Technical report on Creation of a geospatial, time-aware web interface to disseminate EXPAH results – Action 8.1

Authors
M. P. Bogliolo ¹, G. Contino ¹

¹ INAIL Research Division - Dipartimento Installazioni di Produzione e Insediamenti Antropici

June, 2014
1 Executive Summary

Action 8.1 concerning “Integration of data and results by means of GIS techniques”, was aimed at managing the spatial data of the project, to:

- facilitate information exchange among partners;
- support pre-processing of data to be given as input to simulation models;
- optimize ways to analyse and present intermediate and final results having a geographic component;
- develop a visual system to disseminate final information concerning air quality, exposure risk and possible mitigations to stakeholders and to involved population.

The result of the activity carried out is a GIS-based web-mapping system, aimed at providing specialists, stakeholders and population with a simple, while scientifically rigorous way to obtain information about air pollution and people exposure to air pollution in the city of Rome and its surroundings. It combines a geo-spatial visualization approach with easy access to time dimension and to quantitative information.

The core of the system is a GIS, which database contains data and results of the project research activity. They include emission inventory data and derived emission maps for several pollutants used as input to the simulation model, daily maps of Polycyclic Aromatic Hydrocarbons (PAHs), Benzo[a]Pyrene (BaP) and PM2.5 atmospheric concentration for the period June 2011 - May 2012 obtained by simulation modeling, daily indoor and outdoor concentration measurements for various pollutants performed during the experimental campaigns, personal exposure measurements, daily and average exposure maps estimated for the year 2011/2012 and exposure scenarios for the year 2020. Geospatial ancillary information coming from public sources at the European national and local levels have also been added to contextualize the maps and to support interpretation.

A completely open-source approach was used, using open-source toolsets, open ancillary datasets, and open standards for web publication services. INSPIRE compliant coordinate reference system and web services standards have been used.

A set of web mapping applications query the web services to produce a set of interactive time-aware thematic maps.

Finding effective ways to communicate risk for human health, and environmental determinants for it, is a topical and challenging task: the web mapping system developed can be thought as a prototype of a possible model to disseminate scientific results on these items: it provides a sight into impacts of air pollution on people living and working in a big city, conveying information about the overall exposure, its spatial pattern and levels at specific locations.
2 Introduction

Multidisciplinary studies like EXPAH, aimed at characterizing and investigating causes and effects of phenomena happening and evolving on a given spatial domain, can greatly benefit of Geographic Information Science, taking advantage of its “holistic” approach: it facilitates discover and analysis of spatial patterns and of spatial relationships among factors and determinants affecting the phenomenon.

A geographical point of view can also provide a visual approach to information sharing and communication, exploiting the intuitive and immediate interpretation provided by maps. A geographical platform accessible from the web can hence address the need of quick and viable routes from research to stakeholders. Indeed, results achieved by research investigating risks for human health and environmental determinants for it, are of great and urgent interest not only for the scientific community but also for decision makers, to translate scientific findings into policies, and for exposed population, since awareness promotes healthy behaviors. Difficulties in finding simple and straightforward ways to communicate research findings could delay adoption of effective actions.

The web mapping platform developed for the EXPAH Project has been planned using the above considerations as a reference, and with the goal of creating a prototype for dissemination of scientific results about air pollution and population exposure, able to provide stakeholders with a simple tool to analyze spatial and temporal characteristics of air pollution over a big city, and to have a sight into impacts on people living and working in it.

Creating a web mapping platform involved the conversion of data produced by the different project activities into a homogeneous, organized, structured and self-describing dataset: this work consisted in designing, creating and populating a Geographic Information System (GIS).

The GIS was used as a platform to pre-process and format the data, to organize and manage all the geospatial data concerning the project, both coming from the project as a result, and from external sources as ancillary information. The GIS was also used to plan the web-maps. Its database acts as the GIS server data store.
3 The Geographic Information System

Creating a web mapping platform for the EXPAH project involved the conversion of data produced by the different project activities into a homogeneous, organized, structured and self-describing dataset: this work consisted in designing, creating and populating a GIS.

In a multidisciplinary study, the GIS specialist is required to build a coherent model to manage very different data types, and to emphasize spatial information of data that are sometime mainly handled by a different point of view, e.g. along the timeline.

In order to support the GIS development, a Geographic data handling baseline Guide (Bogliolo, 2011), was distributed among the project teams during the first year of the project, containing basic information on geo-spatial concepts, requirements and indications about aspects needing shared choices. Spatial reference, coordinate accuracy, metadata, spatial models, data formats, possible issues and ways to overcome them, warnings about ancillary data already in use, availability of validated ancillary datasets, were the main topics. A questionnaire was also included, aimed at collecting information about data owned, needed, expected to be produced, and possible geo-processing and display strategies of interest.

However, most partners could not take into account the guide and the questionnaire: main reason is the fact that each unit of a multidisciplinary research has its own way of processing, storing, displaying data and presenting results: these aspects are determined by the specific approach to data analysis, that is not necessarily referred to a geographic coordinate system. As a consequence, emphasizing the geographical aspect of data that could take advantage of it, by adding, validating, fixing the geospatial information, and studying strategies to group the Project datasets into a coherent and interoperable model was an important phase of the GIS development.

Result of this activity is a GIS that collects and handles all data produced by EXPAH for which a geographical dependence could be identified. Data were stored in a geographical database: a geospatial point of view was used to organize each dataset, while preserving dependency on time. A set of ancillary data coming from open/public sources was also added, to provide the geographic context to the project results and to support interpretation. Main features of the GIS are described in the following chapters.

3.1 EXPAH-GIS architecture

The HW/SW architecture of the EXPAH GIS is depicted in Figure 1.

Data processing, organizing and managing were performed on Desktop GIS software platforms, mainly ESRI ArcMap 10.0 and the opensource Q-GIS (QGIS Development Team, 2013). ArcMap was used for planning the web maps as well.
The ToolsUI-4.3 Java application from UCAR/UNIDATA was used to explore the netCDF datasets and edit their metadata. Exelis IDL and ENVI 4.7 image processing software were used respectively to edit netCDF data coordinates and to help raster processing.

The web section of the GIS is located on two virtual Windows Servers having public IP address. This section relies completely on open-source tools and standards.

The HW/SW architecture of the EXPAH GIS is depicted in Figure 1.
Data processing, organizing and managing were performed on Desktop GIS software platforms, mainly ESRI ArcMap 10.0 and the opensource Q-GIS (QGIS Development Team, 2013). ArcMap was used for planning the web maps as well.

The ToolsUI-4.3 Java application from UCAR/UNIDATA was used to explore the netCDF datasets and edit their metadata. Exelis IDL and ENVI 4.7 image processing software were used respectively to edit netCDF data coordinates and to help raster processing.

The web section of the GIS is located on two virtual Windows Servers having public IP address. This section relies completely on open-source tools and standards.

The data-store is a geographic database built on the PostgreSQL/PostGIS DBMS. However, the raster datasets were stored directly in the file system, since they are easier published in this form than when stored in the DB.

Data have been published as standard OGC (Open Geospatial Consortium) Web Services using GeoServer 2.4.0 (GeoServer, 2014) as the GIS server, running under Apache Tomcat 7.

The other internet server hosts the Javascript web applications, that exploit the services to produce the web maps. The applications are based on OpenLayers, jQuery and ExtJS/GeoExt libraries.

A dedicated web page of the EXPAH website (http://www.ispesl.it/expah/expahwebgis.asp) gives access to the web maps through a set of links that re-direct to the corresponding Javascript applications, accompanied by a short description of each map.

### 3.2 Input datasets

Only information that could be georeferenced was considered for the GIS. The following EXPAH datasets resulting from the research activity have been included in the geographical database:

- Indoor and outdoor measurements of pollutants concentration at experimental sites: they include data obtained from sampling carried out during the experimental campaigns at different microenvironments (schools, offices, private houses, bus, cars) (Cecinato et al., 2013; Gatto et al., 2013); data acquired by continuously monitoring Regional Environmental air quality Network (REN) stations, and personal exposure measurements from samplers worn by volunteers (Gatto et al., 2013, Gherardi et al., 2013). Data were acquired on different time windows between April 2011 and October 2012, with different time steps (daily, day/night and weekly averages). General information about the experimental sites was included. Datasets are the results of EXPAH Actions 3.2 and 3.3: they were provided as spreadsheet tables, grouped by pollutant. Most sites were only located by address.
LIFE+ Project EXPAH. Creation of a geospatial, time-aware Web interface to disseminate EXPAH results

- Characteristics of the meteorological stations which data have been used in the Project to feed the meteorological pre-processor of the air pollution simulation model. The dataset is a list of the main features of the meteorological laboratories set up for the project plus general information about the meteorological monitoring stations which data integrated the EXPAH dataset. Action 3.4 was responsible for information provided.

- Average emission inventory data for diffuse (municipal level), linear and point sources: they were provided by Action 4.1, coming from other institutions (e.g. point sources – ISPRA05) or as result of Action 4.1 research activity (disaggregation of diffuse emissions at the municipality level, emission calculation from traffic levels). Information was formatted as vector georeferenced datasets.

- Gridded year 2009 average emission data for various pollutants obtained by integration of all available emission inventories (Radice et al., 2012). These data served as input to the chemical transport model simulations. This dataset is the result of Action 4.4: it was provided as a set of netCDF files (Rew and Davis, 1990). Including both geographic and projected coordinates, but incomplete georeferencing metadata.

- Gridded daily concentration data for PAHs, BaP and PM2.5 in the first layer of the atmosphere, calculated by simulation modeling (Gariazzo et al., 2013, Silibello et al., 2014). Data refer to one year of simulations (June 2011 – May 2012) and were provided in netCDF files. Km units were used for coordinates; no georeferencing metadata were included. Annual and seasonal averages were also added. Data is a subset of Action 4.5 research activity results.

- Gridded daily levels of exposure to PAHs, BaP and PM2.5 for the target subjects (young and elderly people) (Gariazzo et al., 2014). They refer to the same period (June 2011 – May 2012) as simulations and were provided on the same grid and in the same format. Annual and seasonal average exposure levels were also added. Data result from Action 5.4 research activity.

- Gridded yearly and seasonal averages of total PAHs toxicity equivalent concentrations (TTECs). PAHs TTEC expresses the PAHs concentration as the sum of the concentrations of the PAHs congeners taken into account in the study, normalized to the BaP toxicity level. Values were calculated in the frame of Action 5.4 activity (Gariazzo et al., 2014). NetCDF archive format was used to store data.

- Annual and seasonal average gridded levels of exposure to PAHs, BaP and PM2.5 for the target subjects (young and elderly people) estimated for year 2020 on the basis of two emission scenarios: the first corresponding to application of current legislation (CLe), and the second obtained by adding the additional reduction measure of complete substitution of biomass combustion for domestic heating with natural gas (CLeM) (Silibello et al., 2014). These data result from the research activity of Action 7.2 and were provided as netCDF files.

Several datasets were added to the EXPAH ones, in the GIS database, to be used as ancillary information, to give a geographic context to EXPAH results, to support data analysis and to help users in map interpretation. Action 8.1 was responsible for exploring,
downloading and processing these data. A choice towards public/open datasets was made: all data are freely available from public/official organizations at the world/European/national/local level. A list is given below:

- Main roads network: source is the DB-Prior10K database from Intesa-GIS Project, a partnership among different Italian government entities (Intesa-GIS, 2007);
- Administrative boundaries at various levels and census tracts: source is the Italian national Institute for Statistics (ISTAT, 2013);
- Traffic limitation areas: open data provided by the Mobility Agency of Rome (Rome Mobility Agency, 2013).
- Digital Elevation Model at 30 m and 90 m resolution: the 30m DEM comes from the ASTER Global Digital Elevation Model Version 2(GDEM2): it is a METI and NASA product, available through different interfaces (U.S.G.S. LP-DAAC, 2014); the 90 m DEM is extracted from the SRTM digital elevation data originally produced by NASA, processed by the CGIAR Consortium to fill data voids (CGIAR-CSI, 2008);
- Shaded Relief rendering elevation maps: this dataset was obtained by processing the 30 m ASTER GDEM;
- Land Use - Land Cover maps from the European CORINE 2006 DB (EEA, 2007);
- Selected land use types extracted from a vegetation map (Camera di Commercio, Industria e Artigianato della provincia di Roma, 1979) made available through the Roma province WebGIS;
- Year 2011 demographic data for the target population (0-14y and over 65y): by census tract: source is the Italian national Institute for Statistics (ISTAT, 2013);
- Schools of the Rome municipality (20011 update): tabular data of the Italian Ministry for Education (MIUR, 2011a, MIUR, 2011b);
- Public hospitals for the Rome city municipality: from various public and commercial sites, edited and integrated;
- OpenStreetMap vector data ((Haklay and Weber, 2008; OpenStreetMap, 2014).

### 3.3 Data pre-processing

Most datasets underwent processing to make them suitable for GIS and web mapping requirements.

When spatial information was not available, it was created. Geocoding was necessary for most pollution and meteorological monitoring sites, for schools and hospitals: geocoding from address was further refined by visual inspection on Google Earth and Google maps. An automatic geocoding tool was created using the Google Maps APIs to automate geocoding of schools and hospitals. Geocoding errors due to unknown addresses were manually edited and fixed. Pollutant measurements carried out on cars and bus have been georeferenced by extracting the traveled route on the road network dataset (traffic information was transferred as attributes to the segment) and attributing measurements to the whole line. Personal exposure measurements were georeferenced through the living
house of the volunteers, assumed to be the “center of gravity” of their activities in consideration of their age (elders and children). Coordinates of the REN stations and of the public meteorological stations were downloaded from the owner website and verified by visual inspection.

Attribute tables containing general information and pictures about the measurement sites were built starting from descriptive information sheets prepared by the teams responsible for data acquisition.

Measurements and descriptive attributes of the monitoring sites were linked to geometry through a site code.

Most datasets of the EXPAH project are multi-temporal. Maximum detail is one day for most of them. The only exception regards personal exposure measurements, that in some cases, have been taken separately for night and day. In order to make these measures comparable with the other data, concentration obtained by summing the day and the night values was used to symbolize personal exposure levels, while retaining the original values in the database. This implicates errors (underestimates) for the few cases in which the day or the night value was not available. A disclaimer about that was included in the Information page about the maps: the actual time span of the data is available through tables.

Date timestamps were used all over the GIS. Each pollutant measurement was attributed with one timestamp for daily measures and with two timestamps (time interval) for weekly averages (PAHs data): in this latter case the daily value is the same for every day of the interval. Each experimental site was attributed with intervals corresponding to the measurement campaigns it was part of, and with a list of the pollutants sampled during that interval.

Most experimental data reported PAHs concentration for each of the evaluated congeners, and as the sum of them. The two teams in charge of data acquisitions evaluated a different set of congeners, and simulation modeling took into account only 4 of them. So, in order to make all data comparable, two new values of total PAHs were calculated: one expresses the sum of the congeners evaluated at all sites, and one is the sum of the 4 congeners taken into account by simulation modeling.

As ancillary data are concerned, subsets were extracted from big datasets, using the domain of 1km resolution simulations as clipping domain. Administrative boundaries have been obtained by generalization of the census tracts.

Demographic data were used for two objectives: Vector density maps for total population and for the age ranges of interest (young and elderly people) were used to produce maps of densely populated areas (two levels) to be shown on the web maps; raster maps representing number of residents (separated by age range and gender) in each of the 1 km cells coinciding with the simulation modeling cells were provided to teams dealing with
exposure, to estimate population exposure across the domain. These two datasets were set starting from demographic data by census tract updated to the last (2011) census.

Suitable land-use classes to be shown on the web maps were obtained by selecting and remapping LULC classes.

Shaded relief images are an effective way to display relief onto maps. They were calculated from DEM setting a suitable zenith and azimuth illumination angle.

A deep processing was necessary for NETCDF data. NetCDF is a very flexible archive format, able to store many different types of grid datasets; projected coordinates can be added together to- or without- geographic coordinates. However, the format lacks tight standards for georeferencing metadata (see for example UNIDATA, 2011), even in the most complete “Convention” (version) used in the atmospheric community (CF - Climate and Forecast). This fact still prevents a full interoperability of this format in GISs. EXPAH data came in the older COARDS convention; simulation modeling outputs had only projected coordinates, expressed in km, and no georeferencing information; some descriptive coordinate specifications were missing as well. File editing was necessary to convert coordinate units from km to m (through an IDL program) and to add georeferencing metadata and specifications (using GDAL tools (OSG Foundation, 2013) to create a metadata template, and ToolsUI-4.3 for editing). New netCDF files could be correctly loaded by the desktop GIS SW, but not by GeoServer. To get around the problem, we exported the netCDF datasets in GeoTiff. Some files contained a single grid dataset, that was simply exported in a single GeoTiff file. The multi-temporal model simulations were contained in 366 layers (one year of daily simulations) netCDF files: each simulation was exported first in a multiband GeoTiff and then in a set of 366 GeoTiff files. These latter were named after a timestamp of the date the maps refers to: this format made the files ready for web publication (see next paragraph).

All datasets have been converted to the projected coordinate reference system (CRS) (UTM32N/WGS84) chosen for the GIS.

Metadata have been used to document geospatial datasets. They were prepared as XML documents containing the mandatory information required by the INSPIRE Metadata Implementing Rules (EC-JRC, 2013), using the metadata editor of the EU INSPIRE geoportal.

### 3.4 Outline of the EXPAH-GIS

A vector representation model was used for all experimental measurements, and for many ancillary datasets. A raster model was used to represent parameters continuously varying across the domain, i.e. the gridded concentration, emission and exposure fields, DEMs, Shaded relief images.
The layout of the geospatial database derives from the definition of a conceptual model for the information structure. Information was organized by theme, identifying 4 thematic sections. The Cartography section contains information providing the geographical context to Project results: it contains ancillary data from external sources. The Atmosphere Properties section is devoted to information acquired by direct measurements: it includes location and description of the chemical and meteorological sites and the pollutants concentration data acquired at the various microenvironments and with personal samplers. The Emissions section contains information that describes the emission context of the study domain: it includes the average emission maps for year 2009 and the point, diffuse and linear emission inventory data used to produce them. The Modeling section groups the outdoor concentration information coming from simulation modeling. The Health Risk section deals with aspects that contribute to define impacts and risks for human health: it collects the exposure maps, both corresponding to the 2011/12 outdoor concentrations and to the 2020 scenarios, the TTEC PAHs maps and other ancillary data specifically included to help interpretation of exposure levels (e.g. population distribution and density, schools, hospitals).

This outline was used as the basis for the geographic database design, and then further detailed defining attributes of each datasets and populating database tables with information. The database was first developed in ESRI environment (ESRI File Geodatabase format) and then exported in a PostGres/PostGIS database, where the thematic organization was replicated using different schemas: one for each theme (Figure 2).

Raster datasets were stored in the ESRI geodatabase during the developing phase, but moved to the file system when exported to the Internet Server, since publication is easier in this format than from the DB.

The GIS design included the choice of the CRS to be used for data and maps. All data were projected in the UTM system, zone 32N, on the WGS84 datum. This latter is considered to be coincident with the INSPIRE (European Union, 2007) mandatory reference datum, ETRS89, at the scales of representation used in the study (less than 1:10000).

The main GIS geographic domain was chosen to be coincident with the horizontal domain of the 1 km resolution simulation modeling: It is an area 60x60 km² wide, centered on the city of Rome. The actual extension of the web maps is larger, since some datasets extend outside this domain: this happens for few chemical and meteorological sites and for some emission inventory data: for this reason few ancillary information providing the basic geographic context was extended accordingly.

The reference scales for data display were defined in function of the resolution and detail of the Project datasets: since the main data are given on cells of 1 km resolution, the main reference scale was set to 1:250000. However, the chemical measurement sites have
been located with much higher accuracy: sometimes they are only few hundred meters of distance each other; furthermore, the user can be interested in inspecting the actual context of the experimental sites (e.g. its position with respect to roads, buildings, green areas). For these reasons, high resolution ancillary datasets (e.g. OpenStreetMaps, schools, residential areas, urban parks) were included, and web maps have been planned to allow high resolution display (about 1:20000).

Figure 2. Example of the PostgreSQL/PostGIS geodatabase organization.
4 The Web maps

4.1 Map planning and data publication

To grant easy access to- and readability of EXPAH results, a careful planning of the web maps was required. Every map was designed based on answers given to the question: “Which information can a map communicate widely, promptly and better than other media?”. Contents, arrangement, rendering symbology and strategies, layer names, labels, user interaction and functionalities were all defined in this planning phase accordingly to the goals identified for each map. Map design was performed on desktop GIS platforms, setting up local maps simulating the maps to be published. As a consequence, an ArcMap version of the maps and of the database is available and was released on DVD.

Vector and raster datasets have been published by GeoServer as OGC (Open Geospatial Consortium) compliant Web Services: WMS 1.3.0, WFS 2.0.0, WCS 1.1.1. These versions are compliant with the INSPIRE directive requirements (EC, 2014).

Raster and vector layers were styled for publication as WMSs using the Style Layer Descriptor (SLD) XML-based markup language.

To enable users 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 set so as to optimize the spatial mapping at most dates, using statistics of the whole datasets. To enhance little differences, double-hue scales (e.g. blue – white/white – red) were used. Since in these color ramps the threshold color between two hues (e.g. white) can be perceived as a warning level, its position in the scale was carefully selected, assigning it to concentrations corresponding to Italian regulatory values if available, or to threshold values taken from international literature. For example, in Italy BaP is regulated according to a daily value of concentration, based on one year average, not above 1 ng/m$^3$. EXPAH demonstrated that as this value was not exceeded during the study time interval, it is exceeded during cold months, both daily and as average of the heating period (period of the year when domestic heating is on). To enable users to point out this aspect, the threshold color (yellow) marks the BaP concentration range centered at 1 ng/m$^3$; to keep the total PAHs and BaP maps comparable, the threshold color (white) of the total PAHs concentration color scale was set to be coincident with values repeatedly associated to BaP concentrations of 1 ng/m$^3$. A similar approach was used for PM2.5.
OpenStreetMap (OSM) was used as the basemap. Directly using the OSM raster tiles, that are projected onto the Web Mercator CRS, involves all the published Web Map Services to be projected onto the same CRS when the web map is created. This is well accomplished by GeoServer, but returned bad results for the simulation modeling grids due their low (1km) resolution, causing mis-alignments with the vector layers, which geometric transformation was instead accurate. The problem could be overcome by resampling the raster files to a higher resolution: however this would led to a huge increase of the dimension of the multi-temporal datasets (366 grids of 60x60km2 for each pollutant) that are queried in almost real-time by the multi-temporal functionality of the web maps (described below). Therefore, we prepared a suitable version of the OSM map, downloading the OSM vector data for our geographic domain: these were stored in the Postgres DB and projected to the UTM32N/WGS84 CRS. The new version was SLD-styled using style similar to the original OSM, and served as tiled WMS through GeoServer.

Every web service containing time-varying information was time-enabled: it allows the service to be queried by time. The “netCDF” extension of GeoServer enables direct publishing of multi-temporal netCDF data. However, as stated in the previous paragraph, the netCDF files in output from the simulation models could not be read by Geoserver even after pre-processing. Consequently, the GeoServer “ImageMosaic” extension was used to publish the multi-temporal GeoTiff datasets as time-enabled Web services.

4.2 Web mapping applications

A set of Javascript web applications, mainly based on the OpenLayers library, consume the Web Map Services exposed by GeoServer to create web maps that allow the user to access information about the EXPAH project via an interactive map viewer.

Maps have been designed to be interactive, with functions like zoom, pan, layer selection, query, pop-ups, and time-aware, that is queryable by date. Maps are also conceived to be used and interpreted at two levels: the main approach is very simple and intuitive, to make maps accessible to non-specialists, like people living in the study area: no GIS knowledge, neither specific software or plug-ins are required to access the maps; a short user guide is available directly from the web map page; quantitative data are also available, through labels, boxes and pop-up tables, targeted to the EXPAH partners, the scientific community and the technical staff of decision makers.

Information shown in the web maps is a subset of the EXPAH dataset included in the GIS: even if maps also carry quantitative information, the main approach to their interpretation is visual, i.e. qualitative: showing datasets that are highly spatially correlated each other is of little use: this is the case, for example, of the one-year daily exposure levels: these datasets are closely correlated with the daily concentration values, and hardly distinguishable from them at a visual inspection.
Three types of maps have been designed to present results of the EXPAH project: opened in sequence, they lead the user through the study carried out, providing a sight into the air pollution outline of the city of Rome, its relationships with emission sources, its dependence on seasons, allowing to explore the level of agreement between models and measures, exposure levels for residents, their spatial pattern and values at specific locations. Maps can be accessed from the EXPAH-GIS web page at http://www.ispesl.it/expah/expahwebgis.asp.

Maps have a common layout that includes a short map description that appears while map is loading, and can be retrieved later from the main map menu. This latter also includes a user guide, the table of contents and some buttons to interact with the map. A bar on top of the map is the “time bar” containing tools that allow map navigation along the time line. The main study area is shown in the maps as a blue frame. Quantitative data for the grid datasets are provided in a “value” box displaying the value at the cell under the mouse pointer, it is continuously updated while moving the mouse. Pop-up tables show quantitative data at the measurement sites when they are queried by location.

4.2.1 EXPAH Experimental campaigns map

The first map (Figure 3) illustrates the chemical and meteorological monitoring activities carried out during the project, and outlines the emission context of the area. Both EXPAH sites and continuous monitoring stations have been included. Symbols identify the site type and the sampled microenvironment.

Location query displays general information, images, web-links. Time dimension is represented by the succession of experimental campaigns: all campaigns carried out, included the first inter-comparison one, are included. For each campaign, only sites that acquired data during that campaign are displayed. Details about measured pollutants and the actual duration of the campaign at the site can be retrieved in pop-up tables by querying the sites by location.

The emission context of the area is illustrated by the set of gridded emission maps for year 2009 obtained by Action 4.4 and used as input to the simulation model. Maps for various pollutants can be interactively displayed on the main map, selecting from a drop-down menu (Figure 4). Quantitative data are available. The main point emission sources of the emission inventory are also displayed: summarized emission data for them are available through query tables. The map can be found at: http://www1.ispesl.it/webgis/general_info.html.
Figure 3. Screenshot of the web map illustrating general information about the EXPAH experimental campaigns.
Figure 4. Display of the PM2.5 average emission map (year 2009) on the map of the EXPAH experimental campaigns.

4.2.2 PAHs and PM2.5 multi-temporal concentration maps

Results of air pollution characterization for year 2011/2012 are reported in two concentration maps: one concerns PAHs (total PAHs and BaP) (Figure 5 and Figure 6) and the other PM2.5.

They illustrate the spatial and temporal distribution of these pollutants in the first layer of the atmosphere through the outdoor concentration fields resulting from simulation modeling and the measurements taken at the experimental sites. Information to be communicated by this map was identified as:

- spatial pattern of concentrations;
- reliability of simulation, through relationships with measurements;
- seasonal and daily variability of concentration (both in value and in pattern);
- relationship between outdoor and indoor concentrations at different microenvironments, to figure the variation of exposure between indoor and outdoor staying.
Comparison of concentration values across space, time and across datasets is allowed by the use of a single color scale for each pollutant (Figure 5); quantitative comparison is granted by availability of numeric values. Communication of indoor/outdoor relationships was obtained enabling symbolization of sampled microenvironments by outdoor values, indoor values, and by an indoor-outdoor bar-diagram (Figure 6) that provides a prompt glance of the indoor/outdoor ratio and so of the infiltration coefficient.

The map is dynamic along the timeline, to display multi-temporal contents of the data: a set of temporal tabs along the “time bar” provide the user with two ways of time selection: calendars of every experimental campaign, and a date box to enter a date, with two arrows to go forward or backward continuously. These tools allow temporal navigation of the map. For each date, the web application sends an AJAX request to GeoServer, that returns data for the correspondent time step/interval. Raster and Vector resources are retrieved together as WMSs and a new map is produced depicting the situation at the selected date.

![Figure 5. Screenshot of the multi-temporal web map illustrating BaP daily outdoor concentration for a given day, both as retrieved by simulation modeling (grid data) and measured at sampling sites.](image-url)
4.2.3 Children and elders exposure map

Results of exposure modeling and calculation are reported in the “Children and elders exposure to PAHs and PM2.5” map (Figure 7 and Figure 8). It shows exposure levels to PAHs, B[a]P and PM2.5 corresponding to the 2011/2012 concentration fields. The map aims to highlight:

- differences between the heating and the non-heating period of the year: to this aim, yearly and seasonal (November 1st –March 31st and April 1st – October 31st) averages...
maps have been published rather than daily maps; users can easily select different periods of the year and load corresponding maps ((Figure 7 and Figure 8).

- role of infiltration in reducing the pollutant intake in target people (spending large part of the day indoors); this information is conveyed by the possibility to interactively select and compare exposure levels with outdoor concentrations for each average period. As in the previous map, a single color scale characterizes each pollutant, representing exposure concentrations and outdoor concentrations in all time averages, to make all data inter-comparable.

- spatial distribution of actual impacts: exposure is calculated cell by cell as the concentration a given person of the target groups is exposed to, when spending the day in the given cell according to a mean time scheme: it expresses a sort of potential exposure since it does not take into account the actual presence of people in the cell. Adding ancillary information such as land use and demography to the map is a simple way to qualitatively locate higher impact areas.

![Figure 7. Screenshot of the children elders exposure to PAHs and PM2.5 web map: average children exposure levels to PM2.5 for the non-heating period are displayed; main roads and traffic limitation zones are displayed as well.](image)

**Figure 7.** Screenshot of the children elders exposure to PAHs and PM2.5 web map: average children exposure levels to PM2.5 for the non-heating period are displayed; main roads and traffic limitation zones are displayed as well.
Figure 8. Screenshot of the children elders exposure to PAHs and PM2.5 web map: average children exposure levels to PM2.5 for the heating period are displayed; ancillary information about population distribution, schools, hospitals, traffic limitation zones is displayed as well.

Elderly people spend most time at home; children at home and at school. So, resident population density maps (total population, and young and elderly population) allow to locate areas in Rome where the potential impact of air pollution on the two targets is higher. Places where these areas overlap high individual exposure levels, are high impact areas, i.e. areas of Rome at higher risk, where more cases concerning acute or long term effects could be expected. Similarly, schools are places where potential impact on children is high (school was ranked 2 among places where children spend most time). Hospitals concentrate sensible targets. So, both schools and hospitals are high risk spots, when located in cells where mean exposure levels are high (Figure 8).

On the contrary, as long as risk for human health is considered, wild natural areas (forests, woods, pastures, sea, lakes) are very low potential impact areas, since people is not expected to spend time there. In these areas the individual exposure levels remain “potential” and do not result in any risk for human health. These land cover types have been used to mask exposure maps. The map also includes urban parks, open air leisure and sport places: at these places actual exposure is outdoor exposure of people spending
free time there, and non-heating period exposure can be more important than heating period exposure. Traffic limitation zones have been also included in the map to allow users to qualitatively examine relationships of air pollution concentration and consequent exposure levels with present traffic policies.

Total PAHs Toxicity Equivalent (outdoor) Concentration data have been included in this map, and can be interactively selected and loaded by the user. Indeed, this parameter can more directly express the environmental determinant of PAHs risk to human health.

A geocoding tool was added to this map, targeted to people willing to inspect exposure level at sites of interest: when an address is inserted, the map locates the site and a window shows the exposure value (Figure 9). The exposure map can be found at: http://www1.ispesl.it/webgis/exposure_map.html.

Figure 9. Screenshot of the exposure map illustrating results of a geocoding query.

4.2.4 Year 2020 children exposure map

The last map illustrates the 2020 exposure scenarios results. Some features are in common with the previously described exposure map: yearly and seasonal averages are reported in place of daily maps; the same ancillary information layers were added; the same geocoding tool was included (Figure 10).
Main communication goal for this map was identified in the comparison between exposure levels foreseen in 2020 and the current (2011/12) situation. To this aim, only results for children have been reported, since those for elderly people are very similar both in values and in spatial distribution, and do not add information about the item identified as the goal of the map. Comparison is supported by the possibility to interactively select, for a given average period, the 2020 exposure levels (for each of the two emission scenarios investigated, see paragraph 3.2) or the current exposure levels: all maps regarding a pollutant are displayed using the same color scale.

Maps of the % variation with respect to current situation were also included, using double-hue color scales where a 0% variation (no change) was attributed to the threshold (white) color. These maps provide an intuitive look of the expected changes, and of their spatial distribution. Indeed, the map showing the change in exposure levels for the scenario of substitution of biomass combustion for domestic heating with natural gas, clearly displays that the foreseen reduction has a spatial distribution largely coincident with that of residential areas mapped from resident population density (Figure 11). Scenarios map can be found at: [http://www1.ispesl.it/webgis/scenarios_map.html](http://www1.ispesl.it/webgis/scenarios_map.html).
5 Dissemination, feedback

The short time the maps have been on line prevents any evaluation about interest and utility for potential stakeholders. However, combing the geo-spatial point of view with an easy access to time dimension and to quantitative information, revealed to be useful also for the Project teams during the development of the project, for inspection and checking tasks: in fact, GIS based web maps gave a common and shared interface, providing a sight into the data that the single disciplines involved in the study are not able to easily access. This was important for example to identify and explain local artifacts in the simulated fields, through analysis of the geometric relationships with other geographic features, and to evaluate the agreement between model and measurements, thanks to the multi-temporal navigation tools.

A link to the EXPAH Web-GIS page was added to the recent RomAria-Salute website (http://romariasalute.it). This site is a partnership among the Epidemiology Department of
the Lazio region Health Service (DEP Lazio), partner of the EXPAH Project, and the Lazio region Agency for Environmental Protection (ARPA Lazio): its aim is to make available information about Rome air quality and about health effects of air pollution in a simple and attractive way: the overall goal is to reach a wide audience among citizens, promoting a better awareness of risks and virtuous behaviors for pollution reduction and health protection. A simplified version of the total PAHs concentration map, containing simulation modeling results expressed as yearly average, was prepared and directly linked to the RomAria web site to be part of the set of available air quality maps (http://romariasalute.it/?page_id=1323).

Articles describing the EXPAH Web GIS have been submitted to the scientific Committee of national and international Conferences:

- Bogliolo M.P., Contino G., “A geospatial time-aware web interface to deliver information about air pollution and exposure in a big city and its surroundings”. ISPRS/IGU Joint International Conference on geospatial theory, processing, modeling and applications, 6-8 October 2014, Toronto, Canada. Accepted.
- Bogliolo M.P., Contino G., “Un sistema di web mapping multitemporale per divulgare l’informazione su inquinamento atmosferico e rischio per la salute: un prototipo sulla città di Roma”. 18° Conferenza ASITA, 14 – 16 ottobre 2014, Firenze, Italy. Accepted.

6 Conclusions

Finding effective ways to communicate risk for human health, and environmental determinants for it, is a topical and challenging task (WHO, 2013). In this field, it is especially important that research results find a fast and simple path from the scientific community to stakeholders: these latter include both authorities in charge of environmental quality policies and regulations, and people exposed to risk agents.

Maps are a very efficient means to transfer information, providing an intuitive way to interpret complex, multi-thematic data; the possibility to interact with maps catches the attention and invites one to go deeper into information.

The web mapping system developed for the EXPAH project is a prototype of a possible model to disseminate scientific results in this field, providing a sight into impacts of air pollution on people living and working in a big city, and shipping information about the overall exposure, its spatial pattern and levels at specific locations.

The effort was focused on developing a very simple interface to access information in three dimensions (2D spatial and temporal), at two levels of complexity: maps convey
scientific, quantitative content to scientists, regulation authorities and decision makers and provide an intuitive communication medium for other stakeholders, mainly to the public.

The complex reading level revealed its advantages during the development of the project, for data checking and evaluation tasks: each team could analyze the quantitative, temporal and geometric relationships of its own results with those of the others, and with the actual geographic context, highlighting features worthy of further analysis, or artifacts to be fixed.

A feedback from stakeholders requires longer time. Having population among the targets of the maps required paying close attention on possible alarm perception issues: just choosing a different color scale in a pollutant concentration map, drastically changes the message delivered. Some of these aspects were faced and taken into account, e.g. choosing threshold levels coinciding with regulation policy levels when available.

7 References


