

How to use Air Quality Downloads page

Date: June 2024

Document History

Version	Date	Author(s)	Remarks
0.1	2023-06-29	AE	
0.2	2023-09-08	AE	
0.3	2023-11-14	AE	
0.4	2024-04-15	AE	
0.5	2024-05-03	AE	
0.6	2024-06-06	AE	

Contents

Overview	3
1. Air quality Download service page	4
1.1. Filters.....	4
1.2. Download data	7
1.3. Metadata.....	7
1.4. Vocabulary.....	8
1.5. Documentation	9
2. Air quality Download service API	10
2.1. Examples of scripts	13
Python script example.....	13
Postman example.....	13
2.2. Visualization of Parquet files.....	14
3. Parquet schema definition	15

Overview

These guidelines explain how to use the new Air quality Download Service from both web application and API.

- **Web application:** <https://eeadmz1-downloads-webapp.azurewebsites.net>
- **API:** <https://eeadmz1-downloads-api-appservice.azurewebsites.net/swagger/index.html>

Throughout the document, the different options that exist in the web application to download the data in the most useful way for users will be shown in detail, as well as some examples of how to use the API to download without having to use the web application.

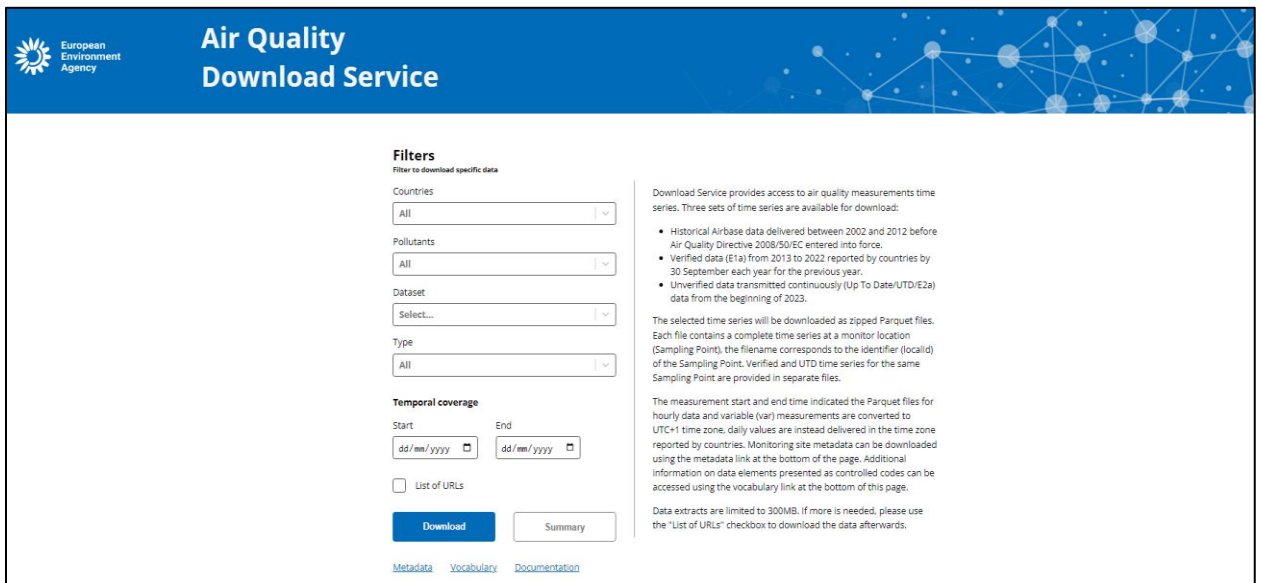
Downloads can be made in two different formats while using the above-mentioned services:

1. **Parquet files:** They contain the complete time series of the corresponding sampling points.
2. **URLs:** The download generates a CSV with the URLs corresponding to the parquet files of each of the desired sampling points.

1. Air quality Download service page

The [download page](#) is divided into two main blocks.

1. **Interface:** This feature allows you to filter and select the desired for downloading directly to your browser (on the left side).
2. **Description of the Download Service:** Detailed information about this service is available on the right side of the screen.



The screenshot shows the 'Air Quality Download Service' page. On the left, there is a 'Filters' section with the following options: Countries (All), Pollutants (All), Dataset (Select...), Type (All), and Temporal coverage (Start and End date pickers, and a 'List of URLs' checkbox). Below the filters are 'Download' and 'Summary' buttons, and links for 'Metadata', 'Vocabulary', and 'Documentation'. On the right, there is a text block explaining the service: 'Download Service provides access to air quality measurements time series. Three sets of time series are available for download: Historical Airbase data delivered between 2002 and 2012 before Air Quality Directive 2008/50/EC entered into force; Verified data (E1a) from 2013 to 2022 reported by countries by 30 September each year for the previous year; and Unverified data transmitted continuously (Up To Date/UTD/E2a) data from the beginning of 2023. The selected time series will be downloaded as zipped Parquet files. Each file contains a complete time series at a monitor location (Sampling Point), the filename corresponds to the identifier (localid) of the Sampling Point. Verified and UTD time series for the same Sampling Point are provided in separate files. The measurement start and end time indicated the Parquet files for hourly data and variable (var) measurements are converted to UTC+1 time zone, daily values are instead delivered in the time zone reported by countries. Monitoring site metadata can be downloaded using the metadata link at the bottom of the page. Additional information on data elements presented as controlled codes can be accessed using the vocabulary link at the bottom of this page. Data extracts are limited to 300MB. If more is needed, please use the "List of URLs" checkbox to download the data afterwards.'

In addition, the download page provides information related to the [metadata](#), a link to list of vocabularies describing definitions that appear in the downloaded files and, finally a link to this documentation.

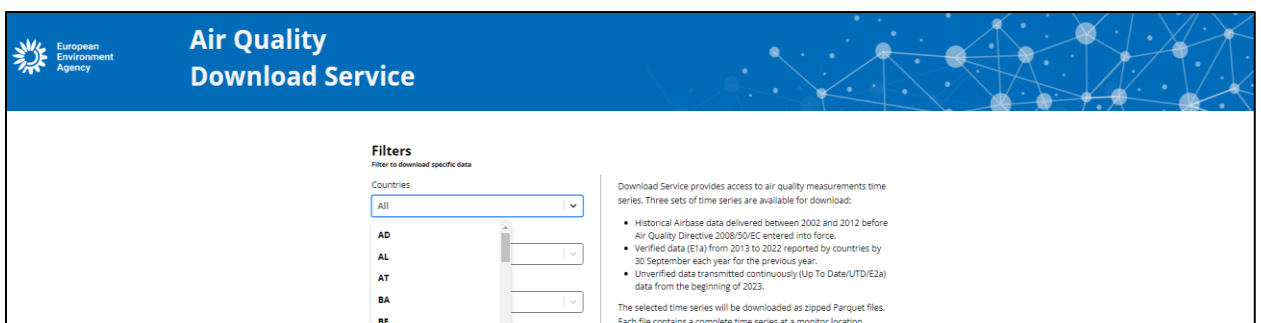
1.1. Filters

In this section, the operation of the filters within the application will be explained. It is important to know that none of the filters are mandatory, so if any of the filters are not filled, the download will be performed on all possible values.

For example, if the pollutants filter is empty, the corresponding files will be download with all the pollutants that exist in the system.

Countries

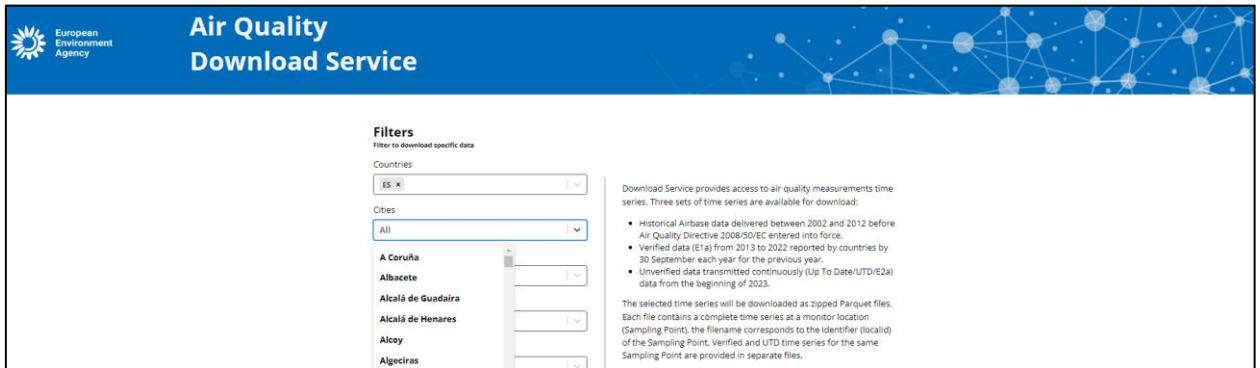
Using this filter allows the selection of one or more specif countries for which the available data will be downloaded. If no country is selected, data for all countries will be downloaded.



This screenshot shows the 'Air Quality Download Service' page with the 'Countries' filter expanded. The 'All' option is selected in the dropdown menu, and a list of countries is visible below: AD, AL, AT, BA, and BE. The rest of the page content, including the service description on the right, remains the same as in the previous screenshot.

Cities

If any of the selected countries contain cities, a new filter will be displayed containing the corresponding cities. Here, as in the country filter, it is possible to select one or more cities, or leave it blank to return all cities in the country.

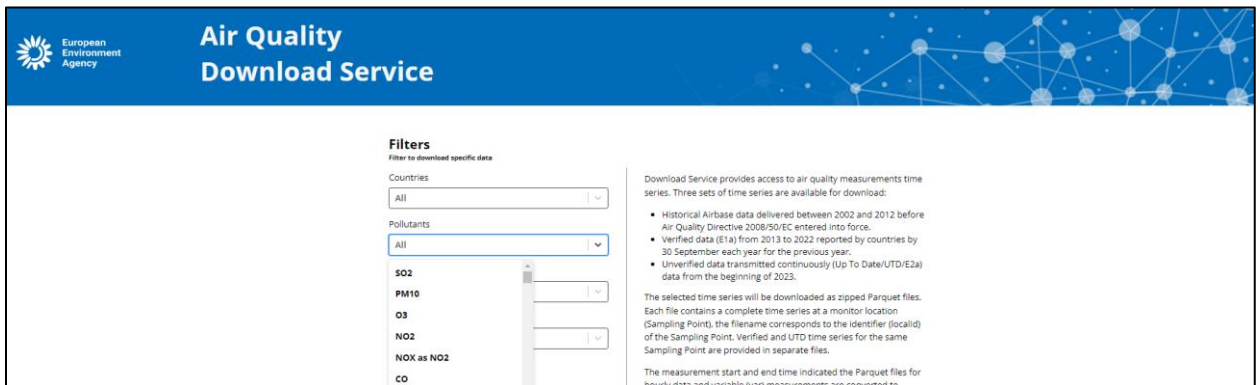


The screenshot shows the 'Air Quality Download Service' interface. The 'Filters' section is titled 'Filter to download specific data'. Under 'Countries', a dropdown menu is set to 'ES'. Below it, a 'Cities' filter is active, showing a list of cities: A Coruña, Albacete, Alcalá de Guadaíra, Alcalá de Henares, Alcoy, and Algeciras. To the right, there is explanatory text about the data sources and download formats.

The information on the geometries of each city has been extracted from the data source available in [Urban audit](#). With this and with the information that contains the coordinates of where each station is located available in this [link](#), the intersection has been made to determine the stations of each of the cities.

Pollutants

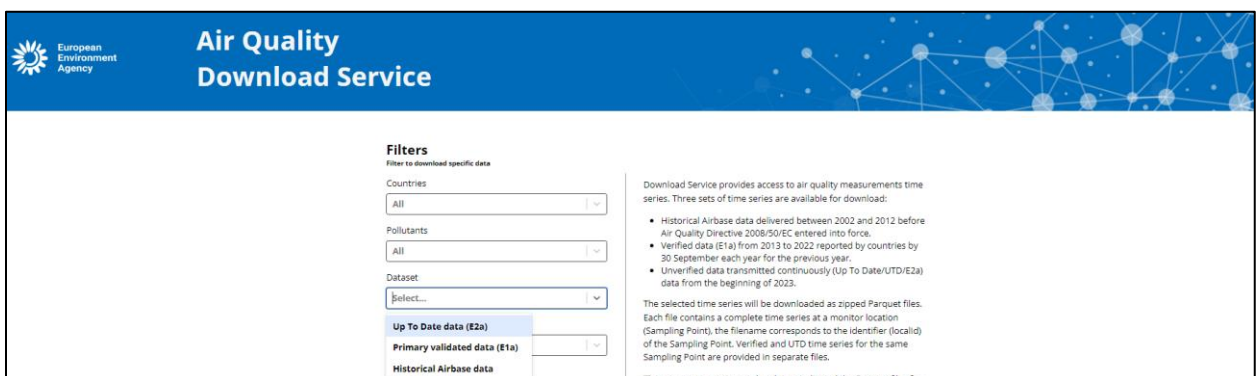
It is possible to filter the information for the desired pollutants. Like the rest of the filters, it is possible to select one or more options, or none and therefore get data for all the pollutants.



The screenshot shows the 'Air Quality Download Service' interface. The 'Filters' section is titled 'Filter to download specific data'. Under 'Pollutants', a dropdown menu is set to 'All'. Below it, a list of pollutants is visible: SO2, PM10, O3, NO2, NOx as NO2, and CO. To the right, there is explanatory text about the data sources and download formats.

Dataset

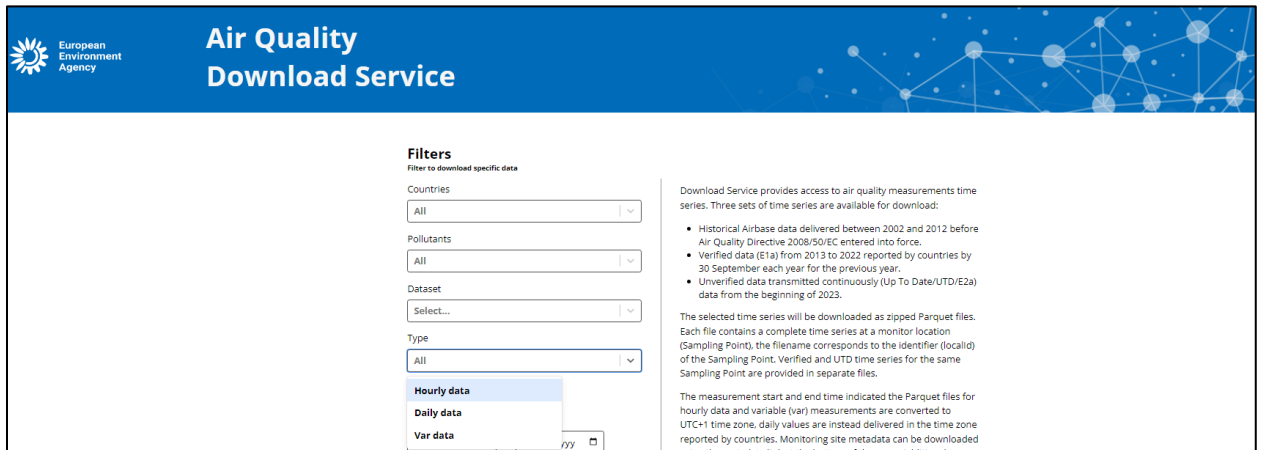
It is possible to filter and download the verified data (E1a) reported by countries, the most recent unverified data transmitted continuously (Up-To-Date/UTD/E2a), or the historical Airbase data delivered between 2002 and 2012. This is possible through the "Dataset" filter and in this case, the value is mandatory.



The screenshot shows the 'Air Quality Download Service' interface. The 'Filters' section is titled 'Filter to download specific data'. Under 'Dataset', a dropdown menu is set to 'Select...'. Below it, a list of dataset options is visible: Up To Date data (E2a), Primary validated data (E1a), and Historical Airbase data. To the right, there is explanatory text about the data sources and download formats.

Type

It is possible to filter and download the data by selecting the type of data. This type represents whether the data collected is obtaining the values in hourly, daily or variable intervals (intervals different than the previous observation such as weekly, monthly, etc.). It is possible to select one option, or none and therefore get data for all the types.



The screenshot shows the 'Air Quality Download Service' interface. The 'Filters' section includes dropdown menus for 'Countries', 'Pollutants', 'Dataset', and 'Type'. The 'Type' dropdown is open, showing options: 'All', 'Hourly data', 'Daily data', and 'Var data'. The 'Hourly data' option is highlighted. To the right, there is explanatory text about the data sets and download format.

Filters
Filter to download specific data

Countries: All
Pollutants: All
Dataset: Select...
Type: All

Hourly data
Daily data
Var data

Download Service provides access to air quality measurements time series. Three sets of time series are available for download:

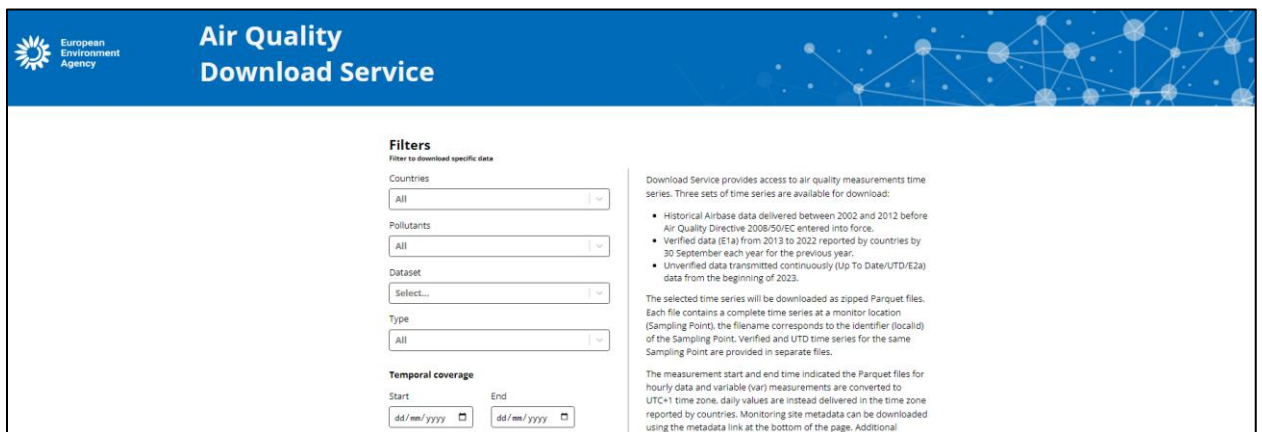
- Historical Airbase data delivered between 2002 and 2012 before Air Quality Directive 2008/50/EC entered into force.
- Verified data (E1a) from 2013 to 2022 reported by countries by 30 September each year for the previous year.
- Unverified data transmitted continuously (Up To Date/UTD/E2a) data from the beginning of 2023.

The selected time series will be downloaded as zipped Parquet files. Each file contains a complete time series at a monitor location (Sampling Point), the filename corresponds to the identifier (localid) of the Sampling Point. Verified and UTD time series for the same Sampling Point are provided in separate files.

The measurement start and end time indicated the Parquet files for hourly data and variable (var) measurements are converted to UTC+1 time zone, daily values are instead delivered in the time zone reported by countries. Monitoring site metadata can be downloaded using the metadata link at the bottom of the page.

Temporal coverage

It is possible to filter and download the data with a specific temporal coverage. For this it is necessary to select the beginning (Start) and the end (End) of the period to be downloaded. In case Up To Date data (E2a) has been selected, it is possible to select a typical download temporal coverage. If no temporal coverage is selected full data will be downloaded.



The screenshot shows the 'Air Quality Download Service' interface. The 'Filters' section includes dropdown menus for 'Countries', 'Pollutants', 'Dataset', and 'Type'. Below these, there is a 'Temporal coverage' section with 'Start' and 'End' date pickers. The 'Start' and 'End' fields are empty, showing the format 'dd/mm/yyyy'.

Filters
Filter to download specific data

Countries: All
Pollutants: All
Dataset: Select...
Type: All

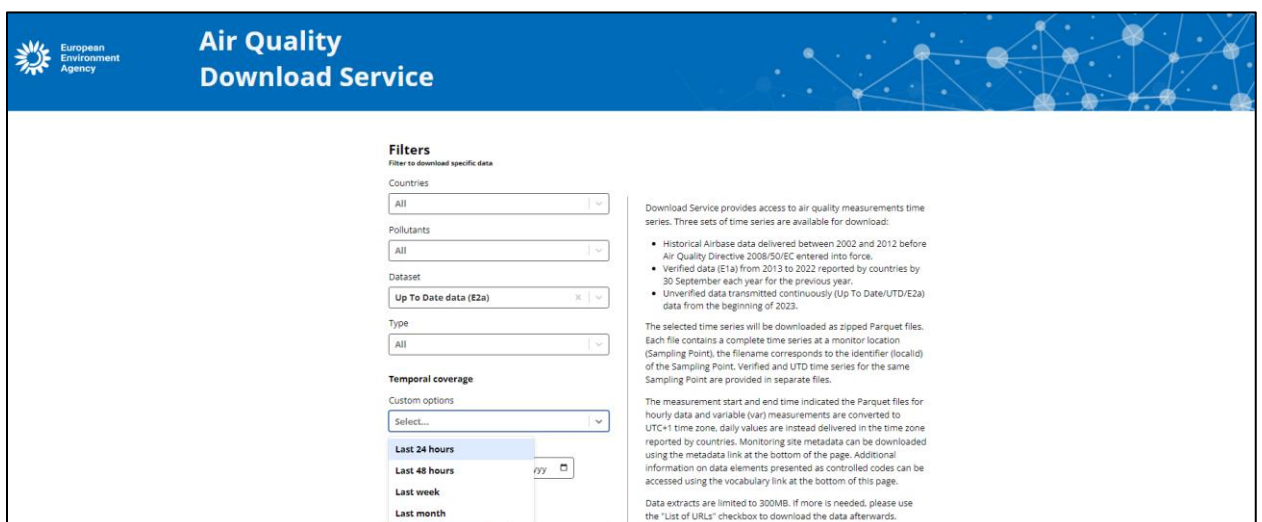
Temporal coverage
Start: dd/mm/yyyy
End: dd/mm/yyyy

Download Service provides access to air quality measurements time series. Three sets of time series are available for download:

- Historical Airbase data delivered between 2002 and 2012 before Air Quality Directive 2008/50/EC entered into force.
- Verified data (E1a) from 2013 to 2022 reported by countries by 30 September each year for the previous year.
- Unverified data transmitted continuously (Up To Date/UTD/E2a) data from the beginning of 2023.

The selected time series will be downloaded as zipped Parquet files. Each file contains a complete time series at a monitor location (Sampling Point), the filename corresponds to the identifier (localid) of the Sampling Point. Verified and UTD time series for the same Sampling Point are provided in separate files.

The measurement start and end time indicated the Parquet files for hourly data and variable (var) measurements are converted to UTC+1 time zone, daily values are instead delivered in the time zone reported by countries. Monitoring site metadata can be downloaded using the metadata link at the bottom of the page.



The screenshot shows the 'Air Quality Download Service' interface. The 'Filters' section includes dropdown menus for 'Countries', 'Pollutants', 'Dataset', and 'Type'. Below these, there is a 'Temporal coverage' section with a 'Custom options' dropdown menu. The 'Custom options' dropdown is open, showing options: 'Last 24 hours', 'Last 48 hours', 'Last week', and 'Last month'. The 'Last 24 hours' option is highlighted.

Filters
Filter to download specific data

Countries: All
Pollutants: All
Dataset: Up To Date data (E2a)
Type: All

Temporal coverage
Custom options: Select...
Last 24 hours
Last 48 hours
Last week
Last month

Download Service provides access to air quality measurements time series. Three sets of time series are available for download:

- Historical Airbase data delivered between 2002 and 2012 before Air Quality Directive 2008/50/EC entered into force.
- Verified data (E1a) from 2013 to 2022 reported by countries by 30 September each year for the previous year.
- Unverified data transmitted continuously (Up To Date/UTD/E2a) data from the beginning of 2023.

The selected time series will be downloaded as zipped Parquet files. Each file contains a complete time series at a monitor location (Sampling Point), the filename corresponds to the identifier (localid) of the Sampling Point. Verified and UTD time series for the same Sampling Point are provided in separate files.

The measurement start and end time indicated the Parquet files for hourly data and variable (var) measurements are converted to UTC+1 time zone, daily values are instead delivered in the time zone reported by countries. Monitoring site metadata can be downloaded using the metadata link at the bottom of the page. Additional information on data elements presented as controlled codes can be accessed using the vocabulary link at the bottom of this page.

Data extracts are limited to 300MB. If more is needed, please use the "List of URLs" checkbox to download the data afterwards.

URLs

Depending on what is desired, it is possible to download only the URLs of the selected files instead of the entire parquet file. To do this, activate the checkbox (tick the box).

List of URLs

1.2. Download data

Once the desired filters have been selected, to download the data, it is necessary to click the "Download" button and after waiting, the files will be automatically downloaded in the browser.

There is the possibility, instead of downloading, to obtain a summary of the number of files and the size of the download by pressing the "Summary" button. It can be helpful to have an idea of the amount of data and files that are going to be downloaded, and if the maximum limit is exceeded (300 MB), download the URLs to access the files as there is a limit to the download. As mentioned above, to download the URLs, activate the "List of URLs" checkbox.

Filters
Filter to download specific data

Countries: ES

Cities: Madrid

Pollutants: SO2

Dataset: Up To Date data (E2a)

Type: Hourly data

Temporal coverage

Custom options: Last 24 hours

Start: 05/06/2024

End: 06/06/2024

List of URLs

Download **Summary**

[Metadata](#) [Vocabulary](#) [Documentation](#)

Download Service provides access to air quality measurements time series. Three sets of time series are available for download:

- Historical Airbase data delivered between 2002 and 2012 before Air Quality Directive 2008/50/EC entered into force.
- Verified data (E1a) from 2013 to 2022 reported by countries by 30 September each year for the previous year.
- Unverified data transmitted continuously (Up To Date/UTD/E2a) data from the beginning of 2023.

The selected time series will be downloaded as zipped Parquet files. Each file contains a complete time series at a monitor location (Sampling Point), the filename corresponds to the identifier (locality) of the Sampling Point. Verified and UTD time series for the same Sampling Point are provided in separate files.

The measurement start and end time indicated the Parquet files for hourly data and variable (var) measurements are converted to UTC+1 time zone, daily values are instead delivered in the time zone reported by countries. Monitoring site metadata can be downloaded using the metadata link at the bottom of the page. Additional information on data elements presented as controlled codes can be accessed using the vocabulary link at the bottom of this page.

Data extracts are limited to 300MB. If more is needed, please use the "List of URLs" checkbox to download the data afterwards.

1.3. Metadata

By clicking on the "Metadata", it is possible to access a new page where there is a summary of all the metadata related to all the different countries. Here it is possible to find an interactive table which shows information on Air Quality Measurements reported within AQ e-Reporting to the Eionet CDR repository.

Direct link:

https://discomap.eea.europa.eu/App/AQViewer/index.html?fqn=Airquality_Dissem.b2g.measurements

Within this page it is possible to download the data and filter by different options in which the user is interested. It must be considered that on this page, it is only possible to download the metadata sent by each country about the networks, stations, sampling points, etc., but it is not possible to download the

data collected by the different stations on different pollutants. For this, it is necessary to use the download page described in this document.

Air Quality Measurements (data flow D)
This viewer shows information on Air Quality Measurements reported within AO e-Reporting

Linked tables + Share Download CSV -

Country	B-G Namespace	Year	Air Quality Network	Air Quality Network Name	Timezone	Air Quality Station Eol Code	Air Quality Station Nat Code	Air Quality Station Name	Sampling Point Id	Air Pollutant	Longitude	Latitude	Altitude	Altitude Unit	Air Quality Station Area	Air Quality Station Type
Andorra	AD.GovernAndorra.AQ	2022	NET-AD001A	Xarxa de vigilancia de qualitat de l'aire del Departament de Medi Ambient - Govern Andorra	UTC+01	AD0942A	942	Escaldes-Engordany	SPO-AD0942A-0001	SO2	1.5391	42.5097	1080	m	urban	background
Andorra	AD.GovernAndorra.AQ	2022	NET-AD001A	Xarxa de vigilancia de qualitat de l'aire del Departament de Medi Ambient - Govern Andorra	UTC+01	AD0942A	942	Escaldes-Engordany	SPO-AD0942A-0005	PM10	1.5391	42.5097	1080	m	urban	background
Andorra	AD.GovernAndorra.AQ	2022	NET-AD001A	Xarxa de vigilancia de qualitat de l'aire del Departament de Medi Ambient - Govern Andorra	UTC+01	AD0942A	942	Escaldes-Engordany	SPO-AD0942A-0007	O3	1.5391	42.5097	1080	m	urban	background
Andorra	AD.GovernAndorra.AQ	2022	NET-AD001A	Xarxa de vigilancia de qualitat de l'aire del Departament de Medi Ambient - Govern Andorra	UTC+01	AD0942A	942	Escaldes-Engordany	SPO-AD0942A-0008	NO2	1.5391	42.5097	1080	m	urban	background

Showing 1-30 of 76312

Filters

Country: [all]

Year: [all]

Air Quality Network: [all]

Timezone: [all]

Air Quality Station Eol Code: [all]

Sampling Point Id: [all]

Air Pollutant: [all]

Longitude: From [] To []

Latitude: From [] To []

Altitude: From [] To []

Air Quality Station Area: [all]

1.4. Vocabulary

From the "Vocabulary" link, users can find a summary of the various vocabularies that are used and appear in the downloaded parquet files. By clicking on the different links, users will be directed to the Eionet Data Dictionary, which provides the definitions for each of these vocabularies. It must be considered that on this page, it is only possible to download metadata (networks, stations, sampling points, ...), but it is not possible to download the data collected by the different stations about the different pollutants. For this, it is necessary to use the download page described in this document.

Vocabulary

Vocabulary	Link
Aggregation Types	https://dd.eionet.europa.eu/vocabulary/daq/primaryObservation
Air Quality Pollutants	https://dd.eionet.europa.eu/vocabulary/daq/pollutant
Concentration Units	https://dd.eionet.europa.eu/vocabulary/uom/concentration
Observation Validity	https://dd.eionet.europa.eu/vocabulary/daq/observationvalidity
Observation Verification	https://dd.eionet.europa.eu/vocabulary/daq/observationverification

Close

EIONET
Data Dictionary

You are here: Eionet» Data Dictionary» Vocabulary

Help and documentation
Datasets
Tables
Data elements
Schemas
Vocabularies
Services
Namespaces

Vocabulary: AQD - Air Quality Pollutants

← Back to set ▾ Exports

Folder	aq (Air Quality Directive e-Reporting)
Identifier	pollutant
Label	AQD - Air Quality Pollutants
Base URI	http://dd.eionet.europa.eu/vocabulary/aq/pollutant/
Registration status	Released 05 Apr 2024 07:50:15
Type	Common
Definition	AQD Decision 2011/850/EU
Version	0.0

Bound elements for concepts

Element	Type	Status
hasProtectionTarget	Vocabulary	Released ----- 21 Jul 2014
mandatoryUnit	Quantitative	Released ----- 27 Jul 2020
measurementEquipment	Vocabulary	Released ----- 21 Jul 2014

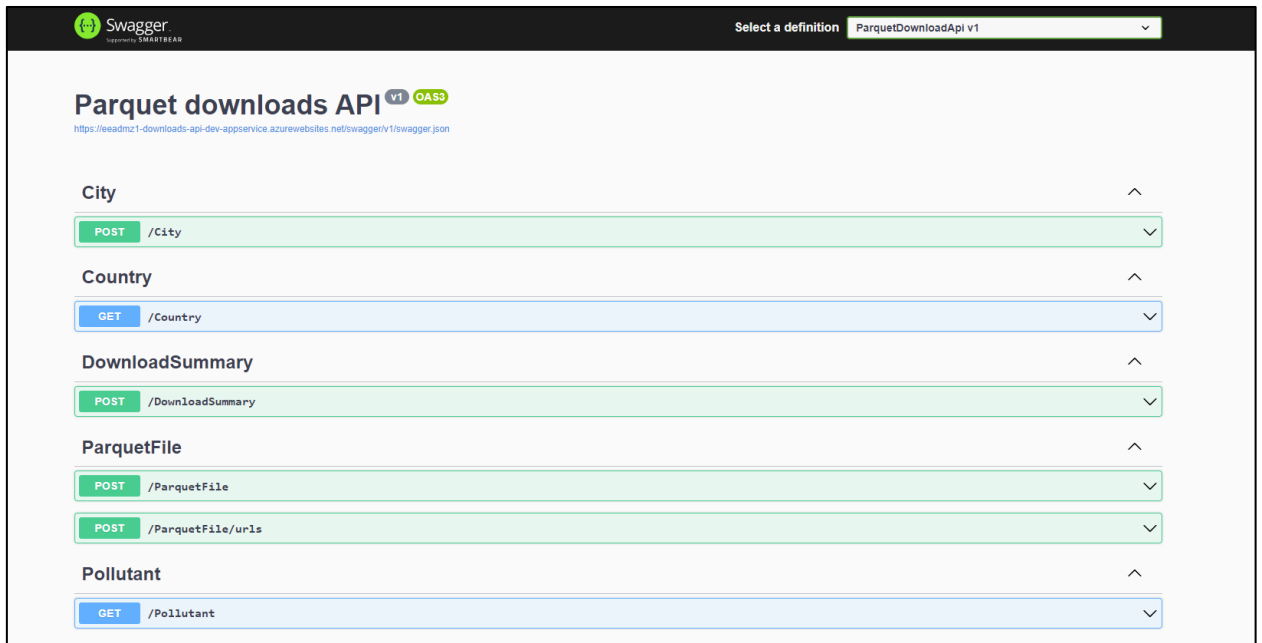
1.5. Documentation

Through the link "Documentation" a new tab opens in the browser accessing this document.

2. Air quality Download service API

There is a possibility of accessing an API to consult the available methods that exist in everything related to downloading files.

- **API:** <https://eadmz1-downloads-api-appservice.azurewebsites.net/swagger/index.html>



Here the desired method is selected. To make a request from the API, it is necessary to press the “Try it out” button of the desired endpoint.

The most interesting endpoints are:

- **ParquetFile:** Download zip file with all the filtered parquet files.
- **ParquetFile/urls:** Download csv file with all the URLs corresponding to the filtered parquets.

Both endpoints have the same parameters to be able to perform filtering and obtain the desired information.



After filling in the values of the request, press the "Execute" button. The values must be filled in as follows:

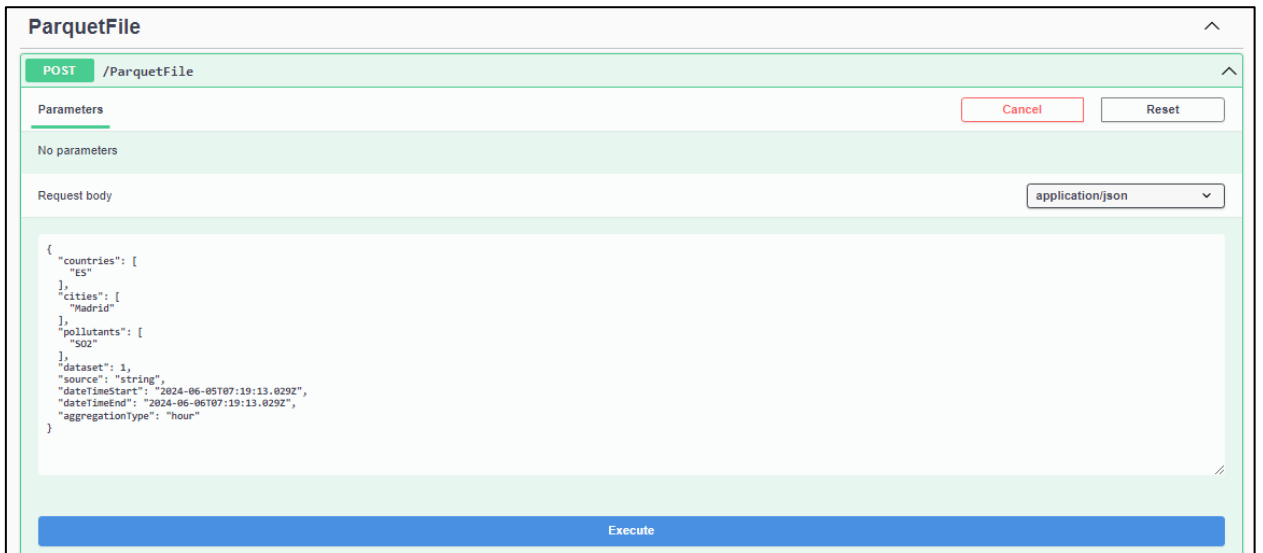
- **countries:** values with country code separated by commas. If no value is entered and an empty list is passed, the filter will not be applied and the data for all the countries will be downloaded. Example: ["IS", "DK", "NO"].
- **cities:** name of cities separated by commas. If no value is entered and an empty list is passed, the filter will not be applied and the data for all cities will be downloaded. Example: ["Berlin", "Madrid"].
- **pollutant:** It is necessary to enter the **URLs** corresponding to each pollutant or de **notation** of each of the pollutants separated by commas. The URL can be obtained by clicking on the different pollutants in this [page](#) and copying the value of the **Concept URI**. The notations can be obtained directly from the previous page. If no value is entered and an empty list is passed, the filter will not be applied and the data for all pollutants will be downloaded. For example: properties: ["http://dd.eionet.europa.eu/vocabulary/aq/pollutant/3", "http://dd.eionet.europa.eu/vocabulary/aq/pollutant/1"] or ["SO2", "CO"].



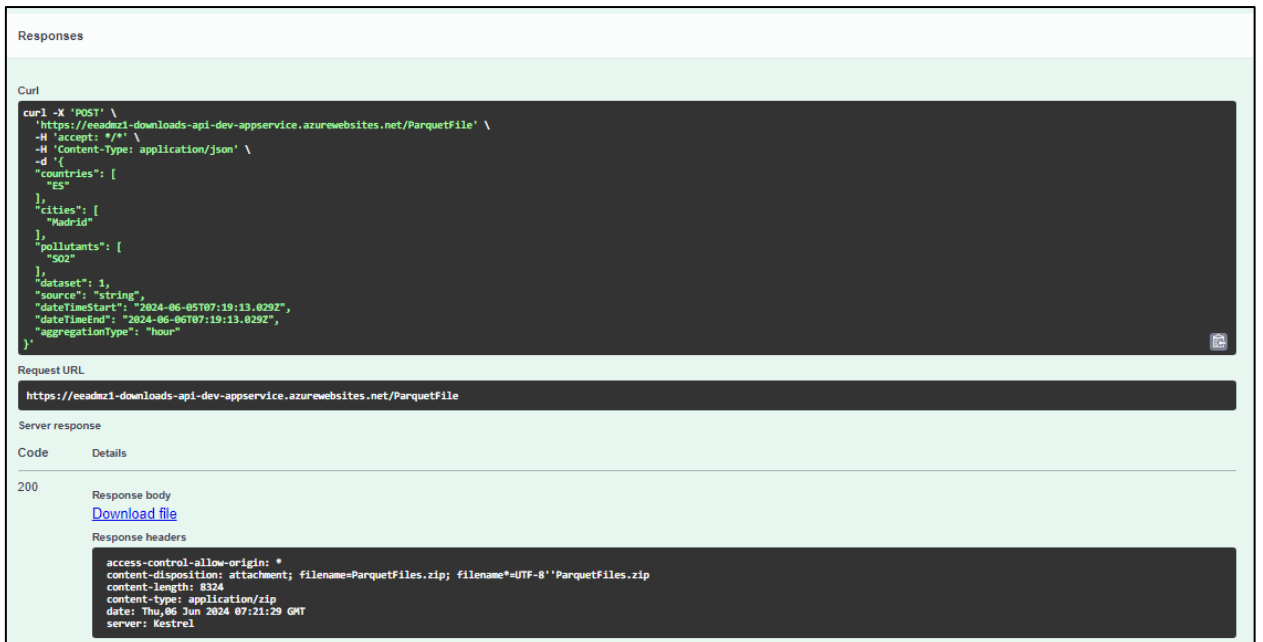
- **dataset:** value of the dataset:
 1. Unverified data transmitted continuously (Up-To-Date/UTD/E2a) data from the beginning of 2023.
 2. Verified data (E1a) from 2013 to 2022 reported by countries by 30 September each year for the previous year.
 3. Historical Airbase data delivered between 2002 and 2012 before Air Quality Directive 2008/50/EC entered into force.
- **source:** optional value that indicates where the download was made from. If the download is made from the official download website, source will contain the value "website". If it is done from the API directly or from a custom script, the parameter can be omitted or enter values such as "API", "Custom script". The objective is to be able differentiate in the system what it has been downloaded from the web and that directly using the API.
- **dateTimeStart:** value that indicates from which date and time user want to filter the download. If parameter is not pa Format: yyyy-mm-ddTHH:MM:SSZ. Example: 2024-05-27T12:00:19Z.
- **dateTimeEnd:** value that indicates until which date and time user want to filter the download. Format: yyyy-mm-ddTHH:MM:SSZ. Example: 2024-05-28T12:00:19Z.

If **dateTimeStart** and **dateTimeEnd** parameters are not included in the request, the filter for the temporal coverage will not be applied and the entire set of data will be downloaded.

- **aggregationType:** represents whether the data collected is obtaining the values:
 1. Hourly data.
 2. Daily data.
 3. Variable intervals (different than the previous observations such as weekly, monthly, etc.)



Finally, click on “Download file” to get a zipped file with all the downloaded parquet files.



2.1. Examples of scripts

Python script example

Below is an example of a Python script to download parquet files from Azure with certain filters. Keep in mind that the downloadPath and the fileName refer to where the downloaded file will be saved, so this is different for different users.

```
import requests

apiUrl = "https://eeadmz1-downloads-api-dev-appservice.azurewebsites.net/"
endpoint = "ParquetFile"
downloadPath = "localPath\\"
fileName = "download_data.zip"

# Request body
request_body = {
    "countries": ["ES"],
    "cities": ["Madrid"],
    "pollutants": ["SO2"],
    "dataset": 1,
    "dateTimeStart": "2024-06-05T07:19:00Z",
    "dateTimeEnd": "2024-06-06T12:07:19.000Z",
    "aggregationType": "hour"
}

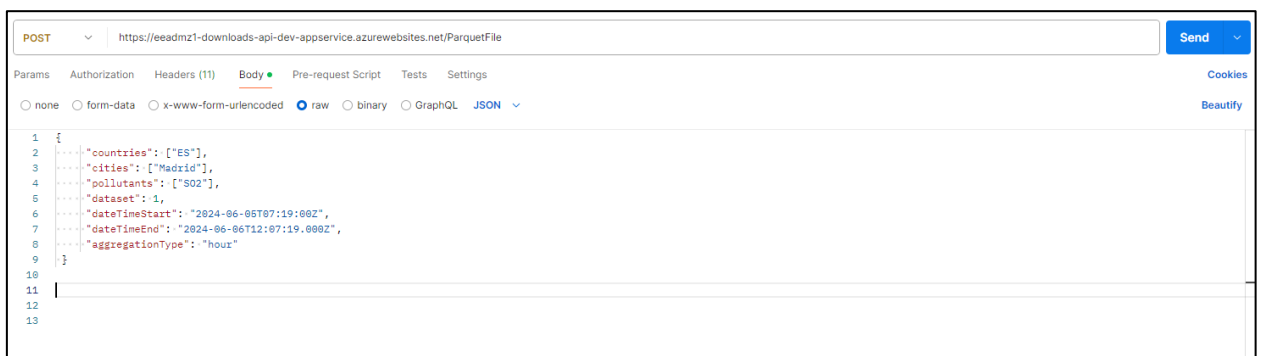
# A get request to the API
downloadFile = requests.post(apiUrl+endpoint, json=request_body).content

# Store in local path
output = open(downloadPath+fileName, 'wb')
output.write(downloadFile)
```

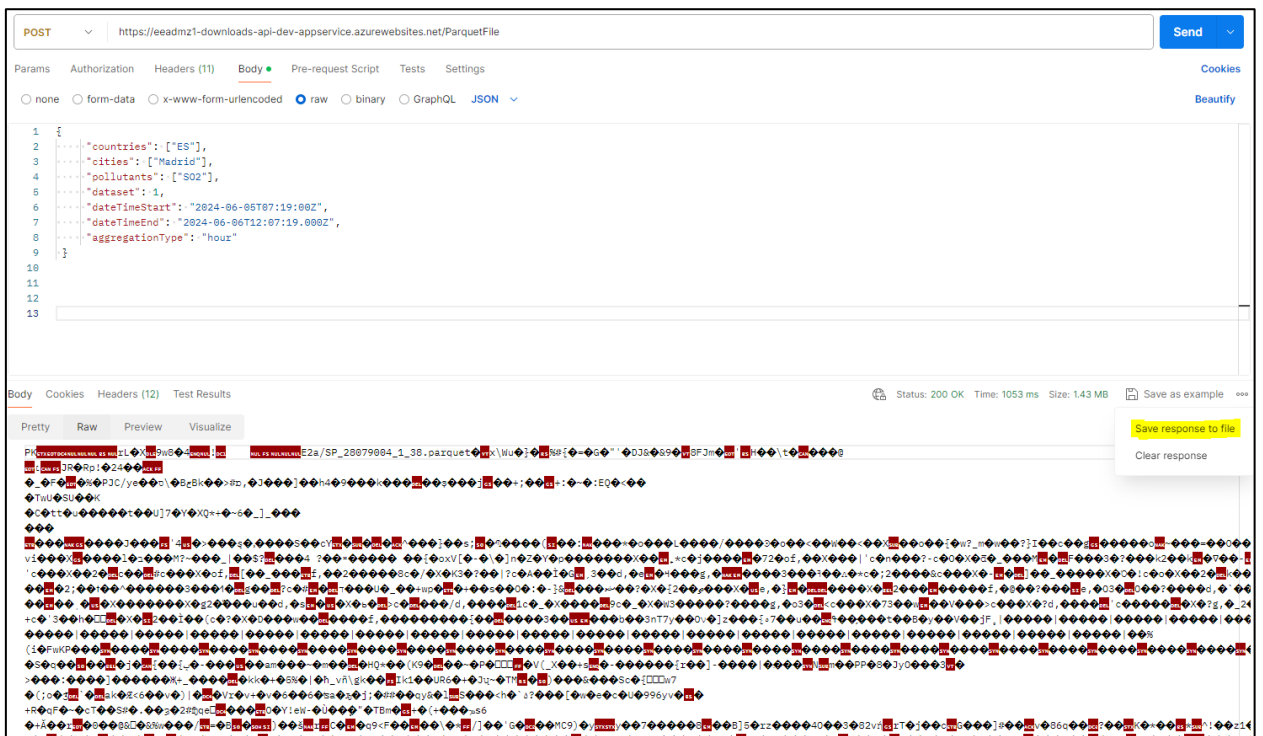
Postman example

In this second example, how to obtain the parquet files through Postman is described.

Once postman is open, the first thing is to configure the POST call and enter the body with the desired information to be downloaded. It is important to set the body type as JSON.



When the call ends, something like the image below can be seen. Next, press the "Save response to file", which gives the option to save a zipped file with the corresponding parquets files.



2.2. Visualization of Parquet files

There are different online pages to be able to view the parquet files or convert them to CSV to be able to read them without any special tool. [Parquet Viewer](#) example of one of them.

3. Parquet schema definition

The Parquet files contain the most relevant information reported by the countries. In the files various attributes can be found and are shown below:

- **Samplingpoint:** Identifier known as “local id” of the sampling point. Has the country code prefix. This identifier is unique to each station and created by national reporters, therefore there is no vocabulary for this parameter.
- **Pollutant:** The pollutant identifier. Find [here](#) more detailed information on pollutants.
- **Start:** Beginning of the time interval in which the information has been reported.
Format: yyyy-mm-dd H:M:S
- **End:** End of the time interval in which the information has been reported.
Format: yyyy-mm-dd H:M:S
- **Value:** Numerical value that represents the measurement obtained for that pollutant in that time interval.
- **Unit:** Unit in which the measurement obtained is represented. Find [here](#) more detailed information on units.
- **AggType:** This is the primary observation used in the measurement and can be hour/day/var. It represents whether the data collected is obtaining the values in hourly, daily or variable intervals (intervals different than the previous observation such as weekly, monthly, etc.). Find [here](#) more detailed information on aggregation types.
- **Validity:** Represents the validity of the measurement reported for the specific pollutant within a specified time interval. Find [here](#) more detailed information on validities.
- **Verification:** Represents the verification status of the measurement reported for that pollutant in that time interval. Find [here](#) more detailed information on verification status.
- **ResultTime:** Represents the date and time in which the information of the file that contained the reported data was generated.
- **DataCapture:** The data capture associated with a primary observation. Percentage of the time for which a sample or observation has been taken.
- **FkObservationLog:** This column has no importance for the end user. This is a column for internal use.