critical sectors that are detected by the UNITS unit as part of their strategic optimization.

To better understand this fact in the three panels of Figure 15 we show the segments V_{ij} between ball *i* and ball *j* whose weight $r_{ij}(d^*,t)$ are $r_{ij}(d^*,t) \geq 0.2$ (left panel), $r_{ij}(d^*,t) \geq 0.5$ (central panel) and $r_{ij}(d^*,t) \geq 0.9$ (right panel) in the time interval t=[12:00, 13:00] of day d^*=12/09/2018. One can notice that by progressively removing links when increasing q, the surviving links, i.e. those with higher weight, are mainly located around the airports. In other words, those around the airports are the links where the traffic load, measured by $\boldsymbol{N_{ij}(d,t)}$ is close to its maximum value, measured by $\boldsymbol{Q(t)}$.

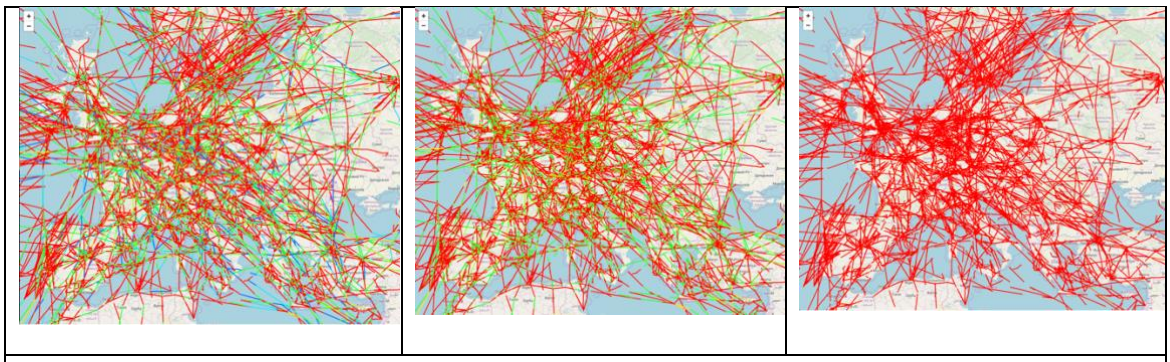


Figure 15: Segments V_{ij} between ball *i* and ball *j* whose weight $r_{ij}(d^*,t)$ are $r_{ij}(d^*,t)\geq 0.2$ (left panel), $r_{ij}(d^*,t)\geq 0.5$ (central panel) and $r_{ij}(d^*,t)\geq 0.9$ (right panel) in the time interval t=[12:00, 13:00] of day d^*=12/09/2018.

The fact that there is not a total superposition between critical sectors and links identified by the percolation approach is not necessarily surprising. In fact, the percolation approach provides a "microscopic" description of the system and therefore captures aspects that a network-wide, macroscopic approach cannot detect.

The percolation approach used in chapter 5 is based on two aspects that cooperate together:

1. on one side we have a thresholding process that allows us to "select" those links ij where the observed number of aircraft present in the considered trajectory segment $N_{ij}(d^*,t)$ is equal or larger than a certain fraction q of the maximal number of aircraft observed over several days $Q_{ij}(t)$ (14 in our case):

$$N_{ij}(d^*,t) > q\, Q_{ij}(t)$$

2. on the other side, the segments thus selected are globally considered, and in particular we detect to which connected component of the network they belong to.
3.

We observe that, by increasing q, the number of connected component in the network starts increasing and such components become smaller in size. After a certain critical value of q, we move from a regime where there is mainly a giant connected component to a regime where there are very many small connected components. We call q_c such critical value.

As such, the existence of q_c allows us to detect those links where having values of $N_{ij}(d^*,t)$ close to $Q_{ij}(t)$ is potentially risky. The risk here consists in the fact that we may end up in having many small connected components in the network. This means that while it is possible to travel without problems within the connected component, traveling outside the connected component is not. In this sense we talk about disruption.

Parameter q can be considered as an order parameter that allows checking the homogeneity of the system with respect to a certain predetermined level of segment capacity. In this respect the critical threshold q_c might assume the role of the critical percentage of the segment capacity that might trigger the fragmentation of the network.

5.3.2 A node percolation approach

We are fully aware of the fact that in ATM the segment capacity (link capacity) is not the real quantity to deal with. Rather, ATM is concerned with a capacity of many segments together, i.e. with sector capacity.

A node percolation approach gives the way by which we can reconcile the microscopic aspect of the network links and the macroscopic aspect where air traffic sectors and their capacity are the relevant objects. The idea is that of constructing a network of air traffic sectors and applying the thresholding procedure used above to the node of this new network, i.e. to the sectors. The relevant order parameter would be related to the sector occupancy or even sector capacity, rather than the link weights considered in the previous section.

Evidently, the sector occupancy is built up starting from the number of aircraft that travel the trajectory segments present in the sector at a certain time interval, and this is exactly how we reconcile the microscopic and macroscopic aspects.

Having proved for the first time that percolation can be effective in ATM when we simply try to export the approach of Ref. [5] from road traffic to the air traffic system, gives us hopes that a percolation approach fully designed for the ATM, as the node percolation approach mentioned above, will be even more able to give us insights about the areas that have the higher probability to cause disruptions. From a link-percolation approach on the navigation balls network we will have to move to a node-percolation approach on the air traffic sectors network.

From a theoretical pint of view, it will also be interesting to investigate how the microscopic order parameter discussed in the previous section relates to the new macroscopic one.

Furthermore, we want to emphasize that the MSTs presented in chapter 4, see Figure 9, are a first example of sector network generated starting from a correlation-based approach.

5.4 Conclusions

The investigation so far performed, although preliminary, yet clearly shows that the percolation approach is able to detect links, characterized by certain q values, which carry a potential risk of disruption in the network. The idea, here, is that the q parameter can be considered as an order parameter that would allow to track the system changing from the case when a large connected component is present to the case when the network gets fragmented. In this context, it is therefore crucial to devise a simple procedure to measure the critical value q_c of the order parameter that marks the phase transition, i.e. the passage from a phase where the network is connected and therefore show a large connected component, to a phase where the network is fragmented. From an operational point of view q_c would indicate the level of occupancy of a segment beyond which disruption becomes an issue, i.e. there is a risk that the network gets fragmented.

Another issue worth of further investigation is the relation between the critical links detected by the percolation approach and the statistically validated di-forks discussed in chapter 3. We have preliminary indications that at least 10% of the statistically validated di-forks are also critical links according to the percolation approach. However, a more systematic study needs to be performed.

Finally, we strongly believe that a full percolation approach to the study of the ATM system need to reconcile the microscopic level -- addressed in this deliverable in terms of link percolation -- with

the macroscopic level given by air traffic sectors. We think that using a node percolation approach on the network of sectors can do this. This is a challenging task that will also put forward interesting theoretical problems.

These are a few issues that we will investigate in the analysis that will be presented in the next deliverable D5.2 (August 2018).

The same analysis as above has also been performed on the navigation balls network obtained starting from the M1 files reconstructed starting from the trajectories that result from the optimization procedure performed in WP3. The qualitative results are similar to those presented here, although the order parameter threshold seem to be lower.

6 Next steps and look ahead

In this deliverable we have presented three approaches for the detection of hotspots in the ATM system:

- Hotspots as places where controllers' actions are taken.
- Hotspots as places where congestion is present.
- Hotspots as places where disruptions are present.

The message we want to convey with this document is that: when considering controllers' actions the di-fork metric discussed in chapter 3 is the right metric to consider; when considering congestion, the complexity metrics discussed in chapter 4 are the right metrics to consider; when considering disruption the right way to address this issue is the percolation approach discussed in chapter 5.

The approaches discussed in chapter 4 and 5 are less mature than the di-fork approach discussed in chapter 3. Next steps will include further refinements of such metrics that will be presented in deliverable 5.2 (August 2018).

When considering the di-fork metric, next steps will involve repeating these analyses directly starting from the M1 file reconstructed from the output of the strategic optimization performed in WP3. Since the computation of the di-forks involve a comparison between planned and flown trajectories, we preliminary need to perform numerical simulations that would reproduce what would have been the flown s9olutions starting from such reconstructed M1 file. To do so, we will take advantage of the capabilities offered by the tactical layer of the ELSA agent based model. However, we are also planning to make a preliminary improvement of the model, by adding a few features that would make it more realistic and apt for the purposes of the present analysis. When this will be done, we will generate surrogate M3 trajectories and we will then compute the di-forks directly starting from the M1 file reconstructed from the output of the strategic optimization performed in WP3. Questions we want to investigate are:

1. Is the total number of observed di-forks smaller than those observed in the present investigation? Are the number of di-forks presumably due to congestion smaller than those observed in the present investigation?

2. Is the geographical localization of di-forks different than in the present investigation, specifically with respect to the borders of the critical sectors?

When considering the complexity metrics, next steps will involve that of looking for metrics that somehow play the role of predictors for complexity. This can be easily investigated by looking at the lagged correlations between metrics and to the Granger causality tests. These are a few issues that we will investigate in the analysis that will be presented in the next deliverable D5.2 (August 2018).

When considering the percolation approach, we want to emphasize that, to our knowledge, in thios document we provided the first empirical evidence that the methodology can be fruitfully applied to the ATM complex system. Next steps will involve:

1. to devise a simple procedure to measure the critical value q_c of the order parameter that marks the phase transition.

2. To investigate the relation between the critical links detected by the percolation approach and the statistically validated di-forks discussed in chapter 3.

3. To reconcile the microscopic level of trajectory segments -- addressed in this deliverable in terms of link percolation -- with the macroscopic level air traffic sectors – to be addressed in terms of node percolation. This would also make possible a direct comparison with the critical sectors detected in WP3 with critical sectors identified in chapter 4.

7 References

[1] C. Bongiorno, G. Gurtner, F. Lillo, R. N. Mantegna, S. Micciché, ***Statistical characterization of deviations from planned flight trajectories in air traffic management***. JATM, **58**, 152-163, (2017).

[2] C. Bongiorno, S. Micciché, R.N. Mantegna, G. Gurtner, F. Lillo, S. Pozzi, ***Adaptative air traffic network: statistical regularities in air traffic management***, presented at the 11th USA/Europe ATM R&D Seminar, 23-26 June 2015, Lisbon, Portugal. http://www.atmseminar.org/seminarContent/seminar11/papers/440 Gurtner 0126150101-Final-Paper-4-30-15.pdf

[3] G. B. Chatterji, B., Sridhar, **Measures for air traffic controller workload prediction**. In: 1st AIAA aircraft, technology, integration, and operations forum. 2001. http://dx.doi.org/10.2514/6.2001-5242.

[4] G. Gurtner, C. Bongiorno, M. Ducci, S. Micciché, ***An Empirically grounded Agent Based simulator for the Air Traffic Management in the SESAR scenario***, JATM, **59**, 26-43, (2017).

[5] D. Li, B. Fu, Y. Wang, G. Lu, Y. Berezin, H. E. Stanley, S. Havlin, ***Percolation transition in dynamical traffic network with evolving critical bottlenecks***, PNAS, 112, 669–672, (2015).

[6] T. Bolic, L. Castelli, L. Corolli, D. Rigonat, ***Reducing ATFM delays through strategic flight planning***, Transportation Research Part E, **98**, 42–59, (2017).

[7] M. Tumminello, C. Coronnello, F. Lillo, S. Micciché, R. N. Mantegna, ***Spanning Trees and bootstrap reliability estimation in correlation based networks***. IJBC, **17**, 2319-2329, (2007).

[8] D. DeLaurentis E.-P. Han E.-P., ***System-of-systems simulation for analyzing the evolution of air transportation***, In: International Council for Aeronautics and Space, Twenty-Fifth Congress. 2006. http://icas.org/ICAS_ARCHIVE/ICAS2006/PAPERS/205.PDF

[9] D. Helbing, ***Traffic and related self-driven many-particle systems***, Rev. Mod. Phys. **73**, 1067, (2001).

[10] S. Ben Amour, T.D. Huy, M. Buynat, ***A percolation based model for atc simulation***, In: INO Workshop 2005. DOI: 10.1109/RIVF.2006.1696412

[11] ADAPT, 2018. *D2.1 Support to modelling activities,* s.l.: s.n.

[12] ADAPT, 2019. D3.1 Flight flexibility and hotspots in the ADAPT solution

-END OF DOCUMENT-