Assessing the spatial representativeness of air quality sampling points

Application of siting criteria and sampling point classification – Task 3 report

Service Request 5 under Framework Contract **ENV.C.3/FRA/2017/0012**

Specific Contract: **07.0203/2018/793545/SFRA/ENV.C.3**

Final report for European Commission -DG Environment

Ares (2018) 4920320

Customer:

Final report for European Commission -DG Environment

Customer reference:
Ares (2018) 4920320

Confidentiality, copyright & reproduction:
This document was prepared for DG Environment. The information herein is confidential and shall not be divulged to a third party without the prior permission of Ricardo Nederland B.V.

Ricardo plc, its affiliates and subsidiaries and their respective officers, employees or agents are, individually and collectively, referred to in this clause as the 'Ricardo Group'. The Ricardo Group assumes no responsibility and shall not be liable to any person for any loss, damage or expense caused by reliance on the information or advice in this document or howsoever provided, unless that person has signed a contract with the relevant Ricardo Group entity for the provision of this information or advice and in that case any responsibility or liability is exclusively on the terms and conditions set out in that contract.

Ricardo Nederland B.V. and Ricardo-AEA are trading names of the Ricardo Group of entities.

Services are provided by members of the Ricardo Group.

© 2017 Ricardo Nederland B.V. All rights reserved.
No parts of this publication may be reproduced distributed, modified and / or made public in any form whatsoever, including printed photostatic and microfilm, stored in a retrieval system, without prior permission in writing from the publisher.

Ricardo reference:
Ref: ED 11492 – Final Task 3

Contact:

Joanne Green, Ricardo.

t: +44 (0)1235 75 3450
e: Jo.Green@ricardo.com

Authors:
Leonor Tarrasón, Claudia Hak, Joana Soares, Håvard Vika Røen, Rune Ødegård, Jo Green and Leif Marsteen

Approved By:
Beth Conlan

Date:
21 December 2020

Table of contents

1 Introduction

This report is part of the project "Assessing the spatial representativeness of air quality sampling points" carried out for the European Commission, DG Environment (specific contract number 070203/2018/793545/SFRA/ENV.C.3, under Framework Contract number ENV.C.3/FRA/2017/0012).

The main purpose of this project is to provide recommendations for assessing and improving the spatial representativeness of the European air quality sampling points. The project is organized into three tasks. The objective of Task 1 is to draft guidance on the assessment of spatial representativeness in the context of air quality, following a tiered approach to account for the different capabilities and resources at EU Member States. The objective of Task 2 is to provide recommendations on how to develop the existing composite mapping platform under FAIRMODE to combine and make available relevant information on air quality concentrations, dispersion and emissions to facilitate the assessment of sampling point representativeness by Member States. Task 3 aims to establish how the criteria for selecting and classifying sampling points of the Ambient Air Quality Directive (2008/50/EC, as amended by Commission Directive (EU) 2015/1480) (AAQD) is applied in Member States of the European Union. The results from Task 3 are to inform the recommendations of Task 1 and 2 and provide a pathway towards harmonized guidance on spatial representativeness that can benefit air quality network design throughout the European Union.

The objectives of Task 3 are as follows:

1. Evaluate the quality of the existing sampling point classification of 'traffic-oriented sites' and 'industrial sites' in a comprehensive overview.
2. Investigate the achievement of the macroscale and microscale siting criteria in Annex III of the AAQD based on existing documentation of methodologies used for air quality sampling point classification and assessing where and why the application in Member States of the criteria for selecting 'traffic-oriented sites' and 'industrial sites' differs.
3. Provide recommendations to facilitate a harmonized application of sampling point selection criteria throughout the European Union.

This report summarizes the key activities and findings carried out in Task 3. The key objective is to assess how the criteria for classifying air quality sampling points is applied in Europe and to provide an overview of how the reported sampling point location relates to the siting criteria established by the AAQD.

It is important to note that this work is not intended to be a compliance checking exercise, but to serve as a means to better understand the different methods employed among countries for siting and classifying stations under the AAQD both for Member States and non-Member States. This understanding of current practices is important in the context of providing recommendations for assessing and improving the spatial representativeness of the European air quality sampling points.

The selection of air quality sampling points in countries follows the AAQD that set limit and target values for the concentration of air pollutants and specify the monitoring and reporting requirements. The AAQD determine the design of the European monitoring network for air quality (AQ). The AQ network of sampling points serves as basis to a) **determine compliance** with AQ limit values and trace progress towards environmental targets, b) carry out **exposure calculations** to estimate the surface area and the length of road where the levels are above the environmental objective, and the total resident population in the exceedance area, and c) carry out **model calibration and validation**. The monitoring network is thus essential for air quality management in Europe.

The specifications for the air quality monitoring network are provided in the different Annexes of the AAQD. Annexes V and IX of the AAQD specify minimum numbers of sampling points per zone, corresponding to population number and pollution level. Annexes III and VIII specify criteria for the location of sampling points. Annex VI specifies reference methods for measurement and criteria for equivalent methods, while the data quality objectives and requirements for quality assurance including requirements for measurement uncertainty are given in Annex I. In addition to these, the Implementing Decision (Decision 2011/850/EU) lays down rules for the reciprocal exchange of information and reporting on ambient air quality (known as e-reporting) and specifies the information on sampling points that is to be reported following the current Guidance on Implementing Provisions for Reporting (IPR Guidance)[1] and the IPR Guidance XML user guide[2].

This report focuses on the **siting** of the sampling points and is therefore mostly concerned with Annexes III and VIII that specify the macroscale and microscale criteria for the location of sampling points. It also focusses on the sampling point **classification,** as established in current guidance on Implementing Provisions for Reporting (IPR Guidance). The sampling point classification system originates from the Exchange of Information (EoI) Decision (EC, 1997) and follows to a great extent the EoI guidance definitions by Garber et al (2002).There are three main types of station according to this classification[3]: a) traffic-oriented, b) industrial and c) background. Traffic stations are located in close proximity to a single major road, while industrial stations are located in close proximity to a single industrial source or industrial area. Background stations are in any location with is neither to be classified as "traffic" or "industrial". In the same site, or monitoring station, different air pollution components can be measured. There is a sampling point per air pollutant measured so that a single monitoring site can include different sampling points. The number and location of the sampling points vary across Europe depending on the air pollutant.

This study focuses on two types of sampling points, namely those classified as 'traffic-oriented' and 'industrial', and on two components: particulate matter less than 10 microns in diameter (PM_{10}) and nitrogen dioxide (NO_2). It also focuses only on the year 2017, as the number of sampling points varies also in time and the year 2017 is also the one used for the Fitness Check of the AQQD[4]. It considers sampling points across the European monitoring network, as reported to the European Environment Agency (EEA) and available at their Central data Repository (CDR), as of 16 January 2020. This study involves a total of 2,007 traffic-oriented sampling points, 1,111 sampling points for NO_2 and 896 for PM_{10}, as well as 905 industrial sampling points, 452 sampling points for NO_2 and 453 for PM_{10}. Table 1.1 shows the total number of sampling points included in our analysis, for different country groups in Europe. The definition of the country groups is provided in Appendix 2. Figure 1.1 shows the location of 'traffic-oriented' and 'industrial' sampling points for nitrogen dioxide in 2017 to illustrate the number and distribution of sites across the Europe.

1 https://ec.europa.eu/environment/air/quality/legislation/pdf/IPR_guidance1.pdf, last viewed 16 January 2020
2 https://www.eionet.europa.eu/aqportal/doc/UserGuide2_AQD_XML_v3.3.0.pdf, last viewed 16 January 2020
3 http://dd.eionet.europa.eu/vocabularyconcept/aq/stationclassification/, last viewed 16 January 2020
4 https://ec.europa.eu/environment/air/pdf/SWD_2019_427_F1_AAQ%20Fitness%20Check.pdf, last viewed 10 December 2020

Table 4.1. Number of sampling points for NO_2 and PM_{10} in 2017 evaluated in this report, per classification type, and for different country groups as defined in Appendix 2.

With metadata reported.	EU27	EU28	EEA33*	EEA39*	All
NO2 - Traffic	971	1036	1097	1109	1111
NO2 - Industrial	417	428	446	452	452
NO2 - Background	1748	1825	1949	1966	1968
Total No2	3136	3289	3492	3527	3531
PM10 - Traffic	787	823	886	895	896
PM10 - Industrial	418	426	446	453	453
PM10 - Background	1852	1883	2069	2090	2092
Total PM10	3057	3132	3401	3438	3441

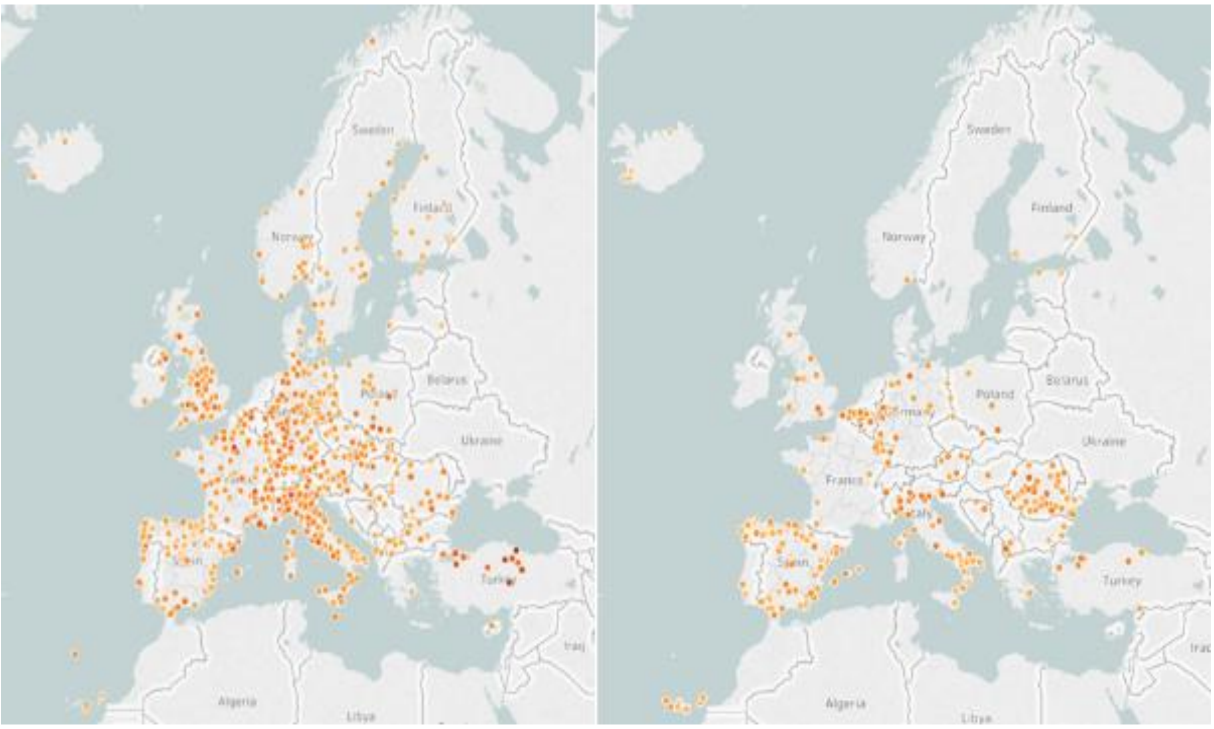


Figure 1.1: Monitoring stations for nitrogen dioxide (NO_2) in 2017. The left panel shows “traffic-oriented sites” while the right panel shows “industrial sites”. Source: EEA.

This report reviews the current location and sampling point classification at all ‘traffic-oriented’ and ‘industrial’ sites across Europe following four different but complementary approaches to understand. The different evaluation approaches are presented in the successive chapters and include:

1) An investigation of the location of all currently reported traffic and industrial sampling points in relation to a specific set of the macroscale and microscale siting criteria in Annex III of the AAQD. The analysis is based on an independent comparison based on satellite images and available Geographic Information System (GIS) mapping information compiled in an interactive web mapping tool that serves as a Monitoring Siting Criteria Viewer (Chapter 2).
2) A review of the metadata reported by countries as part of the IPR reporting obligations in dataflow D with respect to siting criteria in Annex III and sampling point classifications (Chapter 3).

3) A comparison of the reported station classifications with those determined by carrying out an independent cluster analysis on the reported measured air quality concentrations (Chapter 4).
4) A series of consultations with national experts through a web-questionnaire and successive workshop presentations to identify the current documented methodologies used for siting and classifying sampling points.(Chapter 5).

On the basis of the above analysis, conclusions are provided in Chapter 6 to facilitate a harmonized application of sampling point selection and classification methodology and proposals are made to inform further work in Task 1 and 2 to provide a pathway towards harmonized guidance on spatial representativeness that can benefit air quality network design throughout Europe.

2 Monitoring Siting Criteria Viewer

AAQD Annex III, provides criteria for siting monitoring stations for all air pollutants[5] except ozone (O_3), whilst Annex VIII describes the siting criteria for ozone. Siting criteria are given on two different scales: a) a macroscale siting criterion that defines the general position of monitoring stations within a zone; and b) a microscale siting criterion that addresses the immediate vicinity of the monitoring station. Macroscale and microscale siting criteria in Annex III (and VIII for ozone) are provided to ensure representative and consistent monitoring strategies for air pollutants in all zones across Europe.

The macroscale siting criteria are given in Annex III B and aim to ensure that measurements are representative for the average exposure of human population, natural ecosystems and vegetation. With the purpose to protect human health AAQD Annex III B 1 (a) states that the measurements must specifically be made:

- in areas where the highest concentrations in the zone occur to which the population is likely to be directly or indirectly exposed for a period which is significant in relation to the averaging period of the limit value(s); and
- in areas which are representative for the exposure of the general population.

The macroscale siting criteria are therefore closely related to the concept of **station representativeness**. For each type of station, macroscale criteria are given to ensure (where possible) the representativeness of the sampling points. For traffic-oriented and industrial sites that are the focus of this study, Annex III B 1 (b) specifies that the sampling points should be sited in such a way that the air sampled is representative of air quality for a street segment no less than 100 m in length at traffic orientated sites or an area of at least 250 m x 250 m, in the case of industrial sites, where feasible. This is designed to prevent measurements representing very small micro-environments that are not generally representative of population exposure.

However, while the criteria in Annex III B 1 (a) apply for all sampling points, the numerical criteria on station representativeness in Annex III B 1 (b) are applied "where feasible".

The microscale siting criteria are provided in AAQD Annex III C and apply "in so far as practicable". The microscale criteria apply to the vicinity of the sampling point and provide basic requirements to ensure that a sampling point is representative for the area addressed by the macroscale siting criteria. The microscale criteria require that:

- the flow around the inlet sampling probe shall be unrestricted (generally, free in an arc of at least 270° or 180° in the case of sampling points at the building line);
- the inlet sampling height shall be between 1.5 m and 4 m above the ground;
- the inlet probe shall not be positioned in the immediate vicinity of sources, in order to avoid direct intake of emissions unmixed with ambient air;
- the sampler's exhaust outlet shall be positioned so that recirculation of exhaust air to the sampler inlet is avoided;
- for all air pollutants, traffic-orientated sampling probes shall, as far as is practicable, be at least 25 m from the edge of major junctions and no more than 10 m from the kerbside. A maximum distance of 10 m from the kerbside ensures that the highest

[5] These are: Sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter (PM_{10} and $PM_{2.5}$), lead, benzene and carbon monoxide in ambient air

concentrations in the zone are measured. A minimum distance from major junctions helps to avoid the specific influence of stop-go traffic situations.

The microscale siting criteria aim to ensure free air flow around the air inlet as a prerequisite for any representative measurement and avoid situations where the inlet is affected by direct sources instead of being exposed to mixed ambient air.

Any deviation from these criteria must be documented according to AAQD Annex III C (last paragraph), and III D. Annex III D requires also that the site selection for all sampling points is documented and that the documentation is reviewed at least every 5 years to ensure that selection criteria, network design and monitoring site locations remain valid and optimal over time.

Site selection according to the AAQD relies on local expertise and detailed local knowledge of the air quality conditions surrounding the sampling point. Local knowledge is also essential to fully document the site-selection procedures and record information to support network design and the choice of location of all sampling points. Following the requirements in AAQD Annex III D Member States are to document their site-selection procedures and include compass-point photographs of the area surrounding monitoring sites and detailed maps.

With a view to undertaking an initial, high-level review of the current agreement with the macroscale and microscale criteria, the capabilities of using satellite images and available mapping information was investigated. For this purpose, a prototype for an interactive GIS Viewer has been developed that can serve to objectively visualize the position of monitoring sites and determine their distance to emission sources, to help in the application of some of the siting criteria in the AAQD and thus support Member States in their site selection activities. The initial technical development and first results are described in the following sections, as well as conclusions and discussion on future work.

2.1 Technical implementation

The Monitoring Siting Criteria Viewer developed here combines independent emission information and GIS systems to compile information related to the siting criteria of traffic-oriented and industrial sites. It is intended as a support tool for siting criteria of air quality sampling points. Information about the geographical position of sampling points as reported by Member States under e-reporting was combined with available geospatial information on nearby sources. For industrial sites, this involved combining the information from the e-reporting database with spatial information on point sources from the European Pollutant Release and Transfer register (E-PRTR). For traffic-oriented sites, it involved combining the siting of stations with spatial information on roads from the OpenStreetMap (OSM).

The geospatial data has been stored in a common database and visualized with the help of the prototype GIS Viewer. It uses ASP.NET Core[6] as the main web development framework, and a base map provided by Mapbox[7]. In addition, the JavaScript library Mapbox-GL-JS[8] has been used to enable interactivity and customizability of vector maps on the web. The data model was implemented using the open-source relation database engine PostgreSQL[9] with a PostGIS[10] extension for handling spatial data.

6 https://dotnet.microsoft.com, last viewed in June 2019
7 https://www.mapbox.com, last viewed in June 2019.
8 https://github.com/mapbox/mapbox-gl-js, last viewed in June 2019
9 https://www.postgresql.org, last viewed in June 2019
10 https://postgis.net, last viewed in June 2019

Initially, the GIS Viewer has been designed to investigate how at its given location, a sampling point follows the two specific siting criteria under the macroscale and microscale criteria in Annex III that are given below:

- For traffic-oriented sites, sampling probes shall, as far as is practicable, be at least 25 m from the edge of major junctions and no more than 10 m from the kerbside. (Microscale siting, Annex III C).
- For industrial sites, sampling points must be sited in such way that the air sampled is representative of air quality for at least an area of 250 m x 250 m around the industrial source (Macroscale siting, Annex III B 1.(b)).

To investigate the first criteria, the position of all traffic-oriented sites was combined with spatial information on the positioning of roads and road junctions across Europe from OpenStreetMap (OSM)[11]. The position of the roads in OSM is given as a line in the center of the road, which is different from kerbside distance by half of the width of the road. The key functionality developed in the GIS Viewer allows calculating the distance from all traffic-oriented sites to the closest road kerbside, with the help of a position pointer. For all given sampling points, the GIS Viewer automatically calculates the distance to the kerbside of the nearest road by using the distance to the center of the road and adding half of the road width as a proxy. If the distance to the closest road center is less than 10 m plus half of the width of the road, the site is shown in blue as a sign of likely fulfillment of this microscale criteria requirement. If the distance to the closest road center is more than 10 m plus half of the width of the road, the site is flagged in red, indicating that there is need for further investigation of the site position. Similarly, automatic checks can be implemented in the future for establishing whether the distance to the edge of major junctions is above or below 25 m, but this checking functionality is not yet implemented. In both cases, information on the width of the road is important to determine whether the microscale criteria is possibly met. However, a main limitation of using OSM data for this type of analysis is the lack of systematic information on the width of the roads in the dataset. For the current exercise, in places where the width of the road was missing, information on the number of lanes was used to assume a fixed lane-width. Results are shown with two assumptions on the half a road width of either 3 m (Figure 2.2) or 6 m (Figure 2.3). To further improve this, one could apply image road geo-localization to identify the road-width close to the air quality sampling points. This functionality is not implemented in the current version of the Monitoring Siting Criteria Viewer.

To investigate the second criteria, spatial information on the location of industrial facilities as reported by the European countries was combined with data collected at the European Pollutant Release and Transfer Register (E-PRTR)[12]. All sampling points classified as industrial and reported by the countries are placed in the GIS Viewer as red dots while the industrial emission plants reported from E-PRTR are placed as black dots. Using the distance functionality in the GIS Viewer, the user can calculate the distance between any industrial site and any nearby industrial emission source. Users could enable this measurement functionality by clicking on the GIS Viewer, selecting the points as black circles and establishing the distance with a black line. Measured distances are given in kilometers in the top left side of the GIS Viewer. It should be noted however that calculating the distance to nearby industrial sources is not the same as evaluating the representativeness of the industrial sampling points. Further characterization of dispersion conditions are necessary to establish the representativeness required under Annex III.B.1.(b) for industrial sampling points.

11 https://www.openstreetmap.org, last viewed June 2029
12 https://prtr.eea.europa.eu, last viewed June 2019

2.2 Results from the GIS Viewer

The geographical coordinates (longitude and latitude) for all traffic and industrial sites that were registered in the EEA air quality portal[13] as active in 2017 were compiled and visualized in the GIS Viewer. In this context, operational sampling points are defined as those sampling points reporting air quality data in 2017, and thus considered active. Different sampling points may be co-located in the same monitoring site, so that the total number of sites included here is less that the total number of sampling points. The data includes 1,169 traffic-oriented and 559 industrial sites. The position of the monitoring sites is visualized in the interactive tool where it is possible to zoom to a great level of detail and identify the surroundings of the monitoring site from satellite images. Buildings, trees, parks, roads and industries close to the monitoring sites are easily identified in the GIS Viewer, where the functionality to calculate the distance between selected points is already implemented.

Figure 2.1 illustrates the results currently available in the Monitoring Siting Criteria Viewer for traffic-oriented sites. All traffic-oriented sites are depicted as either blue or red dots. Blue dots depict traffic-oriented sites that are that are placed within 10 m of the nearest kerb, while red dots depict traffic-oriented sites that are placed more than 10 m from the kerb of the nearest road.

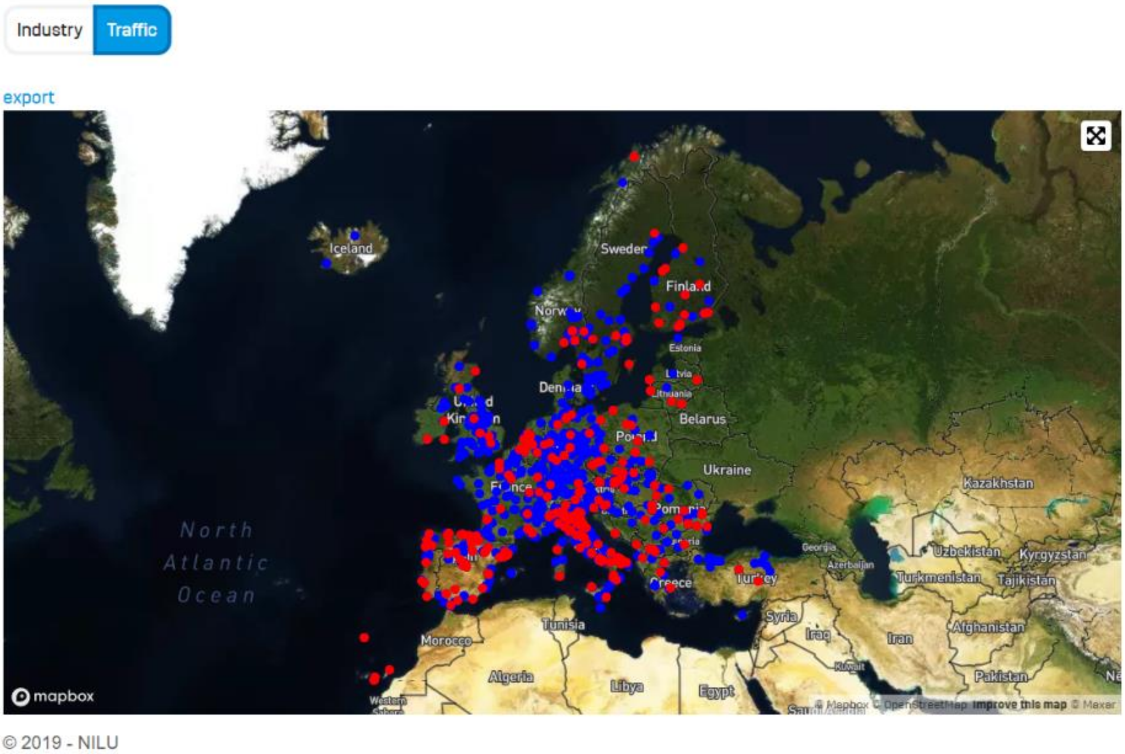


Figure 2.1: Traffic-oriented sites in the Monitoring Siting Criteria Viewer. Blue dots indicate monitoring sites that are placed within 10 m of kerbside from nearest road, while red dots depict traffic-oriented sites that are placed more than 10 m out from the kerbside of the nearest road. All based on the assumption that the width of the nearest road is 3 m.

The GIS Viewer provides support to determine whether the siting of the monitoring sites follows the microscale criteria in Annex III C, but it cannot definitively identify compliance or non-compliance with the Directive. This is because the evaluation of compliance is not the aim

13 https://aqportal.discomap.eea.europa.eu/products/data-download/download-e1a-e2a-for-previous-year/, last accessed in January 2020

of the exercise and, also because there are some important assumptions behind the calculations in the GIS Viewer, as explained below:

1. The first assumption is that the reported position of the sampling point in geographical (longitude and latitude) coordinates as stored at EEA CDR is in fact the correct position.
2. The second assumption is that the position of the nearest road in the OSM system is correct. The same applies for the position of industrial facilities under E-PRTR reporting.
3. The third assumption applies to the width of the road. It is acknowledged that different roads will vary both in terms of the width and the number of lanes, but for the current calculations addressing traffic-oriented sampling points in city areas, it was assumed that the road width is either 3 m or 6 m.

Figure 2.2. shows how the GIS Viewer allows for an independent investigation of the siting of different sampling points but it does not allow for determination of compliance or non-compliance with the Directive. The blue dot in Figure 2.2 is a traffic-oriented site, located according to its reported geographical coordinates. In this case, the sampling point's geographical coordinates are probably wrongly reported, because it is unlikely that the sampling point would be in the middle of the road. In this case, the blue dot should have been flagged as "red", but it is flagged as "blue". This is because it complies with the distance to road criteria. The blue line is the nearest road in the OpenStreetMap dataset, referenced to the middle of the road. For each traffic site (blue dot) a 13 or 16-meter radius around the traffic site is calculated, depicted as the white circle. When the kerbside of the nearest road is inside the white circle, it is assumed that the traffic site is situated in agreement to the Annex III C microscale criteria. In this case, the center line of the nearest road is inside the white circle. So, in the current viewer, the traffic-oriented site is marked as a blue dot. Thus, a blue dot does not necessarily imply compliance with the AAQD.

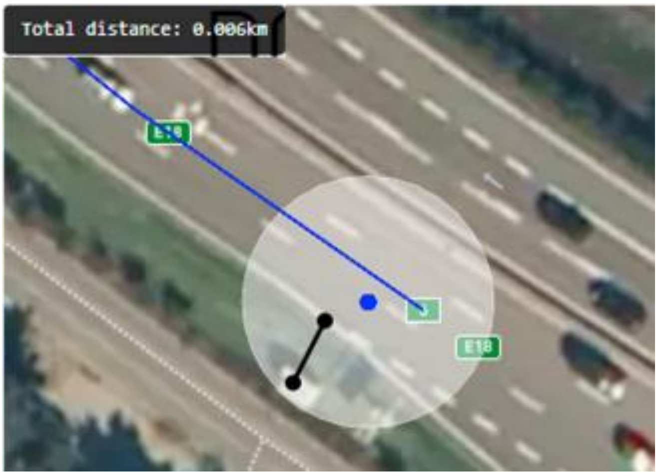


Figure 2.2 The GIS Viewer functionality to calculate the distance from a traffic site to the closest road kerbside as a support tool for the siting of sampling points.

The GIS Viewer has a functionality to check the distance to kerbside and identify possible inconsistencies. In this example, for instance, it allows investigation of the possibility of the sampling point being situated in the white building by the kerbside. The black circles and black line are from the manual measurement facility in the viewer, which the user can enable by clicking on the map. As indicated in the figure, the distance between the two black dots is 6 m (or 0,006km) which would indicate that if the sampling point was placed in or on the white building in the picture (first black dot) it would be placed 6 m from the kerbside of the nearest road (second black dot). In this case, the monitoring site would be placed correctly, and the reporting of geographical coordinates is what needs to be corrected.

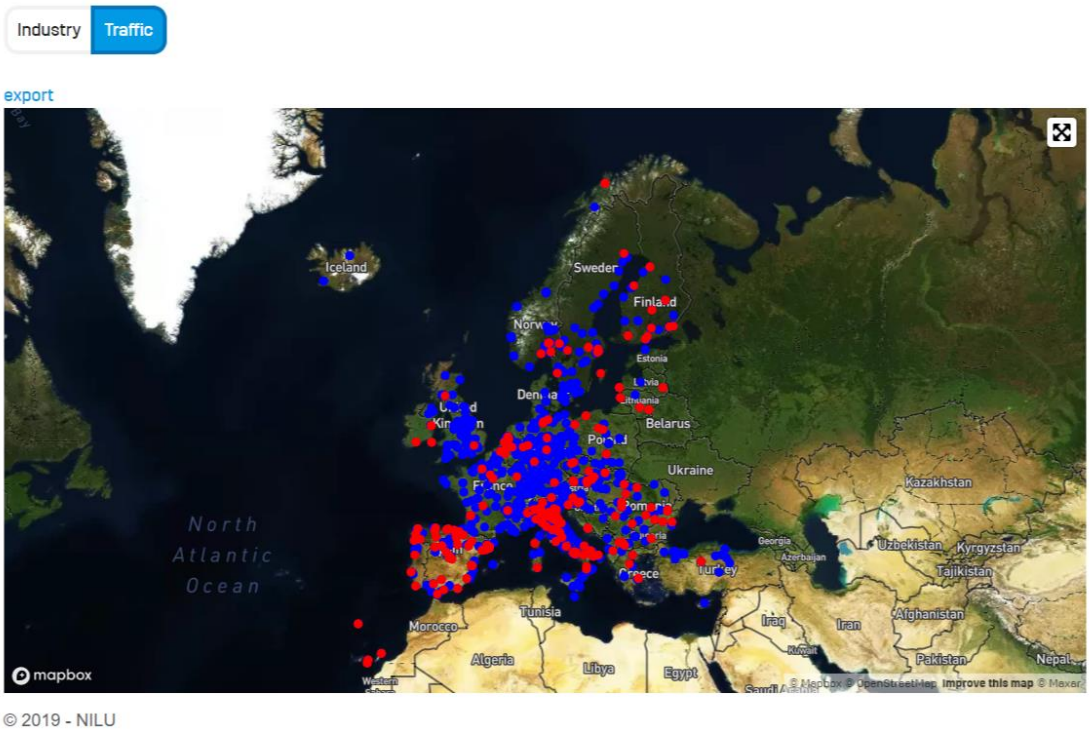


Figure 2.3: Traffic-oriented sites for NO_2 and PM_{10} in the Monitoring Siting Criteria Viewer. <u>Blue dots</u> indicate monitoring sites that are placed within 10 m of kerbside from nearest road, while <u>red dots</u> depict traffic-oriented sites that are placed more than 10 m out from the kerbside of the nearest road. All based on the assumption that the width of the nearest road is 6 m.

Results in Figure 2.3 assume that the width of the nearest road is 6 m. Since assumptions on the road width from the OSM system introduce some uncertainty in the results of this independent check, we have chosen to combine the results from Figure 2.2 and 2.3 to provide an estimate of the total number of traffic-oriented sites that may not be located less than 10 m from kerbside of main roads despite the requirement from Annex III C microscale siting criteria. Our analysis indicates that around 20% of all traffic oriented sampling points may be located more that 10 m from kerbside. It is interesting to note that these results are consistent with the analysis in Chapter 3 and Chapter 4, where sampling points in Austria, Finland, Italy, Poland, Romania, Spain and Sweden are identified as possible “outliers” with respect to siting criteria aspects. However, the recognized assumptions on the road width, these results are only indicative and need to be further validated by the countries in question.

Figure 2.4: Industrial sites for NO_2 and PM_{10} in the Monitoring Siting Criteria Viewer. Red dots are the reported industrial monitoring sites, while black dots show the position of industrial plants reported under E-PRTR.

Figure 2.4 shows the position of industrial monitoring sites depicted as red dots, while the black dots show the position of industrial plants reported under E-PRTR. As explained above, the Viewer includes a distance calculating functionality option that allows measurement of the distance between industrial sampling points and nearby industrial sources. This is a useful first step to help in the analysis of fulfillment with the macroscale criteria in Annex III. B. 1 (b). For industrial locations, the macroscale siting criteria in Annex III B. requires for the air sampled to be representative of air quality in an area of at least 250 m x 250 m. Agreement with this macroscale criteria requires an evaluation of the spatial representativeness of the industrial sampling point. The current GIS Viewer is designed to facilitate an evaluation of the spatial representativeness by using the criteria of proximity to industrial sources, but it does not provide an evaluation of sampling point classification. It can be further developed to include small scale industrial facilities not currently in E-PRTR and improved by linking it to information on dispersion conditions that will be useful to calculate site representativeness.

2.3 Conclusions on the application of siting criteria

The GIS Viewer developed here provides an objective independent method to support the application of the siting criteria of air quality sampling points in a qualitative manner. It is not recommended to use the Viewer for quantitative evaluation because the results rely strongly on the accuracy of the data gathered. In particular, the viewer relies on the accuracy of the reported position of sampling points given in geographical coordinates, on the accuracy of the road map descriptions in the OSM system and on the reliability of the reporting of the position of industrial facilities under E-PRTR reporting. A major limitation for widely using this Monitoring Siting Criteria Viewer to quantitatively consider the application of some of the microscale and macroscale siting criteria in Annex III is the lack of information on the width of the roads in the OSM dataset, which is crucial for determining the distances to kerbside as required by the AAQD.

However, the GIS Viewer could be a useful way for Member States to easily investigate the position of sampling points and consider to what extent they may follow the macroscale and microscale siting criteria in Annex III. It is useful because it provides the reported location of sampling points and allows a calculation of the distance to nearest roads and nearest reported industrial sites.

The automatic independent checks for traffic-oriented sites show that about 20% of all sampling points may be located more than 10 m from the kerbside of the road. This appears to be the case particularly for sampling points in Spain, Italy, Romania, Poland, Sweden and Finland, correlating with sampling points that have also been identified as “outliers” in Chapters 3 and 4 on aspects relating to siting criteria. However, given the above-recognized caveats, notably assumptions made on road width, these results are only indicative and need to be further validated by the countries in question. Nevertheless, these indicative results are an indication of the potential capabilities of this simple GIS Viewer to support, check and improve the siting of sampling points in accordance with the AAQD.

Although beyond the scope of this project, the Monitoring Siting Criteria Viewer could be greatly improved in several ways for use by countries to support their siting activities:

- Use road geo-localization to apply image recognition of the road width close to the air quality sampling points.
- Developing further the measuring facility to include distances to junctions, bus stops, roundabouts, etc. for traffic-oriented sites
- Including dispersion condition information and additional industrial emission sources to help determine the station representativeness for industrial sites.

Such an GIS Viewer could be part of the checking capabilities to support monitoring design and simplify siting and classification activities at national level. It would be possible for countries to use it to check whether metadata information on monitoring sampling points have been introduced correctly in e-reporting. In this context, the viewer would profit from an additional feature that could show the available reported metadata information when clicking on a specific sampling point.

3 Review of reported metadata information

The sampling point location and classification must be reported as part of the metadata reporting requirements under the Implementing Decision (Decision 2011/850/EU). Member States are to report metadata information on the monitoring stations and sampling points in accordance with the Guidance for the Implementing Provisions for Reporting. This metadata information is intended to give insight to details regarding the sampling points and to facilitate the interpretation of the measurement data.

The structure of the data to be reported by countries to the Central Data Repository (CDR) operated at the EEA is laid down in Annex II of the IPR Decision and further specified in the IPR Guidance XML user guide (EEA, 2018).

This chapter reviews the metadata information on all sampling points officially reported by EU Member States and other European countries in dataflow D to the CDR. The metadata information that has been considered is summarized in Table 3.1. This is a selection of metadata information related to sampling point siting location and classification. This metadata review aims to extend the evaluation of sampling location carried out in Chapter 2 and provide further insight on the application of the siting criteria following the macroscale and microscale criteria in AAQD Annex III.

Table 3.1: Metadata information selected with focus on sampling point location and classification. References to the EU AQ legislation which requires the metadata to be reported are given in parentheses.

Mandatory - All	Mandatory - Traffic	Mandatory, where available	Voluntary
Inlet height (AAQD, Annex III. C.and 2011/850/EC II. D ii.19)	**Building distance** (AAQD, Annex III.C and 2011/850/EC, II. D.ii.20) .	**Station information** (2011/850/EC, II. D.ii.27)	**Dispersion local** (IPR Guidance, XML user guide D5.2.11.1 pp 199)
Altitude (2011/850/EC,II.D.ii.26)	**Kerb distance** (AAQD Annex III.C and 2011/850/EC, II.D.ii.21)	**Main sources** (2011/850/EC,II.D.ii.23)	**Dispersion regional** (IPR Guidance, XML user guide D5.2.11.8 pp 203)
Latitude, Longitude (2011/850/EC,II.D.ii.26)	**Distance to major junction** (AAQD Annex III.C and 2011/850/EC, II.D.ii.29),	**Spatial extent of representative area** (2011/850/EC,II.D.ii.16).	**Traffic emissions** (IPR Guidance, XML user guide D5.1.5.3. pp 127)
Classification of the area (2011/850/EC,II.D.ii.28)	**Traffic volume** (2011/850/EC,II.D.ii.30)	**Mandatory, where available - Traffic**	
Station classification (2011/850/EC,II.D.ii.22) .	**Mandatory - Industrial**	**Heavy duty fraction** (2011/850/EC, II.D.ii.31)	
	Distance from source (AAQD, Annex III.B.1.b and 2011/850/EC, II.D.ii.24)	**Traffic speed** (2011/850/EC, II.D.ii.32)	
	Industrial emissions (IPR Guidance, XML user guide D5.1.5.5. pp 128)	**Street-canyon - Width of street** (2011/850/EC,II.D.ii.33)	
		Street canyon - Height of building facades (2011/850/EC, II.D.ii.34)	

The columns in Table 3.1 are organised following the terminology used in the guidance IPR Guidance XML user guide (EEA, 2018) to distinguish between “Mandatory”, ”Conditional” “Mandatory, where available” and “Voluntary” metadata. The selected metadata information includes 22 different parameters:

- Six parameters of “Mandatory” metadata, that is required to be reported for all sampling points.
- Six parameters of “Conditional” metadata, which means metadata that is mandatory to report depending on the sampling point classification, either these are traffic-oriented (4 parameters, referred to in the table as “Mandatory-Traffic”) or industrial samplings points (2 parameters, referred to as “Mandatory-Industrial”).
- Seven parameters of “Mandatory, where available” metadata that applies for either all sampling points (3 parameters) or conditionally only to traffic-oriented sampling points (4 parameters, referred to as “Mandatory, where available-Traffic” including 2 parameters for a subset of street canyon sampling points).
- Three “Voluntary” metadata parameters are also included in the selection because such metadata on emissions and dispersion conditions is in many cases relevant to determine the representativeness of the sampling points.

3.1 Approach

The present analysis of metadata combines the information from all NO_2 and PM_{10} sampling points that were operative[14] in 2017. Officially reported metadata for all sampling points were compiled from EEA’s CDR in May 2020. Metadata from a total of 6,972 sampling points was compiled for analysis. This includes 2,007 traffic-oriented sampling points, 905 industrial sampling points and 4,060 background sampling points. The number of sampling points for which metadata was analyzed, for different classification types and country groups is given in Table 3.2. Note that above 90% of all sampling points considered are inside EU28, so that the analysis results for all sampling points would be mostly dominated by the situation in the European Union and the United Kingdom.

Table 3.2 Total number of sampling points for which their metadata has been analyzed. The definition of the different country groups are given in Appendix 2.

	EU27	EU28	EEA33*	EEA39*	All
Total sampling points	6193	6421	6893	6965	6972
Total traffic-oriented sampling points	1758	1859	1983	2004	2007
Total industrial sampling points	835	854	892	905	905
Total background sampling points	3600	3708	4018	4056	4060

The metadata was sorted per country and parameter and all values/entries were carefully reviewed in terms of their completeness. The information was also checked to establish whether the reported information was plausible. This completeness and plausibility evaluation was done for each country. To make the evaluation comparable across Europe percentages were used with respect to the total number of sampling points in each country. It should be noted that metadata information is reported by sampling point, not by monitoring site. A monitoring site may measure different air pollutants or air quality components, in which case, the same monitoring site will include different sampling points – one for each component. This

[14] Operative sampling points are defined here as those sampling points reporting air quality data in 2017, and thus considered active. Other definitions of operational sampling points were also tested and resulted in similar conclusions although different results on the total number of sampling points analyzed.

is the reason why the number of sampling points is used as reference in the percentage analysis in this evaluation. The evaluation was not automated at this stage.

In order to calculate the total number of sampling points for which plausible metadata is reported, the number of sampling points with no metadata information available were subtracted from the total number of sampling points. This process highlighted the fact that currently there is no homogeneous way to report "non-available" data across European countries. When data is not available for a specific parameter, the currently reported metadata includes values such as -9,999, -999, -100, -99, 999, 9,999 or "NULL". In many cases, only blanks were reported. Also "0" was in many cases used to indicate non-available metadata.

There are metadata information requirements that involve numerical values, such as distances and heights. This type of metadata can be evaluated in terms of its plausibility and the evaluation has allowed the identification of some obviously questionable data. However, the plausibility of metadata reported for some other metadata requirements - such as Station classification, Area classification, Dispersion local, Dispersion regional, Main emission sources - cannot be assessed properly without local knowledge of the monitoring sites and sampling points.

The analysis of the twenty-two metadata information parameters in Table 3.1 focuses on the number of sampling points with reported metadata deviating from the requirements in legislation. It is based in the following percentages:

- percentage of sampling points with **not reported data** (including. values which are identified as lacking demanded information, e.g. -999, blanks, etc.) for each specific metadata information parameter and with respect to the total number of sampling points in each country,
- percentage of sampling points with reported **unit errors** (i.e. missing or wrong unit for the number value reported) with respect to the total number of sampling points reporting the specific metadata information in each country, and
- percentage of sampling points reporting **possible number errors** (i.e. implausible numbers, values higher/lower than required in the AAQD, values equaling 0) with respect to the total number of sampling points reporting the specific metadata information in each country.

It should be noted that "possible error numbers" do not necessary mean that the sampling point metadata is wrongly reported, or the sampling points are inappropriate, but they flag a series of questions for further dialog with the countries on the reported metadata.

3.2 Results of the metadata analysis on siting and classification of sampling points

Results in this section are presented following the structure in Table 3.1, beginning with the analysis of "mandatory" metadata for all sampling point types, followed by mandatory metadata for either traffic or industrial sampling points. The analysis of the country reports for metadata that is "mandatory, where available" follows for those parameters required for all sampling points and for traffic sampling points. At the end of the section, the analysis of reported "voluntary" metadata information is presented.

In all figures in this chapter, countries are presented in alphabetical order following the abbreviation glossary list from Eurostat, as given in Appendix 2. In the figures, EU Member

States are presented first from the left and other European countries follow to the right, organized as EU28, EEA33, EEA39 and the rest of countries.

3.2.1 Mandatory metadata for all sampling points

Six metadata parameters mandatory for all sampling points are evaluated here. These are: Inlet height, Altitude, Latitude, Longitude, Area classification and Station classification.

3.2.1.1 Inlet height

Directive 2008/50/EC Annex III, Section C, in the version amended by Directive 2015/1480, requires that "in general, the inlet sampling point shall be between 1,5 m (the breathing zone) and 4 m above the ground. Higher siting may also be appropriate if the station is representative of a large area and any derogations should be fully documented".

Inlet height is a metadata parameter which is affiliated with an inlet height unit, i.e. a reasonable unit needs to be reported in conjunction with the number value reported as inlet height. Following the IPR Guidance XML user guide, such units should be meters "m". A source of uncertainty or deviation from requirements in the reported metadata is related to situations when the inlet height unit is missing, even though one may expect that the unit should be meters "m".

Most countries report a number value for the inlet height. However, some of the reported values are so-called "no number values". "No number value" in this context refers to values like -999, -100, -99, -9,999, 999, 9,999, which can be interpreted so that the inlet height is not known for the sampling point. These "no number values" were considered in this analysis as "not reported" metadata information and were as high as 24% of all sampling points in Spain and 6% in Italy.

Possible questionable reported numbers for inlet height include reported inlet height "0" (which could also be interpreted as non-available information) or reported inlet height well above 100. With a few exceptions, such as for sampling points on masts or on rooftop locations in cities, these high inlet height numbers indicate a need for further review of the reported metadata. For instance, Denmark reports inlet heights over 150 m at urban traffic sites, while Montenegro reports "0 m" for all sampling points. As mentioned before, these "possible error numbers" do not necessarily mean that the sampling point metadata is wrongly reported, or the sampling points are inappropriately mounted, but they flag a series of questions for further dialog with the countries on the reported metadata.

Another identified problem, although not included in the statistics below, is that some countries report constant inlet heights for all sampling points. For instance, Romania reports an inlet height of "4" for all 207 sampling points; Slovakia reports "4" for all 57 sampling points and Turkey reports "3.5" for all 278 sampling points. For those countries reporting constant inlet heights, the information given seems to be a guess instead of reporting the real inlet height for each individual sampling point.

The share of questionable inlet height metadata reported by the different countries is depicted in Figure 3.1. The results of the analysis of inlet height reported metadata is summarized below for all sampling points across Europe.

- ➔ For 5% of all sampling points in EU27 and EU28, no inlet height was reported or reported as "no value number" The percentage of no inlet height reported is 4% for all countries.

- ➔ For 100% of those sampling points with a number reported, units were correctly reported
- ➔ For 1% of those sampling points with a possible questionable number reported, the value was equal to 0 or larger very high value.

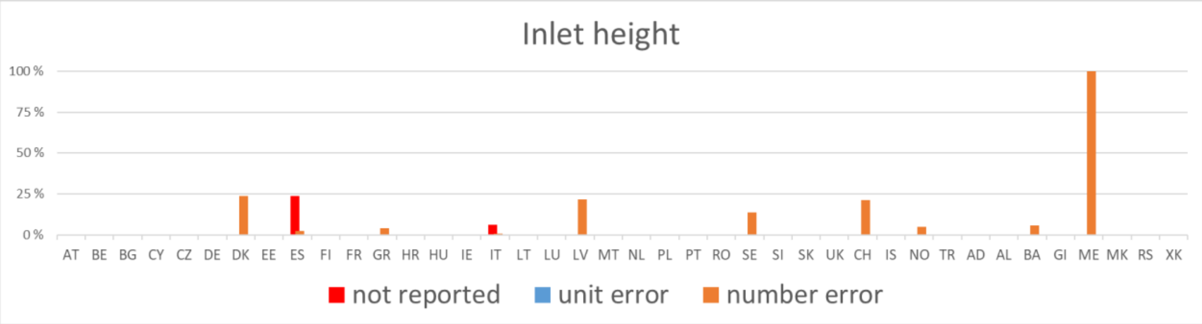


Figure 3.1: Percentage of questionable or missing inlet height metadata reported by countries. Red bars: percentage of sampling points for which no data is reported. Blue bars: percentage of sampling points for which no unit or a wrong unit was reported. Orange bars: percentage of sampling points for which a possible number error was reported.

3.2.1.2 Altitude, Latitude and Longitude

AAQD Annex III Section D prescribes the obligation to document site selection while Decision 2011/850/EU, Annex II. D (ii) (26) specifies reporting the geographical coordinates: altitude, latitude and longitude of the monitoring station.

Following the IPR Guidance XML user guide, the altitude of a monitoring station is to be reported in meters above mean sea level and is mandatory metadata information for all sampling points. The altitude parameter is affiliated with an altitude unit, i.e. a reasonable unit needs to be reported with the number value. Countries report number values for altitude for most sampling points, with the exception of 8% of the sampling points in Iceland. The plausibility of the values reported is not straightforward to assess. An altitude of 0 m or even negative altitudes are possible at coastal sites and in low plains, however, Croatia reports 0 for all its stations.

- ➔ For 100% of all sampling points, a number value was reported for altitude.
- ➔ Only problem identified is the lack of altitude metadata not reported in 8% of the sampling points in Iceland.

The geographical coordinates of monitoring stations are expected to be reported according to the coordinate reference system ETRS 1989 (or WGS 1984 in a transitionary period to 2020). All countries report geographical coordinates, however sometimes the coordinate formats are seen to differ from the required reference systems.

- ➔ For 100% of all sampling points, a number value was reported for both latitude and longitude.

The summary of questionable metadata reporting per countries with respect to the geographical coordinates of sampling points depicted in Figure 3.2 below shows good metadata reporting for these parameters.

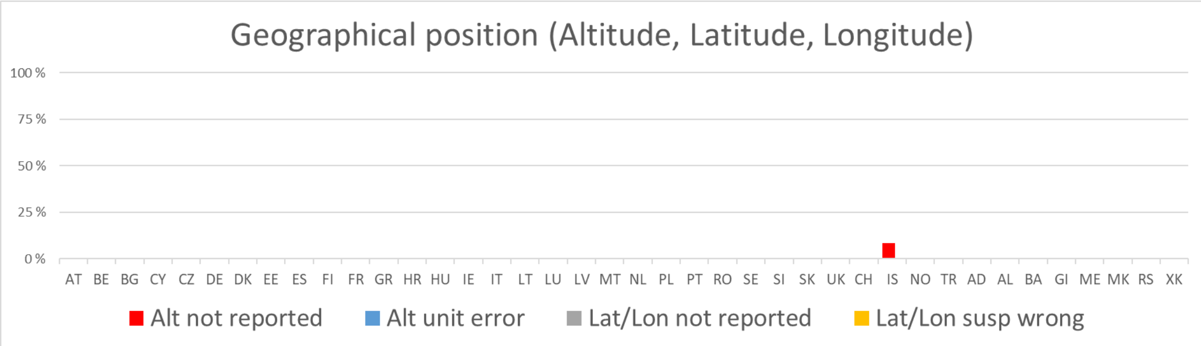


Figure 3.2: Percentage of questionable or missing geographical position metadata reported by countries. Red bars: percentage of sampling points for which no data is reported for altitude. Blue bars: percentage of sampling points for which no unit or a wrong unit was reported for altitude. Grey bars: percentage of sampling points for which no data is reported for latitude and/or longitude. Yellow bars: percentage of sampling points (for which the latitude and/or longitude data is not according to the required coordinate reference systems.

3.2.1.3 Classification of the area

The local area around a monitoring station is classified based on a code list distinguishing between urban, suburban and rural, where rural can (but does not have to) be subdivided into rural-near city, rural-regional and rural-remote.

Following the requirements in Decision 2011/850/EC,II.D.ii.28, all countries have selected an area classification code for each sampling point. To assess the plausibility of the codes without local knowledge of the individual sampling points, it would be necessary to rely on satellite data, for instance by using the GIS Viewer developed in Chapter 2.

➔ For 100% of all sampling points, a valid code was selected as depicted in Figure 3.3.

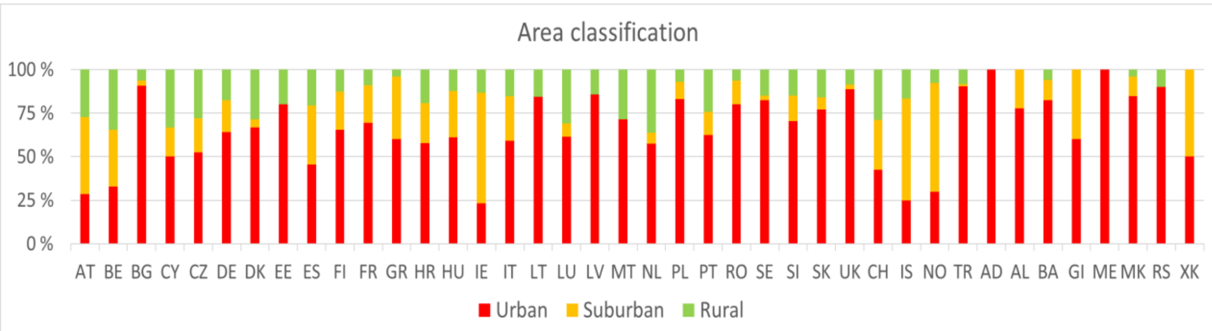


Figure 3.3: Distribution of area classes "urban", "suburban" and "rural" for all sampling points in each country.

3.2.1.4 Station classification

Following the IPR Guidance, monitoring stations are classified in relation to major emission sources relevant for the measurement location. The classification is based on a code list distinguishing between traffic, industrial and background. The classification is made for each sampling point. All countries have selected a code for each sampling point, according to the requirements in Decision 2011/850/EC,II.D.ii.22.

Figure 3.4 summarizes the current distribution of sampling points by station types. It is interesting to note that Andorra and Kosovo have not classified any sampling point as traffic or industrial stations, only background stations. Without having studied the monitoring data from these countries, or having examined the locations of the individual stations, it seems unusual that the urban stations in the three countries are exclusively background stations. Further, there are several countries, which do not classify any sampling points as industrial stations: Albania, Cyprus, Denmark, Gibraltar, Latvia, Montenegro and Sweden.

In summary, for all sampling points:

- ➔ For 100% of all sampling points, a valid code was selected.
- ➔ Two countries outside EU28 (Andorra, Kosovo) have not classified any traffic or industrial stations, only background stations.
- ➔ Thirteen countries (7 EU Member States: Cyprus, Denmark, Ireland, Luxembourg, Latvia, Malta and Sweden; and 7 non-EU countries: Andorra, Albania, Gibraltar, Kosovo, Montenegro, and Switzerland) have not classified any industrial stations for NO_2 and PM_{10} sampling points.

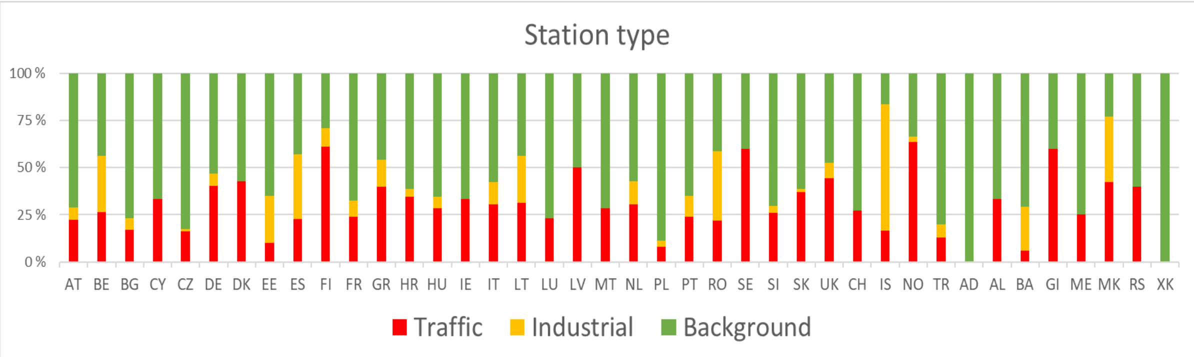


Figure 3.4: Distribution of Station types "traffic", "industrial" and "background" for all sampling points in each country.

3.2.2 Mandatory metadata for traffic sampling points

In this section, four metadata information parameters relevant to characterize the siting and representativeness of traffic stations are analyzed. These are: building distance, kerb distance, distance to major juncture and traffic volume. All these metadata are mandatory for traffic sampling points and related to the application of the macroscale siting criteria and the determination of station representativeness as well as to the application of the microscale siting criteria in the AAQD Annex III.

3.2.2.1 Building distance

The mandatory metadata for traffic sampling points is in fact related to the application of the macroscale siting criteria and the determination of station representativeness as well as to the application of the microscale siting criteria.

The horizontal distance of the inlet to the nearest building (in meters) is mandatory metadata information for traffic stations. It addresses directly the microscale siting criteria in AAQD Annex III C first indent in order to ensure that the flow around the inlet sampling probe is unrestricted.

For 39% of the traffic related sampling points, number values are reported. Montenegro, the Netherlands, Norway and Turkey do not report the building distance. A building distance of

0 m may be real for sampling points at a house wall, although the AAQD establish a minimum distance of 0.5 m. Still, the large percentage of 0 values reported by Greece (in 55% of all sampling points) and Hungary (in 57% of all sampling points) does not appear realistic. Belgium and Romania report the same building distance for all their traffic sampling points, leading to the assumption that the values were not assessed per sampling point. Many countries do not report a unit for the building distance.

The share of questionable metadata on building distance reported by the different countries is depicted in Figure 3.5. The reporting of this metadata parameter is summarized below for all traffic sampling points across Europe:

- ➔ For 29% of all traffic sampling points, no building distance was reported. The percentage is 27% for EU27 and 25% for EU28
- ➔ For 41% of those sampling points with a number reported, no unit or a wrong unit was reported. The percentage is 43% for EU27 and 40% for EU28
- ➔ For 2% of those sampling points with a number reported, the value was equal to 0 (especially Greece, Hungary). The same percentage applies for EU27 and EU28.

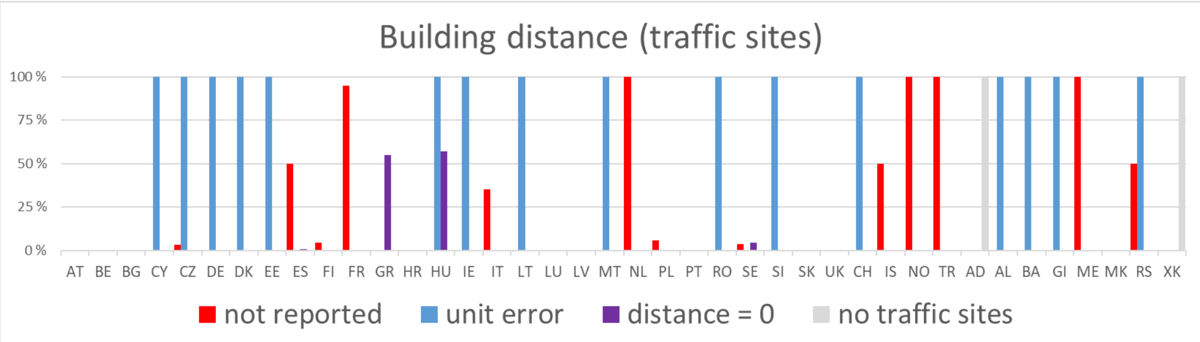


Figure 3.5: Percentage of questionable or missing metadata on building distance reported by countries. <u>Red bars:</u> percentage of all traffic sampling points for which no data is reported. <u>Blue bars:</u> percentage of sampling points for which no unit or a wrong unit was reported. <u>Purple bars</u>: percentage of sampling points for which a value equaling 0 was reported. <u>Grey bars:</u> countries that have not classified any traffic-related sampling points.

3.2.2.2 Kerb distance

The horizontal distance of the inlet to the nearest kerb in meters has to be less than 10 m for traffic stations according to the microscale criteria in AAQD Annex III C, fourth indent. Following the requirement in Decision 2011/850/EC, II.D.ii.21, most countries report values on the kerb distance for most of their traffic stations, except from Turkey and Montenegro. Norway and The Netherlands have also very high percentage of traffic oriented sampling points not reporting the kerb distance. Most of the countries allocated a unit to the distance values. A kerb distance 0 was reported for 45% of the traffic sampling points in Greece and 57% in Hungary, possibly indicating that the sampling point is at kerbside.

In Austria, Bosnia Hercegovina, Bulgaria, Czech Republic, Spain, Finland, Croatia, Ireland, Poland, Romania, Norway, Bosnia and Herzegovina and North Macedonia for more than 25% of the sampling points with values reported, kerb distances larger than 10 m were reported. For some of the countries (France, United Kingdom), the kerb distance of the sampling points reported to be more than 10 m from the kerb is only slightly larger than 10 m (11 m and 12 m are reported), while more than 50 m is reported for some sampling points by Bulgaria, Spain, Finland, Italy, Latvia (500 m) and Sweden. This seems to agree with the conclusions in

Chapter 2 where several sampling points were flagged red with the help of the Monitoring Siting Criteria Viewer.

No unit errors were identified in the reporting of this metadata parameter.

The breakdown of questionable or missing metadata on kerb distance reported by the different countries is depicted in Figure 3.6 and summarized below for all traffic sampling points across Europe:

- ➔ For 26% of all traffic sampling points, no kerb distance was reported. The percentage is 23% for EU27 and 22% for EU28.
- ➔ For 2% of those traffic sampling points with a number reported, the value was equal to 0 (especially large percentage occurrence in Greece and Hungary). The same 2% percentage applies for EU27 and EU28.
- ➔ For 16% of those traffic sampling points with a number reported, the value larger than 10 m (especially high occurrence in Romania, Norway and Bosnia and Herzegovina). The percentage is 17% for EU27 and 16% for EU28.

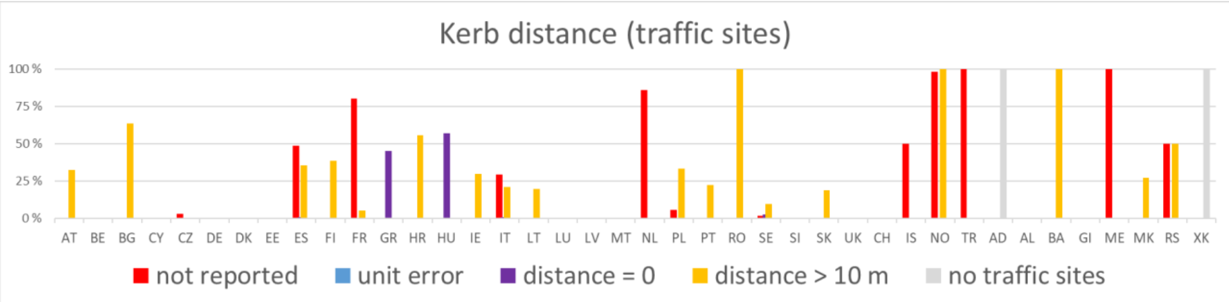


Figure 3.6: Percentage of questionable or missing metadata on kerb distance at traffic sampling points. <u>Red bars</u>: percentage of sampling points for which no data is reported. <u>Blue bars</u>: percentage of sampling points for which no unit or a wrong unit was reported. <u>Purple bars</u>: percentage of sampling points for which a value equaling 0 was reported. <u>Orange bars</u>: percentage of sampling points for which a value above 10 m was reported. <u>Grey bars:</u> countries which have not classified any traffic-related sampling points.

3.2.2.3 Distance to major junction

The distance from a station to the edge of a major junction has to, as far as is practicable, be greater than 25 m for traffic stations according to the microscale siting criteria in AAQD Annex III C, fourth indent. A “major junction” here refers to any junction which interrupts the traffic flow and causes significantly different emissions (stop / go) from the rest of the road. It should be noted that the analysis in Chapter 2 did not allow for an independent evaluation of this metadata parameter as yet, but the GIS Viewer could in the future be easily extended to include the “distance junction” parameter.

Following Decision 2011/850/EC, II.D.ii.29, it is mandatory for traffic stations to report “distance to major junction”, however, most countries do not report this parameter. All countries which report a value, also report a unit. Germany, Spain and Sweden report 0 for individual stations, which suggests that the stations are located within a major junction. Alternatively, the value 0 could be given to indicate that the distance to a major junction has not been assessed. Furthermore, many countries (among them Spain, Latvia, Croatia) report distances below 25 m.

The share of questionable or missing metadata on the parameter “distance to major junction” reported by the different countries is depicted in Figure 3.7 and summarized below for all traffic sampling points across Europe:

- ➔ For 35% of all traffic sampling points, no distance to junction was reported. The percentage is 33% for EU27 and 31% for EU28.
- ➔ For 3% of those sampling points with a number reported, the value was equal to 0 (especially prevalent in Spain).The same percentage applies for EU27 and EU28.
- ➔ For 23% of those sampling points with a number reported, the value was less than 25 m, contrary to microscale siting requirements. The percentage is 25% for EU27 and 23% for EU28.

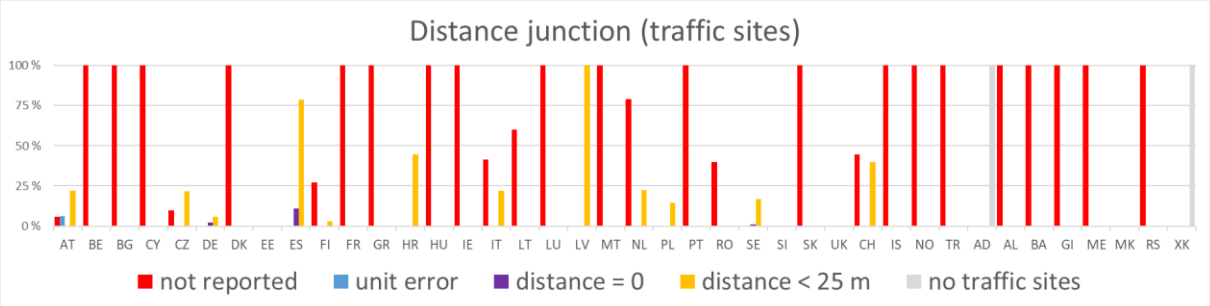


Figure 3.7: Percentage of questionable or missing metadata information on distance to major junction at traffic sampling points. Red bars: percentage of sampling points for which no data is reported. Blue bars: percentage of sampling points for which no unit or a wrong unit was reported. Purple bars: percentage of sampling point for which a value equaling 0 was reported. Orange bars: percentage of sampling points for which a value below 25 m was reported. Grey bars: countries which have not classified any traffic-related sampling points.

3.2.2.4 Traffic volume

Information on sources in the vicinity of sampling points is essential to determine the representativeness of the monitoring site. For traffic-oriented sampling points, traffic emissions are the most relevant sources. Traffic emissions are generally dependent on traffic volume, vehicle composition and traffic speed. Therefore, these metadata parameters contain relevant information to characterize the representativeness of traffic-oriented sampling points.

Following Decision 2011/850/EC, Annex II.D.(ii) (30), it is mandatory for traffic sampling points to report traffic volume. The total traffic volume metadata information corresponds to the annual average daily traffic (AADT) on the adjacent road. However, despite the requirement, about half of the countries do not report these values.

Additional identified deviations were found. For instance, Spain and Croatia reported a traffic volume of 0 for some or all of their traffic stations, which is not realistic for a traffic station. Spain, Finland, Italy and Sweden reported traffic volume below 1,000 AADT for selected stations, which appears too low for a traffic site. The traffic volume values reported by North Macedonia are identical with the values reported for the metadata parameter “Distance to major junction”, which indicates that there is possibly an error in the reporting of either one of the two parameters (most probably in the “Distance to major junction” parameter).

In summary over Europe and as shown in Figure 3.8:

- ➔ For 39% of all traffic sampling points, no traffic volume was reported. The percentage is 38% for EU27 and 36% for EU28.

- ➔ For 4% of those traffic sampling points with a number reported, the value was equal to 0 (Spain and Croatia).The percentage is 5% for EU27 and 4% for EU28.
- ➔ For 3% of those traffic sampling points with a number reported, the value appeared too low (in Spain, Finland, Italy and Sweden). The same percentage applies for EU27 and EU28.

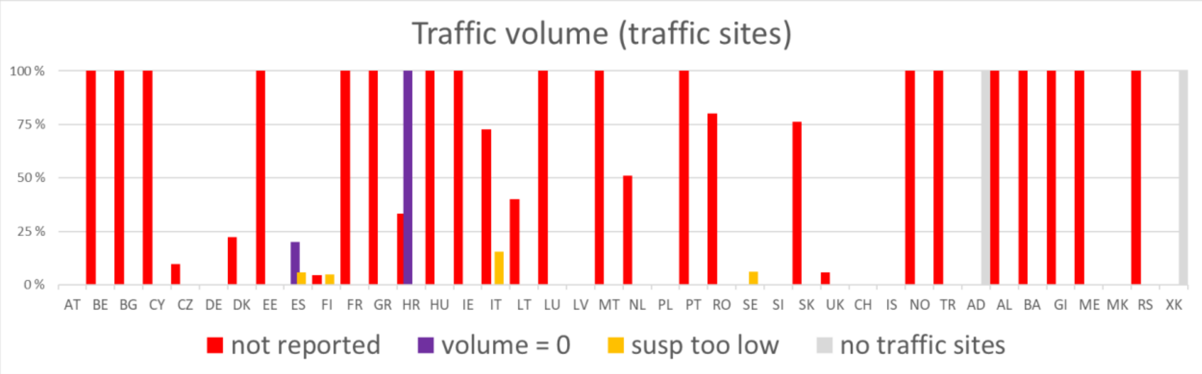


Figure 3.8: Percentage of questionable or missing traffic volume metadata reported on traffic sampling points. Red bars: percentage of sampling points for which no data is reported. Purple bars: percentage of sampling points for which a value equaling 0 was reported. Orange bars: percentage of sampling points for which an unrealistically low value was reported. Grey bars: countries which have not classified any traffic-related sampling points.

3.2.3 Mandatory metadata for industrial sampling points

The macroscale siting criteria for industrial sampling points in AAQD Annex III C (b) requires industrial sampling points to be located to avoid very small micro-environments in the immediate vicinity – such that the air sampled is representative of air quality for at least 250 m x 250 m. To determine an industrial sampling point representativeness, good understanding of the dispersion conditions and the industrial emissions in the vicinity of the sampling point are necessary. For this reason, the mandatory metadata requirements for industrial sampling points evaluated here include the distance to source (as a proxy for dispersion conditions) and industrial emission sources. The two metadata parameters are especially relevant to inform the evaluation of the impact of industry on its surroundings and apply to industrial sampling points.

3.2.3.1 Distance from source

AAQD, Annex III.B.1.b prescribes that “sampling points shall in general be sited in such a way as to avoid measuring very small micro-environments in their immediate vicinity, which means that a sampling point must be sited in such a way that the air sampled is representative of air quality for at least 250 m × 250 m at industrial sites, where feasible”.

Decision 2011/850/EC, Annex II.D.(ii) (24) indicates that for industrial sampling points, the distance from the predominant industrial source or source area must be declared. The distance to the stack in case of point sources or to the nearest edge of a source area in case of spatially distributed emissions shall be provided. According to the XML user guide, the allowed format is numeric integer values with the units in meters. This metadata information is mandatory for all industrial sampling points

Out of the 31 countries which categorized industrial stations, 10 did not report the distance to the industrial source. Croatia reported the wrong unit (t/year) for the distance. France reported distance as 0 for most of their industrial stations, which is questionable, since the industrial site must not be located at the industrial plant according to the AAQD. Very short distances from source were reported for some sampling points by Germany (10-20 m), Spain (0.02-32 m), Finland (4-20 m) and United Kingdom (2 m). Romania reported 500 km from an industrial sampling point, which is a long distance to expect the sampling point to be representative for measuring the impact of the industrial plant. Here it should be noted that the units where reported generally as m, except for the one sampling point in Romania.

The share of questionable or missing metadata on the parameter "distance from source" reported by the different countries is depicted in Figure 3.9 and summarized below for industrial sampling points across Europe:

- ➔ For 44% of all industrial sampling points, the distance from source was not reported. The percentage is 43% for EU27 and 42% for EU28.
- ➔ For 4% of the industrial sampling points, no units or wrong units were reported. The same percentage applies to EU27 and EU28.
- ➔ For 10% of those industrial sampling points with a number reported, the value was equal to 0 (in particular for France). The percentage is 11% for both EU27 and EU28.
- ➔ For at least 23% of the industrial sampling points with a number value reported, the value is questionable because is either too short or too long. The percentage is 24% for EU27 and 23% for EU28.

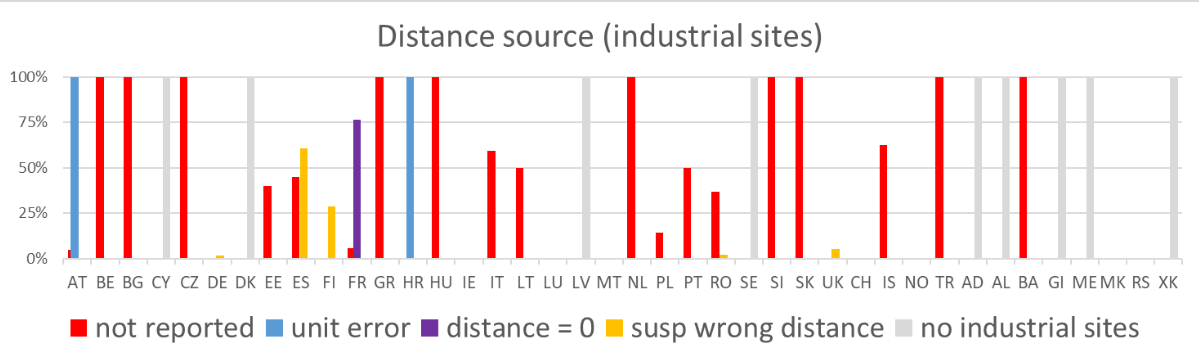


Figure 3.9: Percentage of questionable or missing metadata on distance from industrial source for industrial sampling points. Red bars: percentage of sampling points for which no data is reported. Blue bars: percentage of sampling points for which no unit or a wrong unit was reported. Purple bars: percentage of sampling points for which a value equaling 0 was reported. Orange bars: percentage of sampling points for which a questionable number value was reported. Grey bars: countries which have not classified any industry-related sampling points.

3.2.3.2 Industrial emissions

Decision 2011/850/EC. Annex II.D.(ii) (23) requires specifying what is the main source around the sampling point. For industrial sampling points, and according to the IPR Guidance XML user guide, it is mandatory to declare the emissions from the relevant industrial plant for a representative area of approximately 1 km^2 (with units: t/km^2 year). This element is mandatory for sampling points that are classified as industrial, while for all other sampling point classifications it is voluntary.

There are 31 countries which have classified sampling points as industrial stations but of these, 20 did not report metadata on industrial emissions. France and Norway reported 0 emissions for all sampling points, which is questionable in this context for an industrial station. The metadata information on industrial emissions is largely under reported although it should be pointed out that no reporting unit errors or questionable data beyond the reported "0 emissions" were identified in the present analysis.

In summary, and as depicted in Figure 3.10:

- ➔ For 86% of all industrial sampling points, industrial emissions were not reported. The percentage is 86% in EU27 and 87% in EU28.
- ➔ For 56% of those industrial sampling points with a number reported, the value was equal to 0 (all stations in France and Norway). The percentage is 59% both in EU27 and EU28.

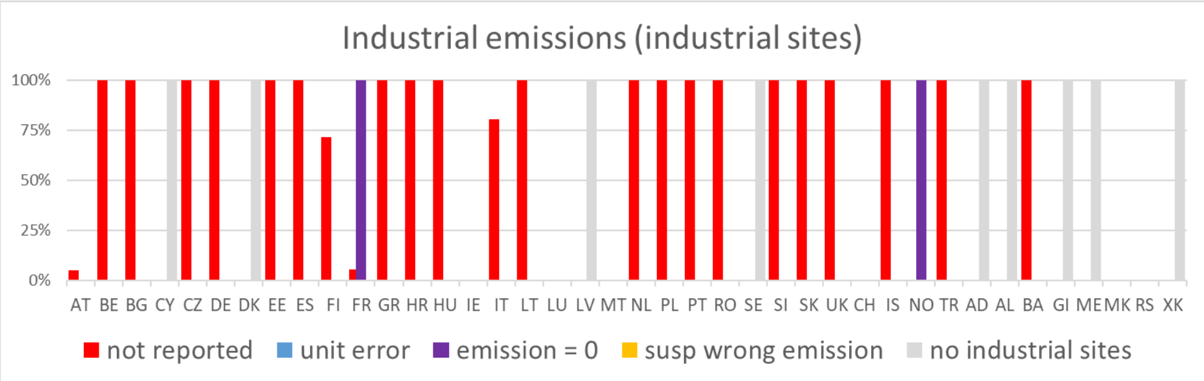


Figure 3.10: Percentage of questionable or missing metadata on industrial emissions for industrial sampling points. Red bars: percentage of sampling points for which no data was reported. Blue bars: percentage of sampling points for which no unit or a wrong unit was reported. Purple bars: percentage of sampling points for which a value equaling 0 was reported. Orange bars: percentage of sampling points for which a questionable value was reported. Grey bars: countries which have not classified any industry-related sampling points.

3.2.4 Mandatory metadata, where available

There are a series of metadata that are mandatory where available. In the present analysis we have focused on seven such metadata parameters. Three of these apply to all sampling points (station information, main sources and spatial extent of representative area), while the other the other four apply to traffic sampling points.

3.2.4.1 Station information

Following Decision 2011/850/EC, Annex II. D.(ii) (27) documentation of station information including maps and photographs should be reported, where available**.** The documentation of this station information can be provided via an external link using a single URL to online maps and photographs of the station which are hosted by the data provider. Only one URL may be provided per station. According AAQD Annex III D, countries must fully document the site-selection procedures for monitoring stations by such means as compass-point photographs of the surrounding area and a detailed map. The sites must be reviewed at regular intervals (maximum 5 years) with updated documentation to ensure that selection criteria remain valid over time. In IPR Guidance XML user guide, page 194, this additional “station information” is categorized, “voluntary, mandatory if available”.

Many countries do not report station information. Exceptions are the Czech Republic, Germany, Spain, Finland, Croatia, Italy, Latvia, Poland, Portugal, Sweden and United Kingdom, Andorra, Gibraltar and Iceland. For Andorra and Gibraltar, the information (maps and compass-point photos) is presented clearly and is easy to find in the provided URLs. Austria has reported this parameter as “update” instead of entering URLs for five stations. It is unclear if this means that they have updated URLs reported earlier or if they are planning to update the information. Iceland reports two different URLs, which both do not seem to contain the required information. Portugal provides URLs which are unclear how to open. Their content could therefore not be assessed. United Kingdom shows a round-view from the stations; however, it is not indicated which compass point directions the individual photos are related to. Some countries provide parts of the required information for some of their stations. It seems that several countries report on a regional basis, such that the demanded information is provided for some regions, but not for other regions.

- ➔ For 63% of all sampling points, no station information was reported. The percentage is 63% in EU27 and 60% in EU28.
- ➔ For at least 5% of all samplings points for which station information was provided, the information given was partial or not accessible. The percentage is 5% in EU27 and 4% in EU28.

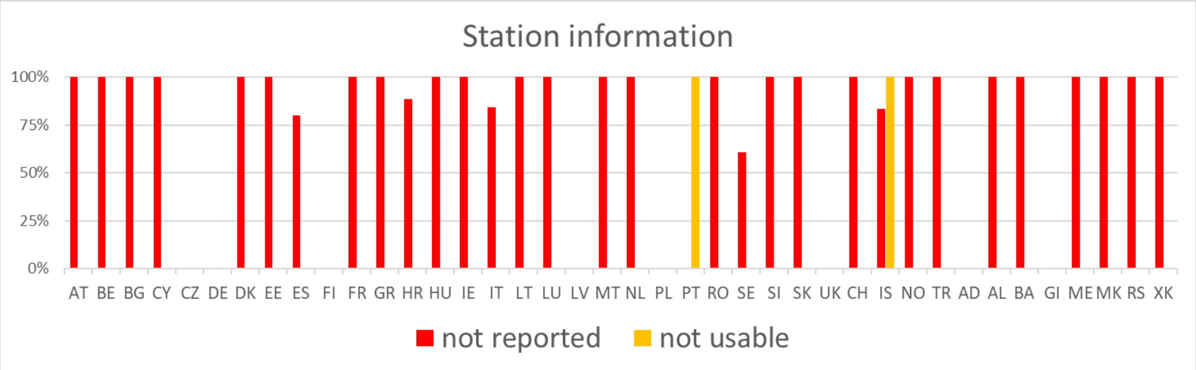


Figure 3.11. Percentage of missing or unusable station information metadata reported. Red bars: percentage of all sampling points for which no data is reported. Orange bars: percentage of sampling points which a URL was reported but partial or not accessible.

3.2.4.2 Main emission sources

Following Decision 2011/850/EC, Annex II.D.(ii) (23), the description of the main emission source for the pollutant specified for the sampling point is mandatory where available. Reporting can be done according to a code list containing 14 source types. It was not feasible at this stage to assess the plausibility of the codes without knowledge of the individual stations. Therefore, the analysis considers data where the data have been reported. It is interesting to note that the main emission sources parameter was reported for 45% of all sampling points, considerably larger than reporting metadata on industrial emissions. It is also noteworthy that the percentage of reported metadata is somewhat larger in EU Member States than for other European countries. The emission metadata information is missing despite the fact that most countries admit such information is available to them (see Chapter 5).

In summary and as depicted in Figure 3.12:

- ➔ For 55% of all sampling points, metadata on "main emission sources" was not reported. The percentage is 50% in EU27 and 52% in EU28.

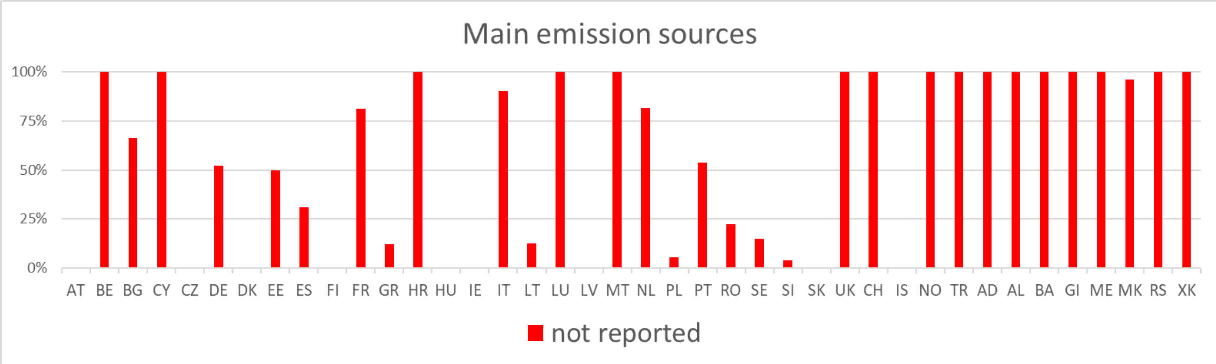


Figure 3.12: Percentage of sampling points for which no data is reported for Main emission sources.

3.2.4.3 Spatial extent of representative area

This metadata parameter on the spatial extent of representative area is essential to the concept of the spatial representativeness of the sampling point. Following Decision 2011/850/EC, Annex II.D. (ii) (16) this parameter is mandatory where available. None of the countries has reported this parameter in the metadata set analyzed here. The information on the spatial extent of representative area is currently not available at EEAs CDR. This is consistent with the general lack of reporting of metadata information related to the characterization of sampling point representativeness and may indicate that the information is actually not available at country level. It is expected that further guidance on how to assess the area of representativeness of sampling points would help to gain access and improve to this information.

➔ For 100% of all sampling points, metadata on the spatial extent of their “representative area” was not reported. The same percentage applies in EU27 and EU28.

3.2.5 Mandatory metadata where available, for traffic sampling points

The four “mandatory where available” metadata parameters considered here for traffic sampling points are: the heavy duty fraction, the traffic speed and for street canyon, the width of the street and the height of the building facades. This metadata information is relevant to characterize the representativeness of traffic sampling points. The last two parameters are particularly relevant to describe the local dispersion situation of traffic sampling points under “street canyon” situations.

3.2.5.1 Heavy duty fraction

Following Decision 2011/850/EC, Annex II.D.(ii) (31) heavy duty fraction is a metadata parameter mandatory to report where available for traffic oriented sampling points. In general, countries that reported values for the heavy duty fraction also reported metadata values for traffic volume. As explained earlier, this information on vehicle fleet composition is an important parameter to characterize air pollution emissions from traffic. However, the number of sampling points where no heavy duty fraction was reported is considerably larger than for traffic volume.

The IPR Guideline requires to report numeric fractional values, e.g. 0.25 equals 25%, however the format used by the countries to report is not always consistent with this requirement. In some cases, the values reported lack a decimal place, in others they seem to be numeric fractional values. Latvia reports values between 578 and 2,350, which may be average daily

traffic counts instead of fractions. The large number of different value formats and units used makes it difficult to interpret the station data regarding their heavy duty fraction.
In all sampling points in Croatia and for a small percentage in the Czech Republic and Sweden, the reported value was equal to 0.

This metadata parameter is more scarcely reported in EU27 than in other countries. Probably, better guidance on how to report the heavy duty fraction and how to use the information to estimate sampling point representativeness may contribute to increase the number of reports on this parameter.

In summary, and as depicted in Figure 3.13:

- ➔ For 68% of all traffic sampling points, no heavy duty fraction was reported. The percentage is 71% in EU27 and 68% in EU28.
- ➔ For 2% of those traffic sampling points with a number value reported, the value was equal to 0 (in Croatia, and for some sampling points in Czech Republic and Sweden). The same percentage applies in EU27 and EU28.
- ➔ For 50% of those traffic sampling points with a number value reported, the value was considered to be in a wrong format. The percentage is 37% in EU27 and 47% in EU28.

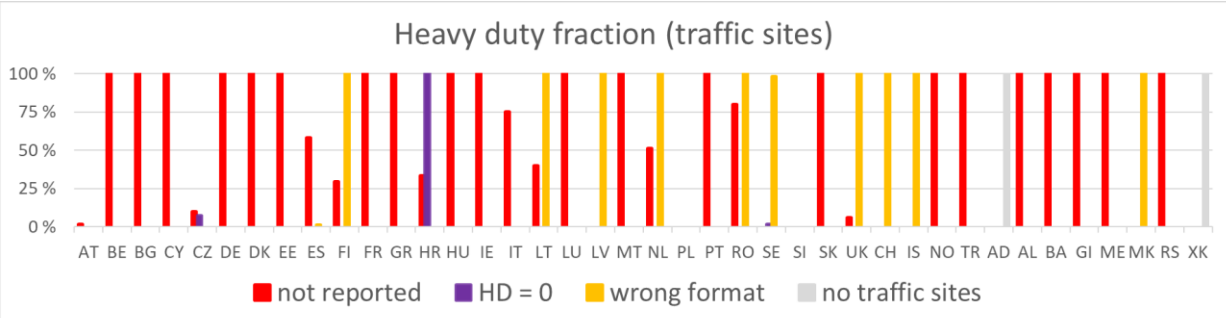


Figure 3.13: Percentage of questionable or missing data on heavy duty vehicle fraction - <u>Red bars:</u> percentage of sampling points for which no data is reported. <u>Purple bars:</u> percentage of sampling points for which a value equaling 0 was reported. <u>Orange bars</u>: percentage of sampling points for which values are considered to be in a wrong format. <u>Grey bars</u>: countries which have not classified any traffic-related sampling points.

3.2.5.2 Traffic speed

The average speed of vehicles in km/h on the adjacent road is a mandatory information where available parameter for traffic stations, following Decision 2011/850/EC, Annex II.D.(ii) (32). However, it does not appear to be clear whether the speed should be assessed by speed measurements or if the maximum allowed speed on the road should be reported. Most countries did not report the average traffic speed. Spain reports 0 for 39% of its sampling points, Croatia reports 0 for all sampling points. An average speed of 0 km/h is not realistic for a traffic site as it would mean that there is no traffic flow in the vicinity of the sampling point. Estonia reports a value of 30 km/h for all its sampling points, Latvia, North Macedonia and Slovenia report a value of 50km/h for all their sampling points. In order to be able to use the data reported for this metadata parameter, there should probably be clear guidance on how to assess the traffic speed.

Figure 3.14 summarizes the questionable or missing data from the reports of countries concerning traffic speed, where:

- ➔ For 66% of all traffic sampling points, the traffic speed was not reported. The percentage is 63% in EU27 and 65% in EU28.
- ➔ For 14% of those traffic sampling points with a number value reported, the value was equal to 0 (some stations in Spain, all stations in Croatia). The same percentage applies in EU27 and EU28.

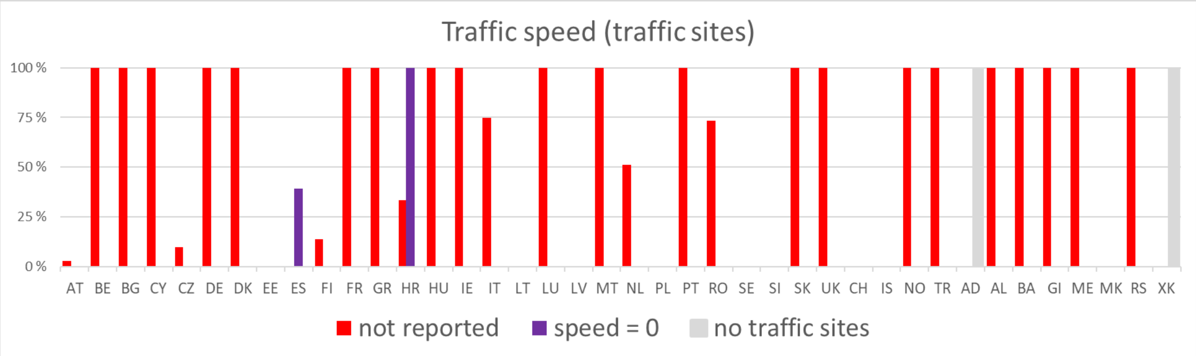


Figure 3.14: Percentage of questionable or missing data on traffic speed. <u>Red bars</u>: percentage of sampling for which no data is reported. <u>Purple bars</u>: percentage of sampling points for which a value equaling 0 was reported. <u>Grey bars</u>: countries which have not classified any traffic-related sampling points.

3.2.5.3 Street canyon – Width of the street

Following Decision 2011/850/EC, Annex II.D.(ii) (33) countries are to report the width of the street for sampling points located in street canyons where the information is available. The width of the street (in meters) at the location of the station in a street canyon, together with the facade height, is an important parameter to assess the local dispersion conditions. According to Table 9 in the IPR Guidance, “street canyon” local dispersion situation apply for traffic sampling points when “continuous/compact buildings are along both sides of the street over more than 100 and the average ratio of height of buildings to width of street is larger than 0,5”.

Only Austria, Czech Republic, Spain, Finland, France, Italy, Latvia, Poland, Slovenia, Sweden, and Switzerland have defined street canyons (within the parameter “dispersion local” – see Section 3.2.6.1). France and Slovenia have defined street canyons but do not report the street width. Sweden reports street canyon width for a large number of location, however the actual positioning of some of these stations may not be compatible with a street canyon dispersion classification due to the structure and height of the neighborhood and these have therefore been considered It seems that the guidance on how to assess the street width could be more detailed in order to be clear whether it is the distance between facades or the distance between the kerbside or the total width of all lanes that should be reported. Still, it is noteworthy that this metadata parameter is indeed well covered by the reporting countries and that there are only occasional lack of reports, as indicated by the low number of red bars in Figure 3.15. The red bars below show the percentage of sampling points at traffic sites in street canyons for which the street width was not reported.

- ➔ Only 11 of 41 countries have categorized some of their traffic stations to be located in street canyons.
- ➔ For 12% of all traffic sampling points located in street canyons, the street width was not reported. The percentage is 9% in EU27 and 9% in EU28.
- ➔ For 1% of all sampling points, the data reported is considered questionable. The same percentage applies for EU27 and EU28.

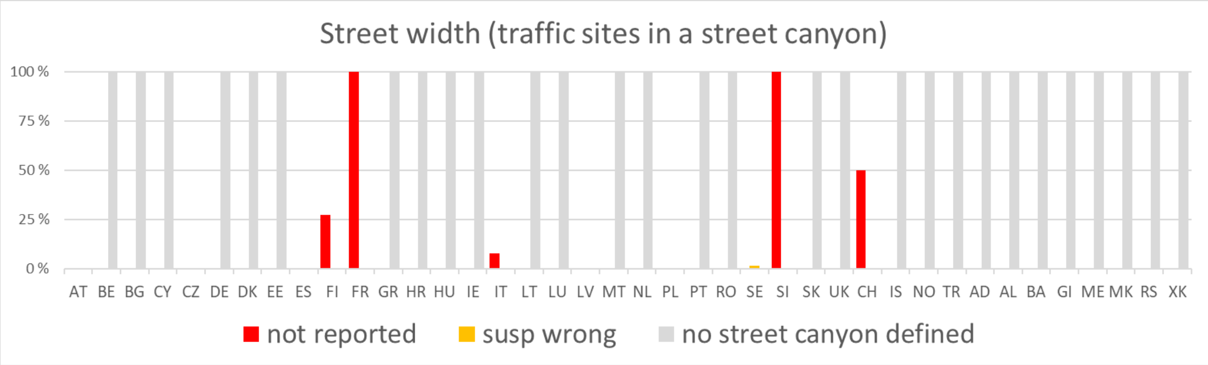


Figure 3.15: Percentage of questionable or missing data on the street width for traffic sites at street canyon. Red bars: percentage of all traffic sampling points in street canyons for which no data is reported for street width. Orange bars: percentage of sampling points for which the reported value was questionable. Grey bars: countries which have not classified any street canyon related traffic sampling points.

3.2.5.4 Street canyon - Height of building facades

Following Decision 2011/850/EC, Annex II.D.(ii) (34) countries are to report the mean height of building facades sampling points located in street canyons where such information is available. The average height of the building facades adjacent to the traffic sampling point (in meters) at the location of the station, together with the street width, is an important parameter to assess the local dispersion conditions. Those countries who reported values regarding the street width also reported facade heights. It is interesting to note that for the 11 countries reporting street canyon sampling points, the reporting of this metadata parameter is well covered, with missing metadata information only for about 8% of the sampling points, even lower (5%) in EU27 and EU28.

Figure 3.16 summarizes the good coverage of metadata reporting on height facades.

- ➔ Only 11 of 41 countries have categorized some of their traffic stations to be located in street canyons
- ➔ For 8% of all traffic sampling points located in street canyons, the height of surroundings facades was not reported. The percentage is 5% in EU27 and 5% in EU28.

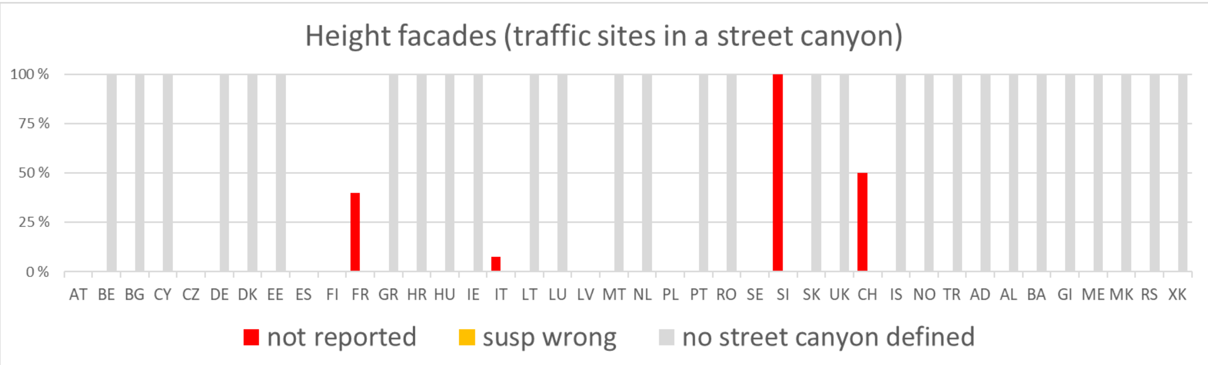


Figure 3.16: Percentage of questionable or missing data for the height of the facades at traffic sites in street canyons. Red bars: percentage of sampling points for which no data is reported. Orange bars: percentage of sampling points for which the reported value was questionable. Grey bars: countries which have not classified any street canyon related traffic sampling points.

3.2.6 Voluntary metadata information

The last three metadata parameters are analyzed here are voluntary information related to dispersion conditions and emissions (dispersion local, dispersion regional and traffic emissions). These are all voluntary parameters to be provided in support to the characterisation of the pollution situation around sampling points and the evaluation of representativeness of sampling points.

The IPR Guidance identifies the legal reference for these metadata as basis for the "Evaluation of representativeness" which is a mandatory requirement where available (2011/850/EC, Annex D. (ii) (17)). Still, the IPR guidance recognizes that "there is as yet no definition of the spatial representativeness of monitoring stations in the AQ legislation" but refers to a series of previous efforts to quantify spatial representativeness and the current going work under the AQUILA and FAIRMODE/frameworks. The present study is in fact intended to contribute to the work under AQUILA and FAIRMODE towards improved guidance on the spatial representativeness of monitoring sampling points.

3.2.6.1 Local Dispersion

The description of local dispersion situations refers to ground level. The local dispersion situation metadata describes the location of the station in relation to nearby buildings and trees which are described using specified codes, distinguishing between "street canyon", "detached buildings", "elevated terrain" and "open terrain". Guidance to assess the local dispersion situation is given in the IPR Guidance and the XML user guide document. The parameter is relevant to better understand the station representativeness of traffic sampling points.

Most countries do not report dispersion conditions, or only for a limited percentage of sampling points. North Macedonia reports an invalid code for 12% of their sampling points. Only Latvia and Slovenia report the local dispersion characteristics for 100% of their traffic-related sampling points.

The reporting of local dispersion classification, where available, is dominated by "detached" conditions, followed by "street canyon", "open terrain" and to a much lesser degree "elevated terrain". The reports on "street canyon" classification maybe biased in some countries, however. Sweden reports for 67 out of 86 traffic sampling points that the local dispersion is comparable to a street canyon, however the actual positioning of some of these stations may not be compatible with a street canyon dispersion classification due to the structure and height of the neighborhood in question. Poland reports 19 of 35 traffic sampling points to be in a street canyon, which again seems to be an over-representation of street canyon sampling points. The countries reporting sampling points with street canyon classification, also report additional conditions such as the street width and the height of the facades, thus providing better description of the traffic sampling point characterization as shown in sections 3.2.6.3 and 3.2.6.4.

For only a very low percentage of sampling points are the dispersion conditions reported as "elevated". Only Finland and France report that some of their stations are located in elevated terrain, which is defined as summit, slope or saddle. Monitoring stations located on towers or high buildings are not considered as "elevated terrain". The definition of "elevated terrain" is somewhat unclear given the requirement for a traffic station to be representative over a street segment of more than 100 m length. Further clarification on what is a meaningful "elevated" sampling point would be desirable.

The percentage of missing metadata information for "dispersion local" information at traffic sites is high (74%), as depicted in Figures 3.17 and 3.18. As seen from the lack of data in

Figure 3.18, further guidance on the classification of dispersion local conditions would be useful. It is expected that such further guidance on the four different local dispersion conditions could help more countries to report this metadata information and make use of it for better establishing the sampling point representativeness area.

In summary:

- ➔ For 74% of all traffic sampling points, no code was reported for “dispersion local”. The percentage is 72% for EU27 and 74% for EU28.
- ➔ Only 2% of all traffic sampling points, the dispersion condition is reported as “elevated” which is an indication of the need for further clarification as to what “elevated” dispersion conditions refers to. The need for further guidance, however, applies to all four classes of the local dispersion classification.

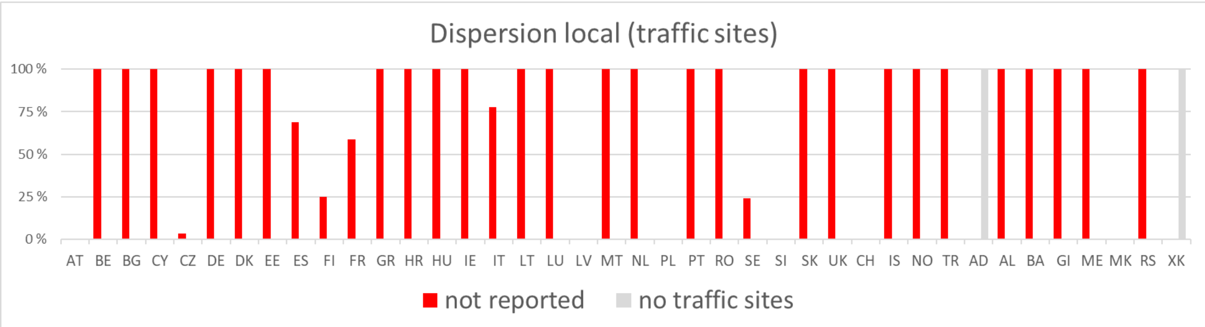


Figure 3.17: Percentage of missing data for “dispersion local” information at traffic sites. <u>Red bars</u>: percentage of sampling points for which no data. <u>Grey bars:</u> countries which have not classified any traffic-related sampling points.

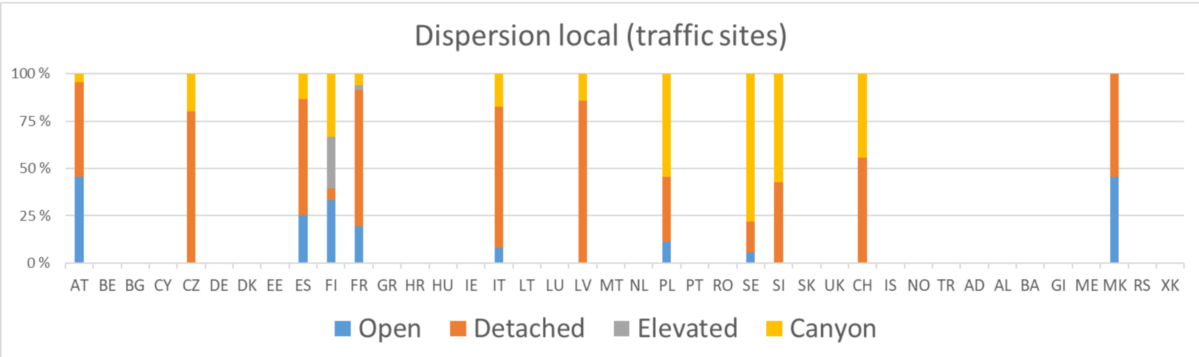


Figure 3.18: Classification of percentage of traffic sampling points with reported "open" (blue bars), "detached" (orange bars), "elevated" (grey bars) and "canyon" (yellow bars) dispersion conditions. Red bars show the percentage of sampling points for which no data was reported. Andorra and Kosovo have not defined any traffic-related sites.

3.2.6.2 Regional Dispersion

Following the IPR Guidance, regional dispersion is classified according to a code list characterizing the surrounding terrain on a larger scale, e.g. alpine, basin, basin-hilly, basin-mountainous, coast-mountainous, coast-plane, hilly, mountainous, mountain-slope, plane, valley-hilly, valley-mountainous.

Regional dispersion was reported by a few countries only. Latvia, Poland and Switzerland reported this parameter for 100% of their sampling points. Another 8 countries (Austria, Bulgaria, Czech Republic, Spain, France, Italy, Sweden and North Macedonia) reported the regional dispersion for some of their sampling points. None of the reported metadata was invalid or undefined,

In summary, as depicted in Figure 3.19

➔ For 74% of all sampling points, regional dispersion was not reported. The percentage is 72% in EU27 and 73% in EU28.

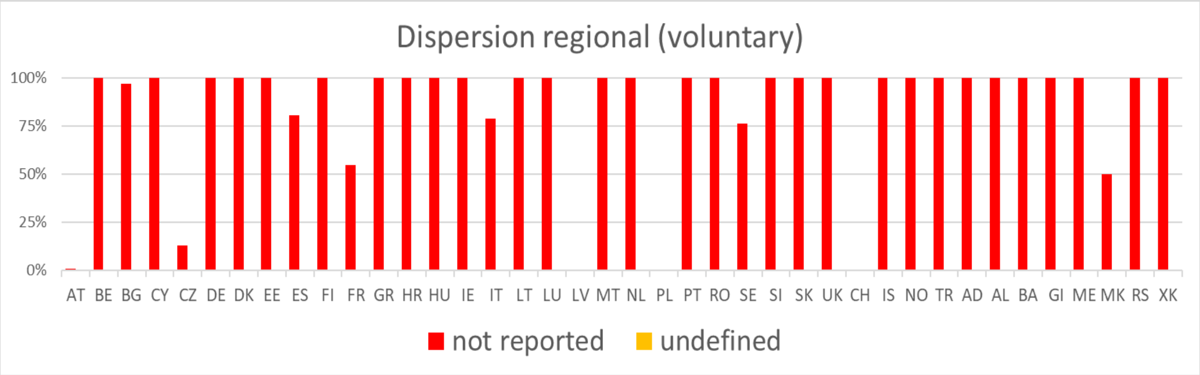


Figure 3.19: Percentage of missing metadata on regional dispersion Red bars: percentage of sampling points) for which no data code was reported. Orange bars: percentage of sampling points for which the data code is invalid.

3.2.6.3 Traffic emissions

Regarding traffic emissions, number values were only reported for a few sampling points by Austria, France, Croatia and Italy. However, all the values reported by Croatia equal 0. For all sampling points with a traffic emission value reported, the reported unit was correct, t/km year.

In summary, as depicted in Figure 3.20:

➔ For 97% of all sampling points, traffic emissions were not reported. The percentage is 96% in EU27 and 96% in EU28.

➔ For 6% of those sampling points with a number reported, the value was equal to 0 (some stations in Austria and Italy, all stations in Croatia). The same percentage applies in EU27 and EU28.

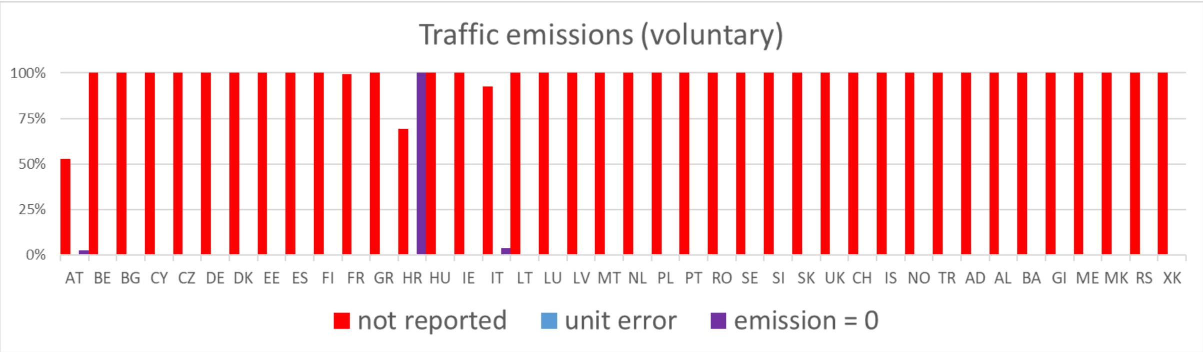


Figure 3.20: Percentage of questionable or missing traffic emissions reported metadata. Red bars: percentage of sampling points for which no data is reported. Blue bars: percentage of sampling points for which no unit or a wrong unit was reported. Purple bars: percentage of sampling points for which a value equaling 0 was reported.

3.3 Conclusions on metadata information

The metadata information on sampling point location and classification that is to be officially reported by countries in dataflow D is intended to give an insight into the location of the sampling point and to facilitate the interpretation of the measurement data. This is very valuable for all users of the actual measurement data independent of the type of application as part of the quality checking process. The metadata information analyzed here serves to inform how countries apply the macroscale and microscale criteria given in AAQD Annex III and provides additional information on needs for further guidance to evaluate the representativeness of the sampling points.

Although the metadata information in dataflow D is mandatory for background, traffic and industrial sampling points, the evaluation carried out here shows that there are important gaps in the reported metadata. The currently reported metadata is far from complete and there are identified inconsistencies in the reporting. In some cases, the reported metadata shows that the sampling points do not comply with the siting criteria established in AAQD Annex III but in most cases, the metadata shows inconsistencies in the choice of units and with respect to the nomenclature for non-available metadata. The reporting of voluntary data related to dispersion conditions and traffic emissions is even more sparse. None of the countries reported metadata on the extent of the area of representativeness.

This section highlights that better guidance on how to determine and use the following parameters could help more countries to report this metadata information and make use of it for better establishing the sampling point representativeness area:

- The four classes for local dispersion conditions;
- The spatial extent of representative area;
- Traffic speed;
- The heavy duty fraction;
- The street width in street canyons.

There are likely different reasons why the reporting of these metadata is incomplete. One reason may be that the actual metadata is not available to the countries, due to either lack of data or lack of guidance on how to compile these data, especially in relation to the representativeness of sampling points. Another possible reason is that the metadata has not been regularly checked and therefore it has received less attention from national experts. To our knowledge this is the first time that a systematic review of the reported metadata has been carried out. Without additional review of the reported metadata, gaps and inconsistencies remain unnoticed.

In order to ensure that the reported metadata on sampling point location and classification become more complete and therefore more useful for the assessment of spatial representativeness, it would be possible to complement the extraction of country reports to EEAs CDR with automated completeness and feasibility check of the metadata such as the one carried out here. The lack of metadata reports of the chosen parameters indicate a need for better guidance on how the individual metadata parameters are to be assessed and reported. In particular, the importance of stating units to the reported values also needs to be pointed out and a common nomenclature for non-available data needs to be adopted.

4 Clustering analysis to evaluate sampling point classification

In this chapter, the reported station classifications are compared to those determined from an independent cluster analysis on air quality measured concentrations. This independent assessment using the temporal variation of the reported air quality measurement data, will provide an indication as to whether traffic-oriented sampling points and industrial sampling points are indeed measuring as intended according to their classification the influence of traffic or industrial sources. The methodology also allows us to identify outliers and to provide some insight into the reasons for differences between sampling point classifications where they occur.

The hierarchical clustering methodology applied here provides a screening tool to analyse the level of similarity or dissimilarity of the reported air concentration data. The clustering analysis is applied both considering all the industrial and traffic stations together (station classification is disregarded) and considering the individual classification types separately. The analysis of the results provides an indication as to whether the current classification of the stations is sufficient to infer the impact of specific near-by emissions, which are the basis of the station classification. It allows the identification of sampling points with specific behaviour that differ from the rest of the sampling points. These different behaviours may be well justified in terms of specific conditions around the sampling points, such as influence from specific sources or specific meteorological dispersion conditions. In this sense, the identification of “outliers” in the clustering analysis is not an excluding exercise but the recognition of a different behaviour that needs to be further investigated.

The current clustering analysis applies for the whole of Europe, therefore, the level of similarity or dissimilarity flagged here applies at a European level. The same analysis can be done on a country basis to assess in more detail the station data behaviour and the grouping of the stations across a single country. The analysis could be used to assess how individual countries set up their monitoring network, including if the classification of the stations is fit for the purpose it was originally intended.

4.1 Methodology

The classification of traffic and industrial stations across Europe was evaluated based on the methodology proposed by Soares et al. (2018) and references therein. The methodology demonstrates the potential use of hierarchical clustering in the analysis of monitoring network data as a means to determine the level of similarity between station data. Hierarchical clustering assumes that the data contains a level of (dis)similarity and groups the station records based on characteristics of the data. Here measurement data observed at traffic and industrial measurement sampling points are analyzed to understand the differences and similarities between stations according to their classification.

The methodology helps us establish the level of similarity (or dissimilarity) of the different air quality measurements across Europe. It requires an air quality dataset covering a long-term time series (preferably, at least one year of measured data). The dissimilarity metrics used here are:

(1) 1-R, where R is the Pearson linear correlation coefficient (Solazzo and Gamarini, 2015),
(2) the Euclidean distance, EuD (Soares et al, 2018), and
(3) multiplication of metric (1) and (2).

The Euclidian distance (EuD) is calculated as indicated in Equation (1)

$$EuD_{x,y} = \sqrt{\sum_{j=1}^{j}(x_j - y_j)^2} \quad (1)$$

where x and y are two different time series, j is j^{th} record of the time series.

The metric based on correlation, (1-R), assesses dissimilarities associated with the changes in the temporal variations in concentration, while the metric based on the Euclidian distance (EuD) assesses dissimilarities on the basis of magnitude of the concentration over the time period of the analysis. The multiplication of these two metrics (1-R) x EuD allows assessing correlations in terms of both time variation and pollution levels.

The clustering analysis is an iterative process. The first step of the analysis consists of producing a dissimilarity matrix cross-comparing all sampling points. This matrix results from calculating the metrics based on the monitoring data available, where every single record of the time series reported by the paired sampling points will be compared. The next step is to identify the pair in the matrix with the highest similarity level and re-evaluate the dissimilarity matrix for all the remaining sampling points. The re-evaluation is done by averaging the latest cluster and the remaining time series' dissimilarity. The averaging method is the general averaging method of Næs et al., (2010). The process is iterated until all points have been paired. This iterative pairing process results in a "cluster" of similarity that includes the different sampling points. Initially, every single time series/sampling point is considered a single cluster. As the hierarchical clustering processes go on, the number of clusters is reduced. The process is completed when the two last clusters have joined so that all sampling points have effectively been clustered.

The hierarchical clustering methodology has been applied to hourly concentration data for NO_2 and PM_{10} for 2017. The annual time series of air quality measurement data are reported by the EEA member countries, including all EU Member States, as well as EEA cooperating countries, and can be accessed via the EEA database[15]. Only time series for sampling points with a minimum temporal data coverage of 75% were included in the analysis. Note that hierarchical clustering requires no gaps in the time series, thus the time series containing data gaps of more than 168 consecutive hours were removed from the analysis, as recommended by Solazzo and Galmarini (2015). Data gaps of 1 to 6 hours duration were also replaced with the linear interpolation between the nearest valid data on either side of the gap and, for data gaps of longer duration, the annual average of the non-gap data was used, again as recommended by Solazzo and Galmarini (2015).

Table 4.1 shows the number of time series available before and after implementing the gap screening procedure. From the total number of time series available after the quality checks were applied, only 675 and 415 traffic sampling points measuring NO_2 and PM_{10}, respectively, had the quality needed for the clustering analysis. By the same token, only 356 and 222 industrial sampling points were available for NO_2 and PM_{10}, respectively.

15 http://aqportal.discomap.eea.europa.eu/products/data-download/, accessed for 2017 as accessed on 9 April 2019

Table 4.1: Sampling points evaluated for clustering analysis suitability as reported by EEA countries for 2017.

Pollutant	Classification type	Sampling points with 75% temporal coverage	Sampling points with 75% temporal coverage after QA/QC	Sampling points after those with > 168 hrs gap are removed
NO_2	Industrial	451	450	356
	Traffic-oriented	868	860	675
PM_{10}	Industrial	287	286	222
	Traffic-oriented	519	514	415

The hierarchical clustering methodology was applied to all traffic and industrial sampling points passing the quality checks, for both types together and for each type individually. The outcome of the methodology indicates the (dis)similarity of stations, at a certain level of the chosen metric, across the whole data set. The higher the dissimilarity level between sampling points or cluster of sampling points, the more dissimilar those pairs are. The dissimilarity level value depends on the metric chosen. The values calculated based on the correlation metric can vary from 0 to 2, with values over 1 meaning that the data samples anti-correlate. The metric values for the other two metrics are varying from zero to a value that strongly depends on the average concentration levels. However, the dissimilarity values are also dependent on the number of sampling points and the time scale of the measurements. The outcome of the analysis also strongly depends on the quality of the reported data, so that random errors in the observations can potentially change the results (Soares et al., 2018).

Dendrograms are 2D representations of the clustering process showing the pattern of linkages between the data series while clustering occurs, as well as their level of dissimilarity. This graphical representation is the easiest way to visualize the results. However, for larger datasets, the dendrograms can mainly be used for identifying large groups of stations, or single stations that cluster on their own very early on during the clustering process, as the dendrogram gets very busy. After plotting the dendrograms, it is possible to identify the dissimilarity level at which sampling points cluster together, distinguishing groups of sampling points. The clusters can then be displayed geographically, colour-coded according to the cluster the stations have been allocated to, and shape-coded according to the station classification. This way of visualization serves as a quick screening tool to envisage how the different sampling points cluster together at different levels of dissimilarity.

The choice of threshold for the allowed levels of dissimilarity is essential to the conclusions driven in the current analysis. The level of dissimilarity chosen for this application is the level where the dendrograms start branching into smaller clusters. The process is carefully balanced choosing the branching that shows a clear division between different groups of stations. On one hand, the highest level of branching shows the cluster of stations that have the highest similarity, within a single cluster, at that level of similarity. On the other, the clustering shows which stations differ the most from their pairs, either for specific meteorological and/or predominant emission sources or possible issues with the data available for those stations.

4.2 Results from hierarchical clustering

4.2.1 NO_2 concentrations

The results for all NO_2 sampling points (both traffic and industrial sampling points) show that the clustering across Europe does not follow national borders but responds to different meteorological and source regimes.

The NO_2 temporal correlation (1-R) analysis for both station types shows that clusters form beyond the borders of different countries. Figure 4.1 shows how the clusters are aggregated according to areas: Iberian Peninsula, central Europe, northern Europe, and Balkans, mainly following the different meteorological regimes. The results also identify two countries where sampling points belonging to the same country are also mostly clustering together, namely Italy and Turkey. This may be an indication of similarities that respond to other reasons than those expected from meteorological regimes and can indicate that the sampling points are possible outliers. Similarly, in Northern Europe, there are four sampling points (in yellow) located in Estonia, Finland, Iceland and Latvia that cluster together at very high levels of dissimilarity with the rest of the Nordic sampling points. The reason for these different behaviour in the four samplings points requires further investigation but is beyond the scope of this study.

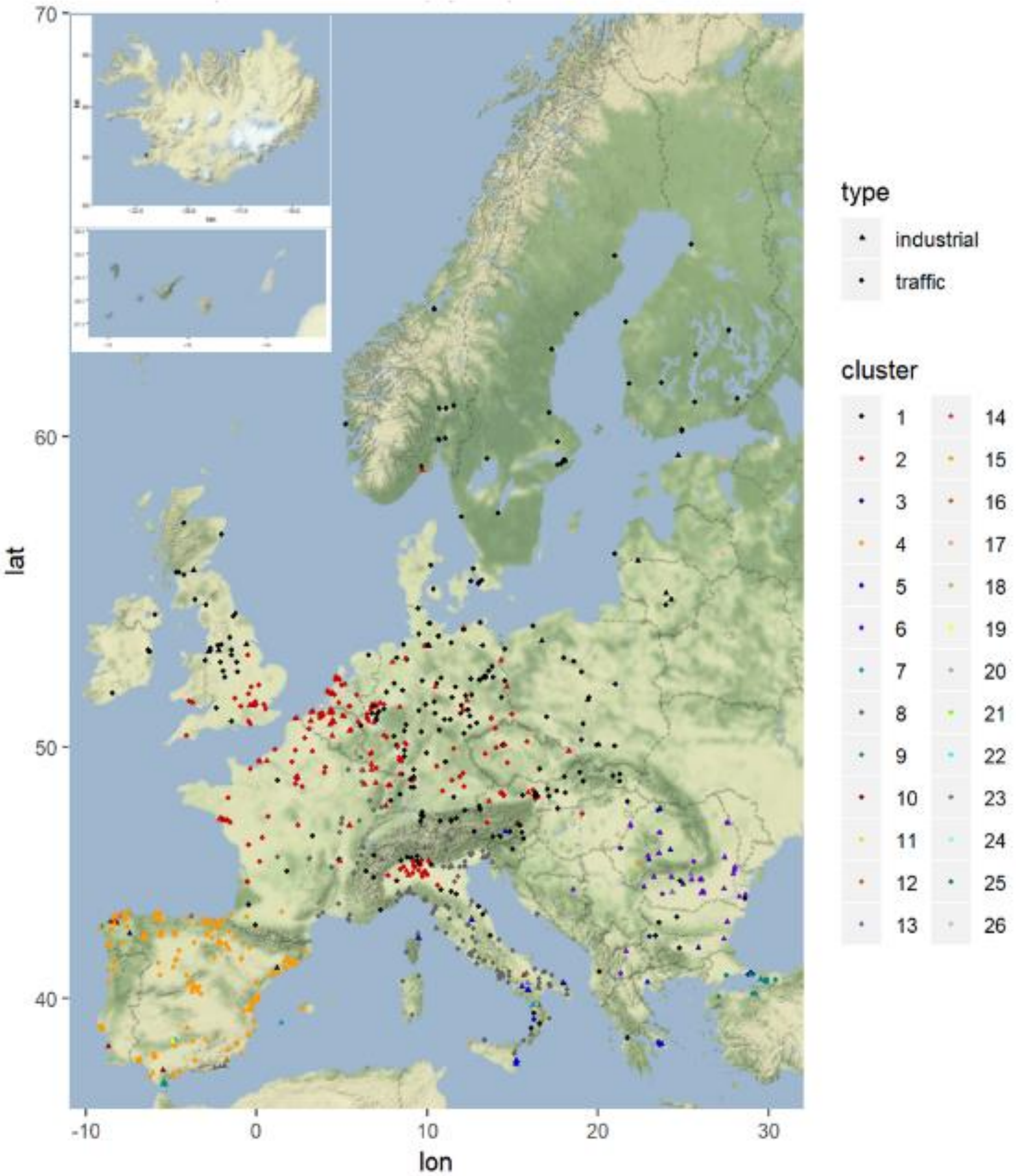


Figure 4.1. Clustering of NO_2 sampling points, color-coded according to the 1-R metric for both industrial and traffic-oriented sites. The clustering shows mainly meteorological regimes.

The analysis based on the Euclidean distance (EuD) metric for NO_2 is provided in Figure 4.2. and shows clustering in terms of emission magnitude. Dissimilarities are due the magnitude of emission sources due to the fact that both industrial and traffic sources are clustered at the same time. Countries such as Spain, Italy, Turkey and Romania present the largest dissimilarity within their country borders. Note such dissimilarities are sometimes due to the difference between traffic and industrial sampling points in the same country while the choice of the threshold shows differences also across Europe,

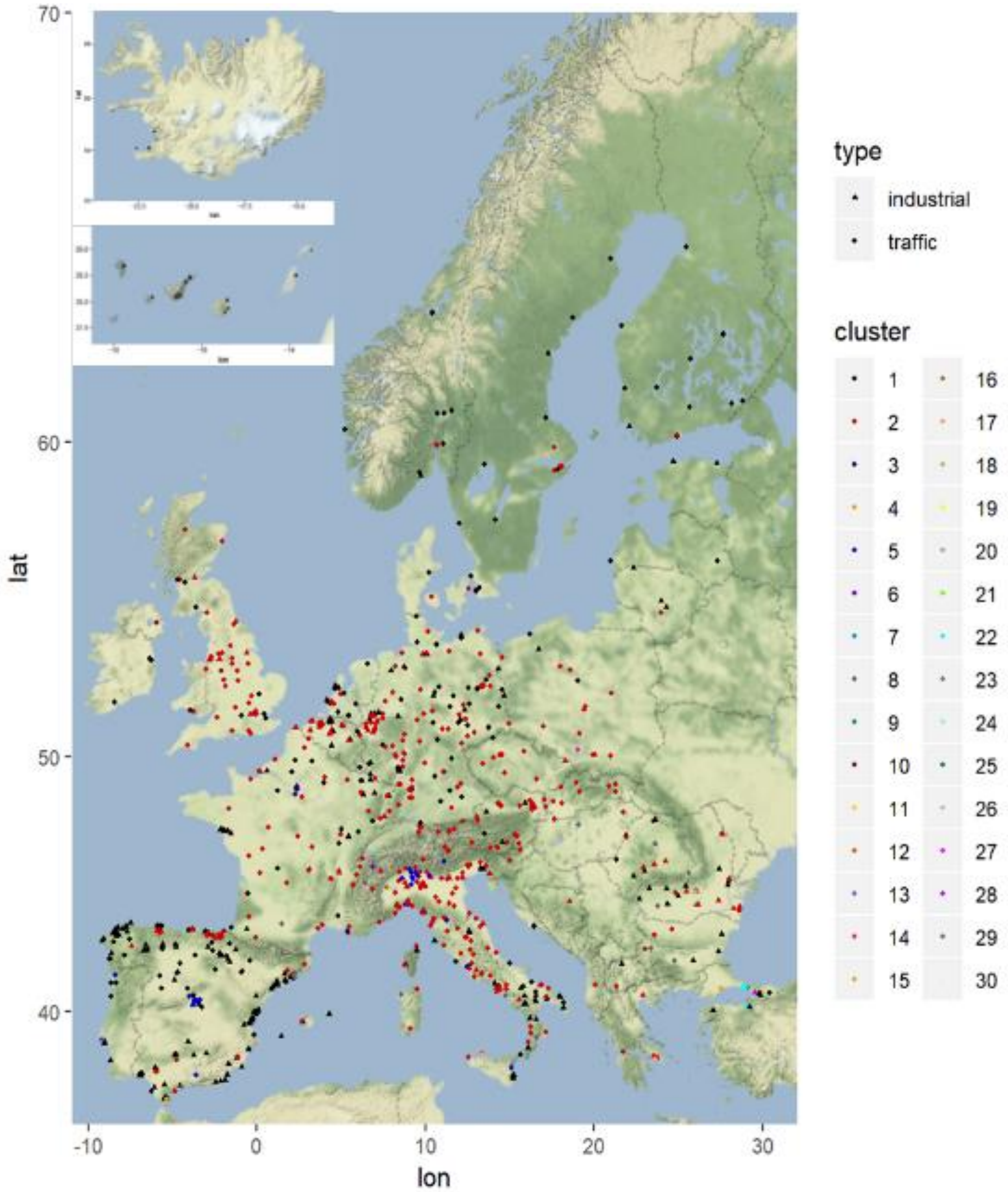


Figure 4.2. Clustering of NO_2 sampling points, color-coded according to the EuD metric for both industrial and traffic-oriented sites. The clustering shows mainly differences related to the magnitude of emission sources.

Figure 4.3 shows the results for NO_2 concentrations for the combined metric (1-R) x EuD that shows the combined effect of meteorological and emission variations. Sampling points in Spain (Gran Canarias) and one in Iceland show large dissimilarity when compared with other sampling points across EEA members when applying the combination of the two metrics. All these sampling points are characterized as industrial, except one. Generally, as levels of dissimilarity decrease, sampling points cluster in accordance with areas associated with meteorological conditions and predominant sources. Figure 4.3. shows different air quality regimes in the Po Valley, Northern Italy, the Netherlands, Belgium and France as well as the western part of the Iberian Peninsula. Again, for this combined metric, the sampling points within Romania are very dissimilar, an indication for the need for further investigation.

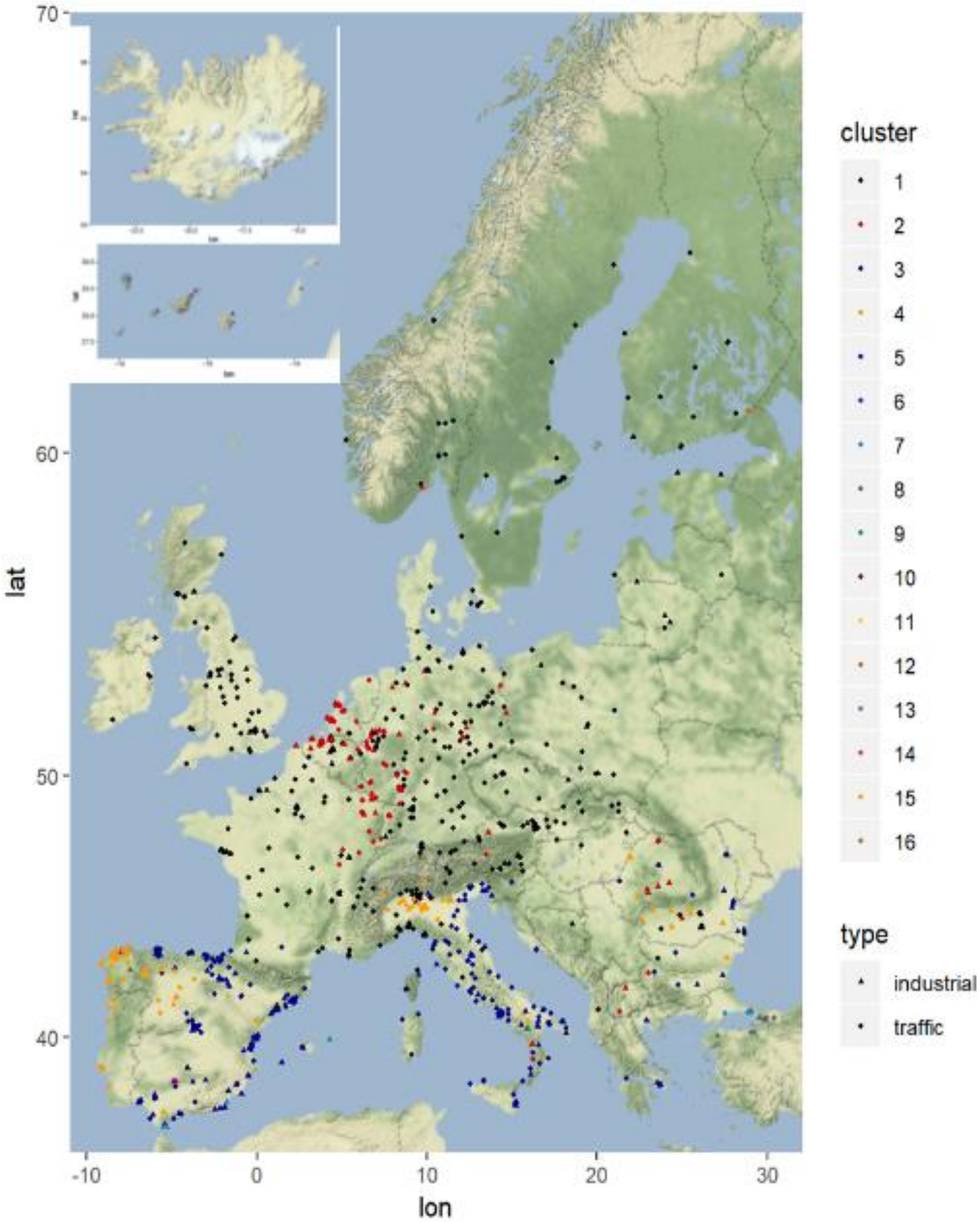


Figure 4.3. Clustering of NO_2 sampling points in 2017, color-coded according to the (1-R) x EuD metric for both industrial and traffic-oriented sites. The clustering shows mainly differences related to the magnitude of emission sources and different meteorological regimes.

4.2.1.1 Traffic sites measuring NO_2 concentrations

The hierarchical clustering analysis has also been carried out separately for industrial and for traffic sampling points. The main purpose of this separate clustering analysis is to assess the accuracy of the application of the siting criteria and to determine if sampling points are indeed measuring traffic sources or industrial sources as intended.

Figure 4.4. shows the results of the correlation metric for traffic-oriented samplings points of NO_2. The NO_2 traffic-oriented sampling points show a high level of similarity across Europe, with only two clusters identified with the current choice of threshold. This implies that most sampling points have a similar behaviour, with similar temporal correlations, and is an indication that the sampling points are indeed measuring traffic sources.

The clustering analysis has also allowed us to identify “outliers” specially in Italy, Romania and in some sampling points in South Eastern Mediterranean countries. These are given in Table A.1 in the Appendix. The reasons for these dissimilarities are difficult to explain without further investigation, especially when most, if not all, of the sampling points reported by a country, like is the case for Italy, Turkey and Romania, are clustering under the same cluster. Moreover, the clustering analysis for the other two metrics show similar conclusions for Italy, Turkey and Romania. These results might point to sampling-points that do not measure the typical values for NO_2-related traffic emissions or the temporal variation expected for a traffic station or that there are problems related with the time span of the reported concentrations. In any case, these results call for further investigation in the identified countries.

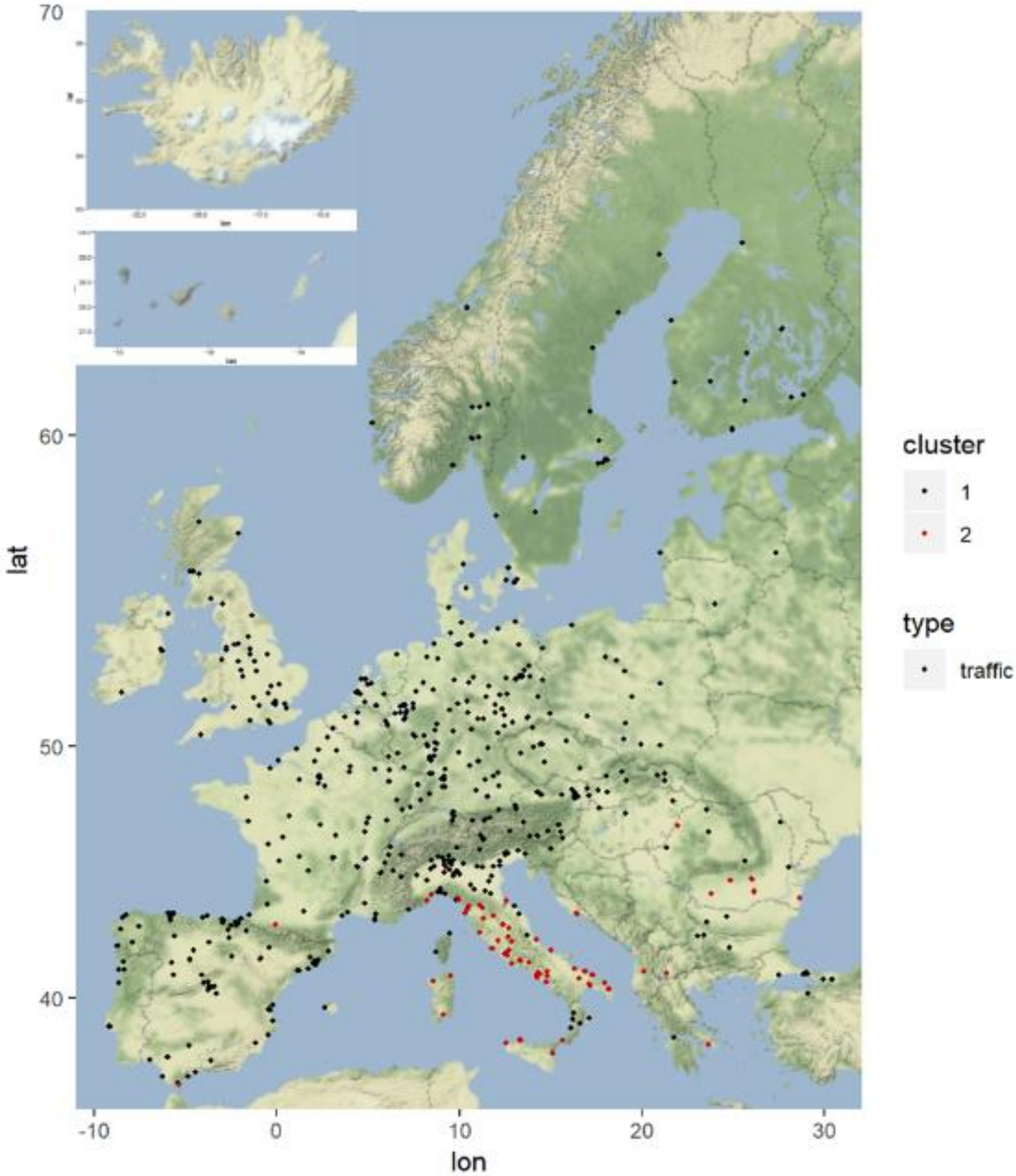


Figure 4.4. Clustering of NO_2 traffic-oriented sampling points for 1-R metric, In red, traffic oriented sites with a pronounced different temporal variation from the rest of European traffic-oriented sites.

4.2.1.2 Industrial sites measuring NO_2 concentrations

The results for NO_2 industrial sampling points provide very similar conclusions to those from traffic-oriented sampling points. For NO_2 industrial sampling points there is also a high level of similarity across Europe, now clustered in 3 different groups. This is an indication that NO_2 industrial sampling points are indeed measuring the influence of industrial emissions.

Figure 4.5 shows the results for the temporal correlation metric for industrial sampling points. The analysis indicates that two groups of stations stand out from the remaining ones: a cluster with Finnish, Estonian, and Icelandic sampling points (cluster 3), and another with several Spanish sampling point and a single one located in Portugal, Iceland, Turkey and Italy (cluster 2). The sampling points in Turkey, Iceland and Spain seem to be the most distinct sampling points when compared to the whole industrial dataset. Table A.2 in Appendix 1 gives more details on the sampling points identification. Results for the analysis based on the EuD metric (not shown) indicate the industrial sampling points located in Turkey and Spain as the most dissimilar in relation to the whole dataset.

Figure 4.6 shows the results based on the industrial sampling points for the combined (1-R)x EuD metric. In this case, the figure shows further dissimilarity across Italy and Romania. Although the identified “outliers” represent a small percentage of the total NO_2 sampling points, the reasons for the different behaviour in these sampling points needs to be further investigated. It could be due to a number of reasons including shifts in the time span of the reporting.

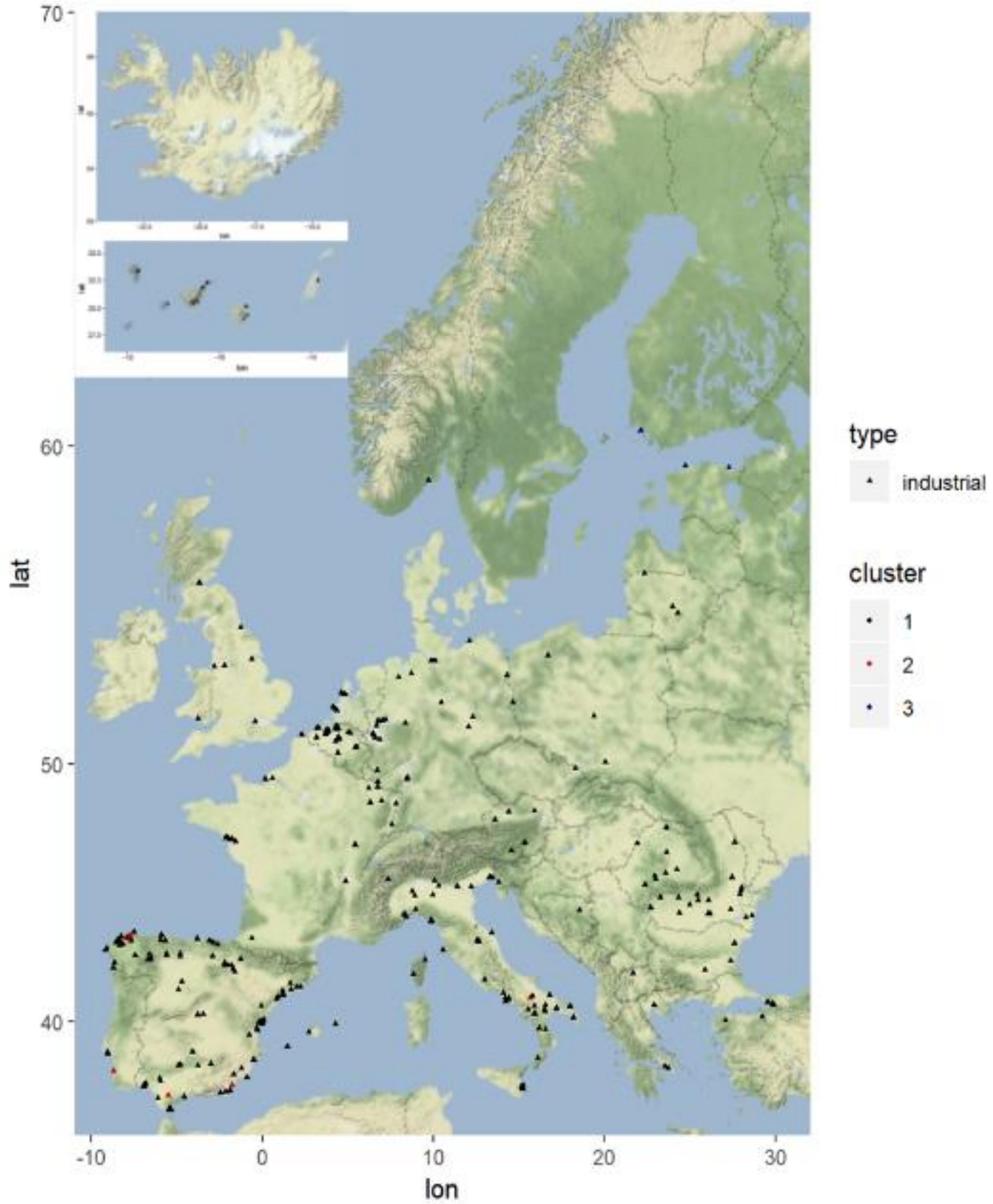


Figure 4.5. Clustering for NO_2 industrial sampling points, color-coded according to the 1-R metric. In red and blue, industrial sites in Northern Europe with a pronounced different temporal variation from the rest of European industrial sites.

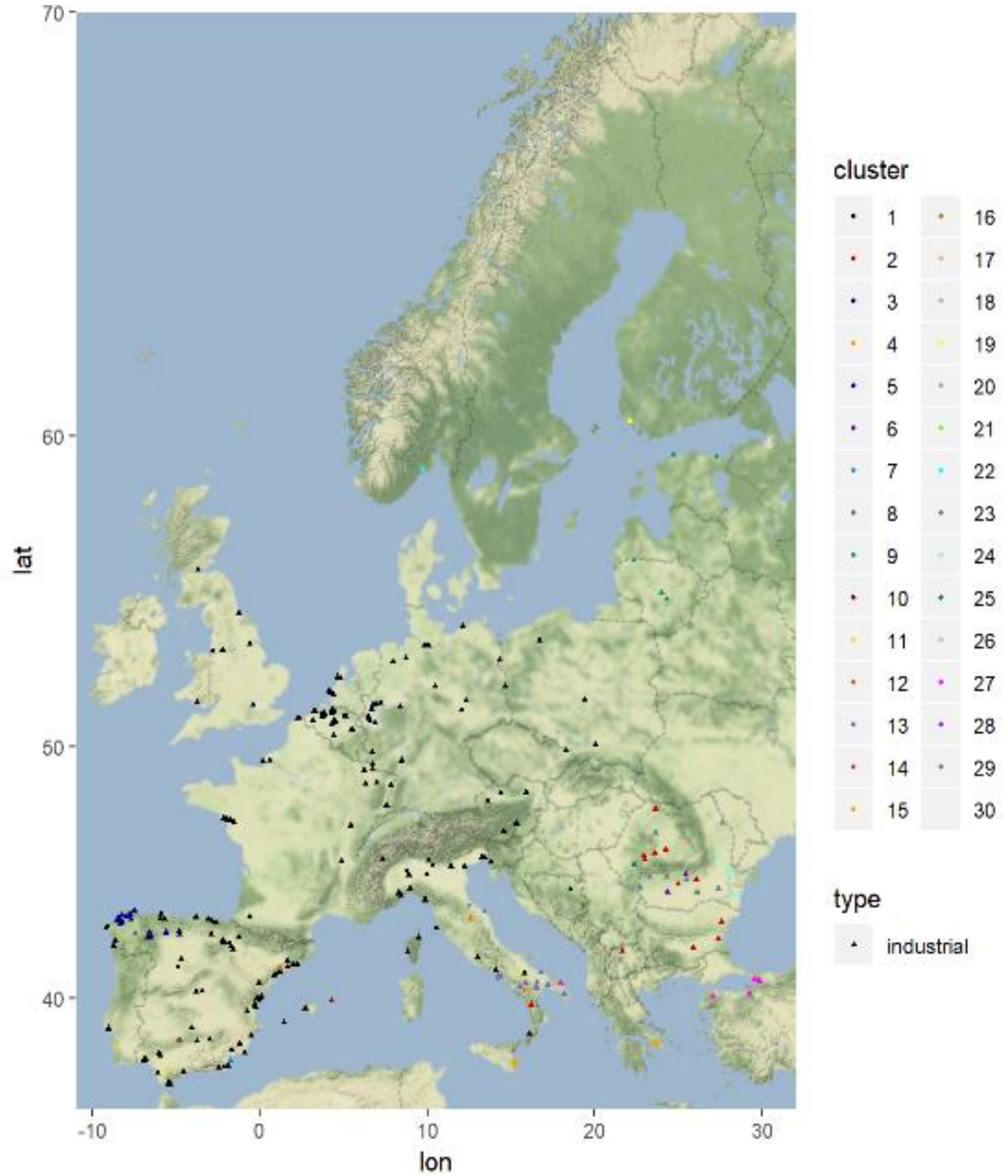


Figure 4.6. Clustering for Spatial distribution for NO_2 industrial sampling points, color-coded according to the (1-R) x EuD metric. In red and orange, industrial sites with a pronounced different magnitude and temporal variation from the rest of European industrial sites.

4.2.2 PM_{10} concentrations

The results of the clustering analysis for PM_{10} traffic and industrial sampling points show higher level of dissimilarity than for NO_2. This is due to the fact that there are a larger number of emission sectors affecting PM_{10} concentrations than NO_2. NO_2 concentrations are mainly from traffic or industrial emissions, while PM_{10} concentrations are also influenced by other sectors such as residential and agriculture.

Figure 4.7 shows the PM_{10} analysis based on the temporal correlation (1-R) metric. Overall, sampling points located in central Europe are the least dissimilar sampling points. Sampling points within countries like Iceland and Turkey, and areas like Gran Canarias, are clustering together. The remaining clusters are divided according to larger areas: Iberian Peninsula, central and northern Europe, and finally clustering according to smaller areas, e.g., three different areas cluster within Iberian Peninsula. This mostly reflects predominate sources and meteorological conditions. In 2017, there was an episode of strong forest fires in Portugal that affected the PM_{10} concentrations and its influence can be seen as a specific cluster over both Spain and Portugal.

Results for the other metrics (not shown) indicate that the higher level of dissimilarity is for sampling points in North Macedonia and Turkey indicating that is it the magnitude of the measured concentrations that drives the differences in these two countries.

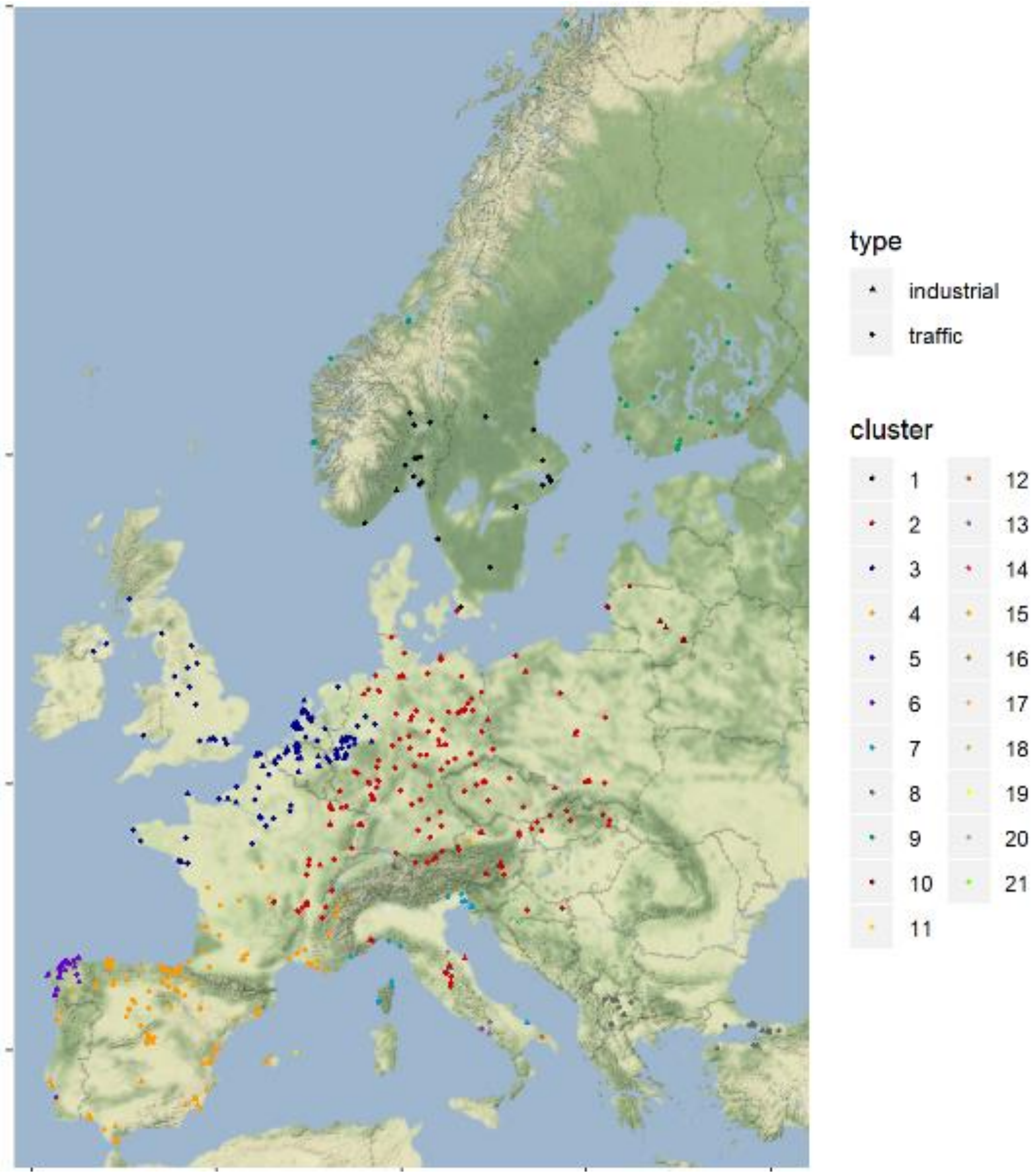


Figure 4.7. Clustering for PM_{10} sampling points, color-coded according to the (1-R) x EuD metric for both industrial and traffic-oriented sites. The clustering shows mainly differences related to the magnitude of different emission sources and different meteorological regimes.

4.2.2.1 Traffic sampling points measuring PM_{10} concentrations

Because of the mixed sources affecting PM_{10} concentrations, it is difficult to classify PM_{10} sampling points in simple terms as industrial or traffic-oriented. The clustering techniques thus do not distinguish clearly between industrial and traffic, as already pointed out by Joly and Peuch (2012).

Figure 4.8 shows the clustering analysis results based on traffic sampling points following the correlation (1-R) metric. The figure clearly shows the different PM_{10} regimes in Northern Europe in contrast to central and southern Europe. Intrusions by Saharan dust could explain part of the dissimilarities in Southern Europe, while sea salt contributions could affect the clustering in coastal areas. Here it is both the meteorological variations and the changes in the emission mixes that drives the clustering. When both temporal and emission magnitude variations are considered, the clustering analysis identifies two main countries that are different from all remaining data. These are North Macedonia and in Turkey, as indicated in Figure 4.9 that shows the Euclidian distance (EuD) metric results for traffic-oriented sampling points measuring PM_{10}.Traffic-oriented sampling points in the Czech Republic, Poland, Slovakia and Sweden are also identified as “outliers” as detailed in Table A.3. in Appendix 1. Further investigation of the reasons behind these dissimilarities is beyond the purpose of this study.

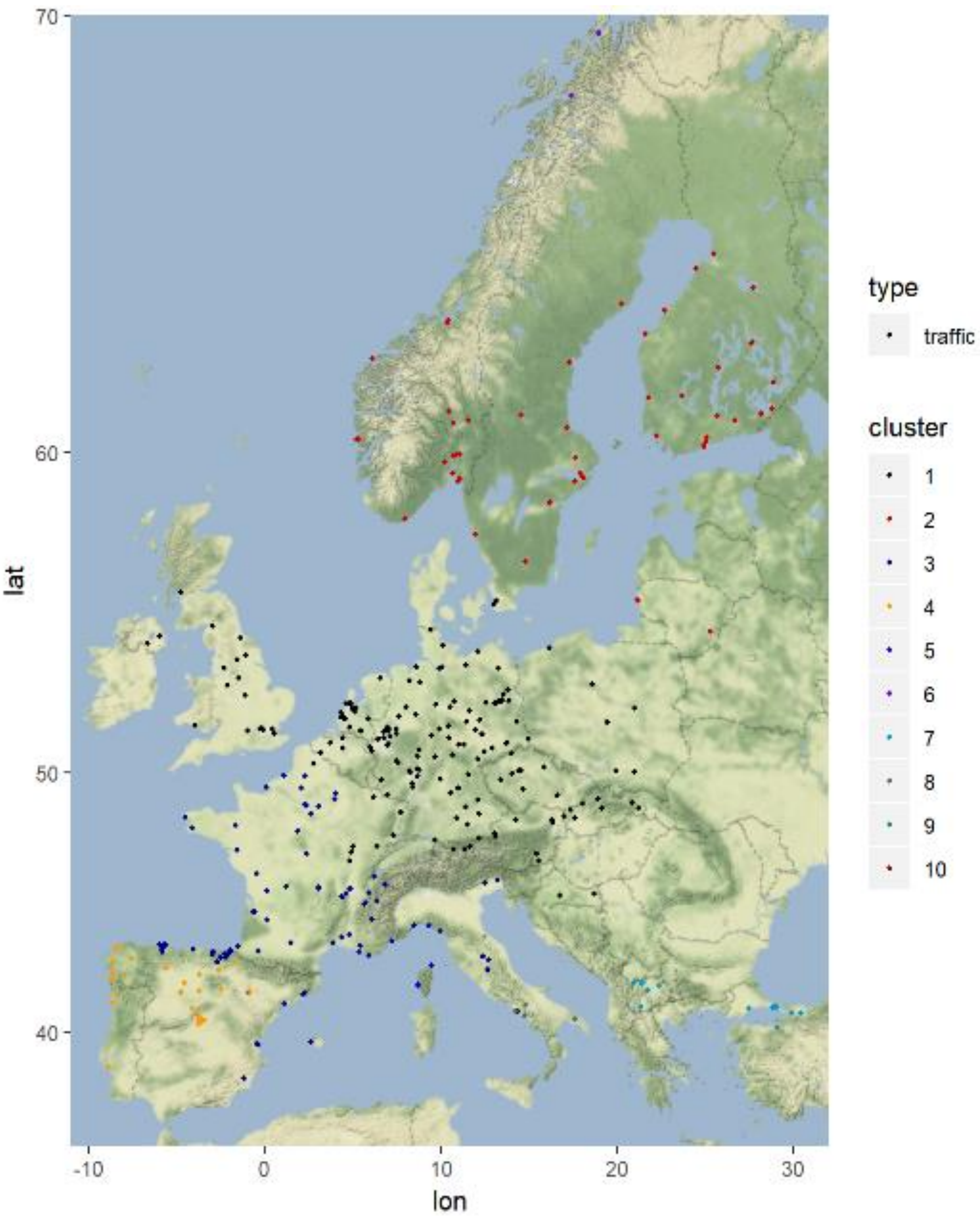


Figure 4.8. Clustering of PM_{10} traffic-oriented sampling points, color-coded according to the 1-R metric. The clustering shows mainly differences related to the temporal variations of emission sources and meteorological regimes.

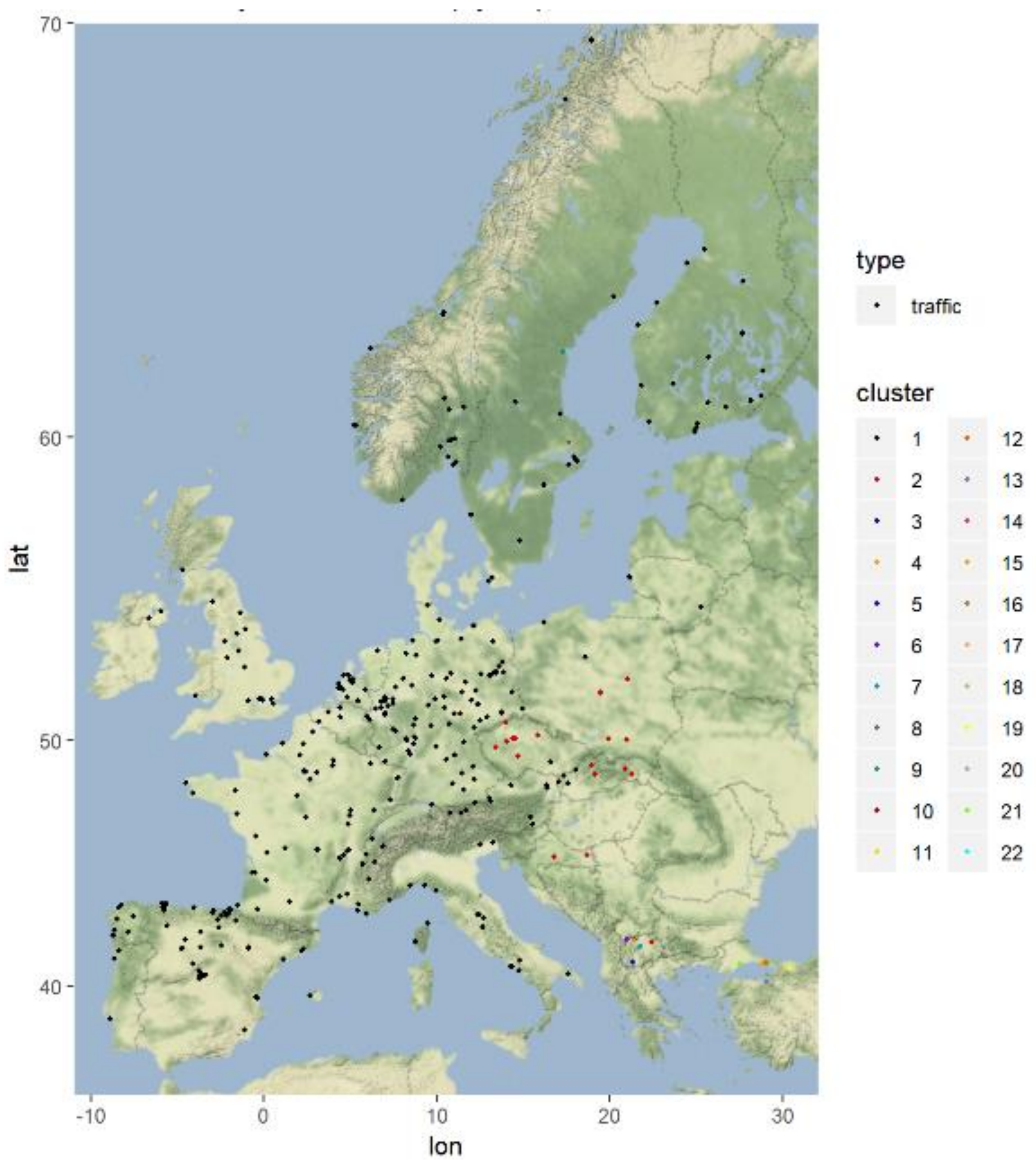


Figure 4.9. Clustering for PM_{10} traffic-oriented sampling points, color-coded according to the EuD metric. The clustering shows mainly differences related to the magnitude of emission different sources affecting PM_{10} at traffic-oriented sites

4.2.2.2 Industrial sampling points measuring PM_{10} concentrations

Figure 4.10 shows the clustering analysis results for the combined metric for all industrial sampling points measuring PM_{10}. The results show again the dissimilarities in North Macedonia and Turkey. These “outlier” industrial sampling points are listed in Table A.4. in Appendix 1.

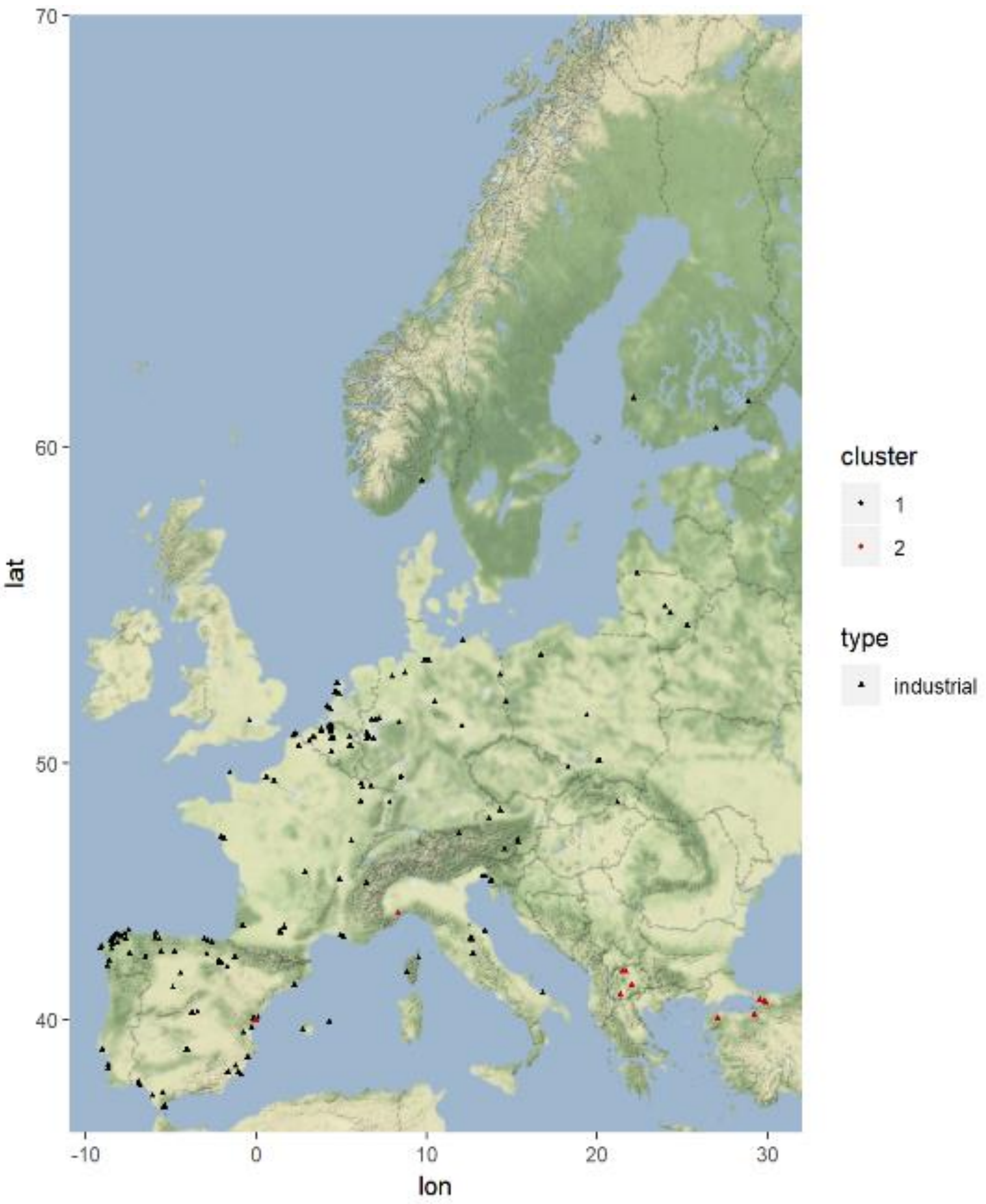


Figure 4.10. Clustering for PM_{10} industrial sampling points, color-coded according to the (1-R) x EuD metric. The clustering shows differences related to the variation of the temporal profile and magnitude of emission different sources affecting PM10 at industrial sites

4.3 Conclusions on sampling point classification

The analysis of the reported time series of hourly NO_2 and PM_{10} air pollutant concentrations in 2017 was carried out for a total of 578 industrial sampling points and 1090 traffic-oriented sampling points. This is only 78% of all the sampling point time series data available at EEA under e-reporting for 2017, as not all sampling points were reported with the necessary temporal coverage to be included. In some countries, the temporal coverage of the reported time series has prolonged gaps so that the concentration data could not be used in this clustering analysis. This is shown in the Figures throughout this chapter where several countries are not represented, and, depending on the pollutant, it is seen that the spatial coverage differs from country to country.

The hierarchical clustering methodology applied here provides a screening tool to analyse the level of similarity or dissimilarity of the reported air concentration data. It allows the identification of sampling points with specific behaviour that differ from the rest of the sampling points. These different behaviours may be well justified in terms of specific conditions around the sampling points, such as influence from specific sources or specific meteorological dispersion conditions. In this sense, the identification of “outliers” in the clustering analysis is not an excluding exercise but the recognition of a different behaviour that needs to be further investigated. The current clustering analysis applies for the whole of Europe, therefore, the level of similarity or dissimilarity flagged here also applies at a European level.

The results for all NO_2 sampling points show that the clustering across Europe does not follow national borders but responds to different meteorological and source regimes. The correlation metric shows the clustering of sampling points in different meteorological regimes while the Euclidean distance metric shows clustering in terms of emission magnitude.

The number of source sectors affecting PM_{10} is much greater that those affecting NO_2 concentrations. Consequently, it is more difficult to discern the explicit influence of source types in the PM_{10} measurement data. The clustering of PM_{10} sampling points data shows more differences due to the influence of multiple sources and different meteorological situations and more clusters are identified. The results from this type of clustering, addressing both industrial and traffic sampling points are valuable to show that the reported monitoring data provide a consistent assessment of air quality throughout Europe. Previously, Solazzo and Galmarini (2015) used European monitoring data for their clustering technique and concluded that the clustering results were showing country borders in some cases. This is not the case anymore in the current 2017 results. Currently, with the choice of a threshold to show differences across Europe, the monitoring network mostly clusters according to meteorology and predominant sources. This indicates that the monitoring network has improved, probably as a consequence of the harmonized quality assessment routines currently implemented throughout Europe.

The hierarchical clustering analysis has also been carried out separately for industrial and for traffic sampling points. The main purpose of this separate clustering analysis is to assess the accuracy of the application of the siting criteria and to determine if sampling points are indeed measuring traffic sources or industrial sources as intended. The exercise proves more useful for NO_2 than for PM_{10} sampling points, as explained below.

The results for NO_2 traffic-oriented sampling points show a high level of similarity across Europe, with only two clusters identified. This implies that most sampling points have a similar behaviour, with similar temporal correlations, and is an indication that the sampling points are indeed measuring traffic sources. The clustering analysis has also allowed us to identify “outliers” in Italy, Romania and in some sampling points in South Eastern Mediterranean countries. The results for NO_2 industrial sampling points provide very similar conclusions. For NO_2 industrial sampling points there is also a high level of similarity across Europe, which is an indication that NO_2 industrial sampling points are indeed measuring the influence of industrial emissions. “Outliers” are identified, in Spain, Iceland, Italy, Romania and Turkey. Although the identified “outliers” represent a small percentage of the total NO_2 sampling points, the reasons for the different behaviour in these sampling points needs to be further investigated. It can be due to a number of reasons including shifts in the time span of the reporting and variations in the implementation of different quality assessment routines.

The results of the clustering analysis for PM_{10} traffic and industrial sampling points shows higher levels of dissimilarity than for NO_2 as expected due to the more diverse emissions sources of PM_{10}, such as residential sources and agriculture. In 2017, there was an episode of strong forest fires in Portugal that affected the PM_{10} concentrations and its influence can be seen as a specific cluster over both Spain and Portugal. Because of the mixed sources affecting PM_{10} concentrations, it is difficult to classify sampling points in simple terms as industrial or traffic-oriented. The clustering techniques will thus not be able to distinguish clearly between industrial and traffic, as pointed out by Joly and Peuch (2012). However, the methodology is able to distinguish stations that, for good reasons or not, are different from all remaining data. This is the case for sampling points in North Macedonia and in Turkey. Further investigation of the reasons behind these dissimilarities is beyond the purpose of this study.

The study can be repeated for specific regions or even at a national level in order to gain further insight on the performance of different sampling points and their classification. To avoid that the flagging of outliers in the clustering analysis reflects only specific events occurred in the year of the analysis, such as influence from wild fires or Saharan dust intrusions, it is recommended to extend the temporal coverage of the time series to be multiannual. This approach could be used for instance as part of the national review of sampling point classification and could help screening which sampling points in the national network have a different behaviour that expected over the last 5-years period.

5 Interaction with national experts

A main purpose of this report is to better understand the different methods employed among countries for siting sampling points and classifying monitoring stations under European legislation. Given the limited information on siting and classification of stations currently provided through metadata reporting, direct interaction with national experts appears to be the most adequate way to gather such information. In fact, the Directive allows for this type of interaction and exchange of documentation to take place. Annex III D requires for the competent authorities to fully document the site-selection procedures and record information to support network design and choice of location for all monitoring stations. It also requires that such documentation be provided to the Commission within 3 months of being requested.

Before engaging in requests for documentation with individual EU Member States, a European-wide communication was established to collect data on the selection of sampling points through a web-questionnaire plus interactions in different workshops. The aim was to understand current practices and information needs regarding site location, as opposed to evaluating compliance. Based on the results from these initial interactions, it is envisaged that further direct communication with national experts will take place during the next stages of the project in order to better adjust the guidance on station representativeness to the practices and needs of countries.

5.1 Web-questionnaire on the Selection of Air Quality Sampling Points

A web-questionnaire titled “On the Selection of Air Quality Sampling Points” was developed to understand how countries apply the siting criteria when selecting their ‘traffic-oriented sites’ and ‘industrial sites’. The web-questionnaire did not intend to check compliance with the siting criteria in the AAQD but to understand the application of the criteria.

It was distributed to the AQUILA[16] and FAIRMODE[17] communities on 4 June 2019. It was organised in four different sections concerning: 1) macroscale and microscale siting criteria, 2) representative area, 3) relevant emissions sources and 4) dispersion situation. Information on all these parameters is requested by dataflow D under the e-reporting and is closely linked to the requirements in the Directive for siting and classification of sampling points. By 21 June 2019, 30 answers had been received from 25 countries (23 EU Member States and 2 EFTA countries). Five countries responded twice via different experts, their responses indicating that sometimes there might be different practices in different regions of the same country. Five EU Member States did not respond while five countries sent duplicate responses from their National Environmental Agencies and from their national air quality reference laboratories.

The responses to the questions in the four different sections are summarised in the following sections. We have chosen to present the results as “percentage of the responses” and not as actual country counts to reflect the fact that there might be different practices in different regions from the same country. The choice of this metric also recognises that five Member States did not answer the questionnaire and five others responded with significant differences in some aspects of the applied practices.

[16] AQUILA is the Network of Air Quality Reference Laboratories, see https://ec.europa.eu/jrc/en/aquila
[17] FAIRMODE is the Forum for air quality modelling in Europe, see https://fairmode.jrc.ec.europa.eu/

5.1.1 SECTION 1: Macroscale and microscale siting criteria

The first block of questions was designed to understand better how countries implement the requirements for macroscale and microscale siting given in Annex III (Annex VIII for ozone) of the AAQD. There were six different questions in this first section. Responses are details below.

Do you have advisory guidance in place to determine the siting of stations in your monitoring network?

- 50% responded that they had advisory guidance in place for the siting of sampling points in their country,
- 50% responded that such advisory guidance was not in place in their country.

In some countries, such as Belgium, the guidance was only available in some parts of the country but not in others, due to the adopted air quality governance approach. Interestingly, Estonia and United Kingdom responded positively to this question only in the responses from the national reference laboratory (NRL) but not in the responses from their national environmental protection agencies, indicating that although guidance is available it might be considered a technical issue by the national environmental agencies responsible for e-reporting.

Of the 14 countries that responded that they have advisory guidance in place for the location of sampling points, four indicated that the advisory guidance is essentially a transposition of the AAQD to national legislation. The countries that responded “NO” to this question referred also to the AAQD as the main reference for determining the location of sampling points. Indeed, most countries understand the text in Annex III B and C of the AAQD as their main reference for the siting of sampling points.

Only four Member States (Finland, France, Latvia and Sweden) provided a link to their guidance documents on-line. A brief review of the provided links to guidance documents shows that they contain detailed description of the macroscale and microscale criteria siting but fall short of developing methods to determine sampling point representativeness. The representativeness of sampling points is an essential requirement of Annex III B - macroscale siting criteria – but the methodology to determine the sampling point representativeness is not indicated. Latvia refers to Spangl et al. (2007) as their main guidance document to establish the representativeness and classification of air quality monitoring stations.

Do you have a compliance–checking system in place in your country to determine the achievement of the macro- and microscale siting criteria in Annex III of the AAQD of your monitoring network?

- 67% of the respondents confirmed that their country has a compliance system in place to determine that the location of their sampling points follow the siting specifications given in Annex III B. (macroscale siting criteria) and Annex III C. (microscale siting criteria). They also provided the link of the responsible institution, which varies depending on each country’s air quality governance and encompasses national and regional reference laboratories well as environmental agencies and ministries.
- 33% answered that they did not have a compliance-checking system in place.

How does your country check that the kerb-distance for traffic stations is according to the requirements in the Directive (less than 10 m)?

According to Annex III C traffic oriented sampling inlets shall, as far as is practicable, be at least 25 m from the edge of major junctions and no more than 10 m from kerbside. The answers to the questionnaire indicate that there are three main methods applied: 1) direct at site length measurements; 2) use of Google Earth and Global Positioning System (GPS) in combination with photographs; and 3) a combination of both 1) and 2). In some cases, the respondents do not specify the method employed but refer to it as the responsibility of the national reference laboratories.

- 60% mentioned at site measurements as the main method to check the distance to the kerb for traffic stations.
- 20% mentioned that they also use Google Earth and photographs from satellite data to check the microscale criteria.
- 20% simply mentioned that checking the compliance with the siting criteria in Annex III is the responsibility of the national reference laboratories.

The responses indicate that establishing the location of traffic stations relies primarily on local knowledge and in-situ measurements, and that in many countries the locations are further checked by the national reference laboratories.

How does your country check that for industrial stations the measurements taken at the sites are representative for an area of 250 m x 250 m (as required by the Directive)?

This question addresses the macroscale siting requirement in Annex III B.1. (b) that is intrinsically connected with the evaluation of sampling point representativeness. The responses from countries indicate that three main approaches are applied at national level to determine the sampling point representativeness: 1) use of modelling calculations or indicative measurements, 2) proximity to sources and 3) expert evaluation. Countries chose between these approaches depending on the availability to either modelling results, dispersion calculations or emission sources.

- 37% of respondents indicate that they use modelling results to determine the representativeness of industrial sampling points. These are Austria, Estonia, Germany, United Kingdom, Greece, Lithuania, Poland, Spain and Slovakia (Belgium is not yet applying this approach but is planning to do so in the future).
- 37% responded that they use expert evaluation to establish the representativeness of the industrial sampling point but without specifying any methodology.
- 11% indicated that they use an estimate based on distance to sources and proximity to emissions.
- The remaining 15% responded that they do not have industrial sites in their country and thus the question is non-applicable.

How often do you review the compliance with siting criteria of your monitoring network?

Annex III D as revised by Directive 2015/1480/EU requires the documentation of site selection to be updated and reviewed at least every five years. 78% of the responses indicate that countries review the compliance with siting criteria of their monitoring network every five years or less.

- 48% responded that they revise the siting of their monitoring network every five years.

- 30% responded that they revise the compliance with siting criteria more often than every five years. Sweden and United Kingdom review the siting of sampling points every three years; Spain, Greece and Croatia do it every two years while Germany, Lithuania and Slovakia do it annually or even more frequently. There are also different practices in the same country: for instance, in Belgium, Brussels reviews their network every single year while in other areas in Belgium, the review can take place every two, three or five years.
- 7% of the respondents indicated that they review their network with less frequency than five years. For instance, seven years in the Netherlands and ten years in Estonia.
- 15% review the compliance with siting criteria “when needed”. This is the case for Denmark, Iceland, Romania and Slovenia.

When was the last time you checked the compliance with the directive macro- and micro-scaling criteria for traffic-oriented and industrial stations?

Most countries have checked the compliance with siting criteria for these type of stations within the last three years, but about 27% of the respondents have not revised the siting criteria for more than five years.

- 63% responded the last year of checking to be less than five years ago with 48% responding that the last year is either 2019 or 2018.
- 18% responded the last year of checking to be more than five years ago (before 2014).
- 19% responded that they have not revised the monitoring network since it was established.

5.1.2 SECTION 2: Representative Area

The ”Representative Area” describes the spatial extent of the area of representativeness for a monitoring station/sampling point derived using geometry features. Understanding what type of guidance is needed to determine the station representativeness of sampling points is the main purpose of the work under Task 1 of this Service Request 5 on “Assessing the spatial representativeness of air quality sampling points”. The main purpose of this block of questions in the web-questionnaire was to understand the existing practices used in countries to calculate representativeness of sampling points.

Is information on the representative area of sampling points/monitoring stations available to you?

- 64% responded that they have information on the representative area of sampling points.
- 36% of the countries responded that information on the representative area of sampling points is not available to them.

If yes, do you use this representative area to assess population exposure and areas in exceedance of limit values?

- 62% of those that have information on the representative area of sampling points use the data to assess population exposure and areas in exceedance of limit values.

- 38% of those that have information on the representative area of sampling points do not use the data to assess population exposure and areas in exceedance of limit values.

This implies that overall, most of the countries (60%) have available information on the representativeness of the sampling point to calculate key indicators for the protection of human health such as population exposure and areas in exceedance of limit values. This information on representative area is however not retrievable in the current version of EEA's CDR system. So, although over 60% of the countries have this information available, it is not part of the metadata evaluated in Chapter 3.

Annex III B. 1 (a) clearly states that the sampling points directed to the protection of human health should provide information both about the areas with highest concentrations and areas representative of the exposure of the general population. Such requirements are difficult to fulfill without information on the representative area, although alternatively, models can be used for determining exposure. In many cases, for the purpose of population exposure calculations, countries do not use sampling point representativeness but modelling results, as indicated below from the responses on methodological approaches.

The respondents were requested to provide a link to the methodology used to determine the representative area of sampling points/monitoring stations in their countries. 28% of the countries, seven in total, have provided links to their methodologies.

- Four countries provided links to their specific guidance for calculating representativeness of sampling points: Austria, France, Finland and Germany.
- Austria's guidance includes a methodology to determine population exposure from sampling point observations
- Finland and Slovenia refer to the FAIRMODE inter-comparison on station representativeness, although Slovenia does seem to have its own guidance and only refers to the FAIRMODE report.
- Belgium provided links to report the methodology to calculate population exposure on the basis of modelling results.
- Latvia provided the link to IPR Guidance.

If no, what are in your opinion the main barriers for the availability of information on site representative area in your country?

- 50% considered that the main barrier to the availability of information on representativeness is the lack of guidance and definition from the European Commission.
- 26% considered that capacity-related barriers also play a role.
- 16% considered that the lack of guidance and definition from national administration is also relevant.
- 8% indicated that data access issues were the main barriers for the availability of information on site representative area.

The need for guidance both at EU and national level is important feedback from the web-questionnaire. Most respondents expressed the need for more clarity in some definitions in Annex III, e.g. "representative of the exposure of the general public" and "period which is significant in relation to the averaging period of the limit values". The few countries that responded with “Other” (like United Kingdom and Germany) indicate that the problem of site representativeness can be circumvented with the use of models. But in this case, guidance for calculating human exposure indicators from model results should be provided too. Most respondents described a need for guidance for exposure modelling calculations where instead of definitions it might be more helpful to give clear examples of what is meant and how to approve different existing procedures.

5.1.3 SECTION 3: Relevant Emissions

The responses from countries indicate that three main approaches are applied at national level to determine the sampling point representativeness: 1) use of modelling calculations or indicative measurements; 2) proximity to sources and 3) expert evaluation. The purpose of Section 3 questions on relevant emissions is to understand to what extent information on emission sources and proximity to sources is available to countries given their relevance both to determine representativeness and to carry out air quality modelling results.

Is information on relevant emission sources affecting pollutant concentrations at the sampling point/monitoring station available to you?

Information on emissions is generally available as indicated below:

- 83% responded that emission information at the sampling point is available to them.
- 17% responded that the emission information is not available to them.

If yes, what information do you rely upon to determine the emissions with predominant influence on pollutant concentrations at the sampling point/monitoring station?

The responses to this question indicate that there are four main approaches to determine what emissions affect the air pollutant concentrations at different sampling points based on: 1) national emission inventories; 2) high resolution emission data; 3) air quality dispersion modelling; and 4) expert estimates**.**

- 38% use proximity to emissions sources to determine the predominant influence at sampling points – with emissions available from the national emission inventories reported under the National Emission Ceiling Directive (2016/2284)[18].
- 33% use proximity to emissions sources to determine the predominant influence at sampling points – with emissions available from high resolution emission inventories.
- 21% use source allocation results from dispersion modelling results.
- 8% indicate that they rely on expert estimates without specifying the methodology.

If no, what are, in your opinion, the main barriers for the availability of information on emission sources affecting sampling points/monitoring stations in your country?

For the few cases where emissions information is not available, the main reason is the lack of coordination to access local high resolution emission data existing only in particular areas (municipalities) in the country.

[18] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016L2284&from=EN, last viewed January 2020

Do you have a common countrywide methodology to identify what emissions predominantly influence pollutant concentrations at sampling point/monitoring stations?

- 67% respond that there is no common countrywide methodology.
- 33% respond that they have a common countrywide methodology to identify what emissions influence pollutant concentrations at sampling points.

The respondents were requested to provide a link to their country's common methodology. Only United Kingdom, Romania and Slovakia provided a link to the methodology for determining source allocation on the basis of air quality dispersion modelling. No links were provided to reports with guidance on how to determine high resolution emission information.

5.1.4 SECTION 4: Dispersion situation

Information on the dispersion situation is essential to determine the area of representativeness of sampling points and it is required to be reported under the IPR. However, this was the part of the questionnaire that raised more questions for clarification among the respondents. The main purpose of this block of the questionnaire was to identify what approaches are currently used in countries to identify the dispersion situation at sampling points.

Is information on the dispersion situation of sampling points/monitoring stations available to you?

- 67% responded that they have information on the dispersion situation at sampling points.
- 33% responded that information on the dispersion situation at sampling points is not available to them.

If yes, does your country have a guidance system and identified methodology to report on the dispersion situation relevant for the representative area of specific sampling points/monitoring stations?

From the responses, it can be concluded that there are different approaches to calculate the dispersion situation at sampling points involving: 1) analytical software for meteorology; 2) air quality modelling; and 3) IPR code-list. However, a significant group indicates that there is no guidance system in their country but do not specify how the information is made available to them.

- 43% responded that there is no guidance or methodology for determining the dispersion situation at sampling points.
- 26% use an air quality modelling approach.
- 21% use analytical software methods for describing dispersion from meteorological conditions mainly.
- 10% responded that they use only the IPR code-list for reporting.

If no, what are in your opinion the main barriers for the availability of information on the dispersion situation at sampling points/ monitoring stations in your country?

- 36% indicated capacity-related barriers.
- 32% indicated the lack of guidance and definition from the European Commission.

- 9 % indicated the lack of guidance and definition from their national administration.
- 14% indicated data access issues.
- 9% provided no answer.

Do you have a common methodology to characterize local and regional dispersion choosing from the code-list in the IPR user guide?

- 77% responded that there is no common countrywide methodology.
- 23% responded that they have a common countrywide methodology to identify the dispersion situation and link to the IPR code-list.

Only Finland and France provide independent links to the methodology used. Most other countries that provided a link refer only to the IPR Guidance.

5.2 Feedback at Workshops

The initial findings from the questionnaire, the clustering analysis and the metadata analysis, as well as a demonstration of the Monitoring Siting Criteria Viewer were presented at four different workshops with the relevant participants of the AQUILA and FAIRMODE networks as well as members of the IPR working group. In the following a very short summary of the initial feedback from the community is presented to help frame conclusions and recommendations.

The four workshops that took place in the autumn of 2019 were:

a) 18 - 19 September – Nordic Reference Laboratory Meeting in Oslo
b) 2 - 3 October – AQUILA meeting in Vienna
c) 7 - 9 October – FAIRMODE meeting in Madrid
d) 5 -7 November – IPR meeting in Copenhagen

5.2.1 Nordic Reference Laboratory Meeting in Oslo, 18-19 September 2019

L. Tarrasón (NILU) presented the first outcome of Task 3 under the DG ENV project "Assessing the spatial representativeness of air quality sampling points". The discussion focused on the reporting of metadata information and on the reporting of exposure indicators using monitoring data and the metadata information for each monitoring station available from e-reporting.

A key point raised by the representatives of Norway, Sweden and Denmark was that while the national reference laboratories were responsible for the siting and classification of sampling points, they were not involved in reporting under IPR. Reporting is the responsibility of the national environmental protection agencies in the Nordic countries. It involves both the reporting of population exposure results and the reporting of metadata. This separation of roles and responsibilities may contribute to some of the gaps and inconsistencies identified in the reported metadata.

The participants agreed with the conclusions from the web-questionnaire on the need for more guidance and clarification from the European Commission, especially with respect to how to proceed with population exposure calculations and with the evaluation of sampling point representativeness. The Nordic Reference Laboratory participants agreed that cooperation with the modelling community on population exposure and station representativeness would be essential for a successful implementation of the AQQD in this respect.

5.2.2 AQUILA meeting in Vienna, 2-3 October 2019

K. Tørnkvist (NILU) participated in the AQUILA meeting in Vienna and presented there the first outcomes of the project "Assessing the spatial representativeness of air quality sampling points", with focus on Task 3.

The discussion about reporting sampling point siting information at the AQUILA meeting was shorter than in the Nordic Reference Laboratory Meeting and focused more on the station representativeness issue than on the gaps of the reported metadata. This is because, as in the Nordic countries, the AQUILA network is not generally directly involved in the reporting of metadata under IPR, so the analysis of the reported metadata was not so relevant for this community. There was more interest on the conclusions of the clustering analysis showing that the community is doing good work on the quality checking of measurement data.

In future presentations of this work to the AQUILA network it will be necessary to focus on the need for cooperation across the modelling and the IPR reporting communities in order to optimise the design of monitoring networks, and to agree on combined methods to determine sampling point representativeness.

5.2.3 FAIRMODE Technical meeting in Madrid, 8 - 9 October 2019

L. Tarrasón (NILU) presented the first outcomes of Task 3. The discussion focused on the reporting of exceedances using monitoring data and the metadata information for each monitoring station available from e-reporting. Because this metadata is scarce and often inconsistent, it limits its value to provide information on monitoring station classification and/or the possible exposure calculations.

Representatives from Norway, Portugal and Sweden requested detailed and clear guidance on the preferred methodologies in the tiered approach for station representativeness and exposure calculations. They pointed out how difficult it is to provide such guidance for users from different communities involving modellers, network managers and administration experts in charge for e-reporting. The representative from Spain mentioned the importance of intra annual variations in the assessment of spatial representativeness and how in order to obtain robust results, multi-year averages could be appropriate.

The discussion showed that the FAIRMODE community is usually not involved in reporting exposure or station representativeness, despite the advantage of using models to determine these. Representatives from the Spanish Ministry of the Environment observed that the Spanish FAIRMODE team is not involved in e-reporting and used the opportunity to invite the FAIRMODE representative to discuss possible co-operation.

The representative from EEA suggested that the presentations should be also discussed at the next IPR meeting on November at EEA to gather the views from the e-reporting community on metadata reporting. Finally, the representative from the Netherlands requested, access to the interactive GIS Viewer developed to check the position of reported sampling points in their country.

5.2.4 Feedback from the IPR meeting in Copenhagen, 5-7 November 2019

J. Soares (NILU) presented the outcome of Task 3, with special emphasis of the results of the analysis of the reported metadata. The discussion focused on the reporting of exceedances using monitoring data and the metadata information for each monitoring station available from e-reporting. Most participants were interested in learning more on their country's performance on the metadata tests of completeness and consistency, and they are looking forward to seeing more results from Task 3.

Concerning the area of representativeness, there was a recognised need for guidance on how to assess it and how to assess exposure. The participants agreed that there is a need for further guidance on simple things such as “distance to road”, what is "close proximity", what are “predominant sources”, etc. They also proposed possible follow-up activities to find a harmonized way to evaluate spatial representativeness and how to calculate exposure. The work to determine spatial representativeness for exposure calculations and for optimal network design should involve the FAIRMODE & AQUILA communities, even though these communities do not need to be directly involved in e-reporting.

5.3 Discussion on methodologies for selection of sampling points

The responses to the web-questionnaire have been very useful to better understand the different methods employed among countries for siting sampling points and classifying monitoring stations under European legislation.

The responses to the questionnaire indicate that only half of the countries have advisory guidance in place to determine the siting of sampling points. The current guidance contains a detailed description of the macroscale and microscale criteria siting but falls short of developing methods to determine sampling point representativeness. The representativeness of sampling points is an essential requirement of Annex III B - macroscale siting criteria - but the methodology to determine the sampling point representativeness is not well described.

Although 64% of the countries responded that they had available information on the representative area of sampling points, it is noteworthy that such information is not available at EEAs CDR. A key finding from the questionnaire is a request by Member States for more guidance from the European Commission on the methods to determine the sampling point representativeness. The responses from countries indicate that three main approaches are applied at national level to follow macroscale criteria in Annex III B and determine the sampling point representativeness: 1) use of modelling calculations or indicative measurements; 2) proximity to sources; and 3) expert evaluation. The three main approaches currently applied in the countries coincide with the tiered approach proposed in Task 1: Tier 1: expert opinion; Tier 2: proximity to sources based on GIS proxy data; and Tier 3: assessment based on modelling results. Four countries, Austria, France, Finland and Germany, have provided links to their specific guidance for calculating representativeness of sampling points and these are essential references for the development of the tiered approach in Task 1 of this project. It is expected that the recommendations from Task 1 will start to help address this need for guidance.

The siting of sampling points needs to follow also the microscale criteria in Annex III C and the answers to the questionnaire indicate that there are three main applied methods by countries: 1) direct at site measurements, 2) use of Google Earth and GPS systems in combination with photographs, and 3) a combination of both 1) and 2). The GIS Viewer presented in Chapter 2 was designed to support checking of this microscale siting criteria. If developed further, for instance as part of the FAIRMODE composite mapping platform, it could serve as guidance to facilitate the evaluation of sampling points in terms of optimal location and sampling representativeness at national level.

6 Summary and conclusions

The location and classification of all traffic-oriented and industrial sampling points for PM_{10} and NO_2 reported under the AAQD for 2017 has been evaluated in this study. The main purpose has been to assess to what extent the siting and classification of air quality sampling points applied in European countries follows the criteria and specifications provided by the AAQD and the Implementing Decision 2011/850/EC.

The analysis has been carried out on the basis of a series of consultations with national experts, a detailed evaluation of the reported metadata available at EEAs Central Data Repository system and two independent objective evaluations. The two independent evaluations have used: 1) a prototype GIS Viewer to objectively test some of the siting criteria in Annex III of the AAQD on the location of sampling points; and 2) a hierarchical clustering analysis to objectively determine the validity of the current sampling point classification on the basis of the temporal variability of the reported air quality data.

The main findings are summarised for each of these four complementary evaluation approaches. Conclusions are given in general terms and no detailed evaluation is provided for any particular country. Although some of the data compiled allows for individual evaluation, it is important to note that compliance-checking was never the purpose of this study. The results are intended to provide an overview on the current status of the application of the siting criteria and sampling point classification throughout Europe, to enable further harmonized application of these criteria and to inform how the current siting and classification activities can contribute to a better understanding of the spatial representativeness of sampling points.

Finally, it is worth mentioning that this report would not have been possible without the interaction and support of national experts especially those participating in the AQUILA, FAIRMODE and IPR networks. The development of good guidance on sampling point representativeness depends on an increased interaction between these three groups and the combination of their expertise on measurements and models. Such combination of expertise can also help to further address the representativeness and optimization of air quality monitoring networks at European level.

Siting location of sampling points- GIS Viewer

The location of all traffic-oriented and industrial sampling points operational in 2017 for NO_2 and PM_{10} as reported under AAQD has been investigated with an interactive GIS Viewer, also known as the Monitoring Siting Criteria Viewer. The GIS Viewer is currently a prototype that shows the potential of using satellite images and available mapping information on roads and industrial emission data to support monitoring siting decisions. The viewer is intended to support the assessment of the location of air quality sampling points and their distance to given emission sources. It has been designed to objectively determine whether some of the macroscale and microscale criteria as defined in AAQD Annex III are fulfilled for all traffic-oriented and industrial sites currently reported by countries.

The initial results have shown that about 20% of all reported traffic-oriented sampling points may not be located less than 10 m from kerbside of main roads as per the criteria under the AAQD Annex III C (microscale siting criteria). However, current results rely on the accuracy of the reported position given in geographical coordinates, the accuracy of the road map description in the OSM system and the reliability of the reporting of the position of industrial facilities under E-PRTR reporting. Also, the current prototype is limited by given assumptions on the width of the roads. So, these results are only indicative and need to be further validated, preferably in cooperation with national experts.

For industrial sampling points, the macroscale siting criteria in AAQD Annex III B requires for the air sampled to be representative of air quality in an area of at least 250 m x 250 m, where feasible. Agreement with this macroscale criteria requires an evaluation of the spatial representativeness of the industrial sampling point. The GIS Viewer designed under this study intends to facilitate an evaluation of the spatial representativeness by using the criteria of proximity to industrial sources. However, given the limitations in the current metadata and the lack of information on industrial sources other than those reported to E-PRTR, no generalised conclusions on industrial sites have been given at this stage. The GIS Viewer may be made available for further checking of distance to E-PRTR plants by countries, if desirable.

Further discussion with national experts could consider the potential to integrate the GIS Viewer as part of the composite mapping capabilities under FAIRMODE. It would be useful for national experts to support their sampling point siting decisions and check that the metadata information on the location of monitoring sampling points has been introduced correctly in e-reporting. To be useful, the current GIS Viewer needs to be further developed, for example, additional information on the width of the roads and small industrial sources needs to be included. In particular, the tool would profit from an additional feature that could show the available metadata reported when clicking on a particular sampling point and from further linking to data on dispersion and emissions sources from the FAIRMODE composite mapping platform to enable access to relevant information related to the determination of the spatial representativeness of sampling points.

Metadata analysis on siting and classification of sampling points

Metadata information on sampling point location and classification as officially reported by countries in dataflow D under the e-reporting has been compiled from EEA's CDR. It involved a selection of reported metadata for all NO_2 and PM_{10} sampling points for 2017. The metadata parameters that were selected were intended to provide useful insights on the location of the sampling points, their classification and the data used to determine spatial representativeness of sampling points. It was selected to help understanding how countries apply the macroscale and microscale criteria given in AAQD Annex III and provide additional information on the dispersion conditions and emission situation that are necessary to evaluate the spatial representativeness of sampling points. Despite the usefulness of these metadata for any further application of the data from sampling points, the results show that there are important gaps in the reported metadata.

The currently reported metadata is far from complete and there are identified inconsistencies in the reporting by countries. In some cases, the reported metadata shows that the sampling points do not always comply with the siting criteria established in Annex III. In many cases the metadata is missing or shows inconsistencies in the choice of units and with respect to the nomenclature for non-available metadata. The reporting of voluntary data related to dispersion conditions and traffic emissions is even more sparse. Metadata on the extent of the area of representativeness was not available in any of the 6,972 sampling points analyzed in this study. Figure 6.1 summarizes the findings from the metadata analysis, showing the percentage of sampling points with missing or not reported metadata. The most significant gaps in the metadata information is for the spatial extent of representative area, for information on emission data and on dispersion conditions. The gaps in metadata are quite similar in EU27, EU28 and for all countries, since more than 90% of the analyzed metadata originates from EU28.

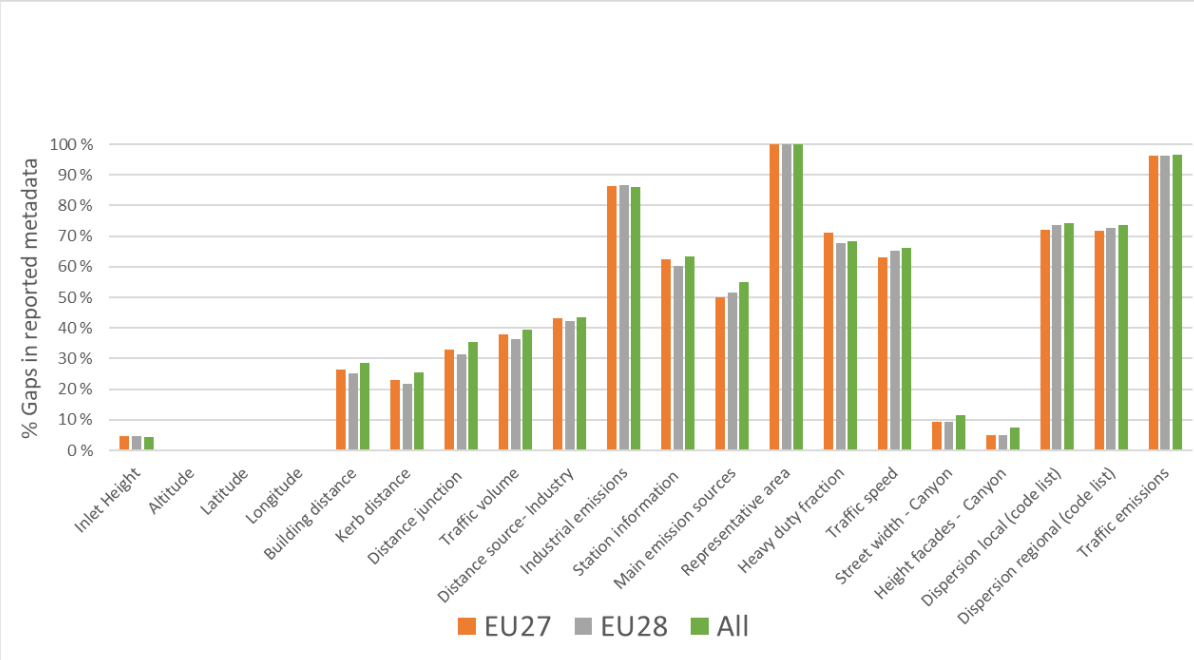


Figure 6.1. Summary of the metadata analysis showing the percentage of missing or not reported metadata for each parameter investigated for EU27, EU28 and all countries included in this analysis.

There are likely different reasons why the reporting of these metadata is incomplete. One reason may be that the actual metadata is not available to the countries, due to either lack of data or lack of guidance on how to compile these data, especially in relation to the representativeness of sampling points. Another reason may be that metadata reporting is less prioritized in the countries than air quality data reporting. It is difficult to explain why national experts respond that information on area of representativeness, emissions and dispersion conditions are available to them, but these are not to be found in the same extent in the metadata reports analyzed here.

The lack of reported metadata on siting and classification of sampling points most probably indicates a need for further guidance. It is expected that improved guidance on how to determine and use of the different parameters could help more countries to report this metadata information. In particular guidance should focus how to determine the spatial extent of representative area and how information on a) the four classes for local dispersion conditions, b) main emission sources, c) the traffic speed, d) the heavy duty fraction and e) the street width in street canyons is to be reported and made use of for better establishing the sampling point representativeness area.

Improvement on the reported metadata on sampling point location and classification under e-reporting would allow better characterisation of the sampling points and facilitate a dialog on the application of sampling point location, classification and representativeness throughout Europe.

Clustering analysis on sampling point classification

A clustering analysis based on the methodology by Soares et al (2018) has been carried out to compare the reported classification of the sampling points with the classification that would be derived from signature of the reported air quality measurements. The hierarchical clustering provides an easy-to-use screening method to analyse the level of similarity or dissimilarity of the reported air concentration data. It allows the identification of sampling points with specific behaviour that differ from the rest of the sampling points. Overall, the clustering technique application here has shown that the current classification of traffic and industrial sampling points is consistent with the reported air quality data.

The results for all NO_2 sampling points show that the clustering across Europe does not follow national borders but responds to different meteorological and source regimes. The clustering of PM_{10} sampling points data shows more differences due to the influence of multiple sources and different meteorological situations and more clusters are identified but the conclusions remain the same as for NO_2.

The hierarchical clustering analysis has been carried out at European level, separately for industrial and for traffic sampling points. The main purpose of this separate clustering analysis was to assess the accuracy of the application of the siting criteria and to determine if sampling points are indeed measuring traffic sources or industrial sources as intended. The results for NO_2 show a high level of similarity across Europe. This implies that most sampling points have a similar behaviour, with similar temporal correlations, and is an indication that the traffic-oriented sampling points are indeed measuring traffic sources and the industrial sampling points are indeed measuring the influence of industrial emissions. Identified "outliers" in Spain, Iceland, Italy, Romania and Turkey represent a small percentage of the total NO_2 sampling points. For PM_{10}, the level of dissimilarity is higher than for NO_2 because there is a larger number of emission sectors affecting PM_{10} concentrations. Because of the mixed sources affecting PM_{10} concentrations, it is difficult to classify sampling points in simple terms as industrial or traffic-oriented. Still, the methodology is able to distinguish sampling points that, for good reasons or not, are different from all remaining data. This is the case for sampling points in North Macedonia and in Turkey. The identification of "outliers" in the clustering analysis is not an excluding exercise but the recognition of a different behaviour that needs to be further investigated. However, further investigation of the reasons behind these dissimilarities is beyond the purpose of this study.

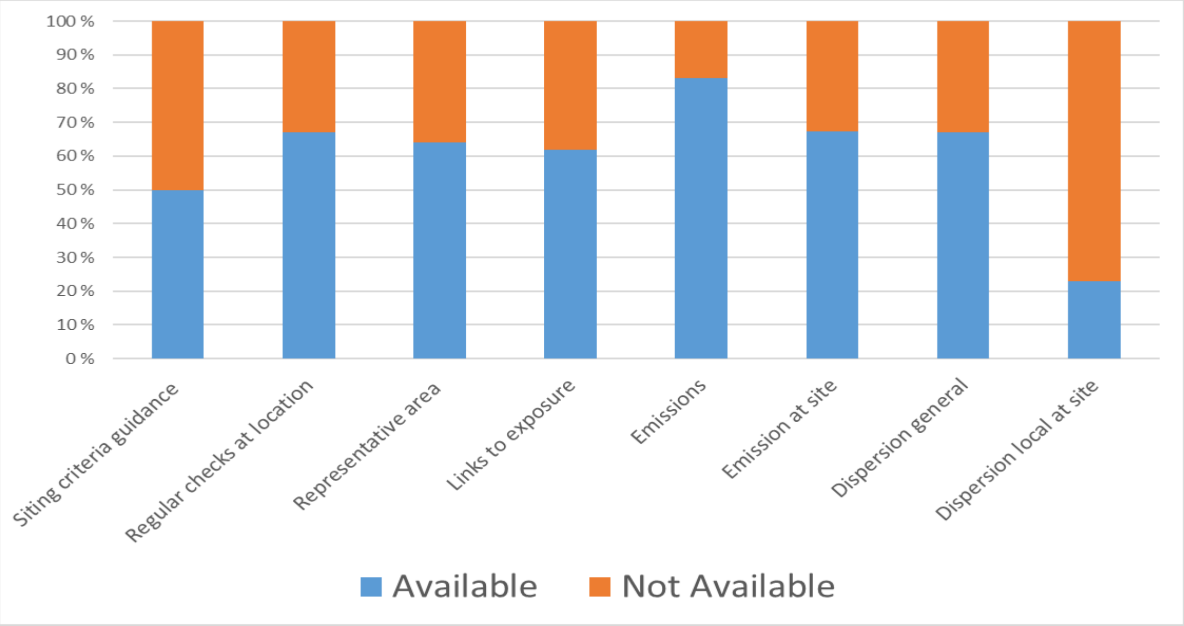
It would be possible to carry out a similar clustering analysis for specific regions at national level in order to gain further insight on the performance of different sampling points and their classification and to understand the reasons behind "outliers". This approach could be used at national level to review the current monitoring design and inform the choice of location of and classification of new sampling points. Further, the method can assist to determine the spatial representativeness of the sampling points and was evaluated further under Task 1 to help develop further guidance in this area.

Methodologies for siting and sampling point classification

The consultations with national experts via a questionnaire and successive workshop presentations have been useful to better understand the different methods employed among countries for siting sampling points and classifying monitoring stations under European legislation. A summary of the responses to the questionnaire is given in Figure 6.2 as an overview of the availability of relevant information for sampling point location and classification.

Only about half of the countries have advisory guidance in place to determine the siting of sampling points. The current guidance contains a detailed description of the macroscale and microscale siting criteria but falls short of developing methods to determine sampling point representativeness. The representativeness of sampling points is an essential requirement of Annex III B - macroscale siting criteria, but the methodology to determine the sampling point representativeness could be better described in guidance documents at the European level.

Figure 6.2. Summary of availability at national level guidance and information for sampling point location and classification. Percentages with respect to total number of countries that responded to the web questionnaire.



With respect to the macroscale criteria in Annex III B, the questionnaire responses from countries have also provided valuable insight on the approaches that are applied at national level to determine sampling point representativeness. There are three main approaches: 1) use of modelling calculations or indicative measurements; 2) proximity to sources; and 3) expert evaluation. Countries choose between these approaches depending on the availability to either modelling results, dispersion calculations or emission sources. These three methods correspond to a large extent to the tiered approach proposed in Task 1 and will be further the elaborated under Task 1 of this project.

With respect to the microscale siting criteria in Annex III C, the answers to the questionnaire indicate that there are three main methods applied by countries: 1) direct at site measurements; 2) use of Google Earth and GPS systems in combination with photographs; and 3) a combination of both 1) and 2). The GIS Viewer presented in Chapter 2 was designed to support checking some of the microscale siting criteria, a system that is currently not yet available to most countries but that complements well the currently used methodologies. Although the GIS Viewer is currently only a prototype with recognised limitations, it could be developed further, for instance as part of the FAIRMODE composite mapping platform, to support evaluation of sampling point location and classification at national level.

Despite the fact that most countries indicate that they have access to relevant information on the spatial representativeness of sampling points, on emissions at national and at site level and on the dispersion conditions around sampling points, the information is currently not available as part of the reported sampling point metadata. The metadata gaps and needs for guidance identified guided further work in the project to provide a pathway for a more harmonized guidance of the calculation of spatial representativeness of sampling points throughout Europe.

7 References

Ambient Air Quality Directive (AAQD): Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. OJ L 152/1

EU (2011) Commission Implementing Decision 2011/850/EU of 12 December 2011 laying down rules for Directives 2004/107/EC and 2008/50/EC of the European Parliament and of the Council as regards the reciprocal exchange of information and reporting on ambient air quality. OJ L 335available at https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:335:0086:0106:EN:PDF

EU (2013) Guidance on the Commission Implementing Decision laying down rules for Directives 2004/107/EC and 2008/50/EC of the European Parliament and of the Council as regards the reciprocal exchange of information and reporting on ambient air (Decision 2011/850/EU). Version of 15 July 2013, last viewed January 2020. Available at: https://ec.europa.eu/environment/air/quality/legislation/pdf/IPR_guidance1.pdf

EEA (2018) User guide to XML & data model for e-reporting v3.4 -Jaume Targa and Tony Bush (ETC/ACM) and Contributions from Rune Ødegård, Francisco Reina, Katharina Schleidt, Barbara Magagna - Optimised for schema version 1.0 https://www.eionet.europa.eu/aqportal/doc/UserGuide2_AQD_XML_v3.4.1.pdf

Joly, M. and Peuch, V.-H. (2012) Objective Classification of Air Quality Monitoring Sites over Europe. Atmospheric Environment, 47, 111-123.

Næs, T., Brockhoff, P. B., and Tomic, O. (2010) Statistics for Sensory and Consumer Science, 6th Edn., John Wiley & Sons, Ltd, Wiltshire, UK, ISBN: 9780470518212, 2010.

Soares, J., Makar, P. A., Aklilu, Y., and Akingunola, A.(2018) The use of hierarchical clustering for the design of optimized monitoring networks, Atmos. Chem. Phys., 18, 6543–6566, https://doi.org/10.5194/acp-18-6543-2018, 2018.

Solazzo, E. and S. Galmarini. (2015), “Comparing apples with apples: Using spatially distributed time series of monitoring data for model evaluation”. Atmos. Environ., Vol. 112, pp. 234-245.

Spangl, W., Schneider, J., Moosmann L. and C. Nagl (2007) Representativeness and classification of air quality monitoring stations. Umweltbundesamt report available at https://www.umweltbundesamt.at/fileadmin/site/publikationen/rep0121.pdf

Appendix 1 – Identified Sampling Point Outliers

Table A.1. Traffic sampling points measuring NO_2 concentrations clustering at 1-R metric in a single cluster (2) representing "outliers" from the traffic-oriented sampling points in the rest of Europe - cluster (1)

Lat	Lon	Sampling point	station	type	cluster
-21.33	55.48	SPO.FR38023_8	STA-FR38023	traffic	2
-21.21	55.3	SPO.FR38024_8	STA-FR38024	traffic	2
43.51	16.45	HR_DOC_TYPE_D_SPO_66	HR_DOC_TYPE	traffic	2
43.09	-0.05	SPO.FR12045_8	STA-FR12045	traffic	2
43.55	16.43	HR_DOC_TYPE_D_SPO_66	HR_DOC_TYPE	traffic	2
41.03	21.34	SPO.MK0038A_00008_10	STA-MK0038A	traffic	2
36.13	-5.35	GIB_SamplingPoint_30	GIB_Station	traffic	2
41.11	20.08	SPO.AL0207A_00008_10	STA-AL0207A	traffic	2
41.86	12.47	SPO.IT1837A_8_chemi_	STA.IT1837A	traffic	2
41.49	13.83	SPO.IT1183A_8_chemi_	STA.IT1183A	traffic	2
43.09	12.44	SPO.IT1182A_8_chemi_	STA.IT1182A	traffic	2
43.55	10.33	SPO.IT1560A_8_chemi_	STA.IT1560A	traffic	2
40.53	17.17	SPO.IT1607A_8_chemi_	STA.IT1607A	traffic	2
43.87	11.11	SPO.IT0945A_8_chemi_	STA.IT0945A	traffic	2
42.95	12.71	SPO.IT1900A_8_chemi_	STA.IT1900A	traffic	2
41.62	13.33	SPO.IT1200A_8_chemi_	STA.IT1200A	traffic	2
40.93	9.5	SPO.IT1309A_8_chemi_	STA.IT1309A	traffic	2
44.15	9.92	SPO.IT1661A_8_chemi_	STA.IT1661A	traffic	2
40.35	18.18	SPO.IT2044A_8_chemi_	STA.IT2044A	traffic	2
44.31	8.47	SPO.IT1145A_8_chemi_	STA.IT1145A	traffic	2
43.46	11.88	SPO.IT0832A_8_chemi_	STA.IT0832A	traffic	2
43.1	12.37	SPO.IT2004A_8_chemi_	STA.IT2004A	traffic	2
39.24	9.11	SPO.IT2056A_8_chemi_	STA.IT2056A	traffic	2
42.56	12.65	SPO.IT1011A_8_chemi_	STA.IT1011A	traffic	2
40.68	14.77	SPO.IT1504A_8_chemi_	STA.IT1504A	traffic	2
42	12.73	SPO.IT0887A_8_chemi_	STA.IT0887A	traffic	2
41.91	12.55	SPO.IT1834A_8_chemi_	STA.IT1834A	traffic	2
44.05	12.58	SPO.IT1044A_8_chemi_	STA.IT1044A	traffic	2
44.05	8.21	SPO.IT2038A_8_chemi_	STA.IT2038A	traffic	2
40.64	17.94	SPO.IT1701A_8_chemi_	STA.IT1701A	traffic	2
42	14.99	SPO.IT1800A_8_chemi_	STA.IT1800A	traffic	2
40.96	17.29	SPO.IT1817A_8_chemi_	STA.IT1817A	traffic	2
40.83	16.56	SPO.IT1818A_8_chemi_	STA.IT1818A	traffic	2
41.57	13.34	SPO.IT0872A_8_chemi_	STA.IT0872A	traffic	2
42.4	12.86	SPO.IT0867A_8_chemi_	STA.IT0867A	traffic	2
41.23	16.29	SPO.IT1822A_8_chemi_	STA.IT1822A	traffic	2
40.85	14.25	SPO.IT0898A_8_chemi_	STA.IT0898A	traffic	2
45.4	9.28	SPO.IT1290A_8_chemi_	STA.IT1290A	traffic	2
43.79	11.23	SPO.IT0860A_8_chemi_	STA.IT0860A	traffic	2
40.85	14.27	SPO.IT1491A_8_chemi_	STA.IT1491A	traffic	2
43.84	10.51	SPO.IT1001A_8_chemi_	STA.IT1001A	traffic	2

40.64	17.95	SPO.IT1618A_8_chemi_	STA.IT1618A	traffic	2
41.13	14.79	SPO.IT0934A_8_chemi_	STA.IT0934A	traffic	2
43.71	10.41	SPO.IT1409A_8_chemi_	STA.IT1409A	traffic	2
41.45	12.89	SPO.IT2174A_8_chemi_	STA.IT2174A	traffic	2
43.77	11.27	SPO.IT0861A_8_chemi_	STA.IT0861A	traffic	2
40.71	8.55	SPO.IT1243A_8_chemi_	STA.IT1243A	traffic	2
41.88	12.51	SPO.IT0828A_8_chemi_	STA.IT0828A	traffic	2
41.8	12.61	SPO.IT1841A_8_chemi_	STA.IT1841A	traffic	2
42.09	11.8	SPO.IT2171A_8_chemi_	STA.IT2171A	traffic	2
40.36	18.17	SPO.IT1932A_8_chemi_	STA.IT1932A	traffic	2
41.11	16.89	SPO.IT1606A_8_chemi_	STA.IT1606A	traffic	2
42.42	12.11	SPO.IT0863A_8_chemi_	STA.IT0863A	traffic	2
43.34	11.32	SPO.IT2184A_8_chemi_	STA.IT2184A	traffic	2
40.96	17.28	SPO.IT2143A_8_chemi_	STA.IT2143A	traffic	2
40.59	17.12	SPO.IT2021A_8_chemi_	STA.IT2021A	traffic	2
41.12	16.87	SPO.IT1621A_8_chemi_	STA.IT1621A	traffic	2
38.11	15.65	SPO.IT1989A_8_chemi_	STA.IT1989A	traffic	2
42.76	11.11	SPO.IT1592A_8_chemi_	STA.IT1592A	traffic	2
44.11	9.96	SPO.IT2229A_8_chemi_	STA.IT2229A	traffic	2
38.11	13.35	SPO.IT1078A_8_chemi_	STA.IT1078A	traffic	2
38.15	13.33	SPO.IT1082A_8_chemi_	STA.IT1082A	traffic	2
38.12	13.33	SPO.IT1552A_8_chemi_	STA.IT1552A	traffic	2
37.52	15.1	SPO.IT1718A_8_chemi_	STA.IT1718A	traffic	2
38.01	12.55	SPO.IT1898A_8_chemi_	STA.IT1898A	traffic	2
40.92	14.79	SPO.IT2227A_8_chemi_	STA.IT2227A	traffic	2
41.08	14.33	SPO.IT1487A_8_chemi_	STA.IT1487A	traffic	2
41.05	14.38	SPO.IT1488A_8_chemi_	STA.IT1488A	traffic	2
40.99	14.42	SPO.IT2218A_8_chemi_	STA.IT2218A	traffic	2
40.98	14.21	SPO.IT2271A_8_chemi_	STA.IT2271A	traffic	2
-20.95	55.31	SPO.FR38018_8	STA-FR38018	traffic	2
28.98	-13.51	SP_35024002_8_8	STA_ES2080A	traffic	2
42.1	11.78	SPO.IT2250A_8_chemi_	STA.IT2250A	traffic	2
44.18	28.64	SPO.RO0131A_00008_10	STA-RO0131A	traffic	2
44.43	26.12	SPO.RO0070A_00008_10	STA-RO0070A	traffic	2
44.44	26.15	SPO.RO0067A_00008_10	STA-RO0067A	traffic	2
44.87	24.85	SPO.RO0098A_00008_10	STA-RO0098A	traffic	2
44.94	26	SPO.RO0175A_00008_10	STA-RO0175A	traffic	2
44.33	23.78	SPO.RO0080A_00008_10	STA-RO0080A	traffic	2
47.03	21.95	SPO.RO0109A_00008_10	STA-RO0109A	traffic	2
44.92	26.04	SPO.RO0179A_00008_10	STA-RO0179A	traffic	2
40.66	35.84	SPO.TR050431_0008	STA-TR05043	traffic	2
41.02	29.1	SPO.TR341131_0008	STA-TR34113	traffic	2
42.47	14.21	SPO.IT1264A_8_chemi_	STA.IT1264A	traffic	2
42.45	14.2	SPO.IT1208A_8_chemi_	STA.IT1208A	traffic	2
37.94	23.65	SPO.GR0030A_00008_12	STA-GR0030A	traffic	2

Table A.2. Industrial sampling points measuring NO_2 concentrations clustering at 1-R metric in clusters (2) and (3) representing "outliers" from the industrial sampling points in the rest of Europe - cluster (1)

Lat	Lon	Sampling point	station	type	Cluster
59.41	27.28	SPO.EE0019A_00008_10	STA-EE0019A	industrial	3
59.46	24.7	SPO.EE0015A_00008_50	STA-EE0015A	industrial	3
60.43	22.1	SPO.FI00460_00008_10	STA-FI00460	industrial	3
64.37	-21.76	SPO.IS0025A_00008_10	STA-IS0025A	industrial	3
64.06	-21.99	SPO.IS0002A_00008_10	STA-IS0002A	industrial	3
64.04	-22.6	SPO.IS0039A_00008_10	STA-IS0039A	industrial	3
64.04	-21.99	SPO.IS0046A_00008_10	STA-IS0046A	industrial	3
36.71	36.22	SPO.TR310222_0008	STA-TR31022	industrial	3
66.08	-17.34	SPO.IS0049A_00008_10	STA-IS0049A	industrial	3
43.22	-8.26	SP_15024005_8_8	STA_ES2050A	industrial	2
43.41	-7.99	SP_15070006_8_8	STA_ES0094A	industrial	2
43.47	-7.84	SP_15070011_8_8	STA_ES1959A	industrial	2
37.25	-1.78	SP_04035002_8_8	STA_ES1547A	industrial	2
36.79	-5.53	SP_11026001_8_8	STA_ES1822A	industrial	2
37.87	-8.72	SPO.PT04003_00008_10	STA-PT04003	industrial	2
43.54	-7.74	SP_27033001_8_8	STA_ES1285A	industrial	2
40.98	15.64	SPO.IT1193A_8_chemi_	STA.IT1193A	industrial	2
28.09	-17.12	SP_38036002_8_8	STA_ES2053A	industrial	2
28.15	-16.52	SP_38006003_8_8	STA_ES1766A	industrial	2

Table A.3. Traffic sampling points measuring PM_{10} concentrations clustering at EuD metric.
.

Lat	Lon	Sampling point	station	type	Cluster
50.07	14.43	SPO.CZ_ALEGA_PM10_86	STA.CZ_ALEG	traffic	2
49.41	14.68	SPO.CZ_CTABA_PM10_41	STA.CZ_CTAB	traffic	2
50.07	14.45	SPO.CZ_AVRSA_PM10_40	STA.CZ_AVRS	traffic	2
50.06	14.54	SPO.CZ_APRUA_PM10_19	STA.CZ_APRU	traffic	2
50.09	14.44	SPO.CZ_AKALA_PM10_41	STA.CZ_AKAL	traffic	2
50.11	14.5	SPO.CZ_AVYNA_PM10_41	STA.CZ_AVYN	traffic	2
50.68	14	SPO.CZ_UULDA_PM10_13	STA.CZ_UULD	traffic	2
49.73	13.4	SPO.CZ_PPLAA_PM10_40	STA.CZ_PPLA	traffic	2
52.23	21.01	SPO_PL0592A_5_001	STA_PL0592A	traffic	2
50.07	14.4	SPO.CZ_ASMIA_PM10_41	STA.CZ_ASMI	traffic	2
50.2	15.85	SPO.CZ_HHKBA_PM10_41	STA.CZ_HHKB	traffic	2
45.48	16.78	HR_DOC_TYPE_D_SPO_43	HR_DOC_TYPE	traffic	2
51.75	19.43	SPO_PL0635A_5_001	STA_PL0635A	traffic	2
49.07	18.92	SPO.SK0039A_00005_10	STA-SK0039A	traffic	2
41.91	22.42	SPO.MK0034A_00005_10	STA-MK0034A	traffic	2
51.76	19.47	SPO_PL0589A_5_001	STA_PL0589A	traffic	2
50.06	19.95	SPO_PL0641A_5_001	STA_PL0641A	traffic	2
42	21.44	SPO.MK0043A_00005_10	STA-MK0043A	traffic	3
48.74	19.15	SPO.SK0214A_00005_10	STA-SK0214A	traffic	2
48.92	20.87	SPO.SK0265A_00005_10	STA-SK0265A	traffic	2
41.99	21.42	SPO.MK0047A_00005_10	STA-MK0047A	traffic	4
41.03	21.34	SPO.MK0038A_00005_10	STA-MK0038A	traffic	5
42	20.97	SPO.MK0041A_00005_10	STA-MK0041A	traffic	6
49.96	14.06	SPO.CZ_SBERA_PM10_40	STA.CZ_SBER	traffic	2
41.71	21.77	SPO.MK0040A_00005_10	STA-MK0040A	traffic	7
48.73	21.26	SPO.SK0267A_00005_10	STA-SK0267A	traffic	2
45.56	18.7	HR_DOC_TYPE_D_SPO_43	HR_DOC_TYPE	traffic	2
50.06	19.93	SPO_PL0012A_5_001	STA_PL0012A	traffic	2
50.02	20.99	SPO_PL0647A_5_001	STA_PL0647A	traffic	2
59.86	17.65	SPO.SE156417_00005_1	STA-SE15641	traffic	8
62.39	17.3	SPO.SE09569_00005_10	STA-SE09569	traffic	9
65.69	-18.1	SPO.IS0008A_00005_10	STA-IS0008A	traffic	10
40.66	35.84	SPO.TR050431_0005	STA-TR05043	traffic	11
39.94	32.88	SPO.TR060231_0005	STA-TR06023	traffic	12
39.93	32.86	SPO.TR060831_0005	STA-TR06083	traffic	12
40.19	29.08	SPO.TR160231_0005	STA-TR16023	traffic	13
40.55	34.97	SPO.TR190231_0005	STA-TR19023	traffic	14
41.07	28.99	SPO.TR340531_0005	STA-TR34053	traffic	15
41	28.84	SPO.TR341031_0005	STA-TR34103	traffic	15
41.02	29.1	SPO.TR341131_0005	STA-TR34113	traffic	15
41.03	29.02	SPO.TR341231_0005	STA-TR34123	traffic	15
41.01	28.95	SPO.TR341431_0005	STA-TR34143	traffic	15
40.99	29.07	SPO.TR341931_0005	STA-TR34193	traffic	16
38.72	35.47	SPO.TR380131_0005	STA-TR38013	traffic	17
40.77	29.94	SPO.TR410831_0005	STA-TR41083	traffic	18
40.77	30.41	SPO.TR540231_0005	STA-TR54023	traffic	19
41.29	36.33	SPO.TR550631_0005	STA-TR55063	traffic	11

39.75	37.01	SPO.TR580231_0005	STA-TR58023	traffic	20
40.98	27.5	SPO.TR590231_0005	STA-TR59023	traffic	21
40.32	36.55	SPO.TR600331_0005	STA-TR60033	traffic	22

Table A.4. Industrial sampling points measuring PM_{10} concentrations clustering at (1-R)xEuD metric.

.

Lat	Lon	Sampling point	station	type	cluster
28.46	-16.26	SP_38038023_10_49	STA_ES2000A	industrial	2
41.98	21.47	SPO.MK0031A_00005_10	STA-MK0031A	industrial	2
28.46	-16.27	SP_38038027_10_49	STA_ES2038A	industrial	2
41.04	21.36	SPO.MK0037A_00005_10	STA-MK0037A	industrial	2
39.95	-0.06	SP_12009007_10_46	STA_ES1877A	industrial	2
41.44	22.01	SPO.MK0044A_00005_10	STA-MK0044A	industrial	2
41.99	21.65	SPO.MK0045A_00005_10	STA-MK0045A	industrial	2
44.37	8.29	SPO.IT1192A_5_nephel	STA.IT1192A	industrial	2
28.46	-16.28	SP_38038021_10_49	STA_ES1976A	industrial	2
28.03	-15.41	SP_35026004_10_49	STA_ES2004A	industrial	2
28.09	-17.12	SP_38036002_10_49	STA_ES2053A	industrial	2
28.15	-16.52	SP_38006003_10_46	STA_ES1766A	industrial	2
28.67	-17.78	SP_38008001_10_49	STA_ES1979A	industrial	2
28.08	-16.56	SP_38017002_10_46	STA_ES1761A	industrial	2
28.46	-16.27	SP_38038025_10_46	STA_ES2020A	industrial	2
28.39	-16.36	SP_38011006_10_47	STA_ES1772A	industrial	2
28.11	-16.58	SP_38017003_10_46	STA_ES1760A	industrial	2
28.69	-17.76	SP_38037001_10_49	STA_ES1980A	industrial	2
28.05	-16.54	SP_38017001_10_46	STA_ES1762A	industrial	2
28.5	-13.85	SP_35017002_10_47	STA_ES1866A	industrial	2
28.38	-16.36	SP_38011004_10_49	STA_ES1756A	industrial	2
28.38	-16.37	SP_38011005_10_49	STA_ES1764A	industrial	2
28.38	-16.37	SP_38011007_10_49	STA_ES2022A	industrial	2
28.46	-16.26	SP_38038010_10_49	STA_ES1131A	industrial	2
27.9	-15.45	SP_35002001_10_A	STA_ES1745A	industrial	2
27.87	-15.39	SP_35002002_10_A	STA_ES1744A	industrial	2
28.03	-15.41	SP_35016011_10_49	STA_ES1451A	industrial	2
40.2	29.21	SPO.TR160421_0005	STA-TR16042	industrial	2
40.03	27.05	SPO.TR170221_0005	STA-TR17022	industrial	2
40.53	34.94	SPO.TR190321_0005	STA-TR19032	industrial	2
36.71	36.22	SPO.TR310222_0005	STA-TR31022	industrial	2
38.74	35.38	SPO.TR380321_0005	STA-TR38032	industrial	2
40.79	29.54	SPO.TR410221_0005	STA-TR41022	industrial	2
40.79	29.52	SPO.TR410521_0005	STA-TR41052	industrial	2
40.75	29.79	SPO.TR411021_0005	STA-TR41102	industrial	2
40.7	29.88	SPO.TR411122_0005	STA-TR41112	industrial	2
41.22	36.46	SPO.TR550521_0005	STA-TR55052	industrial	2
66.08	-17.34	SPO.IS0049A_00005_10	STA-IS0049A	industrial	2

Appendix 2 - List of abbreviations for country codes

This report follows Eurostat country glossary[19] where Member States of the European Union (EU) and other countries have been assigned a two-letter country code, always written in capital letters, and often used as an abbreviation in statistical analyses, tables, figures or maps.

Country Code	Country Name	Country Code	Country Name
AD	Andorra	IS	Iceland
AL	Albania	IT	Italy
AT	Austria	LT	Lithuania
BA	Bosnia Hercegovina	LU	Luxemburg
BE	Belgium	LV	Latvia
BG	Bulgaria	ME	Montenegro
CH	Switzerland	MK	North Macedonia
CY	Cyprus	MT	Malta
CZ	Czech Republic	NL	Netherlands
DE	Germany	NO	Norway
DK	Denmark	PL	Poland
EE	Estonia	PT	Portugal
ES	Spain	RO	Romania
FI	Finland	RS	Serbia
FR	France	SE	Sweden
GE	Georgia	SI	Slovenia
GI	Gibraltar	SK	Slovakia
GR	Greece	TR	Turkey
HR	Croatia	UK	United Kingdom
HU	Hungary	XK	Kosovo
IE	Ireland		

The following groups of countries and their abbreviations are used all through the text:

EU27	= AT+BE+BG+CY+CZ+DE+DK+EE+ES+FI+FR+GR+HR+ +HU+IE+IT+LT+LU+LV+MT+NL+PL+PT+RO+SE+SI+SK
EU28	= **EU27** + UK
EEA33*	= **EU28**+CH+IS+NO+TR (but without Liechtenstein)
EEA39*	= **EEA33*** +AL+BA+ME+MK+RS+XK
ALL	= **EEA39***+ AD+GI

Note that of all countries in the above list only Georgia, GE, did not report any operational monitoring data in 2017.

[19] https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Country_codes, last viewed June 2020

RICARDO

ricardo.com