Danmarks Tekniske Universitet

How Wildfire Season in Canada Affects Air Quality in Europe

Authors
Bruno Brenna Betti - s247230
Ernesto Guzman Saleh - s242930
Luc Pares - s232493
Ngoc Bao Thai - s242504

June 25, 2025

Contents

1 Executive summary

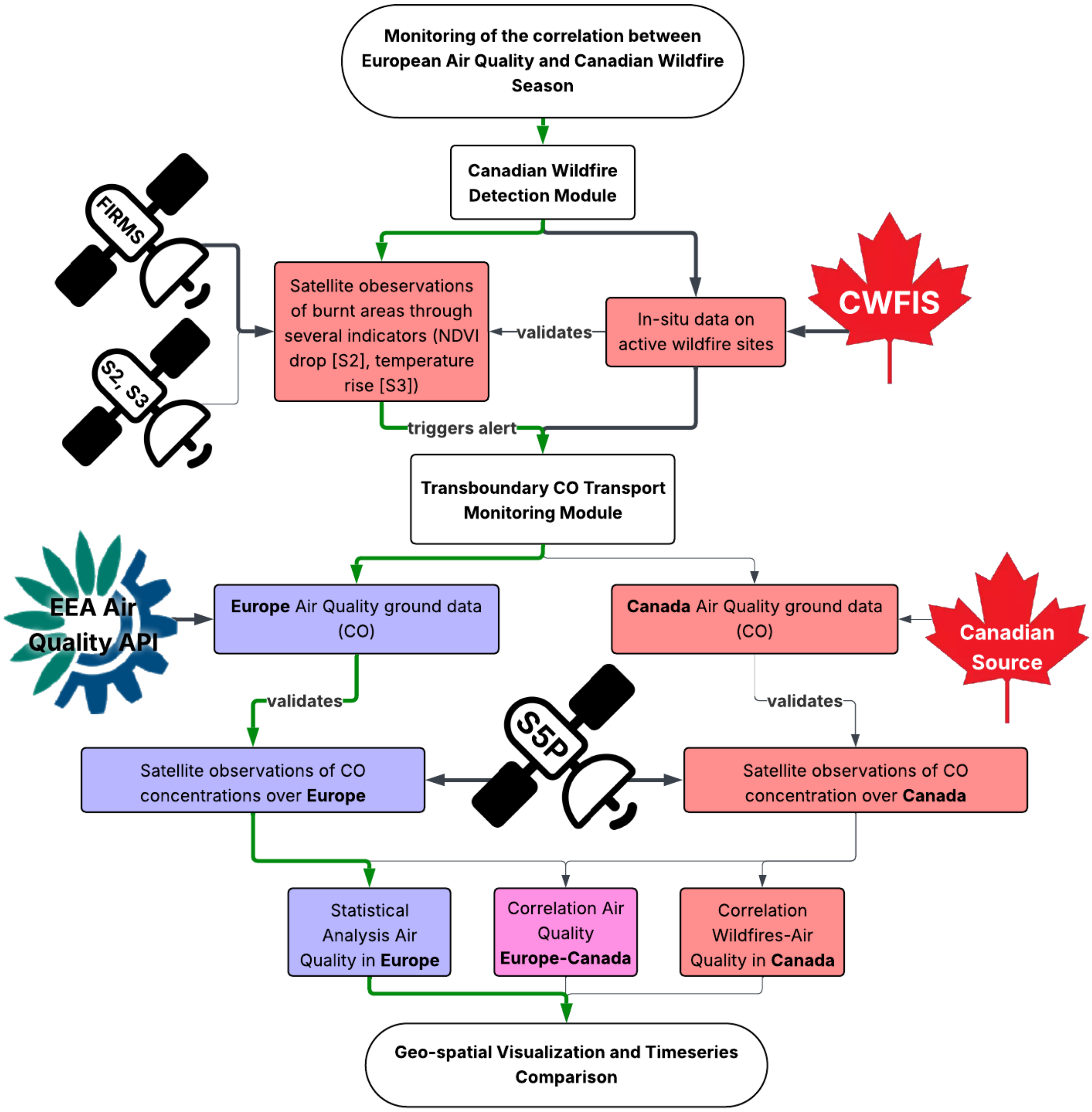


Figure 1: The service workflow. Bold arrows represent what is currently implemented, green arrows what is considered the Minimum Viable Product. Blue is for data or analysis over Europe and red over Canada - purple for both.

The service developed in the present report is meant to monitor transboundary air pollutants plumes originating primarily from Canada during the "wildfire season" (northern summer months), and that travel across the Atlantic Ocean before reaching Europe. By tracking these atmospheric pollutant concentrations, it evaluates the impact of such large-scale smoke events on air quality in multiple European cities. This fits into the scope of the Copernicus Atmosphere Monitoring Service [1], as it would support environmental monitoring efforts and inform public health responses, as well as potentially air traffic safety. As such, the service should be made accessible to governmental agencies e.g. public health- or environment-related to make informed decisions for affected matters such as regulation and announcement for outdoor workers [2]. The service should also be made available to the public for continuous air quality monitoring and regulation of outdoor activities [3]. Related stakeholders could be informed of daily new fires in Canada along with hourly monitoring of pollutants concentrations once the smoke plume arrives in Europe.

The service workflow is summarized in Figure 1. Overall, the service is composed of two modules that aim to:

1. Detect the start of the wildfire season in Canada, monitor its evolution over time, and when it reaches a certain intensity level,

2. Monitor and visualize the resulting CO concentration rise in Europe.

Due to the rapid dynamics of the wildfires, which can dramatically change in intensity within hours, all the data in the services is of real-time or near-real-time nature and is fetched automatically through various application programming interfaces (APIs).

The service's first module, called "Wildfire Detection", collects satellite retrievals from a service the Fire Information for Resource Management System (FIRMS) dataset, using the standard MODIS MOD14/MYD14 Fire and Thermal Anomalies product [4], and validates it against the reported active fire locations in Canada fetched from the Canadian Wildland Fire Information System (CWFIS) [5]. The module should then trigger an alert whenever the summed intensity of detected wildfires reaches over a certain point.

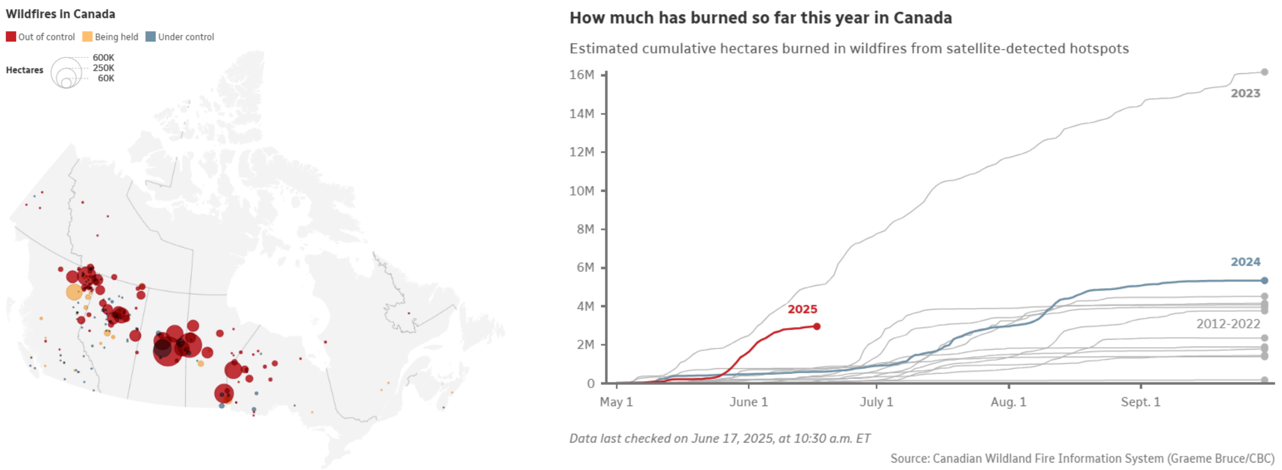
For the second module, it should be noted that at this initial product development stage, only one trace gas - carbon monoxide (CO) - is used as indicator for transboundary smoke from wildfires. The selection of CO is based on the available trace gases measured by the Sentinel-5P (S5P) against the air pollutants measured at ground stations as well as the inert nature of CO, as it is only subjected to physical processes such as dilution, dispersion, and advection in the atmosphere, rather than undergoing significant chemical transformation [6]. It is often treated as a conservative tracer in atmospheric modeling because of its long atmospheric lifetime (approx. 1 month [7]) and limited reactivity under typical ambient conditions [8]. As such, CO concentration is extracted from Copernicus' S5P measurements over all of Europe through the Google Earth Engine API [9], and compared to ground data over specific points, where known and reliable measuring stations are located. The measuring station network used is that of the European Environmental Agency (EEA) Air Quality datahub [10].

As shown by the bold arrows in the workflow (Figure 1), the service in its current state of development is close to the Minimal Viable Product (MVP). Further work could be made so that ultimately, the results can be used to estimate overarching metrics, such as the correlation factor between the various pollutants' concentrations close to the source in Canada, and remotely in Europe, or the lag-time induced by their long-range transport, and thus allow for predictive applications.

The current report presents the preliminary validation of the S5P CO retrievals with $>$20 ground-station measurements across Europe. The current analysis shows that there are 13 stations which have positive correlations (r) between S5P and ground-stations measurements, among which 6 have $r>0.3$. Compared with other published journals attempting to comprehensively validate the S5P CO estimates (with reported $r>0.3$) at site-/city-specific levels using extensive data assimilation pipeline, the current service shows reasonable validation although there are still rooms for improvement especially for locations with negative correlations. However, it should be noted that there are fundamentally differences between the methods to obtain total CO column in S5P algorithm compared to those measured by ground-stations. This uncertainty, along with other sources of errors, e.g. from intrinsic instrumental errors as well as algorithms in matching ground-stations with S5P flight path. This is then followed by possible improvements which can be implemented in further development as well as the draft plan to ensure that the service is fully automated and operational.

2 Introduction: The case and benefit to society

Robust air quality monitoring in Europe has become increasingly necessary after major transboundary pollution events such as the 2023 Canadian wildfires during the summer months which released pollutants into the atmosphere, including fine particulate matter (i.e. $PM_{2.5}$ and PM_{10}) and other compounds such as CO, O_3 and were transported as far as Europe [11] and even China [12]. Satellite data and atmospheric models demonstrated that wildfire smoke could transverse the Atlantic under the right meteorological conditions [13]. Episodes of pollution linked to wildfire smoke pose acute health risks [14] and other services, e.g. flight operations [15]. In May-June 2025, significant wildfires were again observed across Canada. Figure 2 presents the spatial distribution of wildfires (left), and the cumulative total burnt areas (right) update on 17 June 2025 [16].



(a) Wildfires spatial distribution in Canada.

(b) Cumulative burnt areas in Canada.

Figure 2: Statistics on wildfires in Canada (updated 17 June 2025 [16]).

Figure 3 shows the transport of pollutants across the Atlantic from wildfires in Canada to Europe in 2024 [11]. As such, there is increasing concern about air pollution in Europe, suggesting that there is a need to provide validated air pollution data in real time (or near real time) throughout Europe to enable timely alerts, support public health advisories, and inform environmental policy decisions [17]. Monitoring networks, combined with satellite observations (e.g. Copernicus Atmosphere Monitoring Service, CAMS), enhance Europe's ability to respond to and mitigate the effects of distant but impactful pollution events. This study directly supports the CAMS objectives by improving the monitoring and analysis of long-range transport of wildfire smoke and its subsequent impact on air quality in Europe. Furthermore, tapping on the Copernicus data source, it might be feasible to apply the same framework to other regions of the world where transboundary smokes are of concern, e.g. Southeast Asia [18].

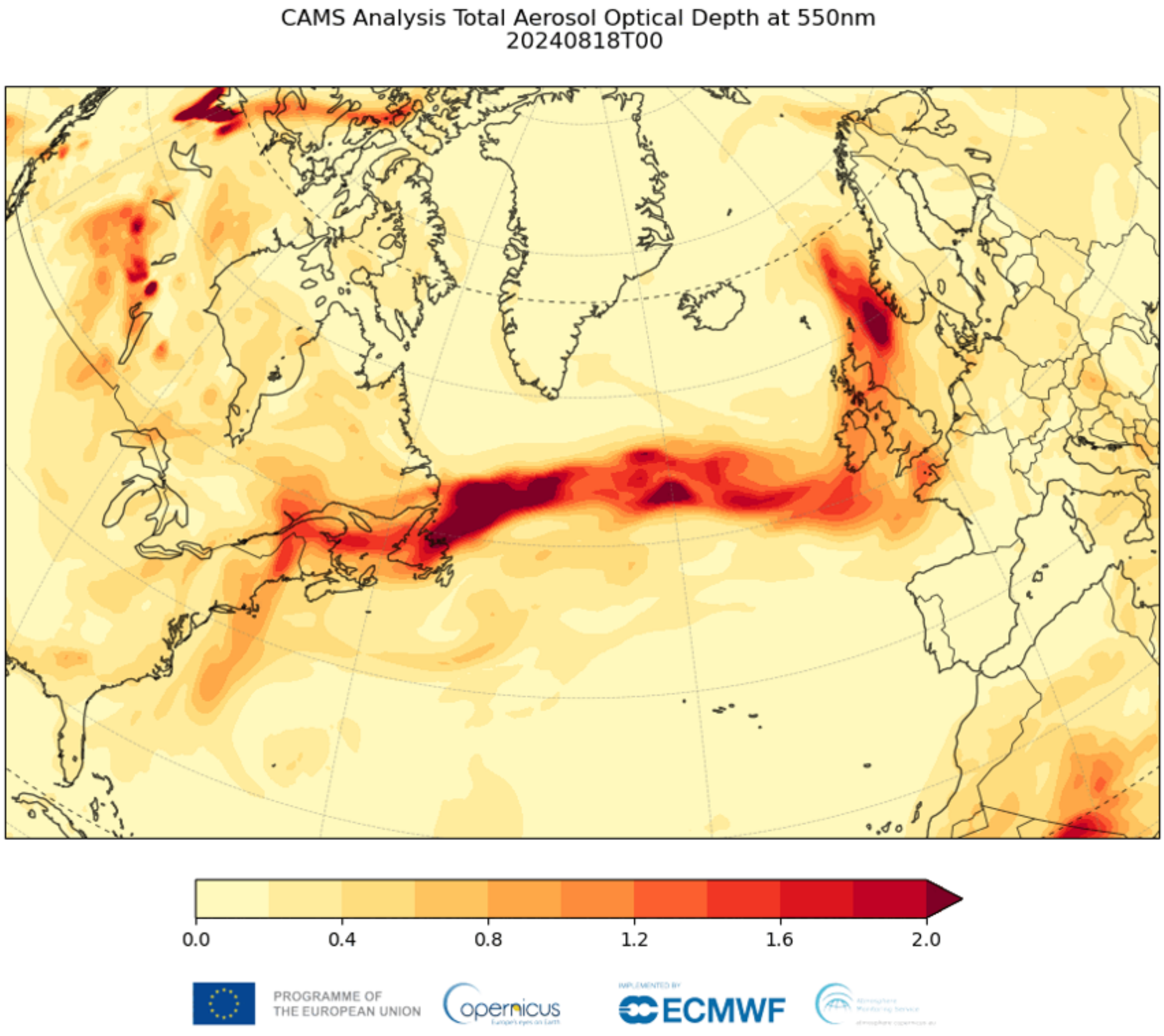


Figure 3: Transboundary pollutants from Canada wildfires to Europe in 2024 [11].

3 Methodology: ATBD (Algorithm Theoretical Baseline Document)

3.1 Instruments & data sources

Table 1: Summary of data source used in the entire service workflow. NA = not available.

Module	Data source	Spatial res.	Temp. res.	Timeliness	Implemented?
Wildfire Detection	FIRMS	1 km	3 hours	Daily	Yes
	CWFIS (in-situ)	NA	Daily	Daily	Yes
	Sentinel-2 MSI	10 m	5 days	5 days	No
	Sentinel-3 OCL	300 m	2 days	2 days	No
CO Transport	EEA (in-situ)	NA	1 hour	1 hour	Yes
	Sentinel-5P NRTI	5.5 x 7 km	3-hours	3-hours	Yes

Table 1 summarizes the data sources for the entire service workflow, with indications of whether the data have been implemented in the current product. The selection of both the spatial and temporal resolution of the datasets used in this service and further described in this section of the report, was driven by the needs of the intended users. Given the dynamic nature of transboundary air pollution events, near-real-time hourly data from ground stations and high-resolution daily satellite observations were deemed essential to enable timely assessments and actionable alerts.

3.1.1 Canadian wildfire detection module

This module primarily requires satellite data to assess where and when wildfires are active in Canada. As such, in the current state of development, the service relies on the Fire Information for Resource Management Systems (FIRMS), but it could be further developed to use Copernicus S2-MSI and S3-OCL instead to ensure its independence, as well as a spatially finer assessment by using several factors, e.g. classifying a pixel as an ongoing wildfire if and only if it shows both an NDVI reduction in S2 images, and a positive temperature anomaly in S3 retrievals. This would also allow a better spatial resolution (10-300m for Sentinel-2 and -3 and 1km for MODIS), but would mean a compromise in the temporal resolution and timeliness of the data (5-2 days vs. 3 hours), as summarized in Table 1.

The estimates from FIRMS are then compared to in-situ measurements reported by the Canadian Wildland Fire Information System (CWFIS), which keeps an up-to-date register of all active wildfire events in the country, as well as the surrounding areas, i.e. Northern United States, including Alaska.

Most importantly, all of the aforementioned data is acquired automatically through their platforms' dedicated APIs to ensure that the service always provides the most timely data possible.

FIRMS retrievals FIRMS provides in 3 hours frequency for active fires worldwide and is updated daily from the Moderate Resolution Imaging Spectroradiometer (MODIS) as well as the Visible Infrared Imaging Radiometer Suite (VIIRS) [19]. The data can be fetched from Google Earth Engine API FIRMS provides geo-located, centered around 1 km fire pixel based on the brightness temperature from channel 21/22 using the standard MODIS MOD14/MYD14 Fire and Thermal Anomalies product [4]. Furthermore, the data include attributes from VIIRS with a finer spatial resolution of 375 m which uses the I-4/5 channel brightness temperature to detect thermal anomalies [20]. Both sources provide a confidence scores, i.e. low, nominal and high, intended to help users gauge the quality of individual hotspots/fire pixels.

Validation – CWFIS wildfire data The CWFIS offers an extensive data collection [21], among which an "active wildland fires" dataset, updated daily. It is compiled from reports of many smaller-scale fire management agencies, including local agencies for each of the 13 provinces and territories that make up Canada. It provides, among other things, the coordinates of the fire, its size in hectares, and its status among the following:

- EX, for EXtinguished, when the fire is out,

- UC, for Under Control, when the fire is contained in a defined area, is not growing anymore, and is therefore bound to be rapidly extinguished,
- BH, for Being Held, when the fire is contained in a defined area but is still growing inside that area,
- OC, for Out of Control, when the fire is growing with no clearly defined maximal perimeter [22].

A very basic (and not very documented) API allows one to fetch this dataset from CWFIS' servers.

3.1.2 Transboundary CO transport monitoring module

There are two main sources of data used in this module, namely:

- Near-realtime S5P satellite products providing CO retrievals through its main instrument TROPOMI's shortwave infrared (2305–2385 nm) spectral band,
- In-situ CO measurements from ground stations affiliated to the European Environment Agency (EEA) Air Quality network, used for validation of the aforementioned retrievals.

Similarly, data from both sources are acquired using relevant APIs to guarantee that the service is fully automated. The validation data acquisition is performed first because the locations of the monitoring stations used had to be known before being able to compare with the values estimated by S5P at these locations.

Validation – EEA ground measurements The hourly ground-level CO measurements were obtained from the EEA Air Quality database [23]. Monitoring stations (7928 in total) were filtered based on several criteria, in order of importance, and thus of application:

- monitoring of the CO concentration,
- up-to-date and continued measurements,
- measurements for 'background' concentrations, as opposed to 'traffic' or 'industrial' measurements, and thus assumed representative of the average exposure of the population of that area,
- similar measurement method for comparability: the most represented method, NDIR (Non-Dispersive Infra-Red), was chosen,
- reasonable population representativity, keeping only the 'urban' stations, i.e. located in densely populated areas such as capitals, as opposed to 'rural' stations,
- 'reasonable' spatial coverage, i.e. good coverage in Western Europe where it is expected to be most affected by the plume of pollutants, along with some stations further east of Europe.

This last criterion was mainly based on the observation of the pollutant plume in 2024 (see Figure 3), which showed a west-to-east dispersion pattern across Europe and primarily affected Ireland, the United Kingdom, the Netherlands, and France, initially making landfall along the Atlantic coast. Following the criteria, a total of 26 measurement stations across 9 European countries were retained for analysis at this stage of service development. The spatial distribution of these stations is illustrated in Figure 4 and is further detailed in Table 2 in the appendix.

S5P retrievals The S5P data are readily available in Level 2 (L2) form, and downloadable through Copernicus' Catalog API [24], as well as Level 3 (L3) which is processed through Google Earth Engine [25]. L3 provides processed daily data files as S5P has a revisit time of approx. 24 hours for the region of interest. The S5P-L2 was processed and mosaicted into one daily file in this L3 version. However, because the satellite images are taken along its flight path as only a snapshot of a place at a given time of day, and given the dynamic changes in pollutant concentration in such critical events across all of Europe, it is crucial to capture the exact time of measure. As such, L2 data are used to extract

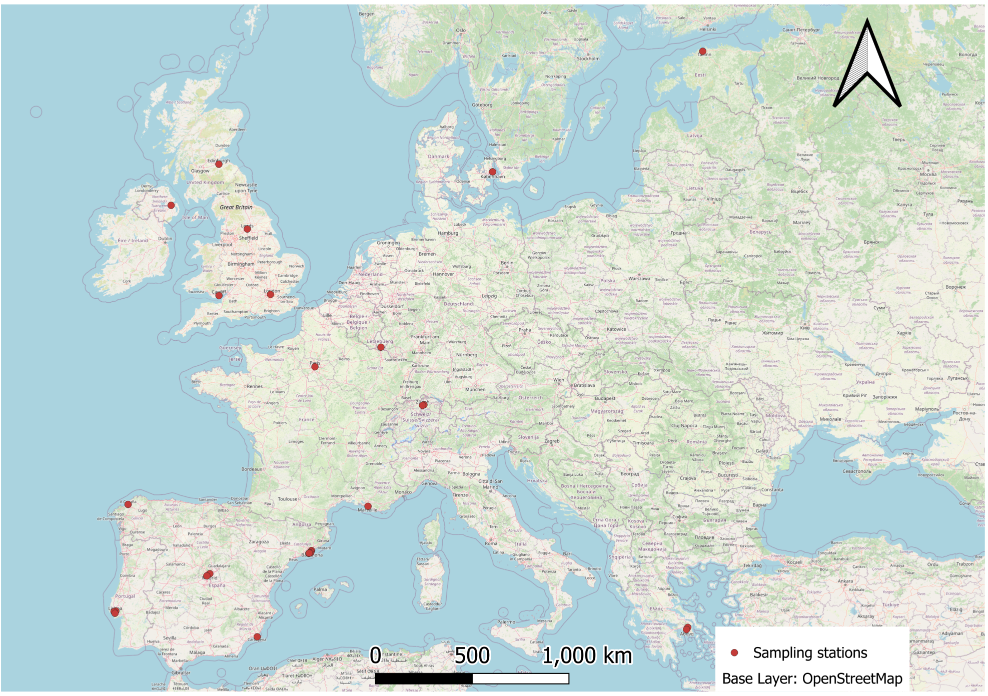


Figure 4: Ground stations across Europe used for satellite data validation.

the exact time and place of the CO estimates. The S5P-L2 product provides near real-time retrievals of the total CO column with a spatial resolution of approximately 5.5 km × 7 km. Since these data represent individual satellite overpasses, the spatial coverage appears as characteristic swaths over Europe, rather than continuous daily maps. This "snapshot" nature requires careful matching with ground-station measurements in time and space. Figure 5 illustrates the S5P-L2 CO data during a pass over Europe on the 20 May 2025.

3.2 Data processing

3.2.1 Canadian wildfire detection module

FIRMS retrievals The FIRMS data only from MODIS is currently implemented in the service. This uses MOD14 (Terra) and MYD14 (Aqua) algorithms based on the principle that fires emit strong thermal radiation, especially in the mid-infrared (MIR) region. The algorithm compares the brightness temperature of a pixel in the 4 µm band to its surrounding pixels and to the 11 µm band (i.e. reference band) to determine thermal anomalies [19]. A pixel is considered a potential fire pixel if its brightness temperature at 4 µm exceeds a nominal threshold of 310 K; however this threshold is dynamically adjusted based on local background conditions, e.g. surface type, time of the day and atmospheric conditions [26].

In the current state of development of the service, no further processing was made. Hence, this would currently only allow for a visual, qualitative comparison with the validation data, as detailed in section 4.2. However, further work could allow for quantitative validation of burnt areas in Canada. To make the output of the processed FIRMS data comparable to the validation data, several steps can be followed:

- To extract the size of a remotely detected wildfire site, the number of pixels composing a single contiguous patch can be counted, then multiplied by an approximation of the area covered by each pixel - in the case of MODIS, each pixel represents 1×1 km. Potentially, a tolerance threshold can be chosen to consider as still "contiguous", some pixels that are separated from the patch by only one or few empty pixels, for robustness to false negatives.

- To extract the location of a site, the coordinates of the centroid of the patch can be computed.

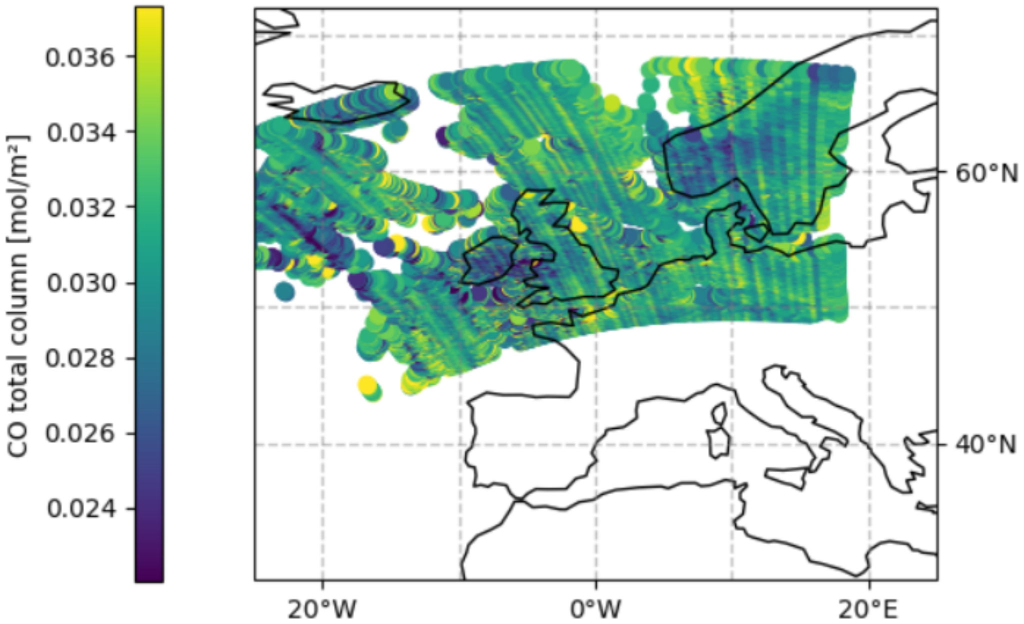


Figure 5: S5P-L2 CO total column estimates over Europe from a single satellite overpass on 20 May 2025, between 12:52 and 12:57 UTC.

- The status of the fire ("under control" or "being held"/"out of control") can also be inferred from comparing two successive frames - if the size of a patch has augmented with time, the fire is still progressing, otherwise it is being extinguished. Such information could be used in a more predictive development of the service.

All these operations could reasonably be implemented in the service, for example through computer-vision, or GIS-based manipulations.

Validation – CWFIS wildfire data The only data processing currently done for the dataset is to uniformize it, i.e. matching the different labelling system for its fires from one local agency to another system namely EX, UC, BH, OC. The initial labelling system quantifies fires in terms of "Prescribed" or "X% controlled", with X ranging from 0 to 100. "Prescribed" fires are triggered voluntarily, typically as a preventive measure to avoid future wildfires to extend past certain areas. They are entirely manmade but should still be included in the analysis, as they contribute to the emission of pollutants. They are therefore classified as "Under Control". To cover the "X% controlled" fires, a simple decision rule is mapped to the whole dataset:

- "100% controlled" fires are classified as "Under Control",
- "X% controlled" fires are classified as "Being Held",
- "0% controlled" fires are classified as "Out of Control".

Finally, all "Extinguished" fires are excluded from the dataset, as only active site-emitters of pollutants are monitored in this analysis.

3.2.2 Transboundary CO transport monitoring module

Validation – EEA ground measurements S5P provides data in the UTC+00 timezone, whereas hourly CO data from ground monitoring stations are provided in the UTC+01 timezone by EEA. Therefore, it was converted to UTC+00 in the service for temporal alignment. No further data processing is needed, as the data is directly provided as hourly values in mg/m^3.

S5P retrievals The data from S5P needed much more work to allow for a meaningful comparison with ground measurements.

The algorithm implemented in the service begins by reading S5P-L2 satellite data, which is not available on an hourly basis but is instead collected during satellite overpasses. These overpasses occur at specific times and do not provide homogeneous spatial coverage across Europe. As a result, each L2 file must be checked to determine whether the geographic coordinates (latitude and longitude) of a given ground station fall within the footprint of the satellite swath.

The algorithm then selects a sampling station, and calculates the distance between the station's coordinates and all valid data pixels in the scene. In theory, the most precise approximation of the geodesic distance between the two points is given by Vincenty's formulae, assuming an ellipsoidal Earth [27]. But the whole process is very costly, not to mention doing it for each pixel of each NetCDF file and for each sampling station. Therefore, for computation-time savings, the distance computed is a rougher approximation of the actual geodesic distance. It is given by the Haversine formula:

$$d = 2r \arcsin\left(\sqrt{\frac{1-\cos(\Delta\varphi)+\cos\varphi_{ss}\cdot\cos\varphi_{pix}\cdot(1-\cos(\Delta\lambda))}{2}}\right)$$

where r is the radius of Earth (taken as 6371.0088 kilometers [28]), φ_{ss} and φ_{pix} are the latitudes of, respectively, the sampling station and the pixel considered, and λ their longitudes. This formula assumes a spherical Earth and is therefore precise only to 99.5% [29] which can mean a lot for distances ranging up to several thousands kilometers, but that level of accuracy is considered sufficient for the purposes of this algorithm.

The pixel with the minimum distance to the station is selected, provided that this distance is below a threshold of 8.9 km, which corresponds approximately to the diagonal of one TROPOMI pixel. The threshold is set considering that S5P TROPOMI pixels are approximately 5.5 km across-track by 7 km along-track. This ensures that the station lies within the area of the pixel selected as S5P usable measurement.

Finally, the value associated to the couple of longitude and latitude determined to be the closest to the sampling station is read and stored, along with the current file's timestamp, in an array for that sampling station. The next station can be considered, and once all stations are treated, the program moves on to the next file. This enables temporal and spatial alignment between satellite and ground-based CO measurements for validation and analysis.

Figure 6 visually summarizes the algorithm.

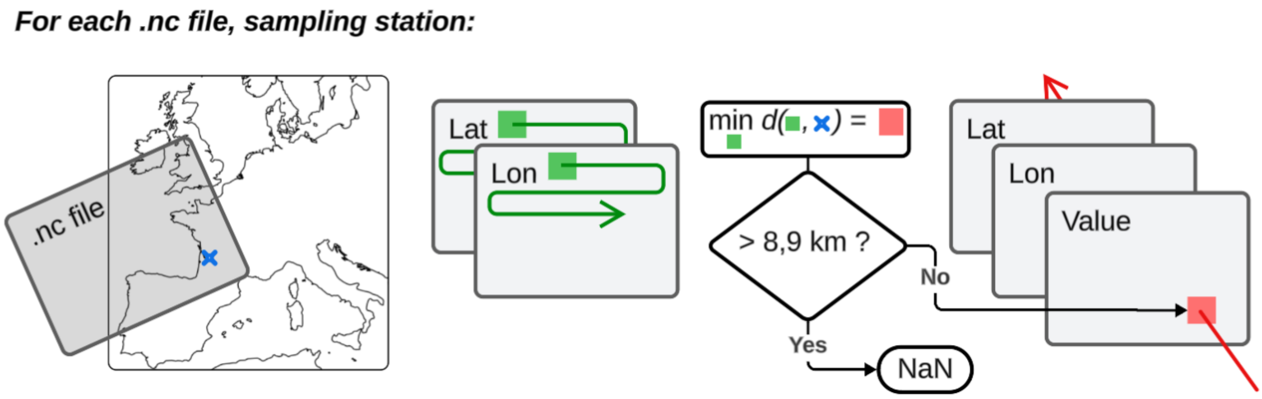


Figure 6: Visual explanation of the extraction algorithm to get the values from all NetCDF files for each sampling station. d corresponds to the Haversine distance function.

3.3 Description of output data format and structure

3.3.1 Canadian wildfire detection module

MODIS retrievals The output from the FIRMS data processing pipeline is a series of raster images at daily frequency with red-highlighted pixels taking the value 1 (red) if they are identified as a thermal anomaly pixel (i.e. an active wildfire), and 0 if not.

One of the most significant burnt areas during the wildfire seasons (i.e. May - June 2025) is shown in Figure 7 where the majority of the fires are concentrated around the province of Saskatchewan (central Canada).

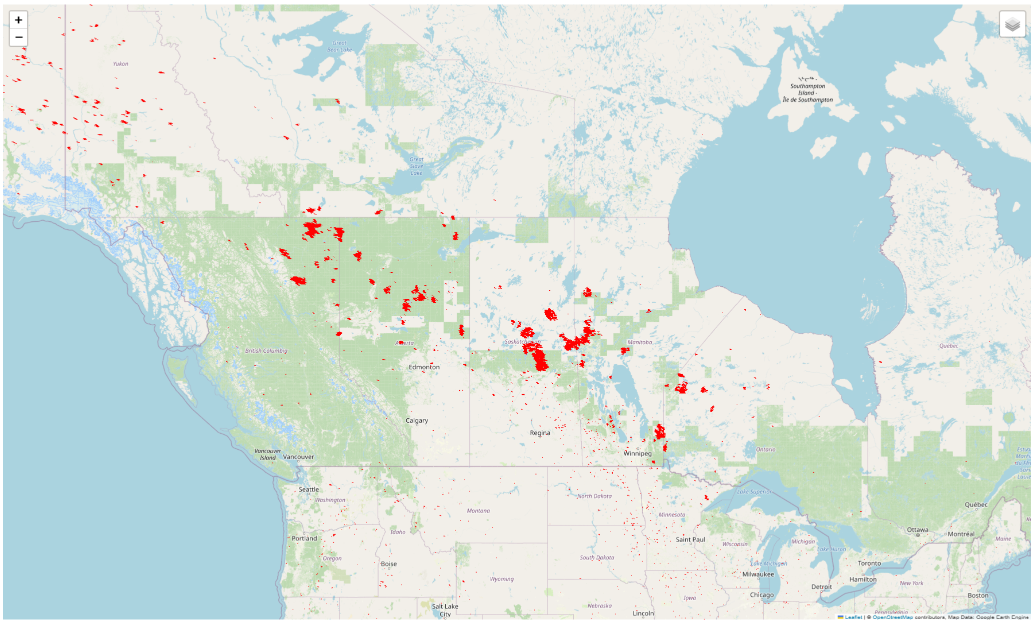


Figure 7: Thermal anomalies retrieved from FIRMS from 1 May to 25 June 2025.

Validation – CWFIS wildfire data The final dataset from CWFIS contains, as of the 24th of June 2025, 613 active wildfire sites. This dataset can then be plotted on a map of Canada, using the hectares burnt by each active fire for its size and its status for color-coding, as shown on Figure 8, and akin to Figure 2. Some small additional data manipulations were made for visualizing purposes: all fires having burnt less than a threshold value of 10 hectares have been artificially set to 10 hectares in order to give them a minimal, distinguishable size, and the final size of all markers is divided by 200 to appear reasonably sized on the map - while still appearing much bigger than the actual burnt area.

Although it has not yet been formally implemented, ideally the validated hotspots, if above a certain threshold, could trigger a warning to the government agencies in Europe to start creating action plans for potential incoming smoke plumes.

3.3.2 Transboundary CO transport monitoring module

Validation – EEA ground measurements The CO measurements obtained through the EEA Air Quality monitoring network are of variable quality. Two stations out of the 26 are selected as examples: Paris and Madrid. The temporal pattern of carbon monoxide (CO) concentrations recorded at these stations during the period of study is showed in Figure 9. It can be observed that some CO concentration records are pretty coarse, jumping from one concentration value to another with no intermediary value in-between, while others offer more details.

S5P retrievals What comes out of the algorithm described above is a collection of timeseries with daily granularity, one for each selected ground station, and each one referencing all the measurements

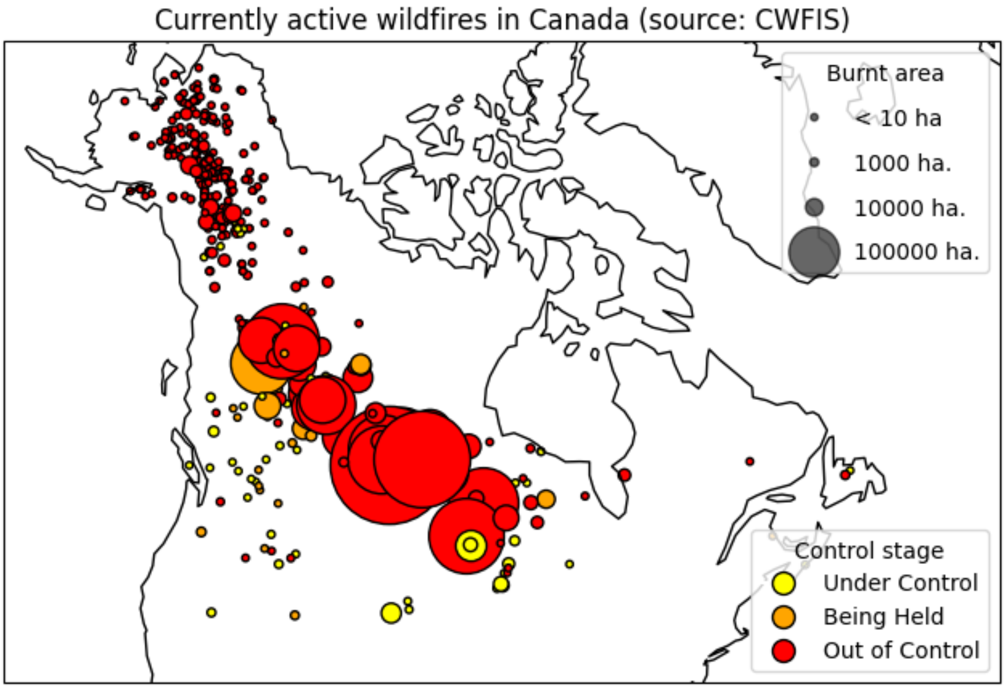
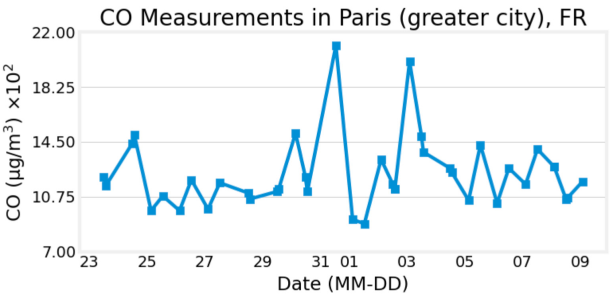
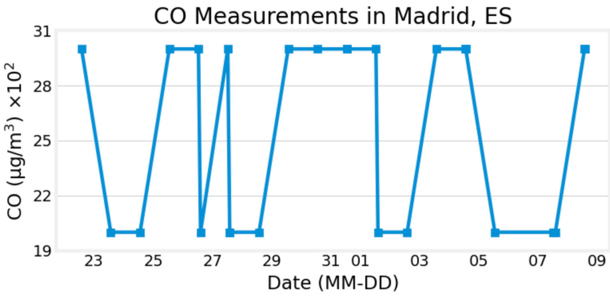


Figure 8: A map of the active fire sites in Canada and surroundings, generated by the service and dated from the 24th of June 2025.



(a) Paris



(b) Madrid

Figure 9: Ground (in-situ) measurement samples from EEA. Resolution and amount of data can vary from one station to another.

made by S5P above these specific locations. The measurements are in mol/m^2 and can be converted to µg/m^3 for better homogeneity with validation values, through the following operation:

$$\rho'_{\mathrm{S5P}}[\mu\mathrm{g/m^2}] = \rho_{\mathrm{S5P}}[\mathrm{mol/m^2}] \times M_{CO}[\mu\mathrm{g/mol}]$$

where $M_{CO} = (16 + 12)\mathrm{g/mol} = 28 \cdot 10^6 \mu\mathrm{g/mol}$ is the molar mass of carbon monoxide.

4 Validation: PQAR (Product Quality Assessment Report)

4.1 Overview

S5P-L2 data provides detailed timing and location of pollutant plumes transported by atmospheric dynamics. An overview of pollution transport is provided in Figure 10, where it is shown that a small wave of CO from Canada reached the west coast of Europe on 1 June 2025 followed by a more intense wave on 4 June 2025 which reached central and eastern Europe by 8 June 2025. From this preview, the origin of the plume appeared to be in central Canada, with propagation across the North Atlantic Ocean, passing over Greenland and first reaching the British Isles. Further validation, presented in section 4.3 results strengthen this observation.

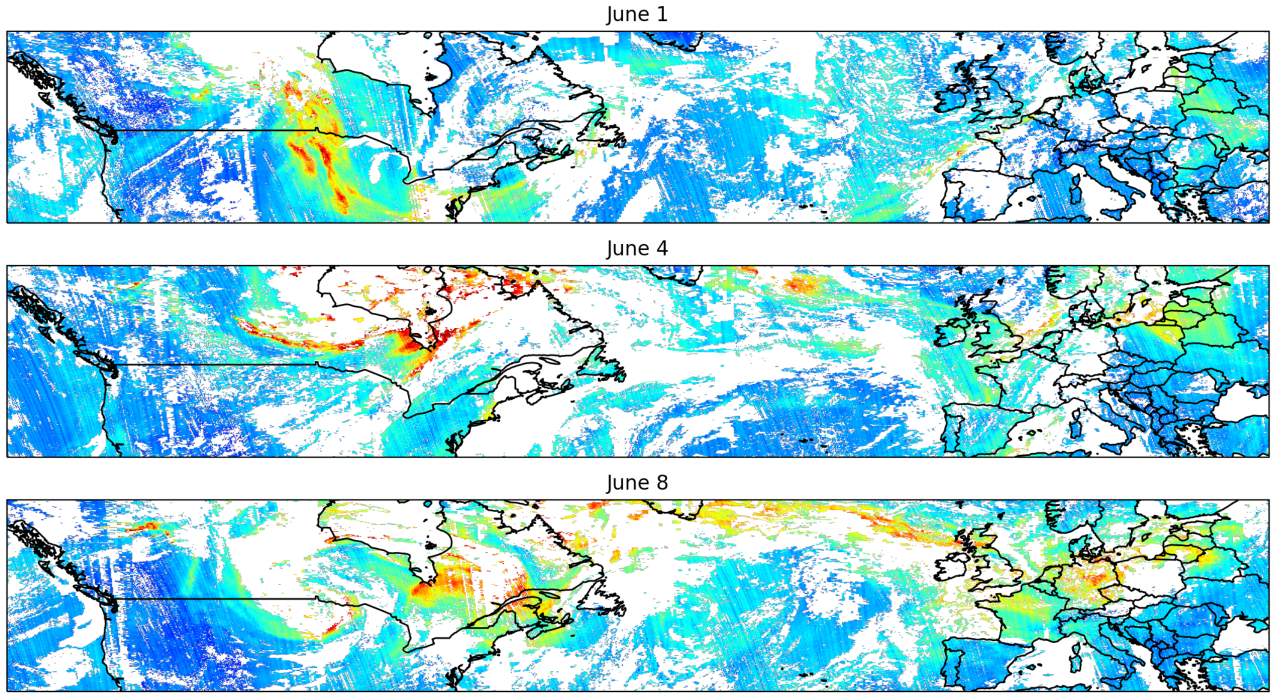


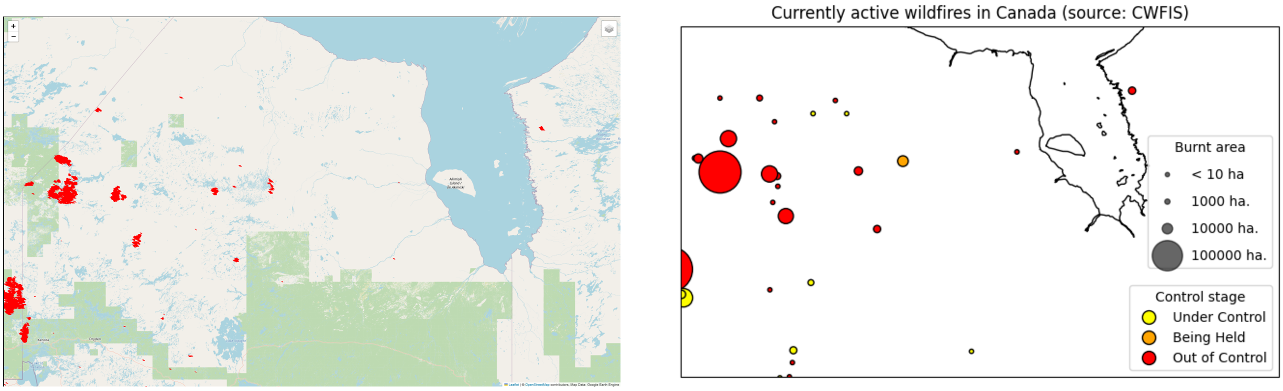
Figure 10: Transboundary CO Transport visualization generated from S5P-L2 data illustrating the CO displacement from Canada to Europe for 2025. Concentration increases from lighter to darker colors, where red is the highest. White regions represent no data (clouds).

4.2 Validation results: Wildfires Detection Module

Figure 11 presents a snapshot of thermal anomalies reported in the FIRMS database against the reported active wildfires from CWFIS on 25 June 2025 near James Bay (Ontario, Canada). As stated in section 3.2, the current state of development of the service only allows for visual comparison between the two datasets, but it can already be seen that both exhibit similar features. Described above in section 3.2.1 is how further processing of the FIRMS data could allow for quantitative validation, through three main indicators: the distance between measured and actual center of the fire, the difference in sizes and the temporal evolution, reflecting the control stage.

4.3 Validation results: Transboundary CO Transport Monitoring Module

The validation process was conducted using in-situ CO measurements from the monitoring stations located at CaWrdiff, Belfast, Leeds, London, Edinburgh, Zürich, Copenhagen, Tallinn, Marseille, Paris, Athens, Luxembourg, Lisboa, Barcelona, Coruña, Madrid, Cartagena, across Western Europe. Assessing the agreement of temporal trend between the two data sources are crucial as it determines the ability of the service to detect sudden changes in the CO concentrations which could be attributed



(a) FIRMS thermal anomalies detection. (b) Currently active wildfires in Ontario (Canada).

Figure 11: Visual comparison between FIRMS and CWFIS results near James Bay (Ontario, Canada) on 25 June 2025.

to the transbourndary smoke from Canadian wildfires. As such, the correlation coefficient is used as the preliminary statistical tool.

The CO density measurements timeseries for a sample of the stations, representative in quality of data, are illustrated in Figure 12. These cities shows positive correlations between ground measurements and S5P estimates of CO. For instance, Figure 12b shows that both sources of CO data in Barcelona agree on the CO increase peaks for 4 and 8 June 2025 which is in agreement with the transboundary CO transport monitoring observations (see Figure 10). This was also observed for Marseille on 9 June 2025. On the other hand, for Luxembourg and other stations, the sudden increase in CO levels detected by S5P was not reflected in the ground sensor measurements. Figure 12d shows the correlation between S5P retrievals and ground measurements of CO for 23 stations. Out of the 26 ground stations studied, three (Lisboa, Coruña, Barcelona) exhibited constant values throughout the time range, making it impossible to extract their correlations with Sentinel-5P data. Furthermore, the stations showing stronger negative correlations appeared to also have significant gaps in their data.

At the current state of the service development, this analysis could be considered as preliminary validation of the S5P CO estimates given that there detect some acceptable/reasonable correlations between S5P and ground measurements, i.e. $r \geq 0.3$. The current correlation is comparable to currently published journals attempting to compare S5P estimates with ground measurements at site-specific/city-specific level with more extensive data assimilation in the validation process [30] [31] [32]. The satellite-ground measurements discrepancy can be explained by the fundamental difference between CO column concentrations and volume concentrations, as well as the influence of more localized environmental conditions. Broader studies have addressed this mismatch by applying linear models that assign weighted values to both satellite and ground-based measurements, with the aim of producing more robust estimates of land surface air quality [32]. This issue is further discussed in section 5.

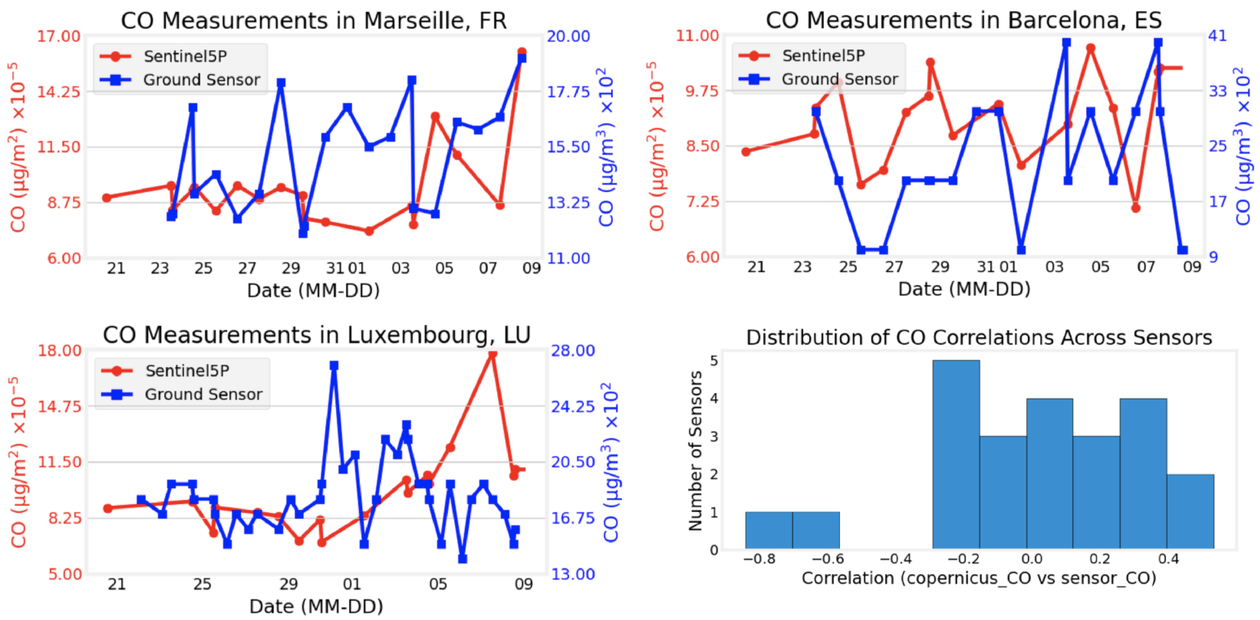


Figure 12: Timeseries for S5P and EEA CO measurements with respective correlations of a) France (0.32), b) Spain (0.53), c) Luxemburg (0.12). Figure 12d shows the correlation between S5P retrievals and ground measurements of CO for the 23 stations.

5 Gap analysis: TRGAD and E3UB (Target Requirements and Gap Analysis Document / User Requirements and End-to-End Uncertainty Budget)

5.1 Product error estimates

There are various sources errors, either intrinsically in the instruments and measuring methods or from numerous assumptions and algorithms used in the current workflow which could lead to large error propagation downstream of the service. These errors are listed below and possible improvements will be discussed later in Section 6.

5.1.1 Intrinsic instrument errors: S5P

The accuracy of satellite-based CO retrievals is influenced by several atmospheric and observational factors. Cloud cover can obscure surface signals, aerosol loading can alter radiative transfer properties, and variations in solar zenith angle can affect the retrieval geometry. Although the S5P-L2 product includes quality assurance (QA) flags to filter unreliable data, these corrections cannot fully compensate for the retrieval uncertainties under certain atmospheric conditions [33].The S5P algorithm to retrieve total CO columns i.e. Shortwave Infrared CO Retrieval (SICOR) [31] has been tested in various scenarios namely clear sky, water clouds, photon trapping, aerosols, cirrus clouds, multiple cloud layers) [34]. This algorithm is also tested for robustness to uncertainties in atmospheric inputs and instrument artifacts [34]. On top of testing the algorithms to derive CO measurement from the received satellite signals, there are routine validations with ground-based total carbon column observing network (TCOON). The main goal of the above tests are to ensure that S5P retrieves total CO column measurements with maximum accuracy of $<15\%$ and a maximum precision of $\leq 10\%$ even for background CO abundance and low surface reflection in the shortwave infrared (SWIR) spectral range [34]. As such, intrinsically, the CO measurements are within error range of $\pm$ 15% consistently 90% of the time. CO measurement by S5P during smoke events could increase up to a factor of 2 (refer to Figure 12) which is larger than the 15% maintained accuracy. This suggests that even though there is large uncertainty in the CO measurements from S5P, this should not affect the detection of smoke plume events although precautions should be taken when interpreting the concentration values. Additionally, due to the nature of the service which uses near real-time (NRTI) satellite data, it uses the L1B reflectance spectra as input to the retrieval instead of further processing of the received spectra [35] which introduces additional sources of errors in the retrieved CO data.

5.1.2 Measurement methods: ground-based vs. satellite

The difference between column-integrated CO measurements from satellite products and publicly available CO concentrations or mixing ratios measured at surface poses as a critical uncertainty and error source. CO measurements from satellite are highly sensitive to the lower troposphere, particularly boundary layer [36]. Furthermore, it should be noted that CO is not uniformly distributed in the atmospheric column [37]. Instead, CO tends to concentrate at specific altitudes depending on the extent of the wildfire events, at times reaching up to the upper troposphere [38]. As a result, there have been numerous attempts at validating satellite and ground measurement CO in various countries e.g. Finland [32], Russia [39], China [31], South Korea [30], as the link between these two sources of data is often based on empirical equations [40]. Otherwise, CO measurements from S5P could be converted to a column mixing ratio (XCO) using surface pressure to calculate the total air column [32]. Although ground-based and satellite both use IR to detect CO, the measurements are not directly comparable and to resolve these differences requires extensive, site-specific monitoring programs. It should also be noted that real-time CO measurements from ground stations are, at times, subject to sensor noises. As such, the validation results presented in this report serves as a preliminary assessment and comparison of ground-based vs. satellite measurements across multiple cities in Europe for initial emergency response.

5.1.3 Geospatial distances: ground-based vs. satellite

Another source of uncertainty arises from the currently implemented algorithm to compare these two data sources. As seen from Figure 6, the spatial mapping of satellite and ground observations, given that the station is covered in the satellite swath at an instantaneous time, depends on distance between the station and all valid data pixels (of resolution approx. 5.5 km across-track by 7 km along-track). This assumes that the 'ground' station located nearby the true station's coordinates measures CO concentration at the areas of approx. 40 km^2. The spatial distribution of CO concentrations are highly depending on locations of the stations, e.g. nearby industries, on-road combustion, rural, mountainous areas, etc. hence this matching of 'assumed' ground station measurement and S5P data could pose a wide variability in error.

5.2 User requirements

This operational service is designed for government agencies as well as the general public. Government agencies would require a preliminary analysis of the situation supported with information of reasonable accuracy in a timely manner (e.g. before smoke plumes hitting certain countries). This means that they would require early detection of Canadian hotspots/wildfires on daily or weekly basis as well as CO measurement in Europe within a reasonable error margin. In Canada, wildfire-prone regions are monitored on a daily basis during fire seasons using an integrated system of satellite and ground-based observations 3.1.1, enabling prompt wildfire detection and response. Furthermore, section 5.1 discusses that S5P retrieves total instantaneous CO column measurements with minimum accuracy of 15% as the error margin could be higher from error propagation discussed in section 5.1, which, at the moment, remains sufficient to detect elevated CO due to smoke plumes as well as the gradual transport of pollution over the Atlantic. This helps support timely action plan to be rolled out for the public. However, data quality and data assessment still remain as the main improvement point of the service.

On the other hand, once the initial action plan is rolled out by the government agencies, the public are eligible to monitor and be informed about the situation closely as the pollution smoke plumes would affect daily-to-subdaily activities. With the latency of S5P data of 3 hours for NRTI data and ground measurements of hourly in Europe (see Table 1), this would suffice for the public to utilize the service for their own usage.

Both stakeholders would require the service to be online and operational. This means that the all data pipeline should be functional and in proper working conditions at all time. Contingency plans, e.g. when one or more sources of data breakdown, should be made. This will be discussed in later sections.

5.3 Gap analysis

Based on the above assessment of the users' requirements, the main gap in the service workflow is identified: further improved data quality control and data assessment as well development of contingency plan in case of pipeline breakdown.

5.3.1 Reducing errors in validation process

This section outlines targeted improvements to reduce errors estimated above through improved data assimilation pipeline as well as enhanced data filtering and quality control. Firstly, longer term and of better-quality-control-state datasets, such as reprocessed (RPRO) and/or offline (OFF) S5P-L2 observations from both Canada and Europe in prior wildfire seasons, could help improve the understanding of temporal time-shift for the pollutant between continents as well as enhance the comparison between S5P and ground measurement data. Secondly, there has been attempts to incorporate simulated CO measurements from chemical transport models (CTMs) using meteorological forcings to improve validation of satellite retrieved CO data [30]. CTMs, e.g. CMAQ [41], CHIMERE [42], are physics-based, three-dimensional models which could couple with weather forecast models such as Weather Research and Forecasting (WRF-Chem [43]) to model the atmospheric chemistry, emission and transport of pollutants. To better constrain the vertical mismatch between CO ground and satellite measurements, models such as WRF-Chem, CHIMERE, and CMAQ can be employed. This would provide an independent set of dataset to bridge the gap between the satellite data (i.e. looking downward) and ground measurement (i.e. looking upward). This would allow for the estimation of the ratio between surface-level and total column values which is critical for translating satellite observations into actionable ground-level air quality indicators [44]. Moreover, these models can model the evolution of pollutant plumes over time and space. Although CO is implemented currently due to its inert nature, should there are more pollutants of concerns, e.g. O_3, $PM_{2.5}$ which are also key indicators of air quality and undergo various chemical transformations throughout transport, CTMs would be crucial in improving the validation process as well as potentially reducing uncertainty and errors of the service [44]. Lastly, the algorithm to match the geo-location of the ground station with the pixel with valid data in the satellite product could be further improved. Instead of finding the nearest valid data point, the ground station measurements could be compared with an area-average CO measurements from the satellite product such that it only covers relevant areas, e.g. only surrounding urban regions. Averaging over an area could potentially reduce the artifacts from results in 1 pixel, hence reducing variance in the estimates.

5.3.2 Contingency plan

As the service should be operational and online 24/7 for close monitoring, there should be contingency plan in case the data pipeline breaks down. As each module in the pipeline contains 2 sources of data, i.e. Canadian Wildfire Detection Module as well as Transboundary CO Transport Monitoring Module including ground-reported and satellite data, should 1 source of data fails, the government agency and public could still be informed of the situation. However, further details on the error estimates should be made known as, for example, missing ground-measurement data would mean satellite data is not validated, or missing satellite data would indicate lack of information of the possible daily transport of pollutants across the Atlantic. In case of missing ground-measurement data, long-term comparison of ground measurement against S5P CO data, as stated above, could aid in establishing an empirical relationship between these 2 sources of data. Hence, the ground measurements could be interpolated based on satellite data. Having this contingency plan in place would enable smooth operation of the service.

6 Draft plan for operational service

6.1 Product frequency and timeliness

The current state of the service includes a workflow that integrates download, process and visualize remote sensing data (i.e. hotspot for wildfires in Canada, S5P CO observations in Canada and Europe) with the finest time resolution available alongside with hourly in-situ CO measurements from 26 air quality stations across Europe. As seen in Table 1, the bottleneck is the S5P data stream which delivers the CO column measurement within 3 hours after sensing for near real-time product and daily for the entire orbit [35]. As such, the service would have a >3-hr update cycle, based on the updated S5P measurements. This >3-hr update cycle should be sufficient for government agencies and the public to make informed decisions. The in-situ CO measurements would be quality-controlled to be used as ground-truth to validate satellite data over Europe. This whole service update takes approx. 30 minutes on a laptop with on a local workstation environment. The final processed data would be presented in visualization dashboard.

6.2 Infrastructure needs

The computational requirements for our temporal framework selection, i.e. mainly required for Copernicus data from 21 May to 7 June 2025, are of 5.242 GB to hold a total of 295 NetCDF files for Europe. The median file size is 21 MB and files over dates tend to be uniform within different time ranges as shown in Figure 13. Since each NetCDF file is processed individually, a reasonable estimate for the required Random Access Memory (RAM) is at least 200 MB, allowing room for local variables during execution. A system with 4 GB of RAM can significantly reduce processing times by minimizing read-out delays. The core of the algorithm involves computing the geodistance using Haversine equation, which can be efficiently handled by a CPU. However, incorporating GPU infrastructure may further enhance performance, as the same operation is applied repeatedly across the NetCDF array elements. For further time analysis within the same region is possible to linearly scale the 5.242 GB of data collected.

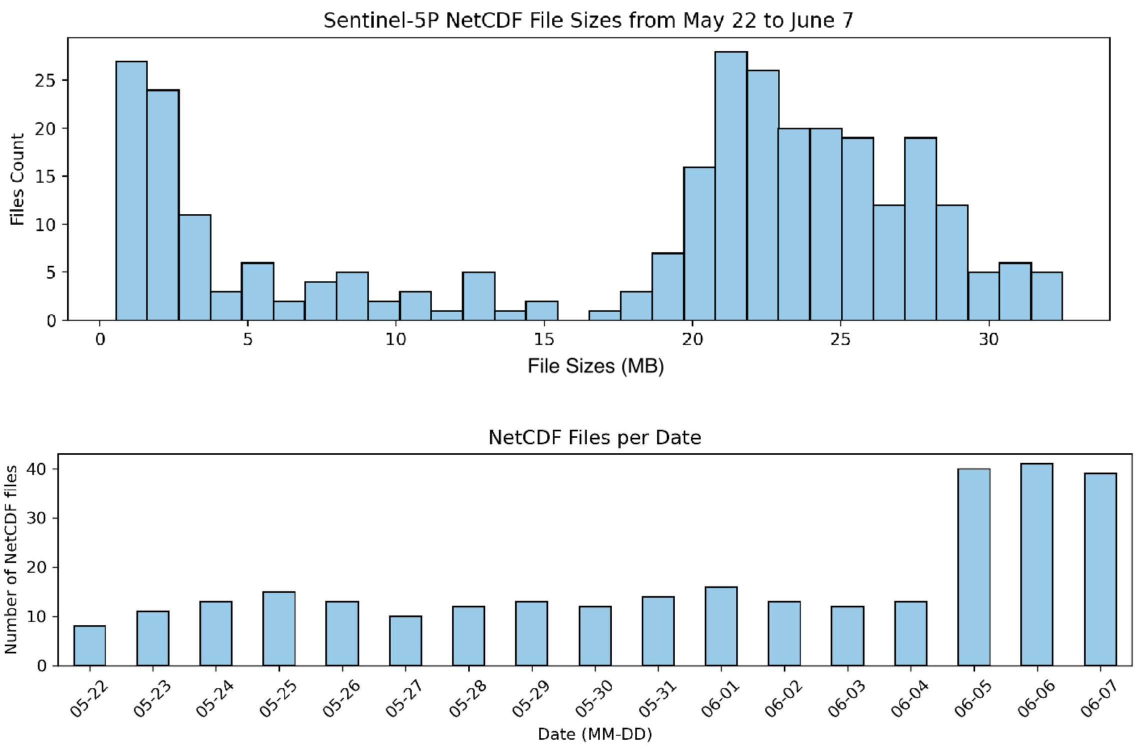


Figure 13: File size distribution showing balance between low and high sizes and number of files for 22 May to 7 June 2025 with uniform distributions before and after 5 June 2025.

For the results reported here, the project is implemented on a local workstation environment specifically configured to support the data processing requirements. The hardware used for processing included an Intel i7 8-core CPU, 32 GB of RAM, and a 1 TB solid-state drive. This setup proved sufficient to handle the daily processing workload across all 26 stations without any significant performance limitations. It should be noted that the proper function for distance calculation in the algorithm has been chosen to reduce the computational time.

The processing workflow is developed in Python, utilizing open-source libraries well-suited to work with geospatial and atmospheric datasets such as numpy, geopy, and pandas. S5P-L2 NRTI CO data could be accessed through manual downloads through the Copernicus web browser or relevant APIs, while ground-based CO observations are obtained from the EEA API from official national air quality networks. Processed data is stored locally in a structured folder system organized by station and date. The workflow scripts are fully reproducible and can be executed manually or automated via standard scheduling tools. Overall, the infrastructure setup validated the practicality of performing near-real-time processing of satellite and ground-based CO data using modest computing resources and openly available software tools.

6.3 Potential future developments

Following intended service workflow in Figure 1, there are several components which could be explored beyond the MVP and currently implemented features to encapsulate the entire service of monitoring the correlation between European Air Quality and Canadian Wildfire Season. This includes:

6.3.1 Canadian Wildfire Detection Module

Observations of burnt areas through other indicators Vegetation and biomass burnt monitoring could be done using Sentinel-2 along with land surface temperature increase from Sentinel-3 provided by Copernicus program. Sentinel-2 could provide high-resolution imagery to assess vegetation conditions and burned severity, enabling a detailed evaluation of vegetation loss at source. It should be noted that, despite high spatial resolution, Sentinel-2 provides surface reflectance of approx. every 5 days, which would not be timely enough to rely on Sentinel-2 alone for this current monitoring service. On the other hand, Sentinel-3 could be utilized to derive surface temperature which enables better locating of potential hotspots and their intensity. It should also be noted that Sentinel-3 has a revisit interval of 2 days which is more frequent than Sentinel-2 and might be sufficient for detection of wildfires in Canada. These two additional satellite products provide other insights for fires at the source and could support the FIRMS dataset derived from MODIS MOD14/MYD14 Fire and Thermal Anomalies product with cadence time of 1 day. This localized analysis near the source region complements the transboundary perspective and enhances the service's capacity for monitoring transbournday smoke transport.

Validation of satellite data In-situ data on active wildfire sites from CWFIS could be used to compare against the satellite observations from MODIS, Sentinel-2 and Sentinel-3. This would increase the confidence and accuracy in detecting significant fire events happening in Canada which could potentially affect air quality in Europe.

6.3.2 Transboundary CO Transport Monitoring Module

Canada air quality Similarly done in Europe, in-situ measurements of air pollutants in Canada could be fetched through an appropriate API to be integrated in the system. This data would then undergo similar quality control process as done for Europe and used as ground-truth to validate satellite observations of air pollutants retrieved from S5P. This can be done in parallel with the processing for Europe at minimal additional computing costs provided that the system uses multiple threads on multiprocessor system.

Additional statistical analysis This includes correlation between 1) wildfires and air quality at source in Canada and 2) correlation, time-shift lag for air quality in Europe-Canada using the validated satellite measurements from both Europe and Canada. This additional features would enable local government in Canada to implement proper action plan to combat wildfires as well government agencies in Europe to gain some lead time in reacting to the potential of pollution plumes arriving.

7 Appendix

Table 2: Sampling Station Information Grouped by Country

Country	City	Lat	Lon	Sampling Station ID
United Kingdom (GB)	Cardiff	51.48178	-3.17625	GB_SamplingPoint_107
	Greater Belfast	54.59965	-5.92883	GB_SamplingPoint_112
	Leeds	53.80378	-1.54647	GB_SamplingPoint_127
	London (greater city)	51.52105	-0.21349	GB_SamplingPoint_205
	City of Edinburgh	55.94559	-3.21819	GB_SamplingPoint_61573
Switzerland (CH)	Zürich (greater city)	47.40290	8.61340	SPO-CH005A_00010_502
	Zürich (greater city)	47.37760	8.53041	SPO-CH001A_00010_500
Denmark (DK)	København	55.70028	12.56140	SPO-DK0045A_00010_100
Estonia (EE)	Tallinn	59.41417	24.64946	SPO-EE0018A_00010_100
France (FR)	Marseille	43.30529	5.39472	SPO-FR03043_10
	Paris (greater city)	48.86212	2.34462	SPO-FR04055_10
Greece (GR)	Athens	37.03084	23.71032	SPO-GR0022A_00010_001
	Athens	37.93200	23.71302	SPO-GR0031A_00010_131
Luxembourg (LU)	Luxembourg	49.59769	6.13759	SPO-LU0101A_00010_100
Portugal (PT)	Lisboa (greater city)	38.76889	-9.10806	SPO-PT03071_00010_100
	Lisboa (greater city)	38.73889	-9.20750	SPO-PT03082_00010_100
	Lisboa (greater city)	38.66361	-9.15778	SPO-PT03083_00010_100
Spain (ES)	Barcelona	41.42608	2.14799	SP_08019054_6_48
	Barcelona	41.38750	2.11510	SP_08019057_6_48
	Barcelona	41.30311	1.99152	SP_08028095_6_48
	Barcelona	41.32177	2.06101	SP_08169009_6_48
	Barcelona	41.31348	2.01382	SP_08301004_6_48
	A Coruña	43.38280	-8.40920	SP_15030027_6_48
	Madrid	40.41917	-3.70333	SP_28079035_6_48
	Madrid	40.32420	-3.87677	SP_28092005_6_48
	Cartagena	37.60180	-0.97620	SP_30016020_6_48

References

[1] Copernicus Atmosphere Monitoring Service, "Homepage," 2025. Accessed: 2025-06-25.

[2] J. Chen and X. Long, "The impact of air pollution on career changes among chinese workers," *Scientific Reports*, vol. 15, p. 88239, 2025.

[3] X. Yin, B. Thai, Y. Tan, S. Salinas, L. Yu, and W. Seow, "When and where to exercise: An assessment of personal exposure to urban tropical ambient airborne pollutants in singapore," *Science of The Total Environment*, vol. 906, p. 167086, 2023.

[4] NASA/LANCE/FIRMS, "Firms: Fire information for resource management system." https://developers.google.com/earth-engine/datasets/catalog/FIRMS, 2025. Accessed: 2025-06-20.

[5] C. F. Service, "Canadian wildland fire information system (cwfis)." https://cwfis.cfs.nrcan.gc.ca/datamart/metadata/activefires, 2022. Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, Alberta.

[6] A. Ding, T. Wang, and C. Fu, "Transport characteristics and origins of carbon monoxide and ozone in hong kong, south china," *Journal of Geophysical Research: Atmospheres*, vol. 118, no. 17, pp. 9475–9488, 2013.

[7] P. Penteado and NASA/JPL-Caltech, "Airs global carbon monoxide over 20 years (2002-2022)." https://airs.jpl.nasa.gov/resources/239/airs-global-carbon-monoxide-over-20-years-2002-2022/, 2022. Accessed: 2025-06-23.

[8] L. S. Jaffe, "Ambient carbon monoxide and its fate in the atmosphere," *Journal of the Air Pollution Control Association*, vol. 18, no. 8, pp. 534–540, 1968.

[9] G. Attard, "An intro to the earth engine python api." https://developers.google.com/earth-engine/tutorials/community/intro-to-python-api, 2025. Accessed: 2025-06-20.

[10] European Environment Agency, "Eea datahub: Environmental data and information." https://www.eea.europa.eu/en/datahub/datahubitem-view/778ef9f5-6293-4846-badd-56a29c70880d, 2025. Accessed: 2025-06-20.

[11] Copernicus Atmosphere Monitoring Service, "Smoke from canadian wildfires reaches europe," 2024. Accessed: 2025-06-17.

[12] Z. Wang, Z. Wang, Z. Zou, X. Chen, H. Wu, W. Wang, H. Su, F. Li, W. Xu, Z. Liu, and J. Zhu, "Severe global environmental issues caused by canada's record-breaking wildfires in 2023," *Advances in Atmospheric Sciences*, vol. 41, pp. 565–571, 2024.

[13] H. Petetin, M. Guevara, S. Compernolle, D. Bowdalo, P.-A. Bretonnière, S. Enciso, O. Jorba, F. Lopez, A. Soret, and C. P. García-Pando, "Potential of tropomi for understanding spatio-temporal variations in surface no_2 and their dependencies upon land use over the iberian peninsula," *Atmospheric Chemistry and Physics*, vol. 23, no. 7, pp. 3905–3935, 2023.

[14] World Health Organization, "Health advice: wildfires in the who european region," 2021. Accessed: 2025-06-17.

[15] BBC News, "Canada wildfires: Smoke chokes us cities as canada burns," 2023. Accessed: 2025-06-17.

[16] CBC News, "Tracking wildfires in canada: Where they are, how much has burned and how it's changing air quality," 2023. Accessed: 2025-06-17.

[17] European Environment Agency, "Air quality in europe - 2022 report," 2022. Accessed: 2025-06-17.

[18] United Nations Environment Programme, "Capacity building workshop: Strengthening regional actions to address transboundary haze and plastic pollution in southeast asia," 2024. Accessed: 2025-06-17.

[19] NASA FIRMS, "Fire information for resource management system (firms)." https://firms.modaps.eosdis.nasa.gov/, 2025. Accessed: 2025-06-25.

[20] NASA Earthdata, "Active fire data attributes for modis and viirs," 2025. Accessed: 2025-06-25.

[21] Canada Wildland Fire Information System (CWFIS), "CWFIS Datamart." https://cwfis.cfs.nrcan.gc.ca/datamart. Accessed: 2025-06-24.

[22] CBC News, "Out of control, held, out — what the stages of wildfire mean," 2023. Accessed: 2025-06-24.

[23] European Environment Agency, "Air quality data download service." https://eeadmz1-downloads-webapp.azurewebsites.net/. Accessed: 2025-06-18.

[24] C. D. S. Ecosystem, "Catalog api." https://documentation.dataspace.copernicus.eu/APIs/SentinelHub/Catalog.html. Accessed: 2025-06-23.

[25] Google Earth Engine, "Sentinel-5p offl co: Offline carbon monoxide." https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S5P_OFFL_L3_CO. Accessed: 2025-06-18.

[26] L. Giglio, W. Schroeder, and C. O. Justice, "The collection 6 modis active fire detection algorithm and fire products," *Remote Sensing of Environment*, 2016. Available at https://www.earthdata.nasa.gov/sites/default/files/imported/Giglio2016_C6_Fire.pdf.

[27] Wikipedia, "Vincenty's formulae." https://en.wikipedia.org/wiki/Vincenty%27s_formulae. Accessed: 2025-06-24.

[28] B. Rouberol, "Haversine 2.9.0." https://pypi.org/project/haversine/2.9.0/#description. Accessed: 2025-06-23.

[29] Wikipedia, "Haversine formula." https://en.wikipedia.org/wiki/Haversine_formula. Accessed: 2025-06-23.

[30] U. Jeong and H. Hong, "Comparison of total column and surface mixing ratio of carbon monoxide derived from the tropomi/sentinel-5 precursor with in-situ measurements from extensive ground-based network over south korea," *Remote Sensing*, vol. 13, no. 19, p. 3987, 2021.

[31] T. Borsdorff, J. A. de Brugh, H. Hu, I. Aben, O. Hasekamp, and J. Landgraf, "Measuring carbon monoxide with tropomi: First results and a comparison with ecmwf-ifs analysis data," *Geophysical Research Letters*, vol. 45, no. 6, pp. 2826–2832, 2018.

[32] T. Karppinen, A.-M. Sundström, H. Lindqvist, J. Hatakka, and J. Tamminen, "Satellite-based assessment of national carbon monoxide concentrations for air quality reporting in finland," *Remote Sensing Applications: Society and Environment*, vol. 33, p. 101120, 2024.

[33] Y. Jung, J. Kim, W. Kim, H. Boesch, H. Lee, C. Cho, and T.-Y. Goo, "Impact of aerosol property on the accuracy of a co_2 retrieval algorithm from satellite remote sensing," *Remote Sensing*, vol. 8, no. 4, p. 322, 2016.

[34] J. Landgraf, J. aan de Brugh, R. A. Scheepmaker, T. Borsdorff, S. Houweling, and O. P. Hasekamp, "Sentinel-5p tropomi atbd carbon monoxide total column retrieval," Tech. Rep. SRON-S5P-LEV2-RP-002, SRON Netherlands Institute for Space Research, 2024. Accessed: 2025-06-23.

[35] J. Landgraf, T. Borsdorff, S. Mandal, *et al.*, "Sentinel-5p carbon monoxide level 2 product readme file, version 2.9." https://sentiwiki.copernicus.eu/__attachments/1673595/S5P-MPC-SRON-PRF-CO%20-%20Sentinel-5P%20Carbon%20Monoxide%20Level%202%20Product%20Readme%20File%202025%20-%202.9.pdf, 2025. Accessed: 2025-06-24.

[36] J. Landgraf, J. aan de Brugh, R. Scheepmaker, T. Borsdorff, H. Hu, S. Houweling, A. Butz, I. Aben, and O. Hasekamp, "Carbon monoxide total column retrievals from tropomi shortwave infrared measurements," *Atmospheric Measurement Techniques*, vol. 9, pp. 4955–4975, 2016.

[37] H. S. Marey, J. R. Drummond, D. B. A. Jones, H. Worden, C. Clerbaux, T. Borsdorff, and J. Gille, "A comparative analysis of satellite-derived co retrievals during the 2020 wildfires in north america," *Journal of Geophysical Research: Atmospheres*, 2024.

[38] NASA Jet Propulsion Laboratory, "Concentration and transport of carbon monoxide from the california wildfires." https://www.jpl.nasa.gov/images/pia12195-concentration-and-transport-of-carbon-monoxide-from-the-california-wildfires/, 2009. Accessed: 2025-06-24.

[39] L. N. Yurganov, V. Rakitin, A. Dzhola, T. August, E. Fokeeva, M. George, G. Gorchakov, E. Grechko, S. Hannon, A. Karpov, L. Ott, E. Semutnikova, R. Shumsky, and L. Strow, "Satellite- and ground-based co total column observations over 2010 russian fires: accuracy of top-down estimates based on thermal ir satellite data," *Atmospheric Chemistry and Physics*, vol. 11, pp. 7925–7942, 2011.

[40] F. Chevallier, M. Remaud, C. W. O'Dell, D. Baker, P. Peylin, and A. Cozic, "Objective evaluation of surface- and satellite-driven carbon dioxide atmospheric inversions," *Atmospheric Chemistry and Physics*, vol. 19, no. 22, pp. 14233–14251, 2019.

[41] K. W. Appel, S. L. Napelenok, K. M. Foley, H. O. T. Pye, C. Hogrefe, D. J. Luecken, J. O. Bash, S. J. Roselle, J. E. Pleim, H. Foroutan, W. T. Hutzell, G. A. Pouliot, G. Sarwar, K. M. Fahey, B. Gantt, R. C. Gilliam, N. K. Heath, D. Kang, R. Mathur, D. B. Schwede, T. L. Spero, D. C. Wong, and J. O. Young, "Description and evaluation of the community multiscale air quality (cmaq) modeling system version 5.1," *Geoscientific Model Development*, vol. 10, no. 4, pp. 1703–1732, 2017.

[42] L. Menut, A. Cholakian, R. Pennel, G. Siour, S. Mailler, M. Valari, L. Lugon, and Y. Meurdesoif, "The chimere chemistry-transport model v2023r1," *Geoscientific Model Development*, vol. 17, pp. 5431–5457, 2024.

[43] S. Archer-Nicholls, D. Lowe, S. Utembe, J. D. Allan, R. A. Zaveri, J. D. Fast, Øivind Hodnebrog, H. D. van der Gon, and G. McFiggans, "Gaseous chemistry and aerosol mechanism developments for version 3.5.1 of the online regional model, wrf-chem," *Geoscientific Model Development*, vol. 7, no. 6, pp. 2557–2579, 2014.

[44] Y. Yang, Y. Gao, X. Zhang, X. Liu, H. Zhang, X. Fan, R. Niu, J. Huang, Q. Fu, M. Xie, *et al.*, "Intercomparison of three coupled meteorology-chemistry models (wrf-cmaq, wrf-chimere, and wrf-chem) for simulating meteorological and air quality variables in east asia," *Geoscientific Model Development*, vol. 17, pp. 2471–2496, 2024.