Spatio-temporal data on the air pollutant nitrogen dioxide derived from Sentinel satellite for France

Hichem Omrania, Bilel Omranib, c, Benoit Parmentierd, Marco Helbiche,e,*

a Urban Development and Mobility, Luxembourg Institute of Socio-Economic Research, Luxembourg
b École Centrale de Lyon, France
c Polytechnique Montréal, Canada
d University of Maryland, National Socio-Environmental Synthesis Center, USA
e Department of Human Geography and Spatial Planning, Utrecht University, Utrecht, Netherlands

Article info
Article history:
Received 11 December 2019
Received in revised form 25 December 2019
Accepted 27 December 2019
Available online 7 January 2020

Keywords:
Air pollution
Remote sensing
Monitoring

Abstract
Monitoring of air pollution is an important task in public health. Availability of data is often hindered by the paucity of the ground monitoring station network. We present here a new spatio-temporal dataset collected and processed from the Sentinel 5P remote sensing platform. As an example application, we applied the full workflow to process measurements of nitrogen dioxide (NO2) collected over the territory of mainland France from May 2018 to June 2019. The data stack generated is daily measurements at a 4 x 7 km spatial resolution. The supplementary Python code package used to collect and process the data is made publicly available. The dataset provided in this article is of value for policymakers and health assessment.

© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Data description

Remote sensing data for air quality monitoring is important for health research [1]. The advantage of remotely sensed air pollution data includes, but is not limited to, large coverage at a useful spatial and temporal resolution. Sentinel 5P is a rather new remote sensing data source but requires downloading and computationally intensive processing that is often a barrier to public use. For illustrative purposes, we focus on generation of a NO2 spatio-temporal measurements across mainland France. The shared workflow repository, however, contains other air pollutants including ozone (O3), sulfur dioxide (SO2), carbon monoxide (CO) (for details, see the GitHub page [2]). The spatial resolution of the measurements (3.5 × 7 km² for all trace gases, except for CO and Methane (CH4) that is 7 × 7 km²) allows observations and mapping of air pollution at a finer scale (e.g. at the scale of an administrative area) (see Table 1 and Fig. 5).

The data consist of a netCDF file containing Sentinel 5P measurements between May 2018 and June 2019, with multiple attributes (e.g., latitude/longitude, WGS84 projection, and date of measurement), allowing both spatial and temporal observations of air pollutants. We cleaned the data using a quality flag parameter (noted by ‘qa_value’ varying between 0 (no data) and 1 (full quality data). We used ‘qa_value’ above 0.5 provided by the Copernicus Sentinel 5 Precursor Tropospheric Monitoring Instrument (S5p/TROPOMI [3]) to filter cloud cover and so to ensure high quality data. The dense measurements performed every 24 hours allow accurate annual averaging at fine-grained spatial resolution as shown in Fig. 1.

The data were also grouped by date allowing temporal assessment of pollutants density and quantitative observations such as monthly pollutants distribution, both spatially and temporally (Fig. 2). Data allow also analysis for weekday variations (Fig. 3).
Table 1
Average tropospheric vertical column of NO$_2$ by administrative area.

| Administrative area                  | Average tropospheric vertical column of NO$_2$ ($\times 10^{19}$ molec/m$^2$) |
|--------------------------------------|--------------------------------------------------------------------------------|
| Corse                               | 1.782                                                                          |
| Ile-de-France                       | 3.923                                                                          |
| Hauts-de-France                     | 3.345                                                                          |
| Nouvelle Aquitaine                  | 1.670                                                                          |
| Normandie                           | 2.445                                                                          |
| Pays de la Loire                    | 1.985                                                                          |
| Centre-Val de Loire                 | 2.208                                                                          |
| Grand Est                           | 2.913                                                                          |
| Provences-Alpes-Côtes d’Azur        | 2.170                                                                          |
| Bretagne                            | 1.922                                                                          |
| Bourgogne-France-Comté              | 2.174                                                                          |
| Occitanie                           | 1.659                                                                          |
| Auvergne-Rhône-Alpes                | 1.951                                                                          |

Note: Row data is publicly accessible on Mendeley repository [4].

Fig. 1. Annual average of NO$_2$ concentrations in France between May 2018 and June 2019.

Fig. 2. Distribution of NO$_2$ measurements collected over France between May 2018 and June 2019.
2. Experimental design, materials, and methods

1) Satellite measurements

Launched in October 2017 by the European Space Agency (ESA), Copernicus Sentinel 5P [5] monitors the density of several atmospheric gases, aerosols, and cloud distributions affecting air quality and climate. The measurements are made by the state of the art instrument called TROPOspheric Monitoring Instrument (TROPOMI). The TROPOMI is a multispectral imaging spectrometer that detects solar radiation reflected or scattered back to space from Earth’s atmosphere and surface. As the spectral fingerprint of each target atmospheric trace gas is known, its concentration can be calculated through the identification of the unique fingerprints of these constituents in different part of the electromagnetic spectrum. Sentinel 5P is able to achieve global coverage every 24 hours, giving access to dense measurements over the entire globe. TROPOMI has more spectral bands than its predecessors: ultraviolet and visible (270–500 nm), near-infrared (675–77 nm), and shortwave infrared (2305–2385 nm). This allows TROPOMI to measure a wider range of atmospheric trace gases such as nitrogen dioxide (NO2), ozone (O3), sulphur dioxide (SO2), methane (CH4), and carbon monoxide (CO). In addition, it observes clouds and aerosols-related parameters, which can be fed into the retrieval algorithms of trace gases [3]. The list of standard S5P/TROPOMI L2 products is given in Table 2.

Overall, the dataset provides:

- geolocated total columns of ozone, sulfur dioxide, nitrogen dioxide, carbon monoxide, formaldehyde and methane–geolocated cloud and absorbing aerosol index
- other products are under development and will made available at a later date. They include geolocated tropospheric columns of ozone, geolocated vertical profiles of ozone, aerosol layer height, ultraviolet index, etc.

All of the mission’s measurements of atmospheric gases and aerosols are ‘column data’, which means they cover the full depth of the atmosphere. For some gases, advanced techniques and algorithms like ozone profile retrieval, the convective cloud differential, and the cloud slicing methods allow to have access to ‘tropospheric column densities’, ‘stratospheric column densities’ and ‘vertical density profiles’. When available, these variables are included in the dataset.

TROPOMI has a very high spatial resolution (3.5 × 7 km² for all trace gases, except for CO and CH4 that is 7 × 7 km²). Further, TROPOMI has an improved signal-to-noise ratio (2–5%) for measurements.
Fig. 4. Spatial distribution of NO$_2$ measurements across France by month.
Fig. 5. Annual average of NO2 measured between May 2018 and June 2019 by administrative area.
under low albedo conditions. Data gaps are documented by Copernicus and can be consulted on the mission page.

2) Processing workflow

For further analysis of TROPOMI data, we produced an aggregated product on a regular grid with spatial resolution of 3.5 × 7 km (0.01 × 0.01 arc degrees). Every orbit that TROPOMI measures has a different spatial distribution of grid cells, depending on the viewing zenith angle at the moment of the observation. For this reason, we resampled each product for the area of interest on this single grid and binned the dataset by latitude/longitude WGS84 projection. The quality of the individual observations depends on many factors, including cloud cover, surface albedo, presence of snow-ice, saturation, geometry etc. In Sentinel 5P, a layer summarizing the different factors affecting the quality of the measurements is provided. This aggregate measure called ‘quality assurance value’ (‘qa_value’) can be used to screen poor quality pixel. This ‘qa_value’ is a continuous variable, ranging from 0 (no date) to 1 (all is well). To filter out errors and problematic retrievals we excluded measurements with a ‘qa_value’ < 0.5 following the Copernicus specifications [3].

Acknowledgments

The authors are grateful to the Luxembourg Institute of Socio-Economic Research (LISER) for funding the internship of Bilel Omrani, École Centrale de Lyon-France and Polytechnique Montréal-Canada, for data collection and processing with Python. The authors would also like to acknowledge the European Spatial Agency for providing the API for the Sentinel 5P Hub. The content is solely the responsibility of the authors and does not necessarily represent the official views of the LISER.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2019.105089.

Conflict of Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] V.V. Putrenko, N.M. Pashynska. The use of remote sensing data for modeling air quality in the cities, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2017, p. 57, 4.
[2] Sentinel-5P. https://github.com/bilelomrani1/s5p-tools. (Accessed July 2019).

Table 2

| Product            | Main Parameter                                | (Planned) released |
|--------------------|-----------------------------------------------|--------------------|
| UV aerosol index   | Aerosol index                                 | Released           |
| Aerosol layer height | Mid-level pressure                           | Released           |
| Carbon monoxide (CO) | Total column                                 | Released           |
| Cloud              | Fraction, albedo, top pressure                | Released           |
| Formaldehyde (HCHO) | Total column                                 | Released           |
| Methane (CH4)      | Total column                                 | Released           |
| Nitrogen dioxide (NO2) | Total, tropospheric, stratospheric column | Released           |
| Ozone profiles     | Total and tropospheric profiles              | Late-2019          |
| Sulfur dioxide (SO2) | Total column                                | Released           |
| Ozone (O3)         | Total column                                 | Released           |
| Tropospheric ozone (O3) | Tropospheric column                        | In development      |
| Ultraviolet (UV)   | Surface irradiance erythemal dose            | In development      |
[3] Copernicus. https://www.copernicus.eu/en. (Accessed 10 June 2019).

[4] Row data is publically available on the following repository. https://drive.google.com/drive/folders/1t5vbQq1g0Uj3a37ScT26JYog43bkL2EWPJ?usp=sharing.

[5] J.P. Veefkind, et al., TROPOMI on the ESA Sentinel-5 Precursor: a GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sens. Environ. 120 (SI) (2012) 70–83, https://doi.org/10.1016/j.rse.2011.09.027.