
A Geospatial Time-aware Web Interface to Inform about Air Pollution and Population Exposure in a Big City and Its Surroundings

Maria Paola BOGLIOLO and Giovanni CONTINO

INAIL – Research Division, Rome / Italy · m.bogliolo@inail.it

This contribution was double-blind reviewed as extended abstract.

Abstract

We developed a GIS-based web-mapping system aimed at providing specialists, stakeholders and population with a simple, yet scientifically rigorous, medium to gain information about air pollution and health risk, over the city of Rome (Italy), combining a geo-spatial visualization with easy access to the time dimension and to quantitative information. The study is part of the EXPAAH (*Population Exposure to PAHs*) LIFE+ EC project, which has the overall goal of identifying and quantifying peoples' exposure to atmospheric Polycyclic Aromatic Hydrocarbons (PAHs) in the city of Rome, and to assess the impact on human health. The system is completely built on open source tools.

1 Introduction

The EC funded the LIFE+ EXPAAH (*Population Exposure to PAHs*) project (<http://www.ispesl.it/expah/index.asp>), whose conclusion is scheduled for June 2014, and which points towards addressing the environmental and health problems caused by emission, dispersion, and transformation of Polycyclic Aromatic Hydrocarbons (PAHs) in the atmosphere. Its overall goal is to identify and quantify exposure amongst children and elderly people to PAHs contained in particulate matter in the city of Rome (Italy), and to assess the impact on human health. One of the project tasks was the development of a GIS aimed at sharing information among partners and disseminating results to the public.

Studies dealing with air quality, population exposure and possible health effects greatly benefit from GIS techniques to discover spatial patterns and multi-factor spatial relationships. Main challenges for the GIS developer include the multi-disciplinary nature of the study, that requires building a coherent model to manage very different data types, and the effort of emphasizing spatial information of data mainly treated along the timeline. These aspects were faced and worked out in a GIS-based web-mapping system.

2 The Geographic Information System

The core of the system is a GIS that collects and handles the time-varying data produced by the project research activity, enhancing the geospatial point of view of every information

source that could take advantage of it, and presenting it through interactive maps. The system is completely built on open-source tools.

2.1 Input datasets

The main data source are the results of the research carried out by the project teams (see the Project Technical Reports at <http://www.ispesl.it/expah/pblTechRep.asp>): they include measurements and model simulations of PAHs and PM_{2.5} concentrations in the atmosphere carried out in the period May 2011 – October 2012. Geographic contextual information was added using thematic layers and basemaps from open/public resources. The main features of the project dataset are summarized below:

- Indoor and outdoor chemical measurements at monitoring sites (air quality network stations, schools, offices, private houses, bus/car tracks, personal sampling), including different time windows and time steps. Data were provided in spreadsheet tables.
- Features of the meteorological stations collecting data that fed the air pollution simulation model.
- One-year daily concentrations of various pollutants in the first layer of the atmosphere, calculated by simulation modeling at the nodes of a horizontal grid, 1 km resolution, over a 60x60 km² domain centered on the city of Rome. Data were provided as netCDF, an archive data format widely used in the atmospheric science community. Km units were used for coordinates; no geo-referencing metadata were included.
- Gridded emission inventory data for the year 2009, provided in netCDF with geographic and projected coordinates, but incomplete geo-referencing metadata.
- One-year daily individual PAHs and PM_{2.5} exposure values for two target subjects (young/elderly), and PAHs toxicity-equivalent concentration; annual and seasonal averages were also provided. Same grid as the simulations, netCDF format.

2.2 Data storage and publishing

Most datasets underwent some sort of data pre-processing. When the spatial and geometric information was not directly available, it was created (e.g. extracting the car/bus tracks from the road dataset; geocoding monitoring sites); suitable ancillary information was extracted from public datasets: for example, population density maps for the total population and two age intervals (0-15 and >65y) were calculated from demographic data (ISTAT 2013); suitable land-use classes were obtained by class remapping of the CORINE 2006 land cover map (EEA 2007). Every multi-temporal vector dataset was attributed with one or two (for time-intervals) *time* fields, where a timestamp temporally locates the feature and its attributes. A vector model (points, lines) was used for ground measurements and for other ancillary data; a grid model for continuously variable parameters (emission inventory maps, concentration maps, exposure maps).

All vector data was loaded in a PostgreSQL/PostGIS database; grid data were stored in external folders as GeoTIFF files. Every dataset was then published as an OGC Web Map Service (WMS 1.3.0) in GeoServer: all time-varying information was time-enabled.

The netCDF extension of Geoserver enables direct publishing of multitemporal netCDF data. However, our experience, confirmed by literature (UNIDATA 2011), demonstrated that the various flavors of netCDF formats, and the lack of tight standard for georeferenced metadata still prevent a full interoperability of this format in GIS tools. Therefore, while

desktop platforms were able to load the files, Geoserver was not, even after pre-processing the data to integrate missing geo-information, and to convert coordinate units. To get around the problem, we exported the one-year daily model simulations netCDF files in GeoTiff, obtaining a set of 366 GeoTIFF files for each pollutant. The Geoserver “ImageMosaic” extension was then used to publish the multitemporal GeoTiff datasets as time-enabled WMS services.

SLD styles were used to apply symbology to the data. To enable the user to compare the data both spatially and temporally, a consistent symbology was applied across datasets and for each dataset across time. Consequently, a single color scale characterizes each pollutant, marking its concentration at ground sites and at simulation modeling cells, at every time step. The number of scale intervals and the corresponding concentration values were chosen so as to optimize the spatial mapping at most dates, using statistics of the whole datasets.

3 The Web Maps

By means of Javascript, Web Map Services, which are based on the OpenLayers library, feed a set of web maps (<http://www.ispesl.it/expah/expahwebgis.asp>) which can be queried, allowing the user to navigate into the results of the project without requiring GIS knowledge, specific software, or plug-ins. These maps are designed to be interactive, with functions like zoom, pan, layer selection, query, pop-ups, and *time-aware*. Each map can be queried both by location, to retrieve quantitative values, detailed information, images and web-links, and by time, using a set of *temporal* tabs, that give access to the time dimension.

A careful planning of map contents, layout and rendering was necessary to grant easy access and readability. Three types of maps have been designed to contain the available information: opened in sequence, they lead the user through the study carried out.

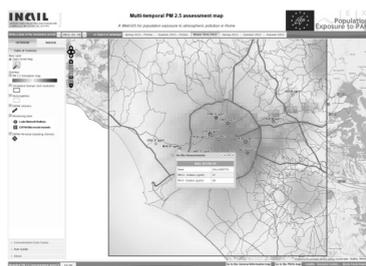


Fig. 1: Screenshot of the multitemporal web map, illustrating the results of the PM2.5 assessment activity

A *General information* map conveys information about the experimental campaigns performed to collect meteorological and chemical measurements. It also outlines the emission context of the area through interactive display of emission inventory maps for various pollutants and of the main point sources. Different experimental campaigns can be selected to display measurement sites for the campaign. Details on each site can be accessed through pop-ups.

Two *Pollution* maps are contained in the results of the pollutants assessment activity (see figure 1). They illustrate the spatial and temporal distribution of total PAHs, Benzo[a]Pyrene, and PM2.5 over the city of Rome. They were obtained through the concentration maps produced by modeling simulations, relating them to the geographic

context and highlighting the spatial-temporal relationships with the concentration measured at the ground. Quantitative data are available, both for ground measurements (through labels and pop-ups), and for simulations (through a *value* box). Ground measures have three display options, based on indoor values, outdoor values and an indoor-outdoor bar-diagram. Calendars for each experimental campaign and a date-box allow temporal navigation of the map. At each date, the map is updated with raster and vector data, depicting the situation at the selected date, with proper symbology. The last set of maps concerns Health. The goal is to provide insight into the impacts of the air pollution on people living and working in Rome and its surroundings. Annual and seasonal average PAHs toxicity-equivalent concentrations and PAHs and PM_{2.5} individual exposure values for the two target subjects (elderly and children) can be selected and combined with ancillary data about the urban context: these later help the user to draw considerations about the overall health risk and its spatial pattern. A geo-location box allows the user to know exposure levels at specific locations.

4 Conclusion and Outlook

The web mapping system of the EXPAH project is the result of the efforts to develop a very simple interface to access information in three dimensions (2D spatial and temporal), at two levels of complexity: it conveys complete, scientific, quantitative content to scientists, and provides an intuitive communication medium for stakeholders and the public. The first aspect revealed its advantages during the development of the project, for data checking and evaluation tasks: each team could analyse the quantitative, temporal and geometric relationships of its results with those of the others, and with the actual geographic context, highlighting features worthy of further analysis.

Finding effective ways to communicate risk for human health, as well as the environmental determinants for it, is a topical and challenging task (WHO 2013). Maps are a very efficient means to transfer information, giving an intuitive way to interpret complex data; the possibility to interact with the maps catches the attention of users, and invites them to go deeper into the presented information. However, big questions arise about how this information is perceived by the user. Merely choosing a different color threshold in a pollutant exposure concentration map drastically changes the message delivered. While developing the application, these aspects were faced and taken into account, e.g. choosing threshold levels coinciding with regulation levels when available. However, going deeper into these aspects was beyond the goal of the project: it is a possible, future goal of the next “GIS for health” study.

References

- EEA (2007), CLC2006 technical guidelines. Technical Report, 17/2007.
- ISTAT (2013), 15^o Censimento popolazione e abitazioni 2011.
http://www.istat.it/it/censimento-popolazione/popolazione-2011_
- UNIDATA (2001), The netCDF users' guide – Limitations.
http://www.unidata.ucar.edu/software/netcdf/docs/netcdf/Limitations.html#Limitations_
- WHO (2013), Health and environment: communicating the risks. F. THEAKSTON Ed., 68 p.