"This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement N°603946

HEALS

Health and Environment-wide Associations based on Large population Surveys

FP7-ENV-2013- 603946

http://www.heals-eu.eu/

D11.1 Report on the development of a probabilistic exposure modelling framework to assess external exposure to chemicals for selected population groups

WP11: Integration of time- and spatially resolved data
Data and model synthesis

Version number 1.5

Lead beneficiary: 4 - USTUTT

Date: 25. April 2016

Nature: R - Report

Dissemination level: PU - Public

Document Information

Grant Agreement Number	ENV-603946	**Acronym**	HEALS
Full title	Health and Environment-wide Associations based on Large population Surveys		
Project URL	http://www.heals-eu.eu/		
EU Project Officer	Tuomo Karjalainen,- Tuomo.KARJALAINEN@ec.europa.eu		

Deliverable	**Number**	11.1	**Title**	Report on the development of a probabilistic exposure modelling framework to assess external exposure to chemicals for selected population groups
Work Package	**Number**	11	**Title**	Integration of time- and spatially resolved data: Data and model synthesis

Delivery date	**Contractual**	M30	**Actual**	M31
Status	Draft ☐		Final ☒	
Nature	Demonstrator ☐	Report ☒	Prototype ☐	Other ☐
Dissemination level	Confidential ☐	Public ☒		

Author (Partners)	Christian Schieberle (USTUTT), Naixin Li (USTUTT), Rainer Friedrich (USTUTT)			
Responsible Author	Christian Schieberle		**Email**	cs@ier.uni-stuttgart.de
	Partner	USTUTT	**Phone**	+49 711 685 87847

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CAAA	Clean Air Act Amendments
CLC	CORINE Land Cover
CLRTAP	Convention on Long-range Transboundary Air Pollution
DEMOCOPHES	Demonstration of a study to coordinate and perform human biomonitoring on a European scale
DWH	Data Warehouse
EC	European Commission
ECMWF	The European Centre for Medium-Range Weather Forecasts
EDMS	Environmental Data Management System
EEA	European Environment Agency
EFSA	European Food Safety Authority
EIONET	European Environment Information and Observation Network
Elfe	Etude Longitudinale Française depuis l'Enfance
EMF	Electromagnetic fields
ETC/ACC	European Topic Centre on Air and Climate Change
EURO-SILC	EU Statistics on Income and Living Conditions
EUROSTAT	Statistical office of the European Union
GLCF/MODIS	Global Land Cover Facility/MODerate-resolution Imaging Spectroradiometer
GTOPO30	Global 30 Arc-Second Elevation
HAF	Harmonised aggregate files
HEF	Harmonised episode file
HSF	Harmonised simple file
JRC	Joint Research Centre
INMA	Infancia y Medio Ambiente
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupation
LAMA	tool for Air exposure Modelling and Assessment
LAT	Latitude
LON	Longitude
MPE	Mean prediction error
MSE	Mean square error

MTUS	Multinational Time Use Study
NUTS	Nomenclature of Units for Territorial Statistics
ORNL DAAC	Oak Ridge National Laboratory Distributed Active Archive Center
PNC	Particle number concentrations
POPD	Population density
RELH	Relative humidity
RMSE	Root mean square error
RRMSE	Relative root mean square error
SEDAC	Socioeconomic Data and Applications Center
SFCR	Roughness of the ground surface
SINPHONIE	Schools Indoor Pollution and Health: Observatory Network in Europe
TEMP	Temperature
UNECE	United Nations Economic Commission for Europe
U.S.EPA	United States Environmental Protection Agency
UTM	Universal Transverse Mercator
WHO	Word Health Organization

1 Introduction

To be able to assess different chemical and physical stressors (selected based on their relevance for the population studies in Stream 5), a probabilistic exposure modelling framework is developed within HEALS. It will be used to derive the distribution of external exposure for subgroups of the population. The progress of the development of this framework is reported in this document.

1.1 Aims and purpose of the modelling framework

We aim to retrospectively assess the external exposures of differently vulnerable population groups to multiple stressors via different exposure routes and pathways. We highlight the importance to also integrate into the framework the uncertainty associated with the measurement and modelling of the various agent-specific potential doses and to integrate across time and space.

Estimates from the individual exposure assessment in the scope of WP9 form the basis and will be combined with population-specific features using data fusion approaches to account for socioeconomic differences within the population groups being analyzed. The difference vary in the sources and range from differing time-activity patterns, describing how people spend their time, over the products used while performing a certain activity, to the specifics of the building conditions of people's location (a so-called micro-environment).

However, the quality of the estimates of the external exposure distributions largely depends on the quality of the underlying data. It is important to be able to estimate the quality and to quantify the uncertainty of the datasets incorporated in the probabilistic framework as all uncertainty will eventually influence the final assessment even though the contributions might differ across exposure variables.

We combine a set of modelling tools into a consistent framework for assessing external exposure of population subgroups, thereby taking into account the temporal and spatial substance group-specific characteristics regarding environmental fate and exposure behavior as well as population subgroup-specific characteristics, such as time-activity patterns, relevant microenvironments (i.e. home, work) and food consumption patterns. The modelling framework partly builds upon existing approaches and models. Additional approaches for modelling group exposure based on time-activity information and activity will be generalized for environmental exposures within the scope of HEALS.

We will utilize models for indoor and aggregate exposure, as well as multi-media models for co-exposure to mixtures of chemicals in the environment that will be further developed and improved to finally estimate the external exposure to environmental stressors.

Data gaps will be covered through data fusion techniques (e.g. using Kalman filters) aiming at maximize the information available by merging available data in a prudent way.

The framework aims to quantify annual exposures via different exposure routes and pathways for the whole lifetime or for critical windows of life for pre-defined population subgroups. The latter will be defined on the basis of a number of indicators such as age, sex, area of residence/work/study, socioeconomic status, and behavioral patterns. The final result of a pathway is always the exposure to one or more health stressors for specific population subgroups or of the whole population in the different countries involved in the population studies carried out in Stream 5.

1.2 Scope and structure of this document

The scope of this document is to inform about the progress of the development of the SES-informed probabilistic modelling framework that can be used to assess external exposure to chemicals. Hence, the focus

of this report is to describe the methodological concept behind the data and model synthesis. A report on the actual application of the framework to population studies in Stream 5 will be delivered one year later in M42.

The structure of the document is as follows. In section 2, we will describe the methodology framework for the study. Section 3 introduces the data and models used in terms of socioeconomics and time-use, along with the data derived from the HEALS Environmental Data Management System (EDMS) developed in WP8 and sensor data collected from WP9. The methodology applied to combine the data sets will be explained in section 4, which forms the basis of the probabilistic framework. We will show preliminary figures that are a result of applying the prototypical framework in section 5. Section 6 concludes the document and gives an outlook of current and future development towards the application of the framework to the population studies.

2 Description of methodology

As the starting point, HEALS WP11 partners scoped and determined the pathways and stressors to be studied in the course of the project (see Table 1). The pathways to be covered include inhalation, ingestion and dermal uptake. The main stressors to be considered are air pollutants, heavy metal in food and water, phthalates, electromagnetic fields (EMF), green space and noise.

Table 1: Stressors and pathways for which the external exposure should be determined in HEALS. In this deliverable, we focus on estimation the exposure via inhalation and ingestion.

Inhalation		**Ingestion**		**Dermal uptake**	**Other**
Outdoor	Indoor	Food	Water	Phthalates, electromagnetic fields	Green and blue space, noise
PM10, PM2.5, NO_2, ozone	PM10, PM2.5, NO_2, radon, mould,	Arsenic, cadmium, Chromium, Lead	Chromium, cadmium, arsenic		

The step afterwards is to incorporate the relevant data for the pathways and stressors to be studied. We utilized the data mined in EDMS from WP10, which covers air pollutant (section 3.2.1 and 3.2.1.4), green and blue spaces (section 3.2.3), food ingestion (section 3.2.4), phthalates (section 0), EMF (section 3.2.6) and noise (section 3.2.7). The data collected from sensors in WP9 are adopted as well, including indoor PM2.5 concentration measured by Dylos and noise data collected by Netatmo. Besides, we integrate a series of modelling tools to generate external exposure data for different chemical and physical stressors (e.g. interpolation method for ambient air pollutant concentration).

According to the description of work, we should take the temporal and spatial substance group-specific and population subgroup-specific characteristics into account when assessing the external exposure to various stressors. These characteristics include socioeconomic variables, time-activity pattern and relevant microenvironments (i.e. home, work), which can be covered by dataset EU-SILC and MTUS. EU-SILC provides annual pan-European detailed microdata on income, poverty, social exclusion and living conditions since 2004 (EUROSTAT 2016, see section 3.1.1). MTUS records the time-activity pattern of the diarists as well as socioeconomic variables and microenvironments for 10 European countries since 1985 (Fisher and Gershuny 2015, see section 3.1.2). A methodology is developed in this deliverable to fuse the two datasets based on common existing variables (e.g. age, income, education level) to generate a comprehensive dataset that inherits the information from both datasets (see section 4.1).

The fused data will be further combined with exposure data (stressor concentration and model parameters data) based on population subgroup-specific characteristics (spatial region, degree of urbanization and SES variable). Since we are not able to know the where the respondents exactly live according to the regulation, only the distribution of stressor concentration and model parameters can be determined. The data after combination will be used as the input data to estimate the probability distributions of the exposure to different stressors for a certain socioeconomic subgroup (see section 5).

Another important issue in WP11 is to develop the method to identify and estimate the uncertainty of the probabilistic exposure assessment. The total uncertainty originates from errors and uncertainties in the input data, the estimation approaches and the methods used to link information throughout the chain. Uncertainty level from each part is explained in this deliverable (see section 3.2.1.2, 3.2.1.3, 4.3 and 5) and a model to assess the integrated uncertainty will be developed within the scope of the project.

The methodological framework will be applied to population studies. In the future work, concentration data of chemicals in air, water, soil and food from WP8, as well as time-activity pattern (sensor-based data

and time-location patterns from agent-based modelling) will be used as the input data (e.g. using sensor data for PM concentration to determine indoor emission source strength, see section 3.2.2.5). In the mean-time, the modeling framework will be refined and applied in the EXHES pilot study to serve as a case study.

3 Data sets and models used

3.1 Socioeconomic data and time-activity patterns

3.1.1 EU Statistics on Income and Living Conditions (EU-SILC)

The European Union Statistics on Income and Living Conditions (EU-SILC) is a reference source for cross-sectional and longitudinal multidimensional microdata concerning income, poverty, social exclusion and living conditions (EUROSTAT 2016). Information covering social exclusion and housing condition (including housing, material deprivation and household income) is collected at household level. Furthermore, education, health and labor information is collected at personal level. The data available are either cross-sectional, which provide comparable and high quality data over a certain time period, or longitudinal, which focus on the changes over time at individual level with limited information on income (eds Atkinson & Marlier 2010). The years and countries covered by EU-SILC are shown in Table 2.

For each year the dataset consists of 4 parts: Household Register (D-file), Personal Register (R-file), Household data (H-file) and Personal Data (P-file). Both register files record only basic variables but more detailed information can be found in the H-file and P-file. H-file and P-file can be combined with register files based on the household and personal identifiers (Household ID and Personal ID).

3.1.2 Multinational Time Use Study (MTUS)

The concept of Multinational Time Use Study (MTUS) was initially raised by Professor Jonathan Gershuny in 1970s to create a multi-nationally harmonised set of time use surveys (Fisher and Gershuny 2015). The MTUS archive located at the Centre for Time Use Research in the Department of Sociology at the University of Oxford includes the following three formats and data coverage is shown in Table 3:

- Harmonised simple file (HSF):

  In each HSF, every row represents the time a diarist spent on 25 activity categories within 24 hours` observation time (diary) (Fisher and Gershuny 2016). Additionally, it records a limited number of demographic and socioeconomic variables.

- Harmonised aggregate files (HAF):

  HAF is similar to HSF, but the activities are classified into 69 categories. Besides, more detailed demographic and socioeconomic information can be achieved.

- Harmonised episode file (HEF):

  Every row of HEF represents a change in the elements of diary report, which include diarist`s identifier, socioeconomic variables, activity categories, location and mode of transport (Fisher and Gershuny 2012).

3.2 Environmental from EDMS

The exposure pathway is defined as the way a person gets into contact with a certain substance (U.S.EPA 2016). In HEALS, we are aiming at assessing the external exposure to stressors via different pathways (Table 1). In this deliverable, we focus on stressors that people are exposed to through inhalation and ingestion. Through inhalation, people might get exposed to air pollutants (PM10, PM2.5 and NO_2) and experience health problem; while via ingestion, people are likely to be exposed to heavy metal contaminants. We also compile the datasets for other stressors (e.g. green and blue space) in this chapter. Exposure to other stressors will be estimated during the course of the project.

Table 2: Years covered by EU-SILC microdata at the time of this study. For most countries data are present for the years 2005 through 2013. However, less data are available for Bulgaria, Switzerland, Croatia and Malta.

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Austria	√	√	√	√	√	√	√	√	√	√
Belgium	√	√	√	√	√	√	√	√	√	√
Bulgaria				√	√	√	√	√	√	√
Switzerland					√	√	√	√	√	√
Cyprus		√	√	√	√	√	√	√	√	√
Czech republic		√	√	√	√	√	√	√	√	√
Germany		√	√	√	√	√	√	√	√	√
Denmark	√	√	√	√	√	√	√	√	√	√
Estonia	√	√	√	√	√	√	√	√	√	√
Spain	√	√	√	√	√	√	√	√	√	√
Finland	√	√	√	√	√	√	√	√	√	√
France	√	√	√	√	√	√	√	√	√	√
Greece	√	√	√	√	√	√	√	√	√	√
Croatia							√	√	√	√
Hungary		√	√	√	√	√	√	√	√	√
Ireland	√	√	√	√	√	√	√	√	√	√
Iceland	√	√	√	√	√	√	√	√	√	√
Italy	√	√	√	√	√	√	√	√	√	√
Lithuania		√	√	√	√	√	√	√	√	√
Luxembourg	√	√	√	√	√	√	√	√	√	√
Latvia		√	√	√	√	√	√	√	√	√
Malta					√	√	√	√	√	√
Netherlands		√	√	√	√	√	√	√	√	√
Norway	√	√	√	√	√	√	√	√	√	√
Poland		√	√	√	√	√	√	√	√	√
Portugal	√	√	√	√	√	√	√	√	√	√
Romania				√	√	√	√	√	√	√
Sweden	√	√	√	√	√	√	√	√	√	√
Slovenia		√	√	√	√	√	√	√	√	√
Slovakia		√	√	√	√	√	√	√	√	√
United Kingdom		√	√	√	√	√	√	√	√	√

Table 3: Data availability of Multinational Time Use Study (MTUS) during several periods from 1985 to to-day. For some years there is only a harmonised simple file (HSF), whereas for others there are also harmonised aggregate files (HAF) and harmonised episode files (HEF) [- = no data available; ○ = HSF only; ● = HSF, HAF and HEF].

	1985-1989	1990-1994	1995 - 1999	2000 - 2004	2005+
Austria	-	●	-	-	-
Denmark	○	-	-	○	-
France	-	-	●	-	-
Germany	-	●	-	○	-
Italy	●	-	-	○	-
Netherlands	●	●	●	●	●
Norway	-	○	-	○	-
Slovenia/ former Yugoslavia	-	-	-	○	-
Spain	-	●	●	●	●
United Kingdom	●	-	●	●	●

3.2.1 Inhalation of ambient air

3.2.1.1 Interpolated air quality maps provided by the EEA

The concentration maps of air pollutants are the important input data for exposure modelling. To interpolate the monitoring data directly, or to build a linear regression model between measurement data and other supplementary information, are the two methods that are commonly used (Horálek et al. 2007). Horálek et al. (2005) took advantage of these two methods and developed the methodology of linear regression models plus interpolation of their residuals. The method was separately applied for rural and (sub-)urban background stations data from AirBase (EEA 2014a). The final interpolation map will be merged from the rural and urban maps based on population density and resampled to 10 by 10 km^2 EEA grid.

Table 4 lists the interpolated maps directly available for PM10, PM2.5 and NO_2 from EIONET (2016). Those maps were generated with the interpolation method developed by Horálek et al. (2005). In this deliverable, the interpolation methodology was applied to generate the PM10 maps from 2000 to 2003 to fill the data gap. For years that interpolated PM2.5 maps were not available, a methodology to derive them from interpolated PM10 maps was applied (see section 3.2.1.3). For NO_2, the interpolation method can be applied to generate maps for the years 2000 to 2012.

Table 4: The data availability of Interpolated air quality maps for PM10, PM2.5 and NO_2 from EIONET [√ =data available from EIONET; - = no data available].

	PM10	PM2.5	NO_2
2000	-	-	-
2001	-	-	-
2002	-	-	-

2003	-	-	-
2004	√	-	-
2005	√	√	-
2006	√	-	-
2007	√	√	√
2008	√	√	-
2009	-	-	-
2010	√	√	-
2011	√	√	-
2012	√	√	-

3.2.1.2 EEA interpolation method applied to PM10 and NO_2

For the years that concentration maps are not available, we followed the interpolation method proposed by Horálek et al. (2005) to generate the PM10 concentration maps. AirBase measurement air quality data, EMEP model data, altitude, metrological parameters and population density were used as the input data to build the linear regression model (see Table 5). The residuals from the linear regression were interpolated in ArcGIS with tool Geostatistical Analyst. Cross-validation was conducted and the results were described by root mean square error (RMSE) and mean prediction error (MPE). This can give indication on the overall uncertainty of the prediction of the model. The figures are given in Table 6 and the interpolated PM10 concentration map of 2003 was shown in Figure 1 as an example.

The interpolated maps have a resolution of 10 by 10 km^2, which decreases the accuracy in exposure assessment. A model to handle the urban increment and street increment of pollutant concentration will be developed to increase the quality of the concentration map at the next stage. It is also noticed that there is a temporal variance of pollutant concentration. Seasonal, even monthly concentration maps will be generated over the course of the project.

3.2.1.3 Methodology to derive ambient PM2.5 levels from PM10 maps

Proper ground-based measurement data for PM2.5 from EEA's AirBase are only available for very recent years and are scarce for years before 2007. They do not qualify for a Kriging approach (Gaussian process regression) as applied for pollutants where measurement data has been collected for longer time periods and a more dense station network is available (e.g. NO_2 and to some extent PM10). Hence, we developed a non-linear prediction model that fuses several data sources and end-up at a consistent estimate of outdoor PM2.5 levels on a 10 by 10 km^2 grid.

Table 5: Input data, resolution and data sources for linear regression model. The data were resampled in ArcGIS to 1 by 1 km^2 grids for further processing.

Input data	Variable	Resolution	Source
Measurement air quality data	Annual average [μg.m^{-3}]	-	EEA AirBase v8 (EEA 2014a)

EMEP model data	Annual average [µg.m^{-3}]	50 x 50 km	EMEP MSC-W (CLRTAP n.d.)
Altitude	Altitude [m]	30 x 30 arcsec	GTOPO30 (ORNL DAAC n.d.)
Meteorological parameters	10 metre wind speed [m.s^{-1}] Solar radiation [MWs.m^{-2}]	0.125° x 0.125° 0.125° x 0.125°	ERA-Interim (ECMWF 2015a) ERA 40 (ECMWF 2015b)
Population density	Population density [inhbs.km^{-2}]	100m x 100m 2.5'x 2.5'	EEA (EEA 2000.) SEDAC (SEDAC 2000)

Table 6: Parameters of linear regression model and the residual kriging variograms (nugget, sill, range) and their statistics (RMSE and MPE) of PM10 annual average for 2003.

Linear regression model plus ordinary kriging of its residuals	Rural area	Urban area
	Coefficient	**Coefficient**
C (constant)	1.0632322	1.995
A1 (ln. EMEP model 2003)	0.9706567	0.5233
A2 (altitude GTOPO)	-0.0002845	0.0004088
A3 (wind speed 2003)	-0.1051993	-0.04195
A4 (solar radiation 2003)	-0.0047149	0.01123
A5 (ln. population density)	0.0124998	-0.01499
Adjusted R^2	0.4205	0.1938
Residaul standard error [µg/m^3]	0.2689	0.2985
Nugget	0.0400	0.0566
Partial sill	0.1213	0.0385
Range	27.267	21.535
RMSE [µg/m^3]	1.077555	1.249963
Bias/MPE [µg/m^3]	1.035442	1.044147

The data sources used are listed in Table 7 and the rationale for using these data sets is as follows: PM2.5 is part of PM10 and we therefore use existing measurement-based interpolated PM10 maps to improve our estimates. As there are regional differences and geographical specifics of PM composition, we use meteorological data (2m temperature, 10m wind speed, 10m U wind component, 10m V wind component, surface roughness, air pressure, cloud cover, relative humidity, and boundary layer height) and land cover data to improve the estimate. Generally, PM2.5 levels are higher in urbanized areas (cf. Torras Ortiz & Friedrich (2013)). Thus we include the JRC's population grid into the estimator to represent urbanized areas by their higher population density.

We use ground-based measurements of PM2.5 to validate our model. Note, that for prediction we left out all measurements of the year under investigation. The results are given in Figure 3 for recent years (2010-2012) and in Figure 4 for years where less measurement data are available (2005, 2003 and 2001). For each year, the figure on the right shows correlation to the limited number of PM2.5 measurement stations that actually were available. The contribution of single predictors in the model is given in the left figure respec-

tively. It is measured as the increase of mean square error (MSE) that would exist if the respective factor was not part of the model. So it becomes evident that, apart from the PM10 level, we can estimate PM2.5 levels when having information on population density (POPD), geographical location (LON and LAT), and meteorological conditions (WIND, RELH, and TEMP) along with the roughness of the ground surface (SFCR). The model performs very well for all years with Pearson's R ranging from 0.86 to 0.96.

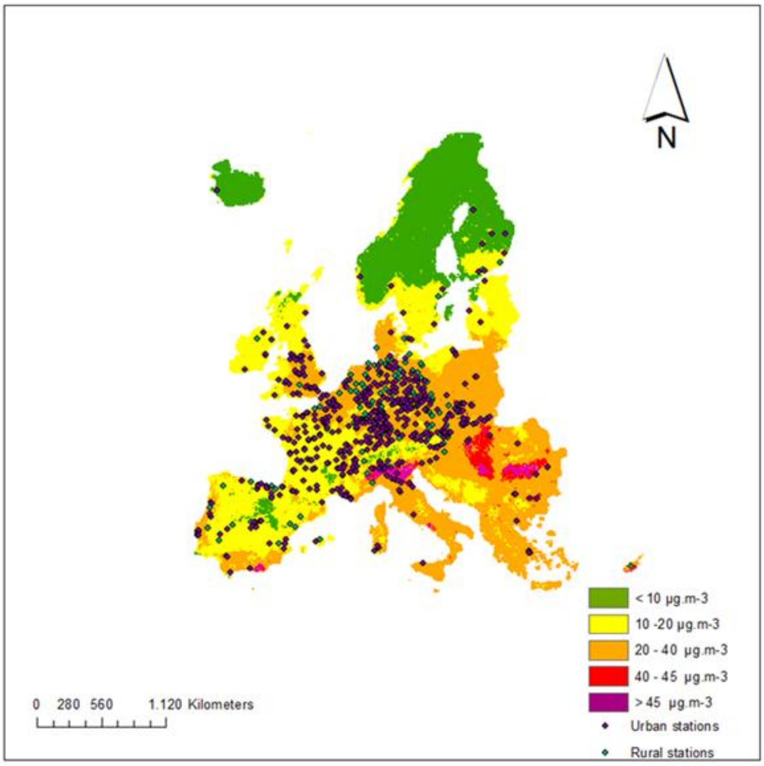


Figure 1: Combined rural and urban concentration map of PM10 annual average of the year 2003 with 10 by 10 km^2 grid resolution after applying the method of Horálek et al. (2005).

3.2.1.4 Earth Observation

Remote sensing represents a relatively new and promising approach for adding spatial information for exposure estimation, particularly in areas far from ground monitor stations providing a synoptic picture of air quality in an urban airshed, including information on sources for isolated events.

To date hundreds of organizations use remote sensing data that are provided by a short list of satellite operators, such as the U.S. National Aeronautics and Space Administration (NASA), the European Space Agency (ESA), the National Space Development Agency of Japan (NASDA), and the Indian Space Research Organization (ISRO). Despite the big potential, only in the very latest years concerted research effort were made to use EO data for air quality studies at local/regional scale.

Although that, to date a review of the published literature reveals relatively few applications in the use of satellite sensing data for urban and regional particulate matter assessment. Works in US (Kondragunta, et al., 2008; van Donkelaar, et al., 2010; van Donkelaar, et al., 2006) used data assimilation techniques to estimate surface PM2.5 concentrations from MODIS and MISR sensors as well as from NOAA's Geostationary Operational Environmental Satellite (GOES) showing good predictive capacity. Others studies have investigated the relationship between total-column Aerosol Optical Depth (AOD) and surface PM2.5 and/or PM10

measurements. Most have developed simple empirical relationships between these two variables (Chu et al., 2003; Engel-Cox et al., 2004; Wang and Christopher, 2003); more recent investigations often have used local meteorological information to better relate AOD and PM2.5 (Gupta and Christopher, 2008; Gupta et al. 2006; Koelemeijer, et al., 2006; Liu et al., 2005, 2007a, b c, 2009; Paciorek et al., 2008; Wang et al., 2010). Most of the literature works focus on estimation of PM2.5 rather than PM10 using moderate resolution satellite sensors such as MODIS.

The methodology we plan to apply in HEALS to derive PM_x concentration at the ground level is based on the calculation of the Aerosol Optical Depth (AOD) from satellite data in the visible spectrum (i.e. 550 nm) as a measure of atmospheric turbidity. Visible light is selected because this is where the difference between Mie and Rayleigh scattering is the largest.

AOD represents the integral of the extinction coefficient due to scattering from the ground to the height of the satellite sensor.

$$AOD = \int_0^{z_h} \sigma_e(z)dz$$

where: σ_e (z) is the height dependent extinction coefficient and z_h is satellite sensor height.

For High Resolution Sensors (HRS) such as Landsat and SPOT, the method to estimate AOD combines two approaches that consider physically independent optical effects in the atmosphere [(Sifakis and Deschamps, 1992; Tanré et al., 1988;)]: (a) contrast reduction by scattering efficient airborne particles in short wavelengths (i.e., visible spectrum); and (b) radiation attenuation that particles engender in longer wavelengths (i.e., thermal infrared spectrum).

This method uses two different satellite images: the reference and the polluted image and it evaluate the variations of the reflectivity of the target between the two. The reference image represents the “baseline” or the “zero pollution” image and for this reason it has to be chosen clearer in terms of cloud coverage and less polluted as possible. The “polluted” image is the image on which the AOD is effectively calculated though the evaluation of the variations of reflectivity of targets respect to the reference.

AOD magnitude is linked to the columnar concentration of the optically effective aerosol. In order to calculate the extinction coefficient, σ_e, of the aerosol that lies close to the ground, AOD has to be divided by the appropriate scale length. Under well-mixed conditions such as the ones around noon, previous experimental works (Sheridan and Ogren, 1999; Wiegner et al., 2006) have demonstrated that the assumption of a homogeneous and almost uniform concentration of pollution loading in ca. the lower half of the mixing layer is realistic and it can be considered reasonable. In this case σ_e can be derived from the AOD taking a known z_b according to the following equation:

$$\sigma_e(0) = \frac{2AOD}{Z_b}$$

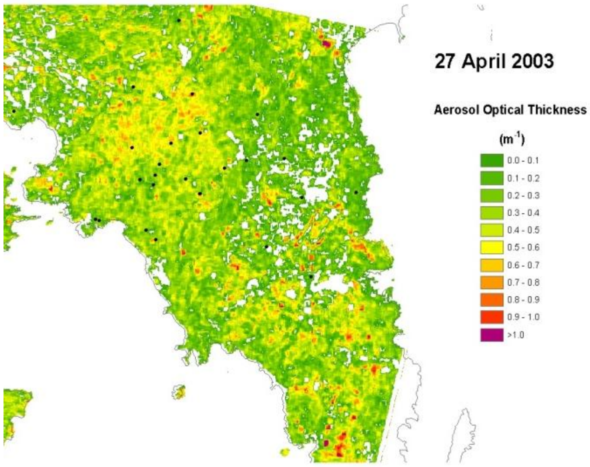


Figure 2: Landsat 5 image (left) and AOD map (right) over Athens (Greece) derived from Landsat 5 sensors

AOD with a high spatial resolution (i.e. resolution of 30 by 30 meters and 10 by 10 meters respectively for Landsat and SPOT) could be retrieved covering a domain as large as 80-100 x 80-100 km^2. AOD represents a *proxy* of PMx concentration and as such it will be then used to estimate PM concentration at the ground level through data fusion techniques described in paragraph 4.4.

3.2.2 Inhalation of air indoors

In the last section, we introduced the methodology to generate the ambient pollutant concentration. However, the totality of exposure to air pollutants consists of the exposure in each microenvironment, which is defined by Gens (2012) as any physical location a receptor spends time in. The exposure to air pollution occurring indoors plays an important role since people spend a large proportion of time indoors. For modelling the PM2.5 levels within indoor environments, we assume the concentration indoors is under steady state and apply the mass balance model from Dockery and Spengler (1981):

$$c_{in} = \frac{c_{out} pa + \frac{\sum_{i=1}^{n} s_i}{v}}{(a+k)}$$

Where

C_{in}	indoor concentration [µg m^{-3}]
C_{out}	ambient pollutant concentration [µg m^{-3}]
p	penetration factor
a	air exchange rate [h^{-1}]
k	decay rate [h^{-1}]
s_i	emission rate for source i [µg h^{-1}]
v	room volume [m³]

In the following section we will give an explanation of the role and contribution of each of the parameters of the model and will present the data sources and uncertainty associated with the individual parameters.

Table 7: Datasets used to develop random forest model. Temporal coverage and spatial resolution refer to the data used to feed into the model. The population density was aggregated to 1 km^2 to cover a more representative area.

Dataset	Variables	Temporal coverage	Spatial resolution
ETC/ACC interpolated air quality maps (ETC/ACC 2008)	PM10	2005	10 x 10 km^2
EEA AirBase v8 (EEA 2014a)	PM2.5, PM10	2000-2012	at measurement station
ECMWF ERA-Interim (ECMWF 2015a)	2m temperature, 10m wind speed, 10m U wind comp., 10m V wind comp., Surface roughness, Air pressure, Cloud cover, Relative humidity, Boundary layer height	2000-2012	1/8° x 1/8°
GLCF/MODIS (Channan et al. 2014)	Land cover type	2001-2012	1/12° x 1/12°
JRC population grid (Gallego 2010)	Population density	2001	aggregated to 1 x 1 km^2

3.2.2.1 Penetration factor

The penetration factor is the fraction of particles in the infiltration air that passes through the building shell (Chen & Zhao 2010). This variable is dependent on geometry, surface materials and pressure drop along the leakage path (Liu & Nazaroff 2001). In this deliverable, the penetration factor is assumed to be invariant over countries. Table 8 shows the value, distribution and source of penetration factor for home and office.

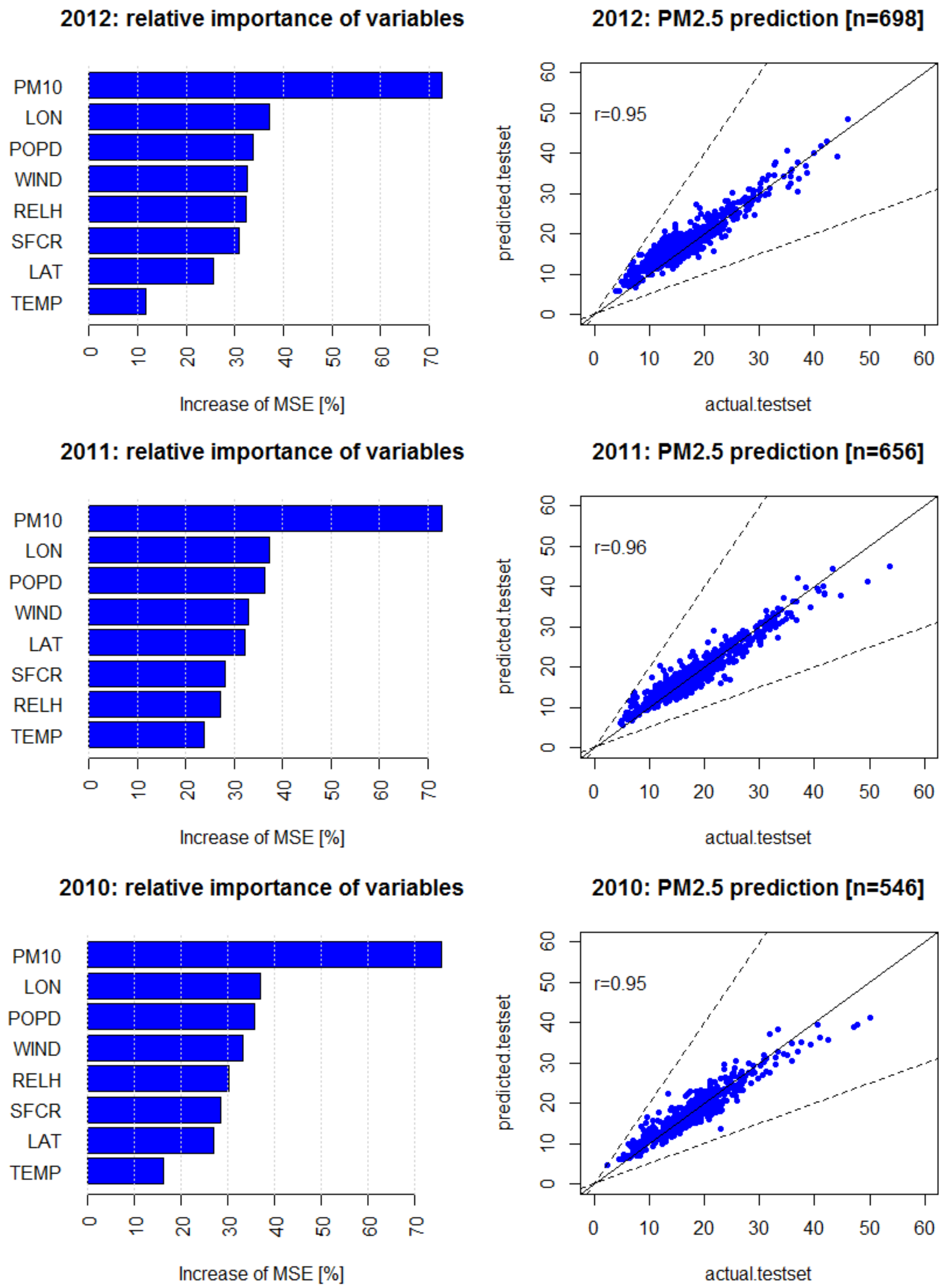


Figure 3: Random forest model of predicted ambient PM2.5 levels in 2012, 2011 and 2010. The right figure shows the correlation with ground-based measurement stations that were not used in the model for prediction, but for validation only. The left figure shows the relative importance (measures as increased mean square error contribution) if one of the listed factors was left out of the model. The variables are: population density (POPD), geographical location (LON and LAT), and meteorological conditions (WIND, RELH, and TEMP) along with the roughness of the ground surface (SFCR).

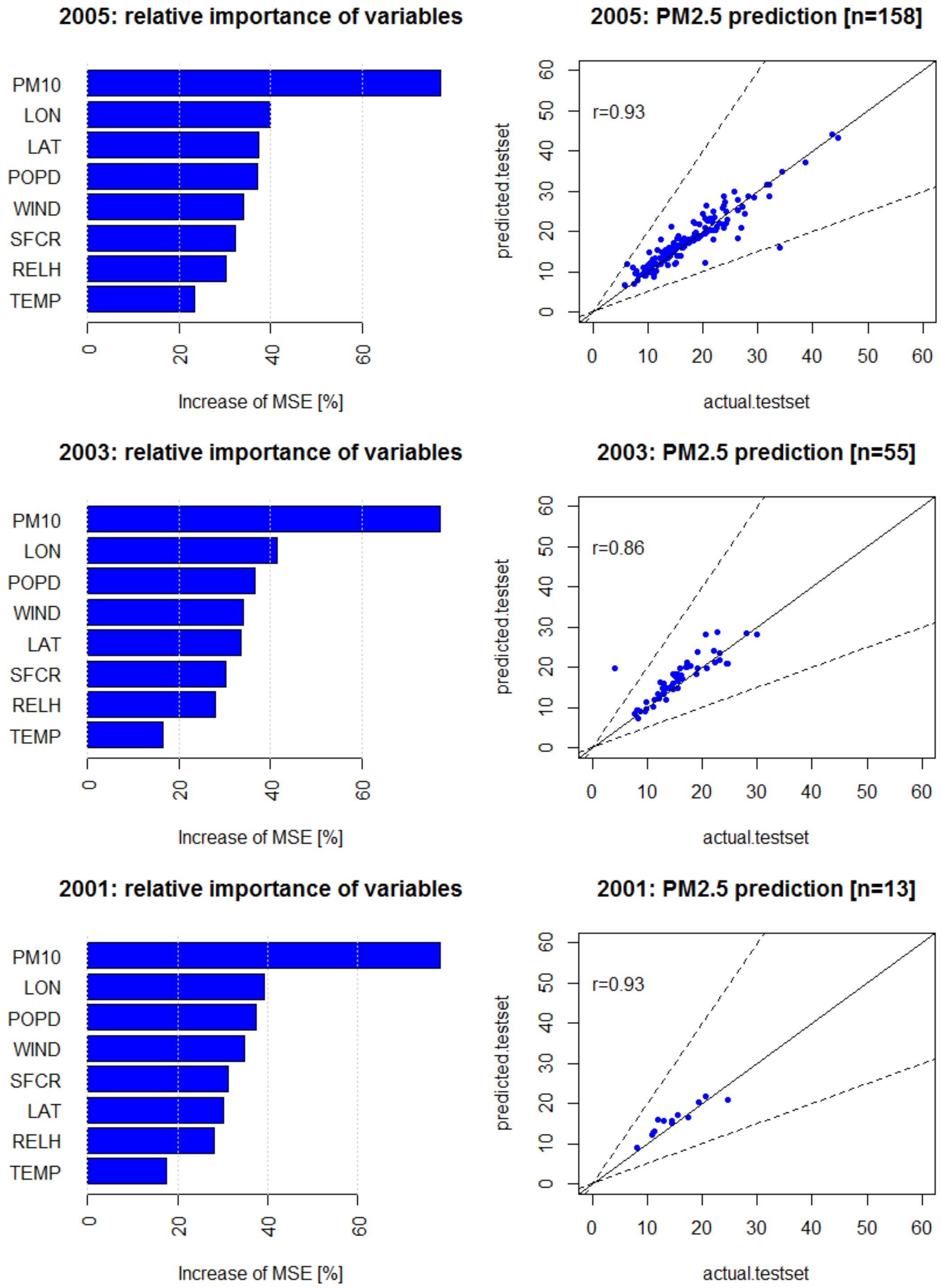


Figure 4: Random forest model of predicted ambient PM2.5 levels in 2005, 2003 and 2001. The right figure shows the correlation with ground-based measurement stations that were not used in the model for prediction, but for validation only. The left figure shows the relative importance (measures as increased mean square error contribution) if one of the listed factors was left out of the model. The variables are: population density (POPD), geographical location (LON and LAT), and meteorological conditions (WIND, RELH, and TEMP) along with the roughness of the ground surface (SFCR).

Table 8: Value, distribution and source of penetration factor for home and office.

Microenvironment	Parameter value	Distribution type	Source
Home	1 ± 0.06	Normal distribution	Özkayna et al. 1996
Office	0.5, 1	Uniform distribution	Gens, 2012

3.2.2.2 Air exchange rate

Air exchange rate defines how often the complete indoor air is exchanged within one hour and is dependent on season, building envelope, ventilation system and pressure conditions (Gens 2012). For residential buildings, Hänninen et al. (2010) made a summary of three studies and revealed a seasonal variation (see Figure 5). For offices and schools, we adopt the data from the projects Officair (2013) and SINPHONIE (2012) (see Table 10).

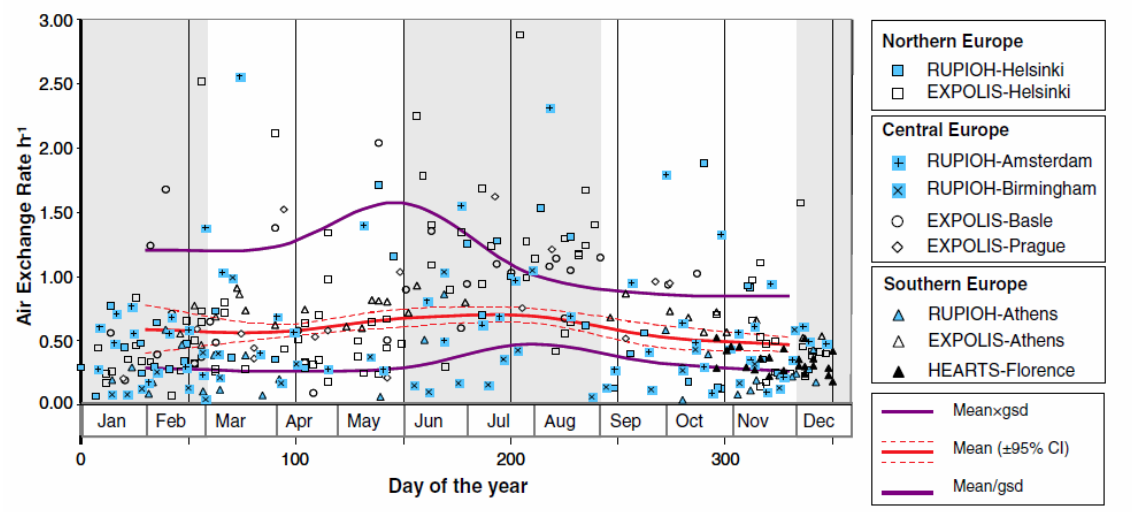


Figure 5: Seasonal variation in air exchange rates in the EXPOLIS and RUPIOH cities and Florence (Hänninen et al., 2010).

3.2.2.3 Decay rate

Decay rate is the loss rate of PM2.5 indoors from processes including Brownian diffusion, gravitational settling and impaction, and the parameter is assumed to be normally distributed with mean 0.39 h^{-1} and standard deviation 0.09 (Gens 2012).

3.2.2.4 Room volume

For residential buildings, we adopt the data from Explolis study (Hänninen et al. 2004) for Greece, Switzerland, Finland and Hungary. For other countries, data of Bulletin of Housing Statistics for Europe and North America 2006 from UNECE (UNECE 2006) are applied. Room volume for office and school are given by project Officair (2013) and SINPHONIE (2012) respectively (see **Table 9**).

Table 9: Value, distribution and source of room volume for home, office and school

Microenvironment	Room volume	Distribution type	Source

Home	Athens: 290 ± 83 [m^3]	Normal	Expolis study,
	Basel: 280 ± 169 [m^3]		(Hänninen et al. 2004)
	Helsinki: 205 ± 83 [m^3]		
	Prague: 233 ± 80 [m^3]		
Office (landscape)	466.8 [44 ; 1680]	Triangular	Officair 2013
Office (cellular)	77.3 [23.4 ; 176.0]	Triangular	Officair 2013
School	Floor area [m^2]:	Uniform	SINPHONIE 2012
	49.3 [43.4 ; 58.2]		
	Ceiling height [m]:		
	3.4 [2.9 ; 3.6]		

3.2.2.5 Indoor background emission source strength

This variable describes the indoor non-ETS PM2.5 sources. Data from the Expolis study for four European cities were adopted (see **Errore. L'origine riferimento non è stata trovata.**). We follow the methodology of the LAMA model (Gens 2012) to assign the data to four geographical regions of Europe (Athens represents Southern Europe, Prague represents Central Europe, Basel represents North-western Europe, and Helsinki represents Northern Europe).

Additionally, the sensor data for indoor air quality (PM2.5) can help us to improve this parameter. In HEALS pilot study (WP9) a low cost sensor for particulate matter, the Dylos, is being used at 150 homes of families across Europe (UK, Greece and the Netherlands) to measure the indoor PM concentration (see section 3.2.9). Simultaneously, their activities, including cooking, vacuum cleaning, dusting, lighting a candle/incense were logged manually by the participants. With the sensor data, it is possible to quantify the impacts of activities on pollutant concentration and determine the indoor emission source strength. The data analysis in WP9 is ongoing and the parameter will be modified when the data is available.

Table 10: Value, distribution and source of air exchange rate for home, office and school.

Microenvironment	**Parameter value**		**Source**
	mean (95% CI)	SD	
Home	Spring: 0.81 (0.60-1.01)	Spring: 0.73	Expolis study,
	Summer: 1.38 (1.03-1.72)	Summer: 1.11	(Hänninen et al. 2010)
	Autumn: 0.77 (0.59-0.96)	Autumn: 0.47	
	Winter: 0.56 (0.41-0.72)	Winter: 0.48	
	Spring: 0.56 (0.35-0.78)	Spring: 0.59	Rupioh study,
	Summer: 1.07 (0.64-1.50)	Summer: 1.14	(Hänninen et al. 2010)
	Autumn: 0.41 (0.28-0.54)	Autumn: 0.43	
	Winter: 0.36 (0.30-0.42)	Winter: 0.21	
	Autumn: 0.36 (0.31-0.42)	Autumn: 0.1	Hearts study,
	Winter: 0.32 (0.26-0.38)	Winter: 0.1	(Hänninen et al. 2010)

Office	(0.38±0.06) Lognormal distribution	Gens 2012
	Naturally ventilated offices:	Officair 2013
	mean: 0.4, range: 0.1-1.8	
	Mechanically ventilated offices:	
	mean: 2.3, range: 0.5-5	
School	Mean: 0.4, range: 0.1-3.7	SINPHONIE 2012

Table 11: Parameters and data source for source strength used in the PM2.5 mass balance model. Only parameters for home microenvironments are available. All values follow a normal distribution.

Microenvironment	Source strength [μg/m^3]	Source
Home	Athens: 4.0 ± 2.8	Expolis study, (Hänninen et al. 2004)
	Basel: 5.3 ±4.9	
	Helsinki: 2.9 ± 3.1	
	Prague: 5.0 ± 8.2	

3.2.2.6 Traffic factors

To determine the concentration in microenvironment “in transit”, we adopt the microenvironment (ME) factors developed by ICF Consulting (ICF) and TRJ Environmental, Inc. (TRJ). The methodology defined three following ME factors (ICF and TRJ 2000):

- Penetration factor (PEN): ratio of the microenvironmental concentration to the outdoor concentration in the immediate vicinity of the microenvironment;
- Proximity factor (PROX: the ratio of the outdoor concentration in the immediate vicinity of the microenvironment (or outdoor microenvironmental concentration) to the measure or modelled ambient concentration;
- Multiplicative factor (MULT): the product of PEN and PROX.

The values of the parameters PROX and MULT are grouped by source category as “Onroad vehicle” and “Major, area, and nonroad vehicle”. The definition of “Major, area, and nonroad vehicle” source was given in 1990 Clean Air Act Amendments (CAAA) as follows:

- Major source: a stationary source or group of stationary sources that emit or have the potential to emit 10 tons per year or more of a hazardous air pollutant or 25 tons per year or more of a combination of hazardous air pollutants (U.S.EPA 2015);
- Nonroad vehicles: includes boats, farm equipment, bulldozers, lawn and garden devices and construction machinery (U.S.EPA 1994);
- Area source: any stationary source that is not a major source (U.S.EPA 2015).

Table 12: Diesel PM microenvironmental factors PEN, PROX and MULT by microenvironments (ICF and TRJ 2000).

Microenvironment	PROX		PEN	MULT	
	Onroad	Major, area, and		Onroad	Major, area, and nonroad

	vehicle	nonroad vehicle		vehicle	vehicle
Car - In vehicle	1.76	1.0	0.63	1.1088	0.63
Bus - In vehicle	1.76	1.0	0.63	1.1088	0.63
Truck - In vehicle	1.76	1.0	0.63	1.1088	0.63
Other - In vehicle	1.76	1.0	0.63	1.1088	0.63
Train/subway - In vehicle	1.76	1.0	0.63	1.1088	0.63
Public garage-Indoors	1.0	1.0	0.75	0.75	0.75
Parking lot/garage - Otdoors	1.45	1.0	1.0	1.45	1
Near road – Outdoors	1.45	1.0	1.0	1.45	1
Motorcycle – Out-doors	1.45	1.0	1.0	1.45	1
Airplane - In vehicle	0.0	0.0	0.90	0	0

Additionally, Gerharz et al. (2011) defines traffic factor as the ratio between the concentrations in microenvironment "in transit" and the ambient background concentration. The parameter is assumed to be normally distributed with mean value of 2.48 and standard deviation 2.13 (Gens 2012).

3.2.3 Green space and blue spaces

According to the study of EEA (2012), green space is essential to our survival and has positive influences on the health of human beings in the following ways:

- Increased physical activity and reduced obesity
- Reduced stress levels and improvements in mental health
- Reductions in noise levels, which can improve mental and physical health
- Improvements in hospital recovery times
- Lower levels of violence and crime and increased social interactions which can also help improve overall well-being.

In HEALS, blue spaces are assumed to have the same positive influences on health of human beings as green spaces. The green and blue spaces layers utilised are from CORINE Land Cover (CLC) and Urban Atlas, which were generated from former EC projects. The CORINE Land Cover map covers the whole area of the European Union, but the resolution is low (25 ha). The resolution of the Urban Atlas land cover map is higher (0.25 ha), but it is available only for large cities in Europe. Figure 6 shows an example of the Urban Atlas land cover of Thessaloniki and Figure 7 shows the green and blue spaces in Europe extracted from the CORINE land cover map.

Table 13 shows the basic information of these CORINE and Urban Atlas land cover maps. The former consists of 44 land cover classes (EIONET 2012). The classes relevant to HEALS are the following:

- 1.4.1 Green urban areas

- 1.4.2 Sport and leisure facilities
- 3.1 Forests
- 3.2 Shrub and/or herbaceous vegetation associations
- 4.1 Inland wetlands
- 5 Water bodies

The 2006 version of Urban Atlas displays 21 classes and the 2012 version displays 27 classes (EEA 2014b). The codes and classes relevant to HEALS are (vector data code in parenthesis):

- 1.4.1 (14100) Green urban areas
- 1.4.2 (14200) Sport and leisure facilities
- 2 (20000) Agricultural areas, semi-natural areas and wetlands
- 3 (30000) Forests
- 5 (50000) Water

Thessaloniki
Forest and Urban green spaces from Urban atlas
(EEA, GMES)

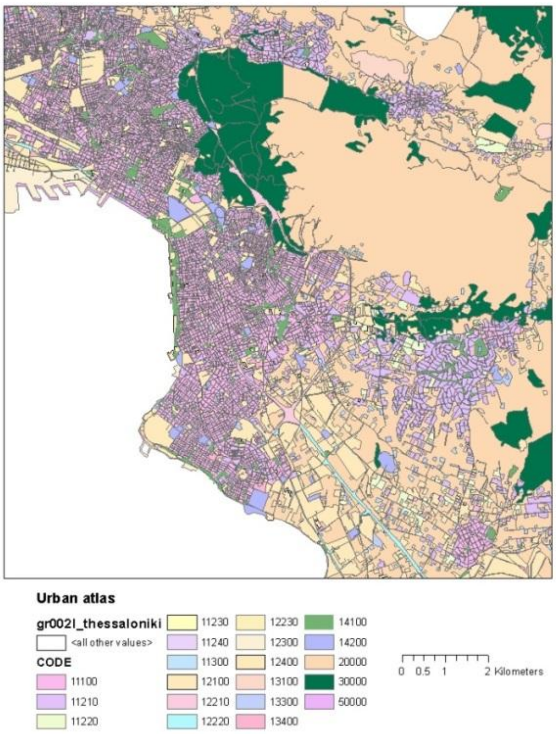


Figure 6: Example of urban atlas map of Thessaloniki. Light and dark green areas are urban green spaces and forests.

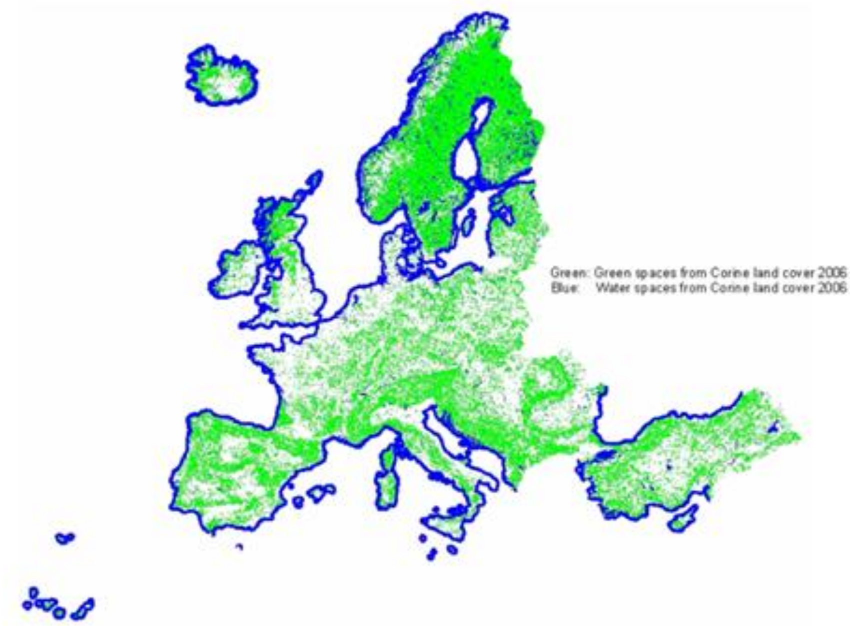


Figure 7: Green and Blue spaces from CORINE land cover map 2006.

Table 13: Basic information, including spatial resolution, coverage, temporal resolution, time span, positional accuracy and minima overall accuracy of green and blue spaces layers in Europe from database Urban Atlas and CORINE.

Dataset	Spatial resolution	Coverage	Time span	Positional accuracy[1]	Min.[2] overall accuracy
Urban Atlas	MMU[3]: 0.25 ha (Urban), 1 ha (Rural) MMW[4]: 10 m	2006: 305 European cities with > 100 000 inhabitants 2012: 695 larger cities in EU28 and EFTA countries	2006, 2012	5 m	80%

[1] Positional accuracy= RMS of control points, accuracy is lower in mountainous areas

[2] Min. overall accuracy= percentage of pixels having right class, minimum over different classes

[3] MMU= minimum mapping unit

[4] MMW= minimum mapping width

CORINE	MMU : 25 ha MMW : 100 m Note: Some countries have higher resolution products	39 European countries	1990,2000, 2006, 2012	100 m	85%

The HEALS green and blue spaces map is combined from CORINE and Urban Atlas to generate higher resolution in urban areas. Both datasets can be downloaded in shapefile format. They were converted into raster imagery with resolution of 50 meters and re-projected from ETRS89 LAEA coordinates to UTM coordinates for further processing. In Urban Atlas, agricultural areas, semi-natural areas and wetlands are in the same class. This is a problem because agriculture areas are considered to have negative effect to health. To solve this problem, the Urban Atlas will be overlaid by the agricultural areas from CORINE dataset to get semi-natural areas and wetlands separated from the agricultural areas. The final combined map contains four classes as green spaces, blue spaces, other areas and no-data areas.

3.2.4 Food consumption patterns and food contamination

Data on food consumption will be gathered from the EFSA databases which are described in the following section. The methodology used for estimating exposure through diet is given in section 4.5.

The EFSA Concise European Food Consumption Database, published in 2008, gathered data on food consumption for adults in Europe according to broad categories. The Concise Database is the first database in Europe containing information from individual dietary surveys from the majority of EU Member States (19 countries).

The EFSA Comprehensive European Food Consumption Database is a source of information on food consumption across the EU. It provides detailed information for a number of EU countries in refined food categories and specific population groups, also partly covering children.

Data are disaggregated per age class as follows for a total of 32 different dietary surveys carried out in 22 different Member States:

- Infants: < 11 months (2 surveys from 2 Member States),
- Toddlers: from 12 up to and including 35 months of age (8 surveys from 8 Member States),
- Children: <10 years (16 surveys from 14 Member States),
- Adolescents: from 10 up to and including 17 years of age (14 surveys from 12 Member States),
- Adults: 18 to 64 years (21 surveys from 20 Member States),
- Elderly: 65 up to and including 74 years of age (9 surveys from 9 Member States) and
- very elderly: from 75 years of age and older (8 surveys from 8 Member States)

FoodEx is a hierarchical system based on 20 main food categories that are further divided into subgroups up to a maximum of 4 levels. It was demonstrated that all data providers were able to classify correctly the large majority of their food to at least the 2nd level of the FoodEx (around 160 categories).

For each country, food consumption data are presented according to the 1st (including 20 categories) and 2nd (including around 160 categories) level of the preliminary FoodEx system; per age class (Infants, Tod-

dlers, Other children, Adolescents, Adults, Elderly and Very elderly); and for the total population and for consumers only. The summary statistics include the total number of individuals and, for each of the first two FoodEx levels, age class, number of consumers, the mean, median and the standard deviation, as well as low and high percentiles. Food consumption statistics are reported both in grams/day and in grams/kg body weight per day, for both chronic and acute consumption.

Summary statistics from the Comprehensive Database have been published for both chronic and acute **consumption**. For calculation of **chronic consumption**, intake statistics have been calculated based on individual average consumption over the total survey period, whereas for acute consumption, statistics have been calculated based on every single reporting day. For example, if subjects in a population had recorded their consumption by means of a 7-day food record, the average intake of each individual over the 7 days was calculated. The average value for each subject was then considered only once when calculating the —chronic average consumption and other statistics related to chronic consumption at population level. On the other hand, —acute consumption figures were calculated using each reporting day independently, and in summing eating occasions for a considered food. All days from each subject (7 days in the above reported example) were used to calculate the —acute average consumption and the other statistics related to acute consumption at population level.

The **EU Menu** project aims at harmonizing data collection on food consumption in Europe. The objective is to provide standardized information on what people **eat in all countries and regions across the EU** in detailed categories and including all population groups. EFSA launched regular calls to MS for data provision. Data should become available at the end of 2017. We are in contact with EFSA to explore the possibility of gaining reserved access to these datasets ahead of time for the purposes of HEALS.

Data on food contamination are only available for WHO regions (more or less the whole Europe) without further spatial disaggregation from the **Gems/food database of WHO.**

From past EFSA reports we can collect food contamination data (average values among all the samples collected at European level) and data on exposure to some contaminants through food at country level. Variation in exposure between countries is influenced by different consumption patterns only, since chemical concentrations in food categories are calculated at a European level.

EFSA is collecting data on food contamination in Europe but they are not available yet to the public. EFSA is in the process of developing a **Data Warehouse (DWH)** that will allow the publication, analysis and distribution, in different formats and at different levels of granularity, of data collected by EFSA. These data include, among others, information on pesticide residues, chemical contaminants, food consumption and chemical hazards. All stakeholders, including the general public, will have access to the data through predefined queries at the level of aggregation decided by EFSA. Not clear when data will be available. Data access policy is under discussion. In relation to our contribution and linkage to the IPCheM portal of the European Commission we should ask for privileged access to these data with the caveat that such data will not be used in their raw form but only after pre-processing and only for exposure modelling purposes in the frame of HEALS.

3.2.5 Phthalates

Phthalates are a family of chemicals that are mainly used as plasticizers and widely found in cosmetics, personal care products, pharmaceuticals, medical devices and paints. People are exposed to phthalates through inhalation, oral and dermal uptake. Evidence has found that phthalates are related to various health effects such as increased risk of allergic disease and asthma.

In HEALS, we have made a literature review on studies that concerned with the phthalates (see Table 14). These studies cover most of the European countries and samples were taken from various vulnerable groups, including infants, children, adolescents, mothers (pregnant women), young men and senior citizens. Samples taken from urine, serum, dust, gas were analyzed and the minimum, maximum, mean and median value can be available for each vulnerable group (not shown here).

Table 14: Studies and datasets for phthalates in Europe and their population group and matrix.

Country	Study (Source)	Population group	Matrix
Austria	Hartmann et al. 2015	Adults, children, senior citizens	Urine
Belgium	DEMOCOPHES 2013	Children	Urine
Bulgaria	Kolarik et al. 2007	-	Dust
Bulgaria	Kolarik et al. 2008	-	Dust
Cyprus	DEMOCOPHES 2013	Children	Urine
Czech Republic	DEMOCOPHES 2013	Children	Urine
Denmark	Bekö et al. 2015	-	Dust
	Langer et al. 2010	-	Dust
	Bekö et al. 2015	Children	Urine
	Copenhagen Mother-Child Cohort (Frederiksen et al.2014)	Children	Urine
	Copenhagen Puberty Study (Frederiksen et al.2014)	Children, adolescents	Urine
	Copenhagen Study on Male Reproductive Health (Frederiksen et al.2014)	Young men	Urine
	DEMOCOPHES 2013	Children	Urine
	Odense Child Cohort (Frederiksen et al.2014)	Pregnant women	Urine
	Mother-Child Cohort (Frederiksen et al.2014)	Children	Urine
France	Dalongeville et al. 2015	-	Gas+PM
	Mercier et al.2012	-	PM
	Elfe (Vandentorren et al. 2011)	Pregnant women	Urine
Germany	Deutschle et al. 2008	-	Dust
	Fromme et al. 2015	-	Dust
	Riechelmann et al. 2007	-	Dust
	Weschler and Nazaroff 2010	-	Dust, gas
	Fromme et al. 2015	-	Gas
	Völkel et al. 2014	Infants, mothers	Urine

	DEMOCOPHES 2013	Children	Urine
	Duisburg Birth Cohort Study (Kasper-Sonnenberg et al. 2012)	Children, mothers	Urine
	ESB (UBA 2012)	Stutdens	Urine
	GerES (Becker et al. 2009)	Children	Urine
	GerES (Koch et al. 2011)	Children	Urine
	LUPE 3 (Fromme et al 2015)	Children	Urine
	Kolarik et al. 2007	-	Dust
Hungary	DEMOCOPHES 2013	Children	Urine
Ireland	DEMOCOPHES 2013	Children	Urine
Italy	Orecchio et al. 2013	-	Dust
Luxembourg	DEMOCOPHES 2013	Children	Urine
Netherlands	Generation R (Ye et al. 2008)	Pregnant women	Urine
Norway	Sabaredzovic et al 2015	Pregnant women	Urine
Norway	ECA (Bertelsen et al. 2013)	Children	Urine
Poland	DEMOCOPHES 2013	Children	Urine
Portugal	DEMOCOPHES 2013	Children	Urine
Romania	DEMOCOPHES 2013	Children	Urine
Slovak Republic	DEMOCOPHES 2013	Children	Urine
Slovenia	DEMOCOPHES 2013	Children	Urine
Spain	DEMOCOPHES 2013	Children	Urine
	INMA (Casas et al. 2011)	Children, pregnant women	Urine
Sweden	Bergh et al. 2011	-	Dust, gas
	DEMOCOPHES 2013	Children	Urine
	PIVUS (Olsén et al. 2012)	Female Seniors, male seniours	Serum
Switzerland	DEMOCOPHES 2013	Children	Urine
United Kingdom	DEMOCOPHES 2013	Children	Urine
	DEMOCOPHES (Exley et al. 2015)	Mothers	Urine

3.2.6 Electromagnetic field

The public is exposed to electromagnetic fields (EMF) generated by an increasing variety of electrical and electronic devices and installations. The rapid increase in mobile telecommunications and the growing range of personal, domestic, commercial and medical equipment have considerably increased the number

of sources of EMF exposure and are significantly changing the level, type and pattern of everyday exposure of the public. In order to properly assess the health risk associated to radiation exposure during different stages of human life in HEALS, a comprehensive estimation of all radiation (ionizing and non-ionizing) exposure events from conception to death is needed. Such a process would lead to the construction of the radiological exposome.

3.2.6.1 Typical EMF exposure levels and EMF levels found in common environments

Most people are exposed to magnetic fields that average less than 2 milligauss (mG), although individual exposures vary. The data shown in the following table are median measurements taken at four different sites for each environment category (Zaffanella et al., 1992).

Table 15. EMF Exposures in Common Environments.

Magnetic fields measured in milligauss (mG)					
Environment	Daily Median* exposure	Top 5th percentile	Environment	Daily Median* exposure	Top 5th percentile
OFFICE BUILDING			MACHINE SHOP		
Support staff	0.6	3.7	Machinist	0.4	6.0
Professional	0.5	2.6	Welder	1.1	24.6
Maintenance	0.6	3.8	Engineer	1.0	5.1
Visitor	0.6	2.1	Assembler	0.5	6.4
SCHOOL			Office staff	0.7	4.7
Teacher	0.6	3.3	GROCERY STORE		
Student	0.5	2.9	Cashier	2.7	11.9
Custodian	1.0	4.9	Butcher	2.4	12.8
Administrative staff	1.3	6.9	Office staff	2.1	7.1
HOSPITAL			Customer	1.1	7.7
Patient	0.6	3.6			
Medical staff	0.8	5.6			
Visitor	0.6	2.4			
Maintenance	0.6	5.9			

*The median of four measurements. For this table, the median is the average of the two middle measurements.

What EMF field levels are encountered in the home?

In a study (Zaffanella et al., 1992) in which spot measurements of magnetic fields were made in the center of rooms in 992 homes throughout the United States, half of the houses studied had magnetic field measurements of 0.6 mG or less, when the average of measurements from all the rooms in the house was calculated (the all-room mean magnetic field). **The all-room mean magnetic field for all houses studied was 0.9 mG.** The measurements were made away from electrical appliances and reflect primarily the fields from household wiring and outside power lines.

3.2.6.2 Radiation Exposure calculation - Methodology

For the purposes of the HEALS study, various radiation sources will be taken into account in order to get a picture of the total radiation population is exposed to. Emphasis will be given is sources such as: a) electrical appliances, b) cell phones, c) power lines, d) UV radiation, c) radiation from medical imaging tests and e) exposure to radon. If over a given period, *T*, a person goes through *n* locations, where the concentration of

the under study exposure factor or environmental stressor under consideration is C_n, spending a fraction of time f_n, then the total personal exposure, E_T, is given by the following equation:

$$E_T = \sum_n f_n \cdot C_n$$

The aforementioned approach, which is commonly used for estimating indoor or outdoor exposure, will be used for calculating radiation exposure as well. Information on the time one spends in a specific location, $\mathbf{f_n}$, will be based on actual time-activity patterns captured by the HEALS pilot campaigns, fused with detailed time-activity data or questionnaires outcomes from previous European studies.

Concentration data, $\mathbf{C_n}$, will be based on direct measurements of a compound taken either from HEALS fieldwork, if available, or previous studies that can be found in literature. Adjustment factors will be taken into account based on one's distance from the emission source (such as the case of assessing EMF exposure from power lines) or inhalation rate that varies for different activities/gender/age (used when assessing exposure to radon).

Sections below focus on how radiation exposure can be calculated depending on the examining source.

3.2.6.3 Electrical appliances

Exposure estimations will be based on the proximity to electric appliances and the frequency of their usage. We will also take into account the total amount of appliances at a nearby distance. Therefore, radiation exposure (per day, month, year) shall be calculated for a single room/office with x electrical devices based on the equation below:

$$E_{devices} = \sum_x f \cdot C_x$$

Where:

f is the time spent indoors (in this case in a house or office) and

$\mathbf{C_x}$ is the magnetic field strength of the electrical appliance x

The strength (magnitude) of a magnetic fields, is expressed in tesla (T) (more commonly in millitesla (mT) or microtesla (μT)) or in gauss, (G) where 10,000 G equals to 1 T. It does not depend on how large, complex, powerful, or noisy the appliance is. The following table, based on data gathered in 1992, lists the EMF levels generated by common electrical appliances commonly found in homes and workplaces. (EPA, 1992; National Institute of Environmental Health Sciences, 2002; World Health Organization, 2002). The field strength around all appliances rapidly decreases the further you get away from them.

Table 16. Typical magnetic field strength of household appliances at various distances.

Sources of Magnetic Fields (mG)*									
	Distance from source (m)					Distance from source (m)			
	0.15	0.3	0.6	1.2		0.15	0.3	0.6	1.2
Office Sources					Workshop Sources				
AIR CLEANERS					BATTERY CHARGERS				
Lowest	110	20	3	-	Lowest	3	2	-	-
Median	180	35	5	1	Median	30	3	-	-
Highest	250	50	8	2	Highest	50	4	-	-
COPY MACHINES					DRILLS				
Lowest	4	2	1	-	Lowest	100	20	3	-

Sources of Magnetic Fields (mG)*

	Distance from source (m)					Distance from source (m)			
	0.15	0.3	0.6	1.2		0.15	0.3	0.6	1.2
Median	90	20	7	1	Median	150	30	4	-
Highest	200	40	13	4	Highest	200	40	6	-
FAX MACHINES					POWER SAWS				
Lowest	4	-	-	-	Lowest	50	9	1	-
Median	6	-	-	-	Median	200	40	5	-
Highest	9	2	-	-	Highest	1000	300	40	4
FLUORESCENT LIGHTS									
Lowest	20	-	-	-	Bathroom Sources				
Median	40	6	2	-	HAIR DRYERS				
Highest	100	30	8	4	Lowest	1	-	-	-
PCS WITH COLOR MONITORS					Median	300	1	-	-
Lowest	7	2	1	-	Highest	700	70	10	1
Median	14	5	2	-	ELECTRIC SHAVERS				
Highest	20	6	3	-	Lowest	4	-	-	-
					Median	100	20	-	-
					Highest	600	100	10	1
Living/Family Room Sources					COLOR TELEVISIONS				
CEILING FANS					Lowest		-	-	-
Lowest					Median		7	2	-
Median					Highest		20	8	4
Highest									
AIR CONDITIONERS					Laundry/Utility Sources				
Lowest		-	-	-	PORTABLE HEATERS				
Median		3	1	-	Lowest	5	1	-	-
Highest		20	6	4	Median	100	20	4	-
					Highest	150	40	8	1
Kitchen Sources					WASHING MACHINES				
MIXERS					Lowest	4	1	-	-
Lowest	30	5	-	-	Median	20	7	1	-
Median	100	10	1	-	Highest	100	30	6	-
Highest	600	100	10	-	IRONS				
COFFEE MAKERS					Lowest	6	1	-	-
Lowest	4	-	-	-	Median	8	1	-	-
Median	7	-	-	-	Highest	20	3	-	-
Highest	10	1	-	-	VACUUM CLEANERS				
DISHWASHERS					Lowest	100	20	4	-
Lowest	10	6	2	-	Median	300	60	10	1
Median	20	10	4	-	Highest	700	200	50	10
Highest	100	30	7	1					
MICROWAVE OVENS					Bedroom Sources				
Lowest	100	1	1	-	DIGITAL CLOCK				

Sources of Magnetic Fields (mG)*

	Distance from source (m)					Distance from source (m)			
	0.15	0.3	0.6	1.2		0.15	0.3	0.6	1.2
Median	200	4	10	2	Lowest		-	-	-
Highest	300	200	30	20	Median		1	-	-
ELECTRIC OVENS					High		8	2	1
Lowest	4	1	-	-	ANALOG CLOCKS				
Median	9	4	-	-	Lowest		1	-	-
Highest	20	5	1	-	Median		15	2	-
REFRIGERATORS					Highest		30	5	3
Lowest	-	-	-	-	BABY MONITOR				
Median	2	2	1	-	Lowest	4	-	-	-
Highest	40	20	10	10	Median	6	1	-	-
TOASTERS					Highest	15	2	-	-
Lowest	5	-	-	-					
Median	10	3	-	-					
Highest	20	7	-	-					

** Dash (–) means that the magnetic field at this distance from the operating appliance could not be distinguished from background measurements taken before the appliance had been turned on.*

At a distance of 30 cm the magnetic fields surrounding most **household** appliances are more than 100 times lower than the ICNIRP (International Commission on Non-Ionizing Radiation Protection) guideline limit of 100 µT (1000 mG) at 50 Hz (83 µT at 60 Hz) for the general public.

3.2.6.4 Cell phones

Radiation exposure estimations will be based on the following equation:

$$E_{phone} = \sum_{n} f_n(s) \cdot SAR$$

Where:

- $\mathbf{f_n}$ is the time spent on phone or VoIP calls and
- **SAR** corresponds to the mobile device's *specific absorption rate*

Nowadays people spend more time using their cell phones for surfing the web, checking social networks or playing games than making phone calls. For example, a UK survey, launched by telecommunications services provider O2 (O2 telecommunications services provider, 2012), indicates that the average smartphone owner spends more than two hours each day using the device. Out of this time, making phone calls with the smartphone was only the fifth most popular use for the gadget (12 min, 6 sec). While capturing mobile usage patterns, it is essential that VoIP calls are also reported. It is inevitable that mobile VoIP will be adopted on mass scale in the next ten years. This will trigger increases in the mobile voice traffic, mainly due to the attractive pricing (sometimes costless) of international calls.

Mobile usage patterns will be retrieved by a) the HEALS time-activity diaries and b) outcomes from previous studies. Outcomes of questionnaires and data from cell phone service providers will be used in order to estimate:
a) how "regularly" study participants use cell phones (the number of calls per week or month),
b) the average length of a typical cell phone call or even

c) the total hours of lifetime use, calculated from the length of typical call times, the frequency of use, and the duration of use.
Variations in age and gender will be also taken into account (OECD, 2013).

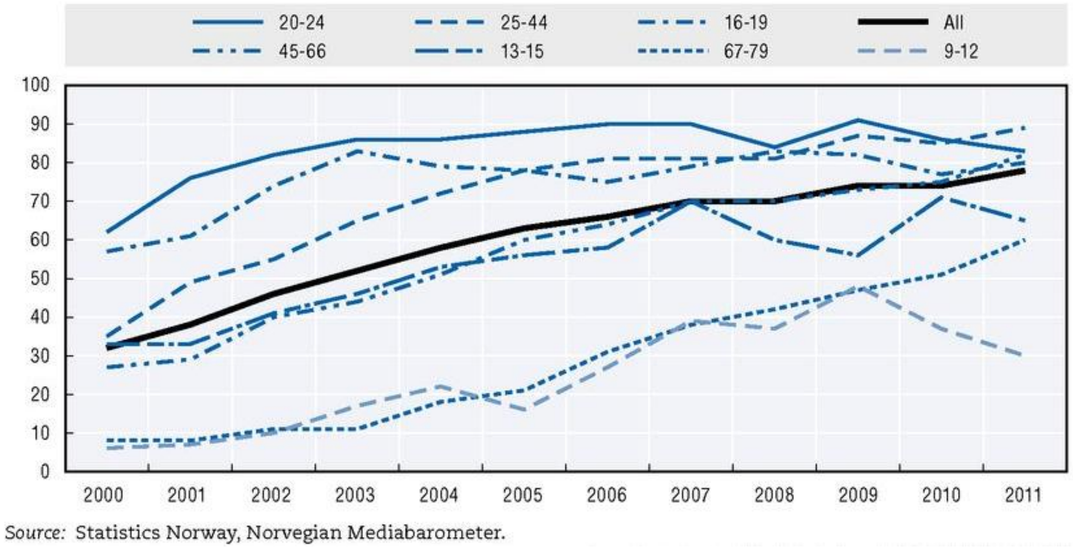


Figure 8. % of individuals in each age category giving/receiving private mobile calls on an average day in Norway.

Specific absorption rate (SAR)
A SAR value is a measure of the maximum energy absorbed by a unit of mass of exposed tissue of a person using a mobile phone, over a given time or simply the power absorbed per unit mass. SAR values are usually expressed in units of watts per kilogram (W/kg) in either 1g or 10g of tissue. SAR provides a straightforward means for measuring the RF exposure characteristics of cell phones to ensure that they are within the safety guidelines set by the Federal Communications Commission (FCC).

In cases where we want to estimate large-scale exposure, mobile penetration will be taken into consideration (mobile cellular telephone subscriptions to a public mobile telephone service per 100 people). The indicator applies to all mobile cellular subscriptions that offer voice communications. Such kind of information clustered per country and year is provided online by the World Bank (The World Bank, 2015).

3.2.6.5 Power lines

For homes near high voltage powerlines, the magnetic field exposure will vary according to the amount of current carried by the powerline and the distance of the home from the powerline. Generally, homes that are more than 50 m from a high voltage powerline are not expected to have higher than typical magnetic fields. For substations and transformers, the magnetic fields at distances of 5-10m away are generally indistinguishable from typical background levels in the home (Kaune and Zaffanella, 1992; National Institute of Environmental Health Sciences, 2002). The figure below shows a range of magnetic field levels measured by the Australian Radiation Protection and Nuclear Safety Agency (ARPANSA) around powerlines and in homes. These are well below the exposure limit in the international guidelines of 100 µT (1000 mG).

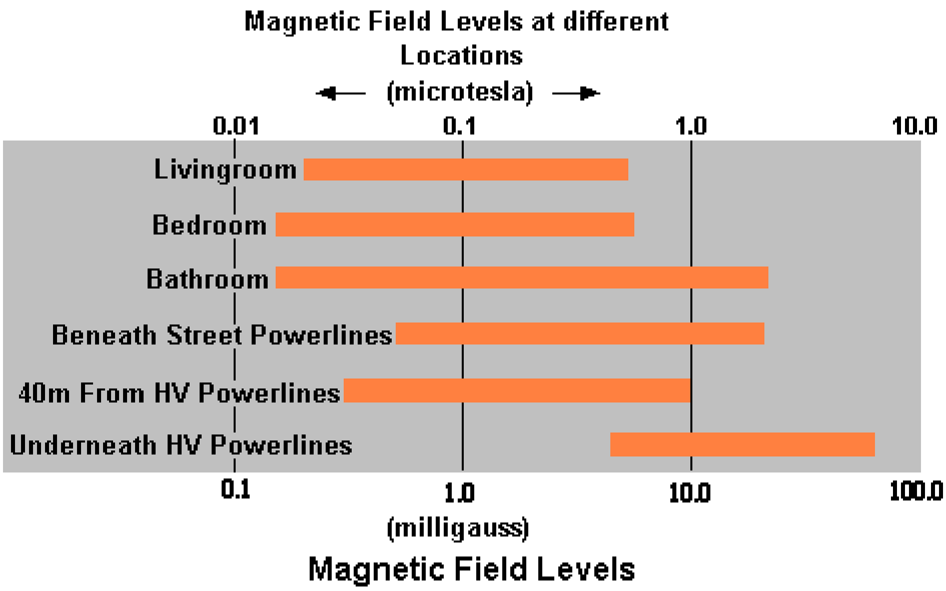
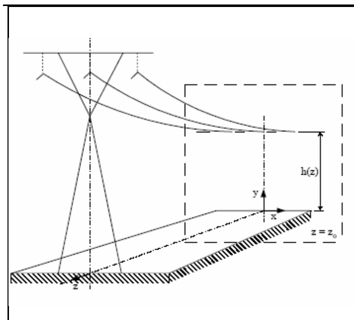


Figure 9. A range of magnetic field levels (measured around powerlines and in homes.

The following table shows the measured values of the magnetic field for different heights from the lowest conductor to ground and different distances to the first pole (Salma, 2006).

Table 17. Zahrani-Saida (Lebanon) Power Lines Magnetic Fields under 220kV, 311.3A.

	Height (m)	26	24	21	21	24.5	28	30.4	25	26.6	31.7
	Distance Correspond-ing to pole 1 (m)	0	40	80	120	160	240	280	320	360	392.4
	Magnetic Field (mG)	8.6	9.9	8.1	6.8	6.3	4.6	4.3	5.7	5	4.4

Based on this example taken from literature, the estimation on the radiation exposure will be made based on the proximity to power lines and the magnetic field levels that correspond to the sport where people live/work.

3.2.6.6 UV

Radiation estimations will be based on a) UV concentrations per region and b) the time spent out in the sun or sun bathing clustered per country/age/gender that will be retrieved from questionnaires and time activity diaries.

$$E_{UV} = \sum_n f_n \cdot C_n$$

Data on UV radiation will be gathered based on the EuroSun (Smith, 1990) atlas of UV irradiation that has been developed by averaging daily UVA and UVB irradiations per month and for four 5-year periods (1988-1992; 1993-1997; 1998-2002; 2003-2007) for Europe as a whole and for individual countries, for UVA, UVB and total UV irradiations, at both absolute and relative levels. UV bandwidth was defined according to WMO standard (WMO no8, guide to meteorological instruments and methods of observation), with UVA

defined from 315 nm to 400 nm; and UVB from 280 nm to 315 nm. Specifically data are available on the SoDa website (SoDa, 2004) with web-based tools to extract data from 1985 for UVA and UVB irradiance for every geographical area in Europe (grid cell of 5km).

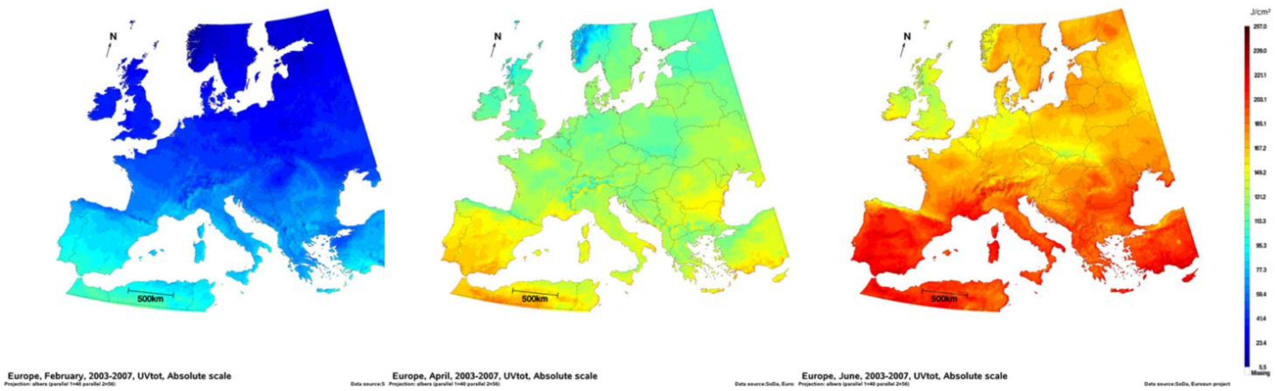


Figure 10. Maps of the averaged daily UV irradiation (in J/cm^2) in Europe. Visual comparison for the periods: i) February, ii) April, iii) June over the period 2003 – 2007.

3.2.6.7 Radiation from medical imaging tests

The table below gives dose estimates for typical diagnostic x-ray and nuclear medicine exams (Lin, 2010).

Table 18. Comparison of Radiation Doses from medical imaging tests and background radiation.

Examination	Radiation dose (mSv)	Time to accumulate comparable natural background dose
Computed tomography		
Sinuses	0.6	2 mo
Head	2	8 mo
Chest	7	2 y
Chest (pulmonary embolism)	10	3 y
Abdomen and pelvis	10	3 y
Multiphase abdomen and pelvis	31	10 y
Radiography		
Extremity	0.001	< 1 d
Chest	0.1	10 d
Lumbar spine	0.7	3 mo
Abdomen	1.2	5 mo
Other		
Mammography	0.7	3 mo
Bone densitometry	0.001	< 1 d
Nuclear Medicine		
Lung ventilation/perfusion	2	8 mo
Bone scan	4.2	1 y, 4 mo
Cardiac perfusion (sestamibi)	12.5	4 y
Fluoroscopy		

Barium swallow	1.5	6 mo
Coronary angiography	5 to 15	20 mo to 5 y

These doses are effective doses (given in mSv and mrem, the SI unit of measure of the effects of ionizing radiation), which are theoretical quantities proposed by the International Commission on Radiation Protection (ICRP) to assess the health risks of low doses of ionizing radiation. The effective dose is used as a way of comparing the risk of a partial body exposure to that due to a whole body exposure (such as that due to background radiation in the environment). A useful way to understand radiation doses from diagnostic examinations is to compare them to average natural background radiation (3 mSv per year). (The Health Physics Society, 2008).

Until today, only a few studies have looked at how much radiation people are exposed to through diagnostic imaging over the course of their lives. By combining the effective dose of radiation exposure for each patient or group of population with information on the frequency with which people (different age/gender/country) are going through image scanning, either in a yearly basis or over the period of their life, we can have a representative picture of the radiation they have been exposed to.

3.2.6.8 Radon

Exposure to radiation can be estimated based on the following formula:

$$\mathrm{E}_{Radon} = \sum_{\mathrm{n}} \mathrm{f_n} \cdot \mathrm{C_n} \cdot \mathrm{inh_{act}}$$

Information on time, $\mathbf{f_n}$, spent indoors (residential building or office) retrieved from questionnaires and time activity diaries. Data on radon radiation will be based either on a) national reports (JRC, 2005) or on b) the European Atlas of Natural Radiation, a project launched by the Joint Research Centre (JRC) of the European Commission (Tollefsen et al., 2014). An additional factor influencing exposure to radon is inhalation rate, *inh*. Different types of activities demand different levels of effort that correspond to different inhalation rates. Detailed description of the activity based inhalation rates is given elsewhere (Sarigiannis et al., 2012). Based on the time-weight contribution of the activities, the level of intensity and the corresponding inhalation rate, adjustment factors will be used for the cases where the under study population is located indoors (at home or at the office).

Table 19. Statistics for European radon surveys, 2005.

Country	Estimated annual mean levels (Bq/m3)	% dwellings > 200 Bq/m3 and <400 Bq/m3	% dwellings > 400 Bq/m3
Albania	NA	NA	NA
Austria	97	8	4
Belgium	48	1.7	0.3
Croatia	68	5.4	1.8
Cyprus	19	0	0
Czech Republic	140	10 - 15	2 - 3
Denmark	53	2.7	0.2
Estonia	60	2 - 2.5	0.3 - 0.5
Finland	120	8.7	3.6

France	63	6.5	2
FYROM	NA	NA	NA
Germany	50	2.5	< 1
Greece	55	2	1.1
Hungary	NA	5.1	0.8
Ireland	89	6	1.5
Italy	70	3.2	0.9
Latvia	NA	NA	NA
Lithuania	55	2.5	0.3
Luxembourg	115	NA	3
Malta	40	0	0
Netherlands	23	0.3	0
Norway	89	6	3
Poland	49	1.6	0.4
Portugal	NA	NA	NA
Romania	45	NA	NA
Serbia-Montenegro*	144	18	4
Slovakia	108	14	11
Slovenia	87	5.5	2
Spain	90	4	2
Sweden	108	6 - 7	3 - 4
Switzerland	77	10	7
United Kingdom	20	0.4	0.1

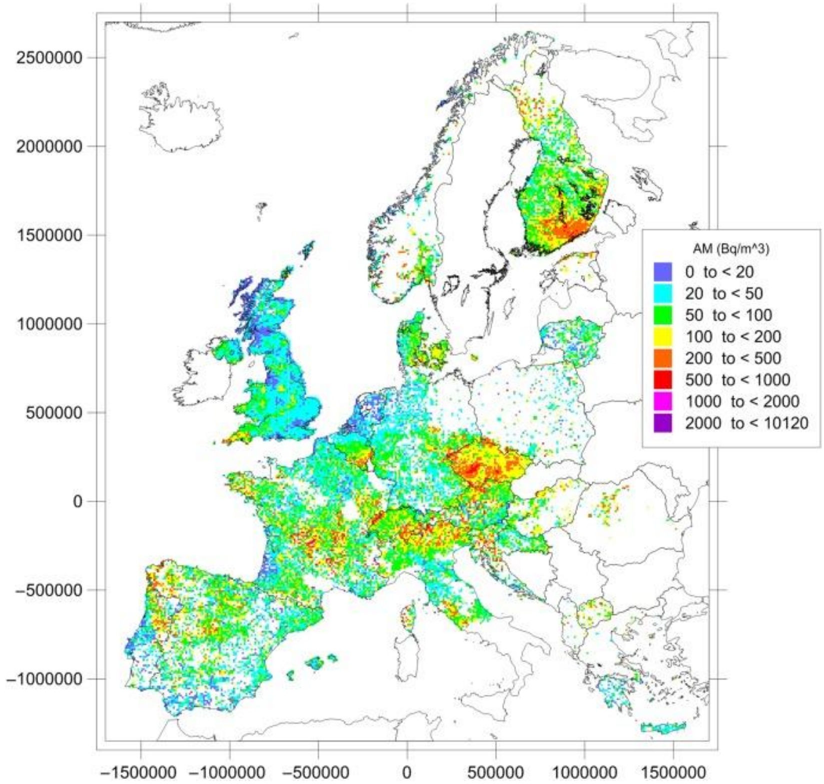


Figure 11. Arithmetic means over 10 × 10 km cells of long-term radon concentration (Bq/m^3) in ground-floor rooms. Data available until May 2014 included – JRC Radon Atlas

Table 20. ICRP Statement or Radon 2010 (Tirmarche et al., 2010) 115 – reference levels.

Location	Reference level (Bq/m^3)	Annual effective dose (mSv)
Homes	0.6	2 mo
Workplaces	2	8 mo
Mines	7	2 y

3.2.7 Noise

Noise is an element that can`t be avoid in modern cities. People exposed to noise may experience harm to hearing, sleep disturbance and even higher chances of liver, heart and blood diseases. Thus HEALS include noise as a stressor to be studied and relevant data are collected.

Noise map is a standard approach to obtain the façade noise levels a person is exposed to. Noise maps are accessible for major cities in Europe. To get the environmental noise level, it is necessary to know the position where a person is located, which can be obtained from GPS. Figure 12 shows the GPS tracks of participant that were collected during eight consecutive days. Different colors are used for the three transport modes: walking, cycling and transport (motorized). In addition, there are GPS track points that are classified as “Place”, which represents the location where a participant is spatially stationary for a period of time.

The noise levels at the façade of the dwelling were determined by plotting the GPS track of a person on the noise map. For locations outside the noise map, an approximate method is developed for estimating the noise levels at the dwelling with the help of Open Street Map (OSM) data. An area of 600 by 600 m^2 OSM data around the dwelling is downloaded, and road traffic flows are assigned to the different road types to calculate the noise level.

Except for the noise map, we also use a small measurement station Netatmo to collect the indoor noise data. Besides, a smartphone App Widenoise was applied. However the usefulness of the App is questionable.

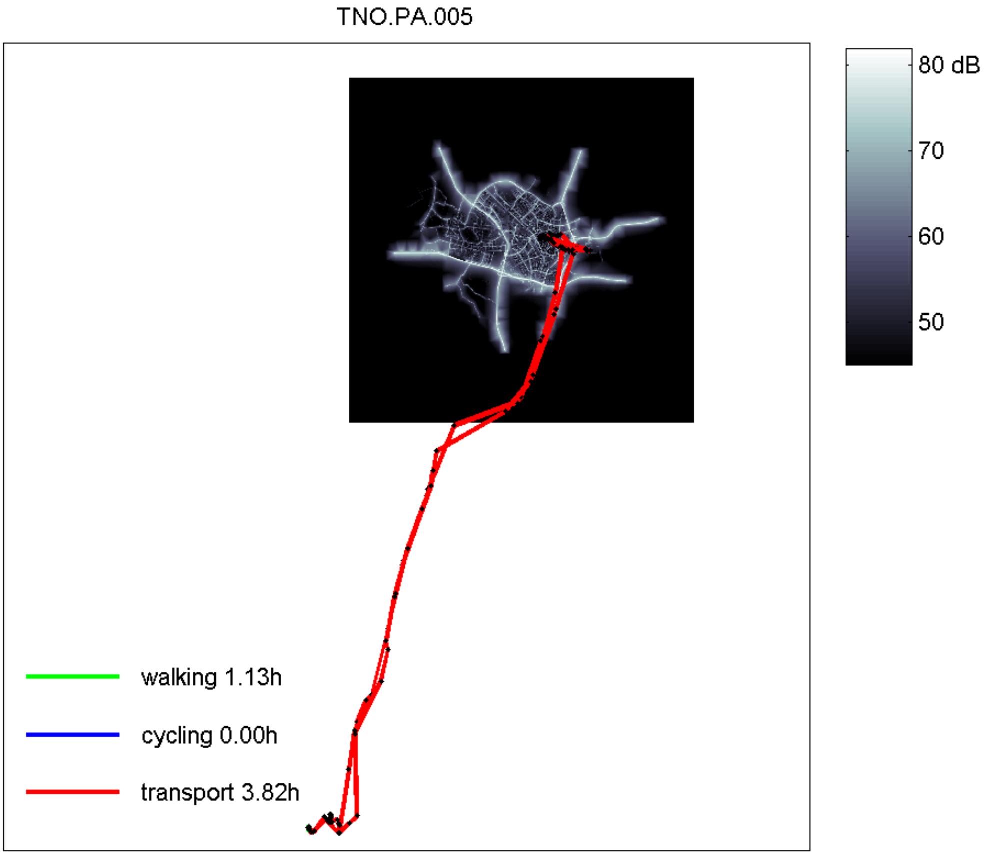


Figure 12: GPS tracks of a participant during eight days. The area is 100 by 100 km^2.

3.2.8 Sensors for time-activity information

The rise of sensors in smart consumer products made the simultaneous widespread collection with multiple sensors possible. The use of these sensors may increase the accuracy of time-activity modelling. In the HEALS pre-pilot study, the validity of using smart consumer products for assessing time-activity data was investigated (to identify whether a person is indoors, outdoors or in transit). Specifically, the accuracy of a

consumer phone app for determining activity patterns (Moves) was compared with a GPS unit and a paper log was investigated.

In the HEALS pre-pilot study sensor data was collected for 28 persons (UK, Germany, Croatia, Greece and the Netherlands) with the Moves phone app, the Fitbit flex, a UV meter, a personal temperature meter and a GPS device. During the test time the volunteers kept a detailed diary on their location and activity.

		Paper Log				
		Indoor	**Outdoor**	**cycling**	**walking**	**motorized**
Moves	**home**	21915	115	0	33	28
	place	26734	2563	27	122	165
	work	4086	34	0	46	17
	unknown	2603	270	0	11	15
	outdoor	0	45	0	0	0
	cycling	5	0	254	17	0
	walking	940	1030	29	751	94
	transport	129	140	90	165	2160
	% correct	**98%**	**1%**	**64%**	**66%**	**87%**

Figure 13: Performance of the Moves app versus the paper log for determining location and activity among 28 volunteers.

The overall accuracy of the Moves app versus the paper log was high (91% correct). However the accuracy for staying stationary outdoors was low (Figure 13). The performance of the GPS device versus the paper log gave similar results (data not shown). Currently, machine learning techniques such as random forest models are being explored to test the potential improvement of the classification by adding additional sensor data (Fitbit, UV and temperature) and other data sources such as time of the day, day of the week, outdoor temperature and rainy day.

These data on time-activity can be broadly applied in combination with exposure levels associated with these locations and activities to assess exposure to a range of stressors.

3.2.9 Sensors for indoor air quality

In HEALS pilot study (WP9) a low cost sensor for particulate matter, the Dylos, is being used at 150 homes of families across Europe (UK, Greece and the Netherlands) to estimate exposure to particulate matter (PM). The Dylos obtains particle number concentrations (PNC) in the range of 0.5 to 20µm in two size bins (above and below 2.5 µm). Outdoor PM levels are usually reported in mass concentrations. At a real home setting the validity of using Dylos to obtain PNC concentrations was tested. This study also formed the basis for calibration and validation of a method to convert the PNC obtained by DYLOS to PM2.5 mass concentration, which can be fused with outdoor PM2.5 levels (section 4).

3.2.9.1 Validation

Side by side measurements were collected for three to five days in the main living space of two homes of volunteers with the Dylos DC1700™ (0.5 to 20µm, two size bins) and the APS™ Aerodynamic Particle Sizer® (0.5 to 20µm, 52 size bins) simultaneously. The volunteers acted according to their normal daily routines. Figure 14 shows the correlation between the Dylos and the APS. The PNC obtained by Dylos correlated well with the PNC obtained by APS for all size ranges (R>0.92). For total particles (>0.5 µm) and for the smaller particles (0.5-2.5 µm), the Dylos particles counts match reasonably well with the APS particle counts for

both volunteers (close to y=x line). However, the APS seems to have slightly higher result than Dylos when particle numbers increase to above 30 particles/cm^3. Visual inspection of the data and activity log indicated that these high particle counts mainly occur during cooking events where the APS counts twice as many particles as the Dylos (data not shown). For larger particles (>2.5 µm), however, the Dylos seems to overestimate the PNC compared to the APS. Currently more data with these same devices are combined with gravimetric sampling (Harvard impactor) at the homes of other volunteers to obtain more validation data.

3.2.9.2 Conversion of the PNC obtained by Dylos to PM2.5 mass concentration

Methods for converting PNC to PM2.5 mass concentration have been described in two publications, both of which are based on simulated sources in experimental setups. The validation study described above was used to corroborate the equations at a real home setting.

The PNC measured by Dylos was converted to PM2.5 mass concentration as follows. First, the PNC (0.5-2.5 µm) measured by APS was converted to PM2.5 mass based on the mean particle density (1.6g/cm^3) and mean aerodynamic diameter by size bin (n=22). Then a second order polynomial equation was used to fit the Dylos measurement on the calculated PM2.5 mass for a random sample of 33% of the data. Figure 15 and Table 21 show this equation together with other two equations obtained from the literature.

The derived model was validated internally with the remaining 67% of the data, resulting in a R^2 of 0.76 between the Dylos PNC and PM2.5 mass concentration estimates based on APS.

The fitted model for converting the Dylos PNC measurement to PM2.5 mass concentration gave results that are comparable to applying a previously published model based on an experimental setting. The model will be used to obtain PM2.5 mass concentrations in the HEALS pilot study. Preliminary conversions for the 40 Dutch homes demonstrate a median modelled PM2.5 mass concentration of 6.7 µg/m^3 (P5-P95: 6-8 µg/m^3). For external validation, gravimetric PM2.5 samples are collected in parallel with the Dylos data among several HEALS pilot study participants.

Table 21: Fitted associations between Dylos PNC and PM2.5 mass estimates in the HEALS study and two published studies.

Study	Fitted relationship	Parameter values	Measurement device
Study from TNO	$Y=a_0+a_1x+a_2x^2$	a_0=4.97 a_1=2.003e^{-3} a_2=1.69e^{-7}	APS 3321
Semple et al. 2013	$Y=a_0+a_1x+a_2x^2$	a_0=11.7 a_1=0.0049 a_2=2e^{-7}	TSI Side Pak AM510
Dacunto et al. 2015	$Y=mx^n$	m=1.09e^{-7} N=2.111	TSI Side Pak AM510

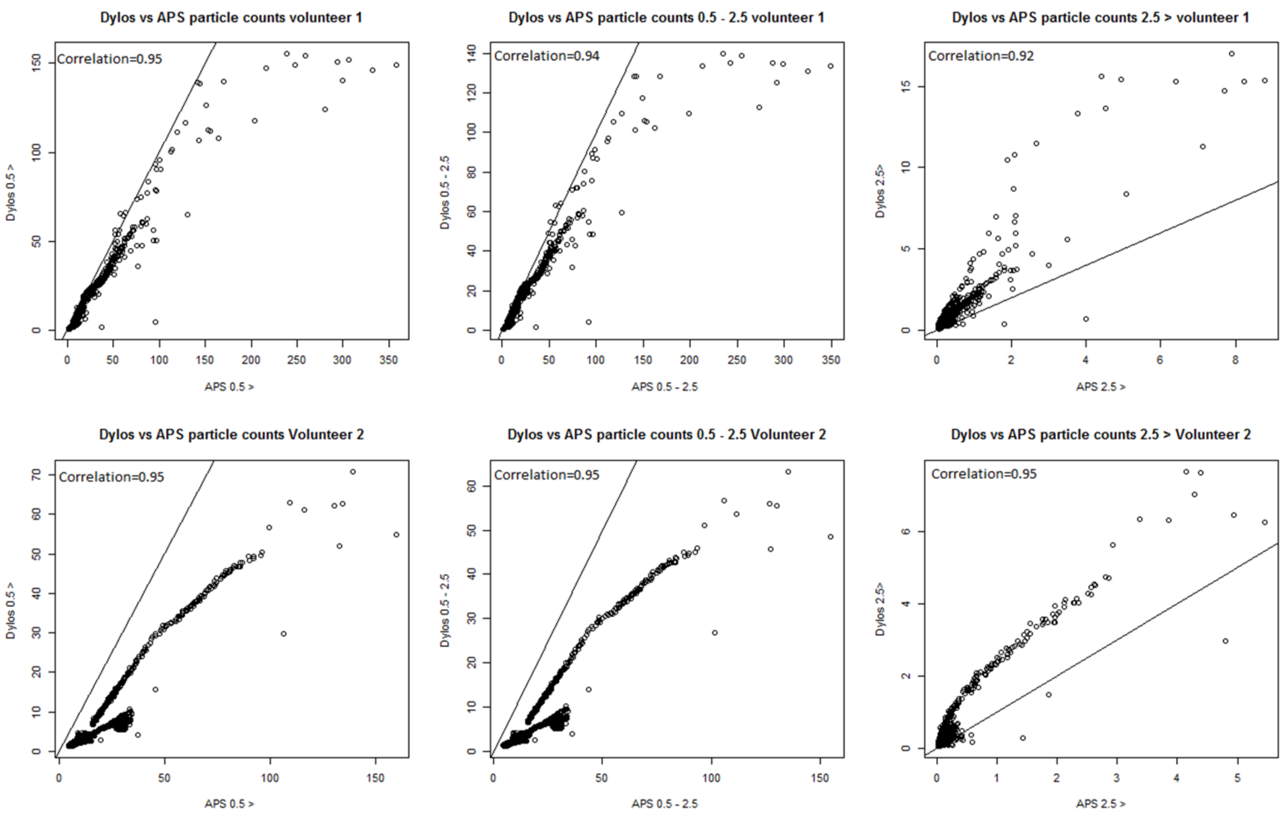


Figure 14: Correlations between the Dylos and APS for three different particle size ranges on all measured days for both volunteers.

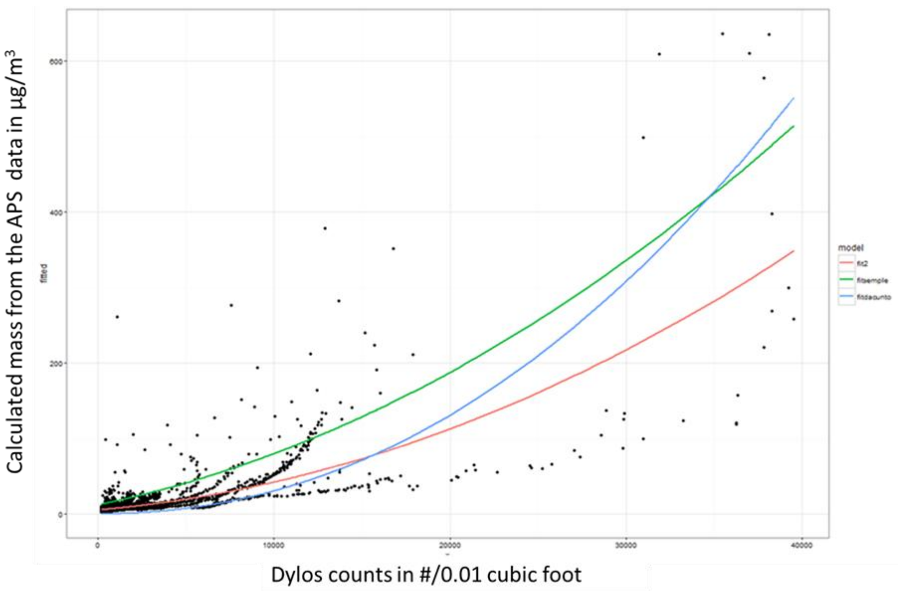


Figure 15: Dylos PNC versus calculated mass from APS (all data points), with the fitted equation on this data (red line) and the equations published by Semple et al. (2013) (green line) and Dacunto et al. (2015) (blue line).

4 Methodology applied to combine data sets

4.1 Combination of socioeconomic data, living conditions and time-activity patterns

As described in section 3.1.1, the EU-SILC data consists of Pan-European comparative information involving income, poverty, social exclusion, housing, labour, education and health at household and individual level since 2004. Due to the framework regulation implemented for EU-SILC, the data collected should stay anonymous and the address of the households and individuals are not given. This obstructs determining the important parameters for estimating exposure to stressors, i.e. mainly the location of the individual. To address this issue we associate probabilities of the geospatial location of an individual based on available information on the region of the residence of the household (variable DB040) and the degree of urbanisation of the individual's location within this region (variable DB100).

MTUS consists of multiple files. The episode file (HEF) holding the time-activity patterns, and the aggregate file (HAF) holding the socioeconomic data will be joined based on identifier variables (including household identifier, person/diarist identifier and diary identifier). The fused data benefits from the format of the episode data, with each row recording the time a diarist spent on a certain activity in a certain location (including indoors, outdoors and in transit). Simultaneously, the fused data inherits more detailed socioeconomic information from the aggregate file compared to the episode file.

We combine EU-SILC and MTUS as each of the individual datasets is too limited in terms of information needed for the exposure assessment. To join the two datasets, some important variables that exist in both datasets will be extracted. The crucial variables to be considered include age, gender, household income, employment status, occupation, tenure status, education level, civil status, car ownership, nationality and region (Table 22). The data after fusion contains the information from both EU-SILC and MTUS, including socioeconomic data, living conditions, geospatial information and time-activity patterns. It will be utilized as the input data for exposure modelling.

Table 22: Variables used to join data of individuals from the EU-SILC dataset and the MTUS dataset.

Variables in EU-SILC	Description	Variables in MTUS	Description
DB020	Country	COUNTRYA	Country of study
DB040	Region	REGION	Region of residence
DB100	Degree of urbanisation	URBAN	Urban or rural household
HH021	Tenure status	OWNHOME	Whether household owns or rents home
HS110	Do you have a car?	VEHICLE	Does household have a private vehicle
HY010	Household income	INCOME	Total household income
PB140	Year of birth	AGE	Age of diarist
PB150	Sex	SEX	Sex of diarist
PL031	Current economic status	EMPSTAT	Employment status
PB200	Consensual union	CIVSTAT	Is diarist in couple?
PE040	Highest ISCED level attained	EDCAT	Harmonised highest level of education
PL051	Occupation	ISCO1	Occupation

4.1.1 DB040 and REGION

Region is a crucial factor to be considered because together with degree of urbanisation (DB100), it helps to geospatially locates the individuals from the dataset. The variable DB040 is recorded based on Nomenclature of Territorial Units for Statistics (NUTS). NUTS were created by the European Office for Statistics (EUROSTAT) as a single hierarchical classification of spatial units used for statistical production across the European Union. Three levels of NUTS are defined (EUROSTAT 2015):

- NUTS1: major socioeconomic regions (see Figure 16).
- NUTS2: basic regions for the application of regional policies (see Figure 16).
- NUTS3: small regions for specific diagnoses.

In EU-SILC, NUTS2 is the lowest territorial level that can be achieved for Czech Republic, Spain, Finland, and France. For other countries, data are only available at NUTS1 or even country level (see Table 23).

Table 23: The lowest territorial unit available in EU-SILC dataset for EU countries.

Territorial unit	Country
NUTS2	Czech republic, Spain, Finland and France
NUTS1	Austria, Bulgaria, Greece, Hungary, Italy, Poland, Romania and Sweden
Country	Switzerland, Cyprus, Germany, Denmark, Estonia, Croatia, Ireland, Iceland, Lithuania, Luxembourg, Latvia, Malta, Nederland, Norway, Portugal, Republic of Serbia, Slovenia and Slovak republic

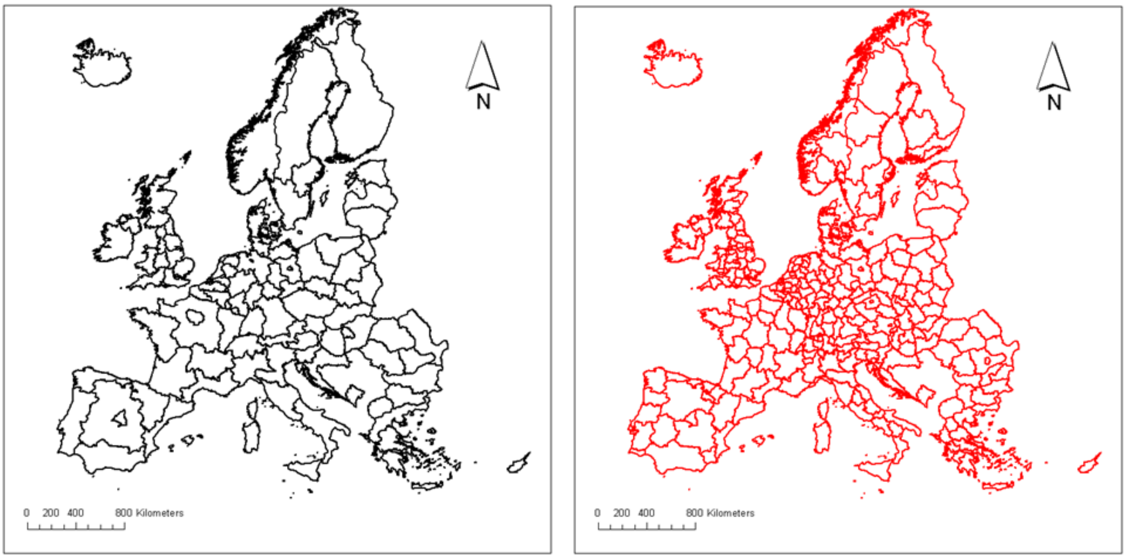


Figure 16: Visualization of Nomenclature of Territorial Units for Statistics at NUTS1 (left) and NUTS2 (right) level. NUTS2 is the lowest territorial units available in EU-SILC.

4.1.2 DB100 and URBAN

Degree of urbanisation is not only a variable to establish the relationship between EU-SILC and MTUS, but also helps to distribute the respondents spatially. DB100 is developed based on population density and defined according to cluster types into three categories (EUROSTAT n.d.):

- Densely populated area: clusters of contiguous grid cells of 1 km² with a density of at least 1 500 inhabitants per km² and a minimum population of 50 000 after gap-filling.
- Intermediate area: clusters of contiguous grid cells of 1 km² with a density of at least 300 inhabitants per km² and a minimum population of 5 000.
- Thinly populated area: grid cells outside high-density clusters and urban clusters.

The Figure 17 displays the three categories of degree of urbanisation in Southern England and Table 24 shows the relationship established between DB100 and URBAN.

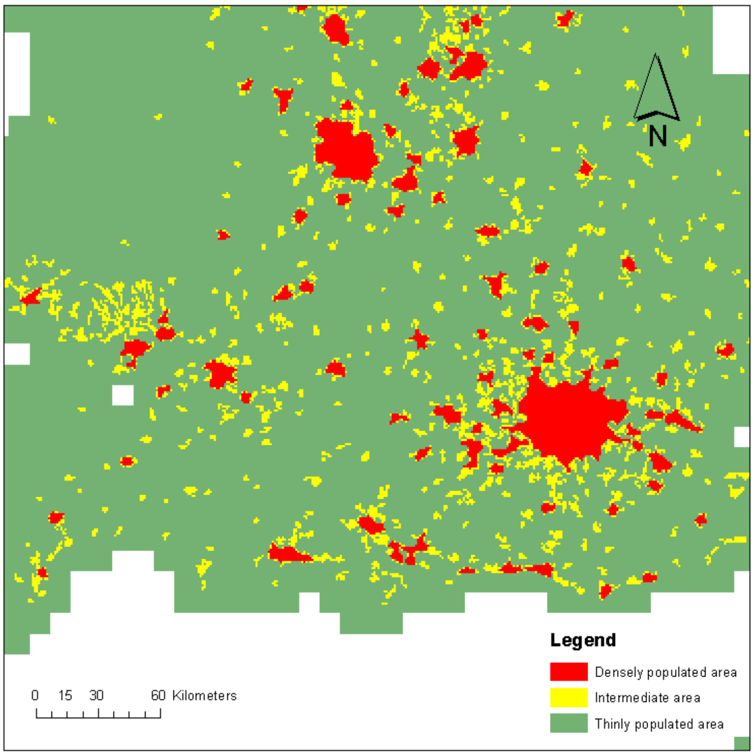


Figure 17: Three categories of degree of urbanization (densely populated area, intermediate area and thinly populated area) defined by EUROSTAT according to type of clusters, taking Southern England as an example.

4.1.3 HH021 and OWNHOME

HH021 and OWNHOME indicate whether a diarist owns or rents accommodation (Fisher et al. 2015). This variable reveals the material deprivation of a respondent to a certain extent. The relationship established between HH021 and OWNHOME is shown in Table 25.

Table 24: Relationship between DB100 and URBAN.

DB100 label	URBAN label
Densely populated area	Urban/suburban
Intermediate area	
Thinly populated area	Rural/semi-rural

Table 25: Relationship between HH021 and OWNHOME.

HH021 label	OWNHOME label
Outright owner	Own (outright or on mortgage)
Owner paying mortgage	
Tenant or subtenant paying rent at prevailing or market rate	Rent
Accommodation is rented at a reduced rate (lower price that the market price)	
Accommodation is provided free	Other arrangement

4.1.4 HS110 and VEHICLE

The car ownership is a factor that largely affects the mode of transport and time-activity pattern of a respondent; hence the car ownership is included in the modelling as an influencing factor. The relationship established between HS110 and VEHICLE is shown in Table 26.

Table 26: Relationship between HS110 and VEHICLE

HS110 label	VEHICLE label
Yes	1 car or motorcycle
	2+ cars or motorcycles
No, cannot afford/ other reason	No
	Animal only
	Non-motorised vehicle

4.1.5 PB140 and AGE

The age of the diarist is an important factor that influences their time-activity pattern. Moreover, the EFSA Comprehensive European Food Consumption Database is disaggregated per age class. This makes age an indispensable variable when applying the fusion methodology.

4.1.6 HY010 and INCOME

Income is deemed to be a factor that influences the time-activity pattern and consumption pattern of a respondent. The variable INCOME from EU-SILC records the annual household income at three levels (lowest 25%, middle 50% and highest 25%). The actual income recorded by HY010 from EU-SILC was processed accordingly into quartiles to join the MTUS data.

4.1.7 PL031 and EMPSTAT

Both variables record the current employment status of the respondents, which affects the time-activity pattern of a diarist. The relationship established between PL031 and EMPSTAT is shown in Table 27.

4.1.8 PB200 and CIVSTAT

These two variables refer to the civil status of the respondents and found to be an important variable that influences the time-activity pattern of a respondent. The relationship established between PB200 and CIVSTAT is shown in Table 28.

Table 27: Relationship between PL031 and EMPSTAT.

PL031 label	EMPSTAT label
Employee working full-time Self-employed working full-time (including family worker)	Employed/self-employed (including military service), full-time hours
Employee working part-time Self-employed working part-time (including family worker)	Employed/self-employed (including military service), part-time hours
Unemployed Pupil, student, further training, unpaid work experience In retirement or in early retirement or has given up business Permanently disabled or/and unfit to work Fulfilling domestic tasks and care responsibilities Other inactive person	Other, including unemployed, looking for work, retired, homemaker, currently attending school, currently on maternity leave and disability retirement/leave

Table 28: Relationship between PB200 and CIVSTAT.

PB200 label	CIVSTAT label
Yes, on a legal basis Yes, without a legal basis	Yes, diarist is in a couple, lives with spouse/partner
No	No, diarist not in a couple

4.1.9 PE040 and EDCAT

These two variables are recorded based on International Classification of Education (ISCED) categories. The relationship established between PE040 and EDCAT is shown in Table 29.

Table 29: Relationship between PE040 and EDCAT.

PE040 label	EDCAT label
Pre-primary education Primary education Lower secondary education	Not completed ISCED level 3
(Upper) secondary education Post-secondary non tertiary education	Completed ISCED level 3 and/or attendance at level 4
First stage of tertiary education (not leading directly to an advanced research qualification) Second stage of tertiary education (leading to an advanced research qualification)	ISCED level 5 or above

4.1.10 PL051 and ISCO1

The occupation is an important factor that probably affects respondent`s time-activity pattern. PL051 is coded according to International Standard Classification of Occupations (ISCO-88) while ISCO1 is based on

ISCO-08. The two systems share the same major groups and the two variables can be directly joined (Table 30).

Table 30: The major groups of PE051 and ISCO1.

Gourp number	Occupation
0	Armed forces occupations
1	Managers, senior officials and legislators
2	Professionals
3	Technicians and associate professionals
4	Clerical workers
5	Service and sales workers
6	Skilled agricultural, fishery, and forestry workers
7	Craft and related trades workers
8	Plant and machine operators and assemblers
9	Elementary occupations

4.2 Linking exposure data to socioeconomic factors and activities

4.2.1 Exposure data and activities

The fused EU-SILC and MTUS data inherits the format of MTUS episode data for time-activity pattern. Each row of the dataset records the time a diarist spent on a certain activity, as well as the microenvironment this activity takes place. MTUS episode data classify the activity into 69 categories and the location into nine microenvironments (variable ELOC). Additionally, the data record if the activity occurs indoors or outdoors (variable INOUT) (Table 31). The activity category and location information make it possible to determine to which stressors diarist might be exposed, as well as the parameters for exposure modelling. Exposure to air pollution (e.g. PM2.5 and NO_2) is assumed to take place for all the microenvironments; exposure to contaminants in food takes place when a diarist eats; exposure to green space mostly takes place when a person stays outdoors (especially for out-of-home leisure); for exposure to phthalates, there are multiple pathways including from cosmetics, personal care products, pharmaceuticals, medical devices and pants (see Table 32).

4.2.2 Socioeconomic data and activities

MTUS data, especially the harmonised aggregate file that contains the detailed socioeconomic information, are only available for a limited number of countries. Even for the countries for which data files are available, not all socioeconomic variables are fully recorded. The lack of data leads to a question how to fill the gaps and impute the time-activity patterns for these countries. It is also noticeable that MTUS data keep track of the time-activity pattern of the respondents only for a single day. However, time-activity patterns of diarists vary within seasons, even within a week. This raises the question how to extrapolate date of one day to larger time periods.

Evidence has shown that socioeconomic status influences people`s time-activity pattern (Ford et al. 1991, Estabrooks et al. 2003). Thus it is reasonable to use socioeconomic variables to predict these patterns. The socioeconomic information covered by the HAF file includes age, sex, household income, tenure status, car ownership, employment status, education level, civil status and occupation. It is found that some of the

socioeconomic variables are highly correlated. To avoid the multicollinearity between predictor variables and simplify the analysis, we apply factor analysis for socioeconomic data to reduce the dimension and extract the important components. Since time-activity patterns change temporally, it is necessary to include “day of week” and “month” as variables.

As described in section 4.2.1, the MTUS codes activities into a wide range of 69 categories. This provides very detailed information on time-activity pattern; however, it also increases the difficulty in data analysis. To simplify the analysis, we aggregate the 69 activities into 27 categories according to the microenvironment where the activity occurs based on variable ELOC and INOUT.

Also, we aimed to reduce the number of socioeconomic variables that we will later use to explain the difference in time-activity patterns. Table 33 shows the result of principle component analysis for the 11 original socioeconomic variables. The analysis extracts six components which explain 72.4% of the total variance. The component 1 extracts 22.1% of the total variance. Income, car ownership, employment status and education level are the dominant variables influence this component. This is followed by component 2 to 6 with variance explained of 12.8%, 11.6%, 9.4%, 8.7% and 7.9% respectively.

Since people behave differently on weekdays and at weekends, we classify the MTUS diaries into “weekday” and “weekend” according to the day of the week. Afterwards we use a random forest model to predict the time spent in each microenvironment with variable “month” along with the scores of the components extracted from the original socioeconomic variables. Since people spend the majority of their time indoors at home and at workplace, we take the model for these two microenvironments on weekdays as examples. The model will be further refined during the course of the project.

Figure 18 shows the predictor importance for the model between socioeconomic components and the time spent at home indoors (including at own home and at other people’s home) on weekdays. Component 1 is the predominant factor that influences the time a diarist spent at home. It is shown that unemployed people, with lower income, without car and with lower education level tend to spend more time at home indoors.

Table 31: Location codes of the fused EU-SILC and MTUS file, which inherit from MTUS episode file. The variable ELOC describes in which microenvironment the activity occurs and INOUT determines whether it takes place indoor or outdoor.

Location variables	Description
INOUT = 1	Inside
INOUT = 2	Outside
INOUT = 3	Travelling
ELOC = 1	At own home
ELOC = 2	At another’s home
ELOC = 3	At workplace
ELOC = 4	At school
ELOC = 5	At services or shops
ELOC = 6	At restaurant, café, bar, pub
ELOC = 7	At place of worship
ELOC = 8	Travelling
ELOC = 9	Other locations

Table 32: 69 activity categories, where the activity takes place and the stressors a diarist might be exposed to.

Activity	**Location (ELOC)**	**Inside or outside**	**Stressors**			
			Air pollutants	**Contaminants in food**	**Green space**	**Phthalates**
Imputed personal or household care	At own home	Inside	√			√
Sleep and naps	At own home	Inside	√			√
Imputed sleep	At own home	Inside	√			√
Wash, dress, care for self	At own home	Inside	√			√
Meals at work or school	At workplace	Inside	√	√		√
Meals or snacks in other places	At own home	Inside	√	√		√
Paid work-main job (not at home)	At workplace	Inside	√			√
Paid work at home (main, second or other job)	At own home	Inside	√			√
Second or other job not at home	At workplace	Inside	√			√
Unpaid work to generate household income	At own home	Inside	√			√
Travel as a part of work	Travelling	Travelling	√			√
Work breaks	At workplace	Inside	√			√
Other time at workplace	At workplace	Inside	√			√
Look for work	At own home, travelling	Inside	√			√
Regular schooling, education	At school	Inside	√			√
Homework	At own home	Inside	√			√
Leisure course or other education or training	At school	Inside	√			√
Food preparation, cooking	At own home	Inside	√			√
Set table, wash/put away dishes	At own home	Inside	√			√
Cleaning	At own home	Inside	√			√
Laundry, ironing, clothing repair	At own home	Inside	√			√
Home/vehicle maintenance/improvement, collect fuel	At own home	Inside	√			√

Activity	**Location (ELOC)**	**Inside or outside**	**Stressors**			
			Air pollutants	**Contaminants in food**	**Green space**	**Phthalates**
Other domestic work	At own home	Inside	√			√
Purchase goods	At services or shops	Inside	√			√
Consume personal care services	Other locations	Inside	√			√
Consume other services	At services or shops	Inside	√			√
Pet care (other than walk dog)	At own home	Inside	√			√
Physical or medical child care	At own home	Inside	√			√
Teach child a skill, help with homework	At own home	Inside	√			√
Read to, talk or play with child	At own home	Inside	√			√
Supervise, accompany, other child care	At own home	Inside	√			√
Adult care	At own home	Inside	√			√
Voluntary work, civic or organizational activity	At own home	Inside	√			√
Worship and religious activity	At place of worship	Inside	√			√
General out-of-home leisure	Other locations	Outside	√		√	
Attend sporting event	Other locations	Inside	√			√
Cinema, theatre, opera, concert	Other locations	Inside	√			√
Other public event, venue	At services or shops	Inside	√			√
Restaurant, café, bar, pub	At restaurant, café, bar, pub	Inside	√			√
Party, reception, social event, gambling	At other's home	Inside	√			√
Imputed time away from home	Other locations	Inside	√			√
General sport or exercise	Other locations	Outside	√		√	
Walking	Travelling	Outside	√		√	
Cycling	Travelling	Outside	√		√	

Activity	**Location (ELOC)**	**Inside or outside**	**Stressors**			
			Air pollutants	**Contaminants in food**	**Green space**	**Phthalates**
Other out-of-doors recreation	Other locations	Outside	√		√	
Gardening/forage (pick mushrooms), hunt/fish	At own home	Outside	√		√	
Walk dogs	Other location	Outside	√		√	
Receive or visit friends	At other's home	Inside	√		√	
Conversation (in person, phone)	At own home	Inside	√		√	
Games (social or solitary), Other in-home social	At own home	Inside	√			√
General indoor leisure	At own home	Inside	√			√
Artistic or musical activity	At own home	Inside	√			√
Written correspondence	At own home	Inside	√			√
Knit, crafts or hobbies	At own home	Inside	√			√
Relax, think, do nothing	At own home	Inside	√			√
Read	At own home	Inside	√			√
Listen to music, iPod, CD, audio book	At own home	Inside	√			√
Listen to radio	At own home	Inside	√			√
Watch TV, DVD, including web streamed content	At own home	Inside	√			√
Play computer games	At own home	Inside	√			√
Send e-mail, surf internet, programming, computing	At own home	Inside	√			√
No activity, recorded travel mode or change of location	Travelling	Travelling	√			
Travel to or from work	Travelling	Travelling	√			
Education-related travel	Travelling	Travelling				
Travel for voluntary/civic/religious activity	Travelling	Travelling	√		√	
Child/adult care-related travel	Travelling	Travelling	√		√	
Travel for shopping, personal or household care	Travelling	Travelling	√			√
Travelling for other purposes	Travelling	Travelling	√		√	
No recorded activity	Other locations	Inside	√			√

Table 33: Component score coefficient matrix for socioeconomic variables. Socioeconomic variables have most important influences on components are marked bold in the table.

	Component					
	1	2	3	4	5	6
Country	0.119	0.076	0.412	**-0.653**	**0.446**	0.279
Sex	0.160	0.377	0.241	**0.565**	**0.555**	-0.273
Age	0.265	**-0.617**	0.375	0.307	-0.207	0.280
Income	**-0.740**	0.045	0.112	0.022	0.029	-0.230
Tenure status	0.484	0.264	-0.190	-0.083	-0.390	-0.332
Urbanisation	-0.196	-0.250	**-0.562**	0.310	0.234	**0.434**
Car ownership	**-0.682**	-0.271	-0.019	-0.017	0.227	-0.190
Civil status	0.185	**0.676**	-0.290	0.062	0.035	**0.406**
Employment status	**0.668**	-0.026	0.371	0.194	0.070	0.037
Occupation	0.422	-0.286	**-0.538**	-0.199	0.292	-0.218
Education level	**-0.613**	0.361	0.168	0.074	-0.241	0.204

Figure 19 shows the predictor importance for the model between socioeconomic components and the time spent at workplace indoor on weekdays. Component 1 is the most important factor that influences the time people spent at work. This indicates that income, car ownership, employment status and education level can be used to estimate people`s indoor working time.

The methodology applied here only predicts time spent in each microenvironment with socioeconomic variables separately. This causes a problem that the sum of the model results of each microenvironment might not add up to a full day. Over the course of the project an improved methodology will be developed to deal with this issue.

4.3 Estimating exposure to air pollutants

As stated in section 4.1, EU-SILC data is demanded to stay anonymous and the address of respondents cannot be accessed by public. This obstructs acquisition of the information of outdoor air quality where the individuals live, which further hinders the estimation of exposure to air pollutants.

The way to compensate data scarcity is to determine the distribution of the air pollutant concentration that the respondents might live in. In EU-SILC, regions (DB040) are determined by NUTS and the lowest territorial unit available is NUTS2. As described in section 4.1.2, degree of urbanisation is classified into three categories defined by EUROSTAT as highly populated area, intermediate area and thinly populated area. Each territorial unit is further classified into three sub-regions according to the degree of urbanisation (DB100). The concentration maps generated in section 2 are in the format of raster with a resolution of 10km by 10 km^2. The concentration data are resampled into 1km by 1km^2 and clipped by the spatial coverage shapefile of each sub-region to extract the distribution of air pollutant concentration.

Here we take Spain as an example. Figure 20 shows the NUTS2 regions of Spain, which are the lowest territorial units available for this country. NUTS2 in is defined as autonomous communities and cities for Spain and the whole country consists of 19 NUTS2 regions. Figure 21 shows the degree of urbanisation defined according to clusters in the typology in Spain. The combination of DB040 and DB100 divides Spain into 57

sub-regions. We then use the spatial coverage of each sub-region to clip the concentration maps of air pollutant and extract the distribution of concentration (see Table 34 and Figure 24).

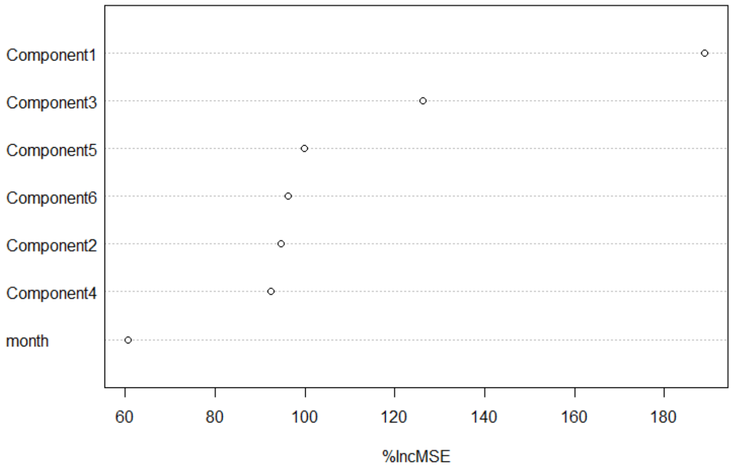


Figure 18: Random forest model of predicted time spent at home indoor on weekdays.

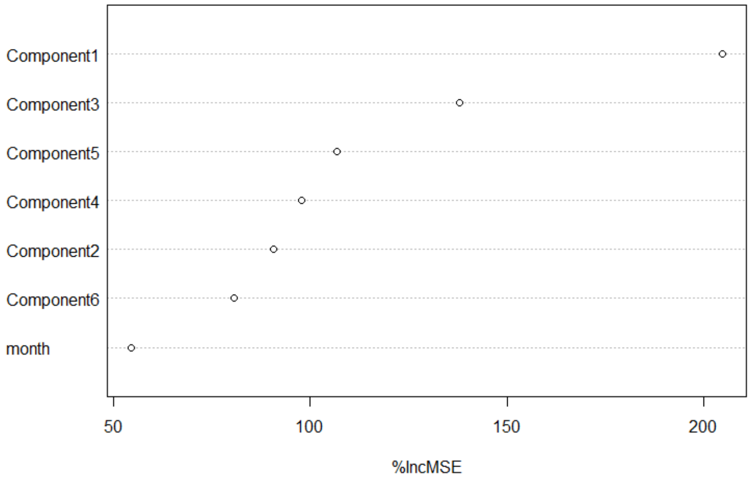


Figure 19: Random forest model of predicted time spent at workplace indoor on weekdays.

The uncertainty of the air pollutant concentration comes from two parts. The first part originates from the process of interpolation. Horálek et al. (2014) applied the leave-one-out-cross-validation, namely withholding one data point and then makes a prediction at the spatial location of that point to estimate the uncertainty and express it with relative root mean square error (RRMSE) for urban and rural interpolation respectively (Table 35). The second part derives from the inaccessibility of spatial location that the respondents live. With the method applied in the deliverable, we can only get the distribution, rather than the exact value of the concentration, which also contributes to the uncertainty in estimation of exposure to air pollutants, especially in thinly populated areas. It is also noticeable that the interpolated concentration maps only generate the background concentration. Over the course of the project, we will also apply the model to handle the urban increment and street increment of pollutant concentration.

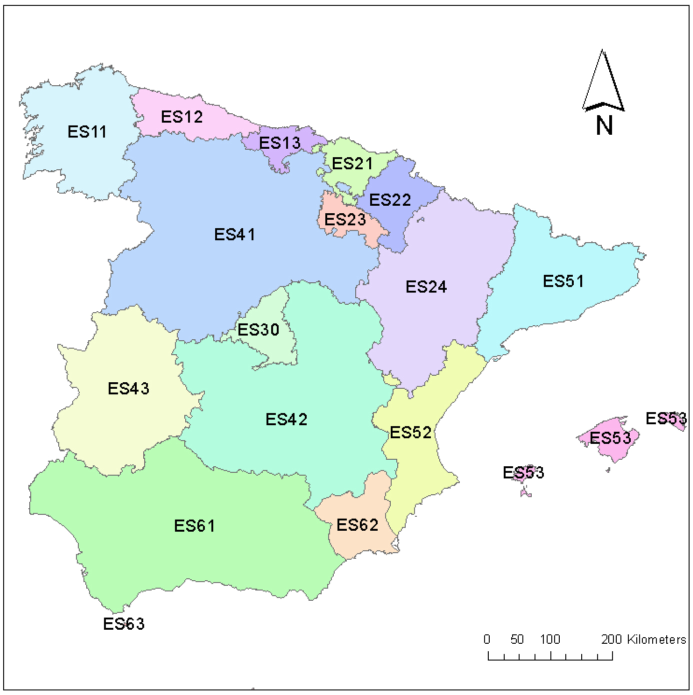


Figure 20: NUTS2 regions of Spain. Spain consists of altogether 19 NUTS2 regions. ES64 (Ciudad Autónoma de Melilla) and ES70 (Canarias) are not presented in this map due to limited space.

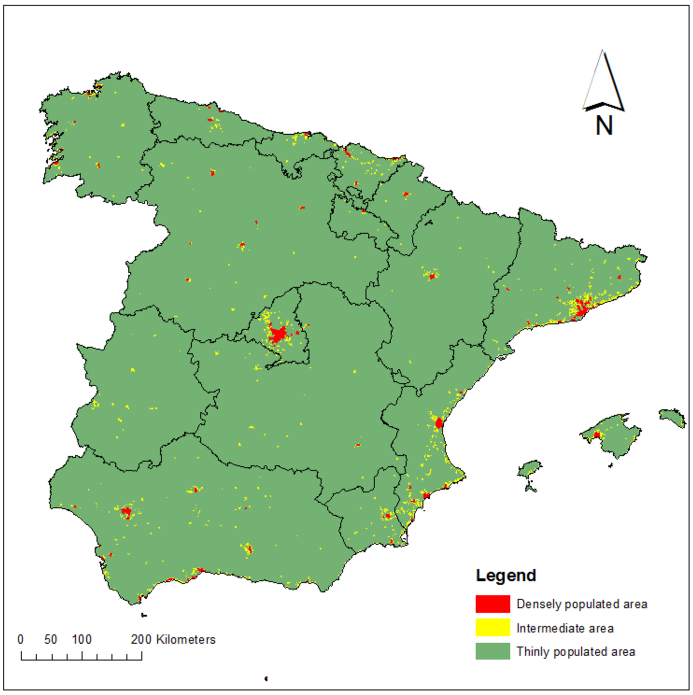


Figure 21: Degree of urbanization defined according to type of clusters in Spain. The three types of clusters in typology are defined by EUROSTAT as densely populated area, intermediated area and thinly populated area.

Table 34: Mean and standard deviation of PM10 concentration for Spain, 2012. The value is given according to the combination of NUTS regions and degree of urbanization.

NUTS ID	**Degree of urbanization**	**PM10 concentration**	
		Mean	Standard deviation
ES11	Densely populated area	12.96	1.53
ES11	Intermediate area	12.11	1.59
ES11	Thinly populated area	10.60	1.71
ES12	Densely populated area	17.85	1.08
ES12	Intermediate area	16.66	1.66
ES12	Thinly populated area	12.95	3.04
ES13	Densely populated area	19.59	1.25
ES13	Intermediate area	16.41	2.50
ES13	Thinly populated area	13.16	3.10
ES21	Densely populated area	18.33	1.73
ES21	Intermediate area	17.43	1.58
ES21	Thinly populated area	15.80	2.12
ES22	Densely populated area	15.70	0.79
ES22	Intermediate area	16.09	2.10
ES22	Thinly populated area	15.00	2.19
ES23	Densely populated area	16.76	0.06
ES23	Intermediate area	16.16	0.85
ES23	Thinly populated area	13.31	2.72
ES24	Densely populated area	18.28	0.62
ES24	Intermediate area	16.36	1.89
ES24	Thinly populated area	13.95	2.72
ES30	Densely populated area	20.55	1.27
ES30	Intermediate area	18.88	2.71
ES30	Thinly populated area	17.50	3.20
ES41	Densely populated area	11.33	0.89
ES41	Intermediate area	11.33	1.38
ES41	Thinly populated area	10.71	1.89
ES42	Densely populated area	17.62	1.25
ES42	Intermediate area	17.88	1.58

NUTS ID	**Degree of urbanization**	**PM10 concentration**	
ES42	Thinly populated area	15.40	2.35
ES43	Intermediate area	15.47	1.02
ES43	Thinly populated area	14.69	1.48
ES51	Densely populated area	20.26	2.07
ES51	Intermediate area	18.74	2.12
ES51	Thinly populated area	16.05	3.35
ES52	Densely populated area	20.38	1.70
ES52	Intermediate area	19.86	1.83
ES52	Thinly populated area	17.63	3.03
ES53	Densely populated area	18.24	0.07
ES53	Intermediate area	18.29	0.90
ES53	Thinly populated area	18.43	0.98
ES61	Densely populated area	21.72	3.97
ES61	Intermediate area	22.69	6.36
ES61	Thinly populated area	19.79	6.06
ES62	Densely populated area	24.13	0.62
ES62	Intermediate area	23.96	1.90
ES62	Thinly populated area	21.42	3.76
ES63	Densely populated area	33.76	0.00
ES63	Intermediate area	34.53	1.09
ES63	Thinly populated area	34.53	1.09
ES64	Densely populated area	25.74	0.09
ES64	Intermediate area	25.74	0.09
ES64	Thinly populated area	25.74	0.09

4.4 Data and model fusion to estimate air pollution burden

Aerosol Optical Depth (AOD) derived from Earth Observation (EO) data represents the starting point of the data fusion methodology applied to derive PM concentration at the ground level. The methodology consists of a tiered approach composed of two consecutive steps. The first step, namely *data fusion*, consists in the fusion between ground-based data concentration of PM and gaseous chemicals as measured by the regulatory monitoring network and/or by ad hoc measurements campaigns properly processed within a chemical transformation model to account for secondary aerosol formation, the extinction coefficient reckoned from EO-derived aerosol optical depth, the mixing layer height and the relative humidity as calculated from meteorological model MM5 (Grell et al., 1994).

Earth Observation data in the visible spectrum (i.e. 550 nm) are used to calculate the optical depth of atmospheric aerosol. However, the optical atmosphere effects of pollution on high spatial resolution EO are more pronounced in certain spectral bands than in other: according to Mie theory (Mie, 1908) highly scattering particles are those with diameters of the order of 0.5 to 2 times the wavelength observed by the sensors. Therefore, the AOD calculated delivers information mainly about aerosols with particle diameter of 0.2 - 1.0 µm. In order to relate the satellite-derived extinction coefficient of fine particles with the ground-measured concentrations of PM_{10}, due account has to be taken of the effect of relative humidity in the lower troposphere. A change in relative humidity leads to an increase of particle radius, because water vapor may be absorbed and condensed on hydrophilic aerosols when atmospheric relative humidity increases. After a certain inflection point (determined experimentally) the effect of the relative humidity on the extinction coefficient σ_e can be approximately represented as an exponential (Day and Malm, 2001; Grant et al.,1999).

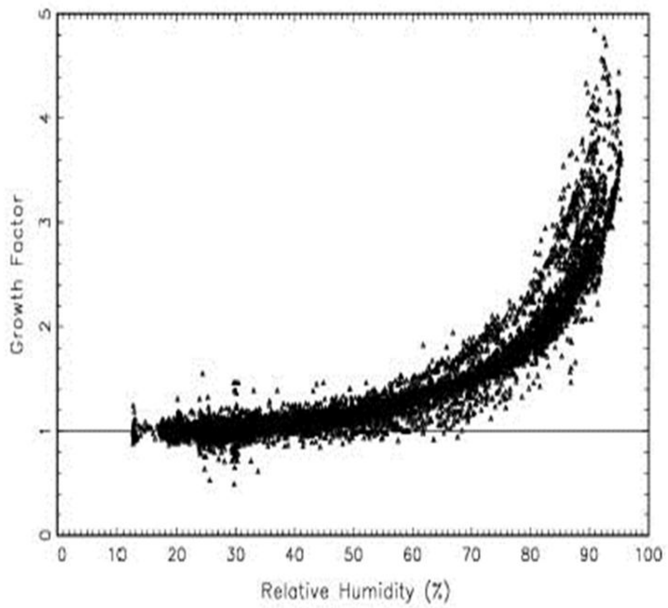


Figure 22: Particle growth as a function of relative humidity (Source of Day and Malm, 2001)

The physicochemical model that was found to best reflect the dependence of PM10 concentration on the ground with the scattering extinction coefficient and relative humidity is:

$$C_{PM10} = a \cdot \sigma_e + b \cdot RH \qquad \text{for RH} \leqslant \text{RH}_o$$

$$C_{PM10} = a' \cdot \sigma_e + b' \cdot e^{K \cdot RH} \qquad \text{for RH} > \text{RH}_o$$

where value of the inflection point for RH, RHo, depends on the prevalent meteorological conditions (in terms of atmospheric humidity and ground-level air temperature) and the coefficients a, b, a' and b' depend on the relative ratio of fine to coarse particles. K is the kinetic constant of fine particle growth due to atmospheric humidity.

In the second step of the data assimilation methodology, the concentration field obtained from the data fusion is used to calculate in a second step, namely “model fusion”, a further refined concentration field. This step is based on the integration of the concentration values resulting from atmospheric chemical transport model with the results derived from the data fusion. The aim of the model fusion process is to

merge on cell-by-cell basis these two classes of information in order to obtain an improved final concentration field that presents the lowest residual error.

Model fusion will be done making use of a Kalman filter (Durrant-Whyte, 2001; Kalman,1960) which is designed to account for the relative error introduced by each of the two computation models. The Kalman gain, K, is estimated as a function of the standard deviation of the PM_{10} estimates from chemical transport modeling (Z_1) and satellite data processing (Z_2) using equations as follows:

$$K = \frac{\sigma_{Z_1}^2}{\sigma_{Z_1}^2 + \sigma_{Z_2}^2}$$

Where σ_{z1} and σ_{z2} are the variances on cell by cell basis of Z1 and Z2 that is the values of PM concentration on each single cell respectively derived from chemical transport model and from EO data

The final PM10 concentration is calculated as

$$C^K = Z_1 + K(Z_2 - Z_1)$$

This technique ensures that data singularities generated from point disturbances of the optical parameters field captured by the satellite sensor are eliminated. At the same time, atmospheric model weaknesses, such as the high dependence of result accuracy on the adequacy of the emissions inventory and boundary/initial values can be dealt with via fusion with the satellite-derived estimates (Sarigiannis et al., 2002, 2004).

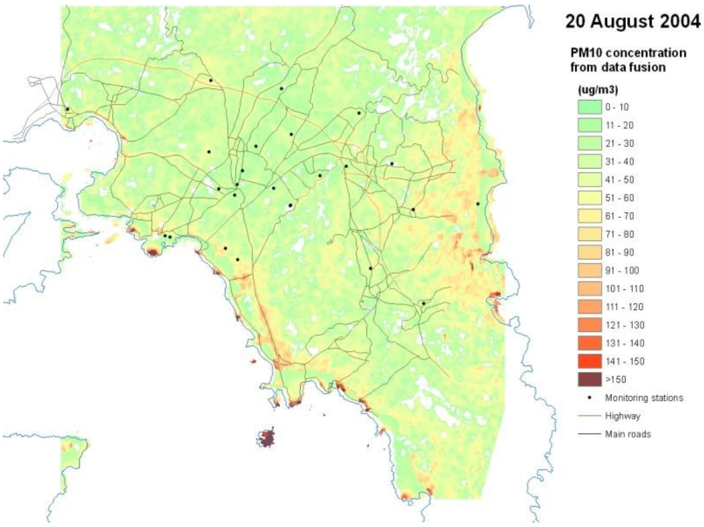


Figure 23: PM0 concentration field at ground level derived from SPOT V sensor (spatial resolution 10 by 10 meters) over Athens through data fusion

4.5 Estimating exposure to contaminants in food

Dietary exposure to chemicals can be assessed at individual level by multiplying the average daily consumption for each food item with the corresponding average chemical content (LB, MB and UB), summing up the respective intakes throughout the diet and finally dividing the results by the individual's body weight.

Dietary intakes is calculated using the following formula

$$I_j = \sum_{k=1}^{n} C_k \cdot L_{k,\,j}$$

Where $\boldsymbol{I_j}$ is the dietary intake j, n is the number of foods in the diet, $\boldsymbol{C_k}$ is the consumption of food k, $\boldsymbol{L_{k,j}}$ is the concentration level of contaminant j in food k.

Dietary exposure to each contaminant of interest is calculated individually using the following formula:

$$E_j = \frac{\sum_{k=1}^{n} C_k \cdot L_{k,j}}{BW}$$

Where $\boldsymbol{E_j}$ is dietary exposure to contaminant j and ***BW*** is the body weight.

Bodyweight used are: 75 kg – Adult male; 65 kg -Adult female; 36 kg – Boys (aged between 3-17); 35 kg – Girls (aged between 3-17).

4.5.1 Food contamination data

Food contamination data are available at European (EFSA) or at WHO European Region (WHO/Gems) aggregation level. No further spatial disaggregation is available.

From past EFSA reports we can collect data on food contamination (average values among all the samples collected at European level) and on exposure to some contaminants through food: the latter at country level. Variation in exposure between countries is influenced by different consumption patterns only. Since chemical concentrations in food categories were calculated at a European level.

Detailed literature review will be carried out to identify possible studies reporting food contamination in specific European regions or countries[5]. For these countries this data will be used after data quality assessment.

In this regard a major source of information on food contamination is represented by the various National Total Diet Studies (TDS). The World Health Organization (WHO) supports Total Diet Studies (TDSs) as one of the most cost-effective means of assuring that people are not exposed to unsafe levels of toxic chemicals through food. WHO and FAO promoted and supported TDSs to assess several contaminants in the diet and many countries endorsed the TDS.

So far approximately 15 countries in Europe had been engaged in the process of performing TDS or TDS-like studies including Belgium, Czech Republic, Denmark, Finland, France, Ireland, Italy, Latvia, Poland, Portugal, Slovak Republic, Spain (Basque Country and Catalonia), Sweden, the Netherlands and the United Kingdom.

Table 35: Uncertainty analysis for interpolated concentration maps of PM10 from EEA for 2005 to 2012. The uncertainty is expressed by relative root mean square error (RRMSE) of leave-one-out-cross-validation for rural and urban areas respectively.

RRMSE (%)	2005	2006	2007	2008	2009	2010	2011	2012
Rural areas	26.0	26.6	23.5	27.2	23.9	22.7	21.1	24.5
Urban areas	20.0	20.9	18.4	22.4	23.0	22.5	20.7	24.5

[5] In some countries contamination data are available (e.g. Catalunia RIBERFOOD)

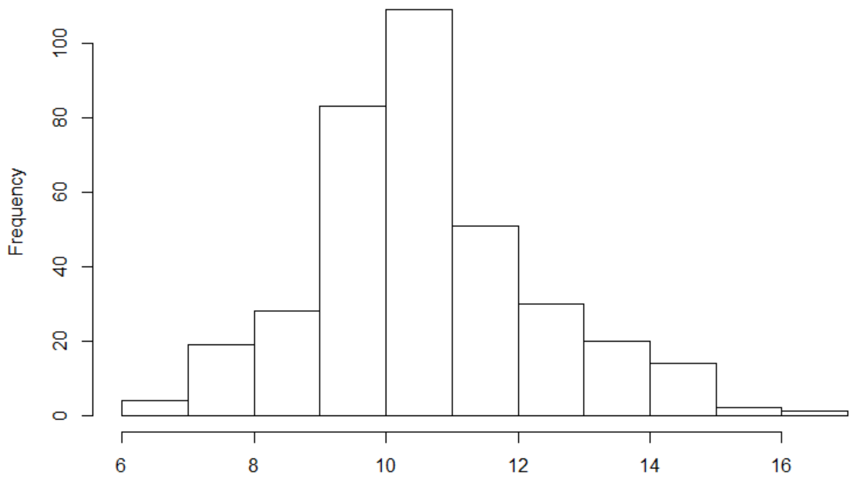


Figure 24: Distribution of PM10 concentration [µg/m^3] for the year 2012 of region ES11, thinly populated area. The uncertainty of estimating exposure partly originates from the limitation of accessing the exact address where the respondents live. We compensate the uncertainty by determining the distribution of pollutant concentration in each sub-region combined with the degree of urbanization.

4.5.2 Food consumption

The consumption data will be derived from the EFSA Comprehensive European Food Consumption Database (Comprehensive database) which was built in 2010 from existing national information on food consumption at the individual level (EFSA, 2011; Huybrechts *et al*., 2011; Merten *et al*., 2011).

Data collected through food frequency questionnaires in past population surveys (when available) and in EXHES will be compared with the EFSA Comprehensive European Food Consumption Database for the same country and used to derive food consumption in relation with other covariates such as SES and location (e.g. seaside vs. mountain and urban vs. rural) so as to derive food consumption patterns which better reflect different population groups in Europe. EFSA considers food consumption as the main driver for the estimation of exposure through food.

One key aspect about dietary exposure assessment is the food classification system. Indeed, the selection of food items to be analysed is based on the information available in existing consumption datasets, often on national level. To assess dietary exposure, an harmonized food classification system is needed to link existing food consumption data with the food contamination analytical data. In Europe, there is a need for a harmonized approach, including a harmonised exposure assessment, to make comparison between countries possible. FoodEx-1 is the suggested food classification system recommended by the European Food Safety Authority (EFSA) as a classification system on pan-European level for its use for exposure assessment.

4.5.3 Exposure assessment from biomonitoring data

Verification of exposure estimates may also be based upon exposure reconstruction from biomonitoring data. Although exposure to the compounds selected occurs through various exposure pathways, diet is expected to be the dominant exposure pathway due to the bioaccumulation and biomagnification (across food web transfer) potential of the persistent compounds. To this aim the INTEGRA computational platform

developed by AUTH could accommodate the assessment of exposure through a reverse dosimetry approach.

4.5.4 Major challenges

It has to be underlined that detailed raw data (especially of food contamination) are rarely freely available as they are owned by the Member States. The HEALS team is in contact with EFSA to make the effort to collect them. Both food consumption and contamination data are available at different levels of detail in different European countries (uncertainty is not always included). Food consumption data are not always disaggregated for socioeconomic status and only in some cases they are spatially resolved within a country. Also, food consumption and contamination data are not always available for the remote past. As previously mentioned a further challenge arises from combining the data due to the use of non-uniform food classification coding systems.

5 Preliminary results of the probabilistic framework

5.1 Air pollution

The following section presents a preliminary assessment of the part of the external exposome that deals with exposure to air pollutants (hereafter only PM2.5 is shown). To show the results of applying the methodology described before we consider the different microenvironments a certain socioeconomic group or subgroup will typically enter and leave during the course of a weekday. The variation of external exposure can be mainly explained due to the combined variation in outdoor air concentration and conditions pertinent in the indoor environments (at work, at home, in shopping areas, etc.).

For this example the model is initialized with the EU-SILC and MTUS data (described in sections 3.1.1 and 3.1.2) that were combined using the methodology described in section 4. This is utilized to determine the time-activity patterns of socioeconomic groups. However, also individual sensor-based data can be used to determine the time-use patterns, as well as time-location patterns from agent-based modelling.

Additionally, we feed the (uncertain) outdoor air concentration maps into the model. The maps were generated as described in section 3.2.1. Parameters that determine indoor air quality (i.e. penetration factor, air exchange rate, decay rate etc.; see section 3.2.1.4 for details) are also fed into the model.

In a next step a certain socioeconomic group is selected and the random variables (i.e. the framework parameters) are sampled from their respective distributions in a Monte-Carlo-like fashion. As a result we are able to determine the probability distributions of the exposure to air pollutants via inhalation for a certain socioeconomic group.

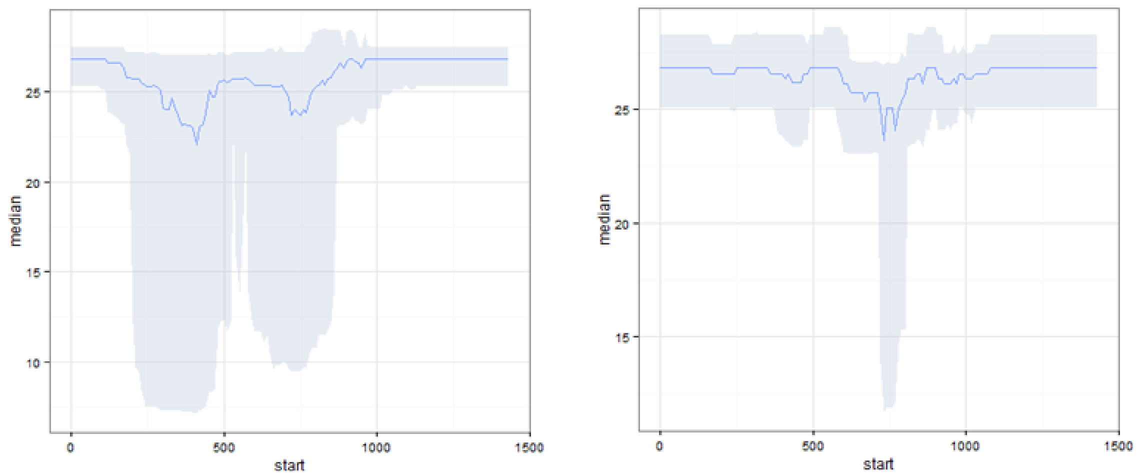


Figure 25: Median and +/- one standard deviation of PM2.5 concentration (y-axis, in µg/m3) over the course of a day (x-axis, in minute of the day). The figure shows the typical exposure of a weekday in summer of women aged 25-34 of (a) high-education and high-income on the left and (b) low-income and low-education on the right.

Consider the example shown in Figure 25 depicting the median (dark blue line) and standard deviation (surrounding blue shade) of the exposure to PM2.5 of a certain socioeconomic group due to the change of microenvironments over the course of the day. The figure shows a typical weekday in summer (June-August) of women aged 25 to 34 years exposed to the air pollution conditions of the year 2003. The left part of the figure shows women of high educational level (tertiary education) and high income (within the highest quartile of household income at country level), whereas the right figure shows women of the same age

range that have attained only a low level of education (lower secondary education) and earn low income (within the lower quartile of household incomes at country level). Clearly, there is variance in the result due to the large group of people selected (involving many countries) and due to the uncertainty about the outdoor concentrations and microenvironment specific parameters but we can spot an obvious difference between the two socioeconomic groups, namely the two peaks on the left and the single peak on the right.

We can identify the reasons for this by having a closer look in the difference of the respective time-activity patterns of the socioeconomic groups. When grouping the activities by microenvironment and plotting the mean time spent within such a microenvironment for the groups mentioned before (women aged 25-34 years) in the same period (June-August) we can plot Figure 26. It is easy to spot that the exposure reduction peaks shown in Figure 25 closely resemble the variation that stems from moving between microenvironments: Clearly, the group of lower income/education people spent on average more time at their own home (depicted in ochre) showing less overall variation whereas within the high-income/education group on average a considerable amount of time is spent at the workplace. Also, the overall amount of travelling (which usually is associated with high levels of exposure) is higher in the high-income/education group. From a modelling perspective there is a generally lower penetration factor and a generally higher air exchange rate associated with typical indoor work environments (a=1.8±1.2 h^{-1} and p=0.75±0.25) compared to typical parameters for home environments (a=0.8±0.4 h^{-1} and p=1.0±0.06). This may be explained by the prevalence of ventilation systems, air conditioning systems and stronger air-tightness in indoor work environments. However, uncertainty is considerably high for these parameters and a very limited number of measurements exist (cf. section 3.2.1.4).

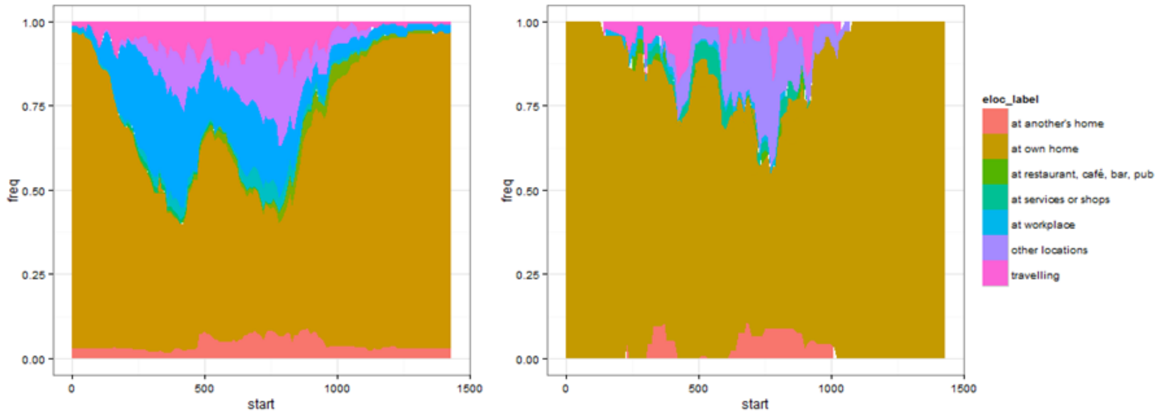


Figure 26: Share of the mean time spent in a certain microenvironment over the course of a typical weekday in summer (June-August) for women aged 25-34 years, whereas the figure on the left shows (a) the high-education and high-income group and (b) the low-education low-income group.

However, there is also a considerable influence of other factors to exposure that is considered in the model. Consider for instance the case of men aged 45 to 54 in Spain. We can, similar to the example given before, distinguish between the two socioeconomic subgroups of low-income/education and high-income/education. The respective exposure levels over the course of a weekday is given in Figure 27.

Comparing the exposure of these two groups the difference is much less apparent than between the former example. The mean time members of each of both groups spend in a certain microenvironment is depicted in Figure 28. One can see that even though low-income/education men spend less time at work compared to their high-income/education counterparts, the difference is less evident than for women. Also, remember that we only selected the both subgroups for Spain during summer. Here, the model as-

sumes different conditions for the estimate of indoor concentration levels involving higher air exchange rates at home environments (a=1.2±1.0 h^{-1}).

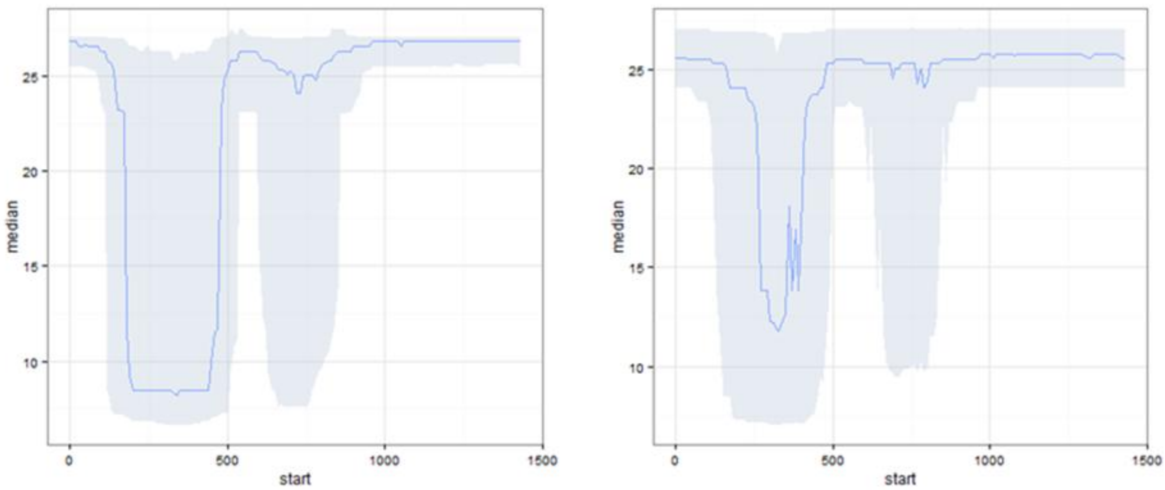


Figure 27: Median and +/- one standard deviation of PM2.5 concentration (y-axis, in µg/m3) over the course of a day (x-axis, in minute of the day). The figure shows the typical exposure of a weekday of men in Spain aged 45-54 of (a) high-education and high-income on the left and (b) low-income and low-education on the right.

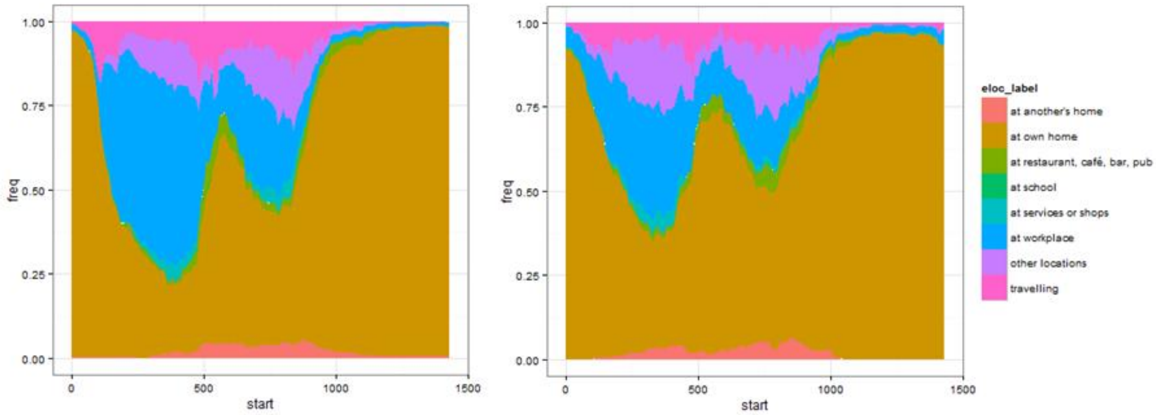


Figure 28: Share of the mean time spent in a certain microenvironment over the course of a typical weekday for men in Spain aged 45-54 years, whereas the figure on the left shows (a) the high-education and high-income group and (b) the low-education low-income group on the right.

5.2 Food intake

For several chemicals including trace metals and POPs food is the most important exposure pathway being responsible for up to 90% of the total body burden. HM and POPs health end points cover a wide range of diseases ranging from neurological, neurodevelopmental disorders, cardiovascular, effects on immune and endocrine systems, impair reproductive function, to renal failure.

To show the results of the applied methodology to estimate exposure to trace metals and POPs through diet we considered France as a case study. The first French Total Diet Study was undertaken between 2000 and 2004 by the French National Institute for Agricultural Research (INRA), in collaboration with the French Food Safety Agency (AFSSA). This led to a comprehensive appraisal of the population’s exposure, including

adults and children, to inorganic contaminants and minerals. In 2006, AFSSA undertook a second TDS, which included a larger number of target substances. The survey included all of the substances which had already been analysed in the first study, in order to determine trends in the monitoring of the levels to which the population was exposed. Other substances were added to the list, either to update knowledge of them or more simply to fill a gap in French and international data. In addition to covering a wider range of substances (more than 400 as against 34 in the first survey), this new study covered all the administrative regions in mainland France (as opposed to 3 major cities for TDS 1).

Eighteen inorganic and mineral contaminants, which had already been screened for in the first study (Leblanc et al. 2005), were analysed: Arsenic (As), Lead (Pb), Cadmium (Cd), Aluminium (Al), Mercury (Hg), Antimony (Sb), Chromium (Cr), Calcium (Ca), Manganese (Mn), Magnesium (Mg), Nickel (Ni), Copper (Cu), Zinc (Zn), Lithium (Li), Sodium (Na), Molybdenum (Mo), Cobalt (Co), Selenium (Se). Ten more elements were added: Silver (Ag), Barium (Ba), Tin (Sn), Gallium (Ga), Germanium (Ge), Strontium (Sr), Tellurium (Te), Vanadium (V), Potassium (K), Iron (Fe).

Given their toxicity and the lack of national contamination and exposure data, persistent organic pollutants (POPs) were analysed:

- 17 congeners of polychlorodibenzo-p-dioxins and polychlorinated dibenzofurans (PCDD/F): TCDD-2378, PCDD-12378, HCDD-123478, HCDD-123678, HCDD-123789, HCDD-1234678, OCDD, TCDF-2378, PCDF-12378, PCDF-23478, HCDF-123478, HCDF-123678, HCDF-234678, HCDF-123789, HCDF-1234678, HCDF-1234789, OCDF;
- 12 congeners of dioxin-like polychlorinated biphenyls (DL-PCBs): PCB-77, 81, 126, 169, 105, 114, 118, 123, 156, 157,167, 189, and 6 congeners of non-dioxin-like polychlorinated biphenyls (NDL-PCBs) (PCB-28, 52, 101, 138, 153,180);
- 16 perfluorinated compounds: the carboxylates (PFOA, PFBA, PFPA, PFHxA, PFHpA, PFNA, PFDA, PFUnA, PFDoA, PFTrDA, PFTeDA) and the sulfonates (PFOS, PFBS, PFHxS, PFHpS, PFDS);
- Brominated Flame Retardants (BFRs): 8 polybrominated diphenyl ether congeners (BDE-28, 47, 99, 100, 153, 154, 183, 209), 3 polybrominated biphenyl congeners (BB-52, 101, 153), and 3 hexabromocyclododecane congeners (HBCD-alpha, beta, gamma).

The samples to be analyzed were purchased between June 2007 and January 2009. In all, 19,830 products were purchased and prepared 'as consumed'. For example, fruit and vegetables were washed. Vegetables (except for vegetables consumed raw), meat and seafood were cooked: braised, pan-fried, grilled, baked, deep-fried, etc., depending on the practices declared by the participants in INCA 2.

Values below the limits of detection or quantification are referred to as censored data. Censored data were processed according to the World Health Organization (WHO) recommendations (GEMS-Food Euro 1995).

For items with a censoring rate of less than 60%, the censored data were replaced by an estimate corresponding to a median or middlebound (MB) assumption: concentrations below the LOD (non-detected substances) were replaced by 0.5 LOD, and concentrations below the LOQ but above the LOD (called traces) were replaced by 0.5 LOQ.

For items with a censoring rate of at least 60%, two assumptions were made about concentrations: the low or lowerbound (LB) assumption, and the high or upperbound assumption (UB). The lowerbound assumption corresponds to a scenario in which non-detected values are estimated to be 0 and the values detected, but not quantified, are estimated to be equal to the LOD. The upperbound assumption corresponds to a scenario in which non-detected values are estimated to be equal to the LOD and the values detected but not quantified are estimated to be equal to the LOQ. The LB scenario represents the minimum possible value, and the UB scenario represents the maximum possible value.

Data on food consumption were gathered from the INCA 2 food consumption study. INCA 2 was carried out by AFSSA between 2006 and 2007 and included 4,079 individual subjects (2,624 adults aged from 18 to 79 years and 1,455 children aged from 3 to 17 years). France was divided into eight major regions (Table 36) covering all the administrative regions as follows: (i) West (Bretagne, Pays de Loire, Poitou-Charentes), (ii) North-West (Basse Normandie, Haute Normandie, Nord-Pas de Calais, Picardie), (iii) Paris Region (Ile de France), (iv) East (Champagne Ardennes, Lorraine, Alsace), (v) Centre-East (Franche-Comté, Rhône Alpes), (vi) South-East (Provence Alpes Côte d'Azur, Languedoc Roussillon), (vii) South-West (Midi-Pyrénées, Aquitaine) and (viii) Centre (Centre, Bourgogne, Limousin, Auvergne).

Table 36: Breakdown of participants in the INCA 2 study by region

Region	Adults	Children
North-West	280	246
East	210	148
Paris Region	255	243
West	295	209
Centre	205	141
Centre-East	229	164
South-West	109	141
South-East	253	152

The field-study phase of the food consumption survey was carried out over a period of 11 months in order to include seasonal variations in eating habits. The individuals recruited kept a record of all they consumed over seven consecutive days and also completed a questionnaire covering anthropometric and socio-economic factors.

Results of exposure to Arsenic, Cadmium, Chromium and Lead through diet are reported in **Errore. L'origine riferimento non è stata trovata.**, Figure 30 and Figure 31 for different genders and age groups, educational status and geographic location. Dietary exposure to trace metals and POPs for toddlers and other children is higher compared to adults; this is explained by the higher food consumption in relation to their body weight. Analysis of exposure in relation with educational status shows higher education adults have slightly higher exposure levels to Arsenic, Chromium and Lead mainly due the higher sea food consumption. The same patterns is not clearly visible with regard to children who show more constant exposure levels in relation to the educational status of their parents. No clear trends are visible with regard to the geographic location.

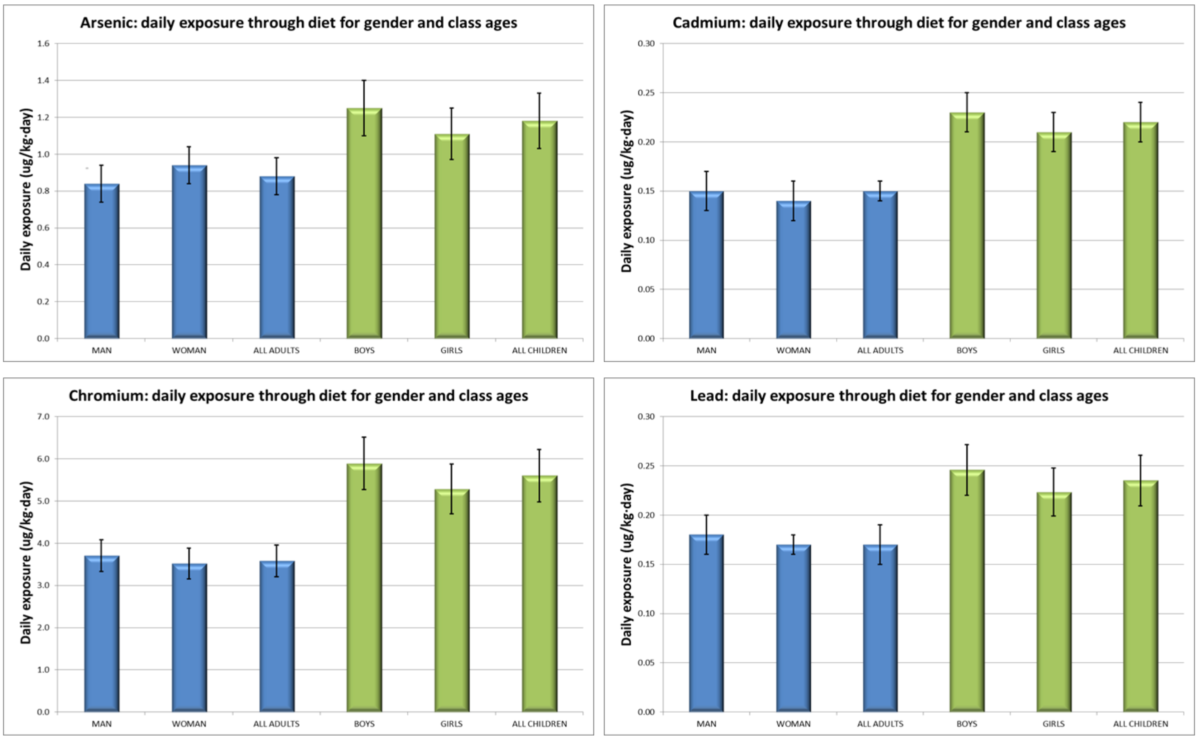


Figure 29: The daily intake (µg/kgBW day) through diet stratified by gender and age class. Data of four heavy metals are shown: Arsenic, Cadmium, Chromium and Lead.

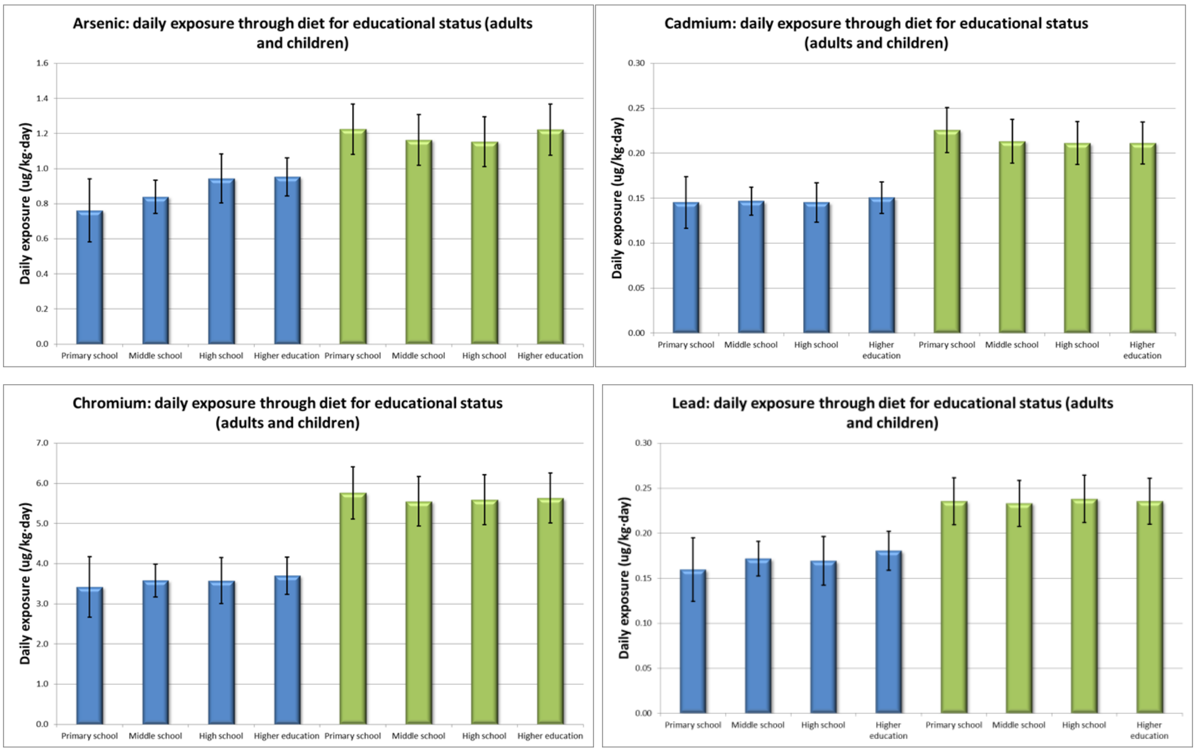


Figure 30: The daily intake (µg/kgBW day) through diet stratified by educational status and age class. Data of four heavy metals are shown: Arsenic, Cadmium, Chromium and Lead.

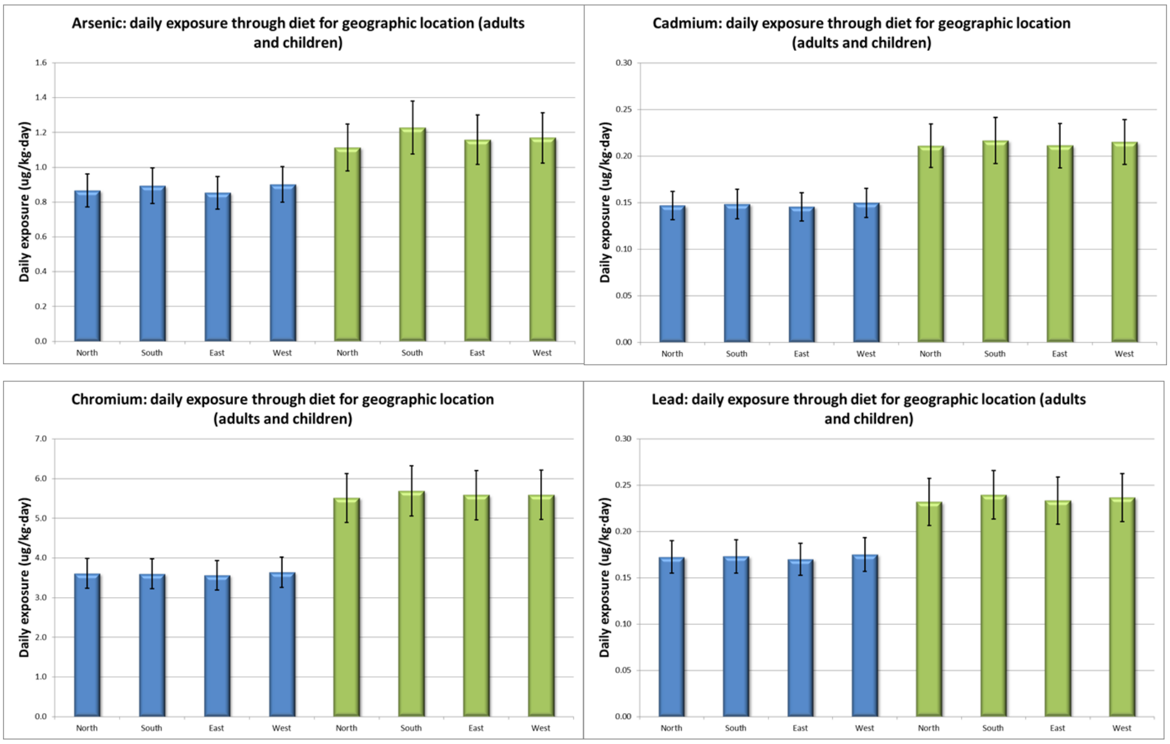


Figure 31: The daily intake (µg/kgBW day) through diet stratified by geographical location and age class. Data of four heavy metals are shown: Arsenic, Cadmium, Chromium and Lead.

Data on food contamination and food consumption disaggregated by age groups and gender are generally available at EU level from literature and freely accessible databases. In general, food consumption patterns are the main drivers for the inter- and intra-individual exposome. Some food items (e.g. fish and seafood) affect significantly the overall dietary exposome of some contaminants (e.g. As, Cd, Hg, Pb, PCDD/Fs+DL-PCBs, 6 NDL-PCBs).

6 Conclusion and outlook

6.1 Towards the application to population studies and EXHES

As mentioned before the aim of the probabilistic framework is to assess the external exposome in a general manner under several different settings and context to foster its reusability. This will be done by using as many individual-level data as possible and gap-fill data from observational datasets and modelled data wherever needed. We use agent-based modelling to allow for a more detailed estimate of the gap-filling methodology while making sure that uncertainty and error propagation is accounted for in a proper manner.

Within HEALS these settings include the individual level assessment, e.g. on data from the pilot and within the scope of EXHES, as well as population-wide assessments by applying it on the case studies of WP14 through WP16. Applying the methodology on the latter will provide quantitative estimates of external exposure for the population and subgroups thereof which can be derived based on their respective socioeconomic characteristics and their vulnerability.

HEALS integrate multiple pre-existing cohort studies from France, Italy, Croatia, Slovenia, Spain, Portugal, United Kingdom, Sweden and Poland. These population studies will allow for a further refinement of the probabilistic modelling framework and serve as a case study for the EXHES pilot study in WP17. Then detailed exposure estimates at quasi-individual level will be available by applying the agent-based modelling methodology developed within Stream 3.

The overall exercise of applying the methodology on the rich synthesized datasets will allow to draw conclusions on the association between environment, health and (epi-)genetic susceptibility at the population level starting from the individual.

6.2 Future improvements of the framework within HEALS

Incorporating the final results of the pilot study of individual sensor-based data collection and fusing them with population-specific features partly described in this report will make them usable to estimate the exposure of population (sub-)groups. This will support the data synthesis described in this report and allows a prospective characterization of the exposome.

We have shown that time-activity patterns derived at an individual scale are difficult to use for extrapolation to broader groups of people when used as such. However, HEALS' novel approach to incorporate agent-based modelling to support the estimation of quasi-individual time-activity patterns has proven useful and will be further developed and incorporated in the methodology described in this report. Gaps in censuses, time-use datasets and other surveys will be filled using this methodology. This will eventually inform the division of exposure across the different socioeconomic groups.

Additionally, we can utilize the upcoming data from EXHES to directly survey the singletons and twins cohort's parents for socioeconomic status at the household level. This will act as a promising means of validation for modelled estimates. A report on the application of the exposure modelling framework to population studies covered in Stream 5 will be delivered one year from the finalisation of this document (D11.2 in month 42).

7 References

Atkinson, AA & Marlier, E (eds) 2010, *Income and living conditions in Europe.* 2010. Available from: http://ec.europa.eu/eurostat/documents/3217494/5722557/KS-31-10-555-EN.PDF/e8c0a679-be01-461c-a08b-7eb08a272767. [17 February 2016].

Becker, K, Göen, T, Seiwert, M, Conrad, A, Pick-Fuss, H, Müller, J, Wittassek, M, Schulz, C & Kolossa-Gehring, M 2009, GerES IV: phthalate metabolites and bisphenol A in urine of German children, *International Journal of Hygiene and Environmental Health*, vol.212, no.6, pp.685-692

Bekö, G, Callesen, M, Weschler, CJ, Toftum, J, Langer, S, Sigsgaard, T, Høst, A, Kold Jensen, T & Clausen, G 2015, Phthalate exposure through different pathways and allergic sensitization in preschool children with asthma, allergic rhinoconjunctivitis and atopic dermatitis. *Environmental Research*, vol.137, pp.432-439.

Bergh, C, Torgrip, R, Emenius, G & Ostman C 2011, Organophosphate and phthalate esters in air and settled dust - a multi-location indoor study, *Indoor Air,* vol. 21, no.1, pp.67-76.

Bertelsen, RJ, Carlsen, KC, Calafat, AM, Hoppin, JA, Håland, G, Mowinckel, P, Carlsen, KH & Løvik, M 2013, Urinary biomarkers for phthalates associated with asthma in Norwegian children. *Environmental Health Perspectives.*vol.121, no.2, pp.251-256.

Casas, L, Fernández, MF, Llop, S, Guxens, M, Ballester, F, Olea, N, Irurzun, MB, Rodríguez, LS, Riaño, I, Tardón, A, Vrijheid, M, Calafat, AM, Sunyer, J & INMA Project, *Environment International,* vol. 37, no. 5, pp. 858-866.

Channan, S, Collins, K. & Emanuel, WR 2014, *Global mosaics of the standard MODIS land cover type data*. Available from: http://glcf.umd.edu/data/lc. [17 May 2015]

Chen, C & Zhao, B 2010, Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor*. Atmospheric Environment,* vol. 45, no.2, pp. 275-288.

Chu DA, Kaufman YJ, Zibordi G, Chern JD, Mao J, Li CC, Holben BN. Global Monitoring of Air Pollution over Land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS). Journal of Geophysical Research: Atmospheres 2003; 108(D21): 4661; doi: 10.1029/2002jd003179.

CLRTAP n.d., *EMEP MSC-W modelled air concentrations and depositions*. Available from: http://webdab.emep.int/Unified_Model_Results/. [12 March 2016].

Dacunto, PJ, Klepeis, NE, Cheng, KC, Acevedo-Bolton, V, Jiang, RT, Repace, JL, Ott, WR & Hildemann LM 2015, Determining PM2.5 calibration curves for a low-cost particle monitor: common indoor residential aerosols, *Environment Science: Process & Impacts,* vol. 17, no.11, pp. 1959-1966.

Dallongeville, A, Costet, N, Zmirou-Navier, D, Le Bot, B, Chevrier, C, Deguen, S, Annesi-Maesano, I, Blanchard O 2015, Volatile and semi-volatile organic compounds of respiratory health relevance in French dwellings. *Indoor Air*. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26010323. [15 March 2016].

Day DE, Malm WC. Aerosol light scattering measurements as a function of relative humidity: a comparison between measurements made at three different sites. Atmospheric Environment 2001; 35: 5169 – 5176.

DEMOCOPHES 2013*, Demonstration of a study to coordinate and perform human biomonitoring on a European scale,* Available from: http://www.eu-hbm.info/democophes. [10 December 2015].

Deutschle, T, Reiter, R, Butte, W, Heinzow, B, Keck, T & Riechelmann, H 2008. *Environmental Health Perspectives*, vol. 115, no.11, pp. 1487-1493.

Dockery, DW & Spengler, JD 1981, Indoor-outdoor relationships of respirable sulfates and particles. *Atmospheric Environment*, vol. 15, no.3, pp. 335-343.

Durrant-Whyte H. Multi sensor data fusion, Australian Centre for Field Robotics, The University of Sydney, 2001.

ECMWF 2015a, *ERA-Interim dataset (January 1979 to present).* Available from: http://www.ecmwf.int/en/forecasts/datasets/era-interim-dataset-january-1979-present. [01 December 2015].

ECMWF 2015b, *ERA-40 (Sep 1957 - Aug 2002).* Available from: http://apps.ecmwf.int/datasets/data/era40-daily/levtype=sfc/. [01 March 2016].

EEA 2000, *Population density GIS data zipped. File format: ESRI grid, raster data*. Available from: http://www.eea.europa.eu/data-and-maps/data/population-density-disaggregated-with-clc2000-1/population-density-gis-data-zipped-file-format-esri-grid-raster-data/population-density-gis-data-zipped-file-format-esri-grid-raster-data. [01 December 2015].

EEA 2012, *Forests, health and climate change*. Available from: http://www.eea.europa.eu/articles/forests-health-and-climate-change/. [26 February 2016].

EEA 2014a, *AirBase - The European air quality database*. Available from: http://www.eea.europa.eu/data-and-maps/data/airbase-the-european-air-quality-database-8. [17 May 2015].

EEA 2014b, *Urban Atlas*. Available from: http://www.eea.europa.eu/data-and-maps/data/urban-atlas . [20 March 2016].

EFSA 2011, Use of the EFSA Comprehensive European Food Consumption Database in Exposure Assessment. *EFSA Journal,* vol 9, no.3, pp 2097-2131. Available from: http://www.efsa.europa.eu/. [01 February 2015].

EIONET 2012, *Corine Land Cover classes.* Available from: http://uls.eionet.europa.eu/CLC2000/classes: [20 February 2016].

EIONET 2016, *Annually interpolated air quality maps of Europe.* Available from: http://acm.eionet.europa.eu/databases/interpolated_aq_maps/index_html. [01 March 2015].

Engel-Cox JA, Hoff RM, Haymet ADJ. Recommendations on the use of satellite remote-sensing data for urban air quality. Journal of the Air & Waste Management Association, 2004; 54: 1360-1371.

EPA. EMF In Your Environment: Magnetic Field Measurements of Everyday Electrical Devices, 1992.

Estabrooks, PA, Lee, RE & Gyurcsik, NC 2003, Resources for physical activity participation: does availability and accessibility differ by neighborhood socioeconomic status?. *Annals of Behavioral Medicine*, vol 25, no.2, pp 100-104.

ETC/ACC 2008, *European air quality mapping 2005 including uncertainty analysis: ETC/ACC Technical Paper 2007/7*. Available from: http://acm.eionet.europa.eu/reports/. [17 May 2015].

EUROSTAT 2015, *NUTS classification (Nomenclature of territorial units for statistics), Overview*. Available from: http://ec.europa.eu/eurostat/web/nuts/overview. [23 March 2016].

EUROSTAT 2016, *Glossary: EU statistics on income and living conditions (EU-SILC).* Available from: http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:EU_statistics_on_income_and_living_conditions_(EU-SILC). [15 March 2016].

EUROSTAT n.d., *Territorial typologies*. Available from: http://ec.europa.eu/eurostat/statistics-explained/index.php/Territorial_typologies . [20 March 2016].

Exley, K, Aerts, D, Biot, P, Casteleyn, L, Kolossa-Gehring, M, Schwedler, G, Castaño, A, Angerer, J, Koch, HM, Esteban, M, Schindler, BK, Schoeters, G, Den Hond, E, Horvat, M, Bloemen, L, Knudsen, LE, Joas, R, Joas, An & Sepai, O. Pilot study testing a European human biomonitoring framework for biomarkers of chemical exposure in children and their mothers: experiences in the UK*, Environmental Science and Pollution Research*, vol.22, no.20, pp. 15821-15834.

Ford, ES, Merritt, RK, Heath, GW, Powell, KE, Washburn, RA, Kriska, A & Haile,G 1991, Physical activity behaviors in lower and higher socioeconomic status populations. *American Journal of Epidemiology*, vol 133, no.12, pp 1246-1256.

Fisher, K & Gershuny, J 2012, *Multinational Time Use Study, User`s Guide and Documentation, Version 5-updated.* Available from: http://www.timeuse.org/sites/ctur/files/public/ctur_report/5715/mtus-user-guide-r5.pdf. [20 March 2016].

Fisher, K & Gershuny, J 2015, *Multinational Time Use Study, User`s Guide and Documentation, Version 7-updated.* Available from: http://www.timeuse.org/sites/ctur/files/9727/mtus-user-guide-r7-february-2015.pdf . [20 March 2016].

Fisher, K & Gershuny, J 2016, *Multinational Time Use Study, User`s Guide and Documentation, Version 9-updated.* Available from: http://www.timeuse.org/sites/default//files/9727/mtus-user-guide-r9-february-2016.pdf . [20 March 2016].

Frederiksen,H, Jensen, TK, Jørgensen, N, Kyhl, HB, Husby, S, Skakkebæk, NE, Main, KM, Juul, A & Andersson AM 2014, Human urinary excretion of non-persistent environmental chemicals: an overview of Danish data collected between 2006 and 2012, *Reproduction*, vol.147, no.4, pp. 555-565.

Fromme, H, Lahrz, T, Kraft, M, Fembacher, AL, Schütze, A, Dietrich, S, Sievering, S, Burghardt, R, Koch, HM & Völkel, W 2015, Exposure to plasticizers in german daycare centers: The LUPE 3 study, *Air Pollution XXIII*, vol. 198, pp. 369-376.

Gallego, FJ 2010, A population density grid of the European Union. *Population and Environment*, vol.31, no. 6, pp.460-473.

GEMS-Food Euro (1995) Report on a workshop in the frame of GEMS-Food Euro, EUR/HFA target 22. Second workshop on reliable evaluation of low-level contamination of food. 26-27 May 1995. Kulmbach, Federal Republic of Germany.

Gens, A 2012, *Modelling the Exposure to Fine Particles and its Impacts on Human Health in Europe*. Available from: http://elib.uni-stuttgart.de/opus/volltexte/2012/7858/pdf/Diss_LK_final_version.pdf. [15 March 2016].

Gerharz, L, Yang, A, Loh, M & Alexandra,K 2011, *LAMA - tooL for Air exposure Modelling & Assessment, Software documentation & User`s guide, Version 1.1.*

Grant KE, Chuang CC, Grossman AS, Penner JE. Modeling the spectral optical properties of ammonium sulfate and biomass burning aerosols: parameterization of relative humidity effects and model results. Atmospheric Environment 1999; 33: 2603 – 2620.

Grell GA, Dudhia J, Stauffer DR. A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-397, IA, 1994:114 pp.

Gupta P, Christopher SA, Wang J, Gehrig R, Lee YC, Kumar N. Satellite remote sensing of particulate matter and air quality over global cities. Atmospheric Environment 2006; 40(30): 5880–5892.

Gupta P, Christopher SA. Seven Year Particulate Matter Air Quality Assessment from Surface and Satellite Measurements. Atmospheric Chemistry and Physics 2008; 8: 3311-3324

Hanninen, OO; Lebret, E; Ilacqua, V; Katsouyanni, K; Kunzli, N; Sram, RJ & Jantunen, M 2004, Infiltration of ambient PM2.5 and levels of indoor generated non-ETS PM2.5 in residences of four European cities. *Atmospheric Environment,* vol.38, no. 37, pp. 6411–6423.

Hänninen, OO; Hoek; G; Mallone, S , Chellini, E, Katsouyanni, K, Gariazzo, C, Cattani, G, Marconi, A, Molnár, P, Bellander, T & Jantunen, T 2010, Seasonal patterns of outdoor PM infiltration into indoor environments: review and meta-analysis of available studies from different climatological zones in Europe. *Air Quality, Atmosphere & Health*, vol.4, no.3, pp. 221-233.

Hartmann, C, Uhl, M, Weiss, S, Koch, HM, Scharf, S & König, J 2015, Human biomonitoring of phthalate exposure in Austrian children and adults and cumulative risk assessment. *International Journal of Hygiene and Environmental Health,* vol 218, no.5, pp. 489–499.

Horálek, J, Kurfürst, P, Denby, B, de Smet, P, de Leeuw, F, Brabac, M. & Fiala, J 2005, *Interpolation and assimilation methods for European scale air quality assessment and mapping, Part II : Development and testing new methodologies.* Available from: http://acm.eionet.europa.eu/docs/ETCACC_TechnPaper_2005_8_SpatAQ_Part_II.pdf. [20 March 2016].

Horálek, J, Denby, B, de Smet, P, de Leeuw, F, Kurfürst, P, Swart, R & van Noije, T 2007, *Spatial mapping of air quality for European scale assessment, ETC/ACC Technical Paper 2006/6.* Available from: http://acm.eionet.europa.eu/docs/ETCACC_TechnPaper_2006_6_Spat_AQ.pdf. [20 March 2016].

Huybrechts, I, Sioen, I, Boon, PE, Ruprich, J, Lafay, L, Turrini, A, Amiano, P, Hirvonen, T, De Neve, M, Arcella, D, Moschandreas, J, Westerlund, A, Ribas-Barba, L, Hilbig, A, Papoutsou, S, Christensen, T, Oltarzewski, M, Virtanen, S, Rehurkova, I, Azpiri, M, Sette, S, Kersting, M, Walkiewicz, A, Serra-Majem, L, Volatier, JL, Trolle, E, Tornaritis, M, Busk, L, Kafatos, A, Fabiansson, S, De Henauw, S & Van Klaveren, J 2011, Dietary Exposure Assessments for Children in Europe (the EXPOCHI project): rationale. methods and design. *Archives of Public Health,* vol. 69, no.4.

ICF & TRJ 2000, *ADDENDUM-Determination of Microenvironmental Factors for Diesel PM.* Available from: http://www.epa.gov/ttn/atw/sab/appendix-b.pdf . [05 April 2016].

JRC. An Overview of Radon Surveys in Europe. In: Dubois G, editor. Radioactivity Environmental Monitoring Emissions and Health Unit, 2005.

Kalman RE. A new approach to linear filtering and prediction problems. Transactions of the ASME – Journal of Basic Engineering 1960; 82 (Series D): 35-45.

Kaune WT, Zaffanella LE. Analysis of magnetic fields produced far from electric power lines. IEEE Transactions on Power Delivery 1992; 7: 2082-2091.

Kasper-Sonnenberg, M, Wittsiepe, J, Koch, HM, Fromme, H & Wilhelm, M 2012, Determination of bisphenol a in urine from mother-child pairs-results from the duisburg birth cohort study, *Germany, Journal of Toxicology and Environmental Health, Part A.* vol. 79, no. 6, pp.

Koch, HM, Wittassek, M, Brüning, T, Angerer, J & Heudorf, U.2011, Exposure to phthalates in 5-6 years old primary school starters in Germany--a human biomonitoring study and a cumulative risk assessment, *International Journal of Hygiene and Environmental Health*, vol.214, no.3, pp.188-195.

Koelemeijer RBA, Homan CD, Matthijsen J. Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe. Atmospheric Environment 2006; 40: 5304-5315.

Kolarik, B, Naydenov, K, Bornehag, CG & Sundell, J 2007, 'Concentrations of phthalate esters' found in homes of healthy and allergic children', *Proceedings of IAQVEC 2007*, pp. 255-262. Available from: http://www.inive.org/members_area/medias/pdf/Inive%5CIAQVEC2007%5CKolarik.pdf. [20 March 2016].

Kolarik, B, Naydenov, K, Bornehag, CG & Sundell, J 2008, The association between phthalates in dust and allergic diseases among Bulgarian children. *Environmental Health Perspectives.*vol.116, no.1, pp.98-103.

Kondragunta S, Lee P, McQueen J, Kittaka C, Prados AI, Ciren P, Laszlo I, Pierce RB, Hoff R, Szykman JJ. Air Quality Forecast Verification Using Satellite Data. Journal of Applied Meteorology and Climatology 2008; 47: 425-442.

Langer, S, Weschler, CJ, Fischer, A, Bekö, G, Toftum, J & Clausen G 2010. Phthalate and PAH concentrations in dust collected from Danish homes and daycare centers. *Atmospheric Environment.*vol.44, no.19, pp.2294-2301.

Leblanc JC, Guerin T, Noel L, Calamassi-Tran G, Volatier JL, Verger P (2005) Dietary exposure estimates of 18 elements from the 1st French Total Diet Study. Food Addit Contam 22(7), 624-41.

Lin EC. Radiation Risk From Medical Imaging. Mayo Clinic Proceedings 2010; 85: 1142-1146.

Liu Y, Sarnat JA, Kilaru V, Jacob DJ, Koutrakis P. Estimating ground-level PM2.5 in the eastern United States using satellite remote sensing. Environmental Science and Technology 2005; 39: 3269-3278.

Liu Y, Franklin M, Kahn R, Koutrakis P. Using aerosol optical thickness to predict ground-level PM2.5 concentrations in the St. Louis area: A comparison between MISR and MODIS. Remote Sensing of Environment 2007a;107: 33-44.

Liu Y, Koutrakis P, Kahn R. Estimating fine particulate matter component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: Part 1 – Method development. Journal of the Air & Waste Management Association 2007b; 57: 1351–1359.

Liu Y, Koutrakis P, Kahn R, Turquety S, Yantosca RM. Estimating fine particulate matter component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: Part 2 – A case study. Journal of the Air & Waste Management Association 2007c; 57:1360–1369.

Liu Y, Paciorek CJ, Koutrakis P. Estimating regional spatial and temporal variability of PM2.5 concentrations using satellite data, meteorology, and land use information. Environmental Health Perspectives 2009;117: 886-892.

Liu, DL.& Nazaroff, WW 2001, Modeling pollutant penetration across building envelopes. *Atmospheric Environment*, vol.35, no.26, pp.4451–4462.

Mercier, F, Glorennec, P, Blanchard, O & Le Bot, B 2012, Analysis of semi-volatile organic compounds in indoor suspended particulate matter by thermal desorption coupled with gas chromatography/mass spectrometry, *Journal of Chromatography A*.vol. 1254, pp. 107-114.

Merten, C, Ferrari, P, Bakker, M, Boss, A, Hearty, A, Leclercq, C, Lindtner, O, Tlustos, C, Verger, P, Volatier, JL.& Arcella D 2011, Methodological characteristics of the national dietary surveys carried out in the European Union as included in the European Food Safety Authority (EFSA) Comprehensive European Food Consumption Database. *Food additives & contaminants. Part A, Chemistry, analysis, control, exposure & risk assessment,*vol. 28, no. 8, pp.975-995.

Mie G. Beiträge zur Optik trüber Medien, speziell kolloidaler Metalllösungen. Annals of Physics 1908; 25(3): 378-445.

National Institute of Environmental Health Sciences NIoH. Electric and Magnetic Fields Associated with the use of Electric Power - Questions and Answers NIEHS /DOE EMF RAPID Program, 2002.

O2 telecommunications services provider. Making calls has become fifth most frequent use for a Smart-phone for newly-networked generation of users [http://news.o2.co.uk/?press-release=making-calls-has-become-fifth-most-frequent-use-for-a-smartphone-for-newly-networked-generation-of-users

OECD. OECD Communications Outlook 2013: OECD Publishing, 2013.

Officair 2013, *OFFICAIR Modelling System.* Available from: http://www.officair-project.eu/index.php?option=com_content&view=article&id=15&Itemid=15. [20 March 2016].

Olsén, L, Lind, L & Lind PM 2012, Associations between circulating levels of bisphenol A and phthalate metabolites and coronary risk in the elderly, *Ecotoxicology and Environmental Safety*, vol.80, pp. 179-183.

Orecchio, S, Indelicato, R & Barreca, S 2013, The distribution of phthalate esters in indoor dust of Palermo (Italy), *Environmental Geochemistry and Health*, vol.35, no. 5, pp. 613-624.

ORNL DAAC n.d. *Global 30 Arc-Second Elevation Data Set (GTOPO30).* Available from: http://webmap.ornl.gov/ogcdown/dataset.jsp?ds_id=10003. [03 March 2016].

Paciorek CJ, Liu Y, Moreno-Macias H, Kondragunta S. Spatiotemporal associations between GOES aerosol optical depth retrievals and ground-level PM2.5. Environmental Science & Technology 2008; 42: 5800–5806

Riechelmann, H, Deutschle, T, Grabow, A, Heinzow, B, Butte, W & Reiter, R 2007, Differential response of Mono Mac 6, BEAS-2B, and Jurkat cells to indoor dust. *Environmental Health Perspectives*, vol. 115, no. 9, pp. 1325-1332.

Sabaredzovic, A, Sakhi, AK, Brantsæter, AL & Thomsen, C 2015, Determination of 12 urinary phthalate metabolites in Norwegian pregnant women by core-shell high performance liquid chromatography with on-line solid-phase extraction, column switching and tandem mass spectrometry, *Journal of Chromatography B*, vol.1002, pp.342-352

Salma AI. Electromagnetic Field from Power Lines. Faculty of Engineering and Architecture Department of Electrical and Computer Engineering, Dr. Farid Chaaban, 2006.

Sarigiannis D, Soulakellis N, Schäfer K, Tombrou M, Sifakis N, Assimakopoulos D, Lointier M, Bossioli E, Dantou A, Saisana M. ICAROS: An integrated computational environment for the assimilation of environmental data and models for urban and regional air quality. Water, Air and Soil Pollution, Focus 2002; 2(5): 641-654.

Sarigiannis D, Soulakellis N, Sifakis N. Information fusion for computational assessment of air quality and health effects. Photogrammetric Engineering and Remote Sensing 2004; 70(2): 235–245.

Sarigiannis DA, Karakitsios SP, Antonakopoulou MP, Gotti A. Exposure analysis of accidental release of mercury from compact fluorescent lamps (CFLs). Sci Total Environ 2012; 435-436: 306-15.

SEDAC 2000,. *Gridded Population of the World (GPW), v3*. Available from: http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-density/data-download.[03 January 2016].

Semple, S, Apsley, A, & Maccalman, L 2013, An inexpensive particle monitor for smoker behaviour modification in homes, *Tobacco Control,* vol. 22, no. 5, pp295-298.

Sheridan, P.; Ogren, J. Observations of the vertical and regional variability of aerosol optical properties over central and eastern North America. J. Geophys. Res. 1999, 104, 16793-16805

Sifakis N, Deschamps PY. Mapping of air pollution using SPOT satellite data, Photogrammetric Engineering and Remote Sensing 1992; LVIII: 1433-1437

SINPHONIE 2012, *Schools Indoor Pollution and Health: Observatory Network in Europe*. Available from: http://www.sinphonie.eu/. [10 December 2015].

Smith A. Noise, performance efficiency and safety. International Archives of Occupational and Environmental Health 1990; 61: 1-5.

SoDa. Solar Energy Services for Professionals - Free WebServices [http://www.soda-is.com/eng/services/index.html]

Tanré D, Deschamps PY, Devaux C, Herman M. Estimation of Saharan aerosol optical thickness from blurring effects in Thematic Mapper data. Journal of Geophysical Research 1988; 93: 15955-15964.

The Health Physics Society. Radiation Exposure from Medical Diagnostic Imaging Procedures, 2008.

The World Bank. Mobile cellular subscriptions (per 100 people) [http://data.worldbank.org/indicator/IT.CEL.SETS.P2]

Tirmarche M, Harrison JD, Laurier D, Paquet F, Blanchardon E, Marsh JW. ICRP Publication 115. Lung cancer risk from radon and progeny and statement on radon. Ann ICRP 2010; 40: 1-64.

Tollefsen, T, Cinelli, G, Bossew, P, Gruber, V & De Cort, M, From the European indoor radon map towards an atlas of natural radiation, *Radiation Protection Dosimetry*, vol. 162, no.1-2, pp.129-134.

Torras Ortiz, S & Friedrich, R 2013, A modelling approach for estimating background pollutant concentrations in urban areas. *Atmospheric Pollution Research*, vol. 4, no.2, pp.147-156.

UNECE 2006, *Bulletin of Housing Statistics for Europe and North America.* Available from: http://www.unece.org/hlm/prgm/hsstat/Bulletin_06.html. [10 December 2015].

Vandentorren, S, Zeman, F, Morin, L, Sarter, H, Bidondo, ML, Oleko, A & Leridon H 2011, Bisphenol-A and phthalates contamination of urine samples by catheters in the Elfe pilot study: implications for large-scale biomonitoring studies, *Environmental Research*. Vol. 111, no. 6, pp. 761-764.

van Donkelaar A, Martin RV, Park RJ. Estimating ground-level PM2.5 using aerosol optical depth determined from satellite remote sensing. Journal of Geophysical Research 2006; 111: D21201, doi:10.1029/2005JD006996.

van Donkelaar A, Martin RV, Brauer M, Kahn R, Levy R, Verduzco C, Villeneuve PJ. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. Environmental Health Perspectives 2010; 118: 847-855.

Völkel, W, Kiranoglu, M, Schuster,R, Fromme, H & HB, Mnet 2014, Phthalate intake by infants calculated from biomonitoring data, *Toxicology Letters*, vol.225, no.2, pp.222-229.

UBA 2012, *Environmental Specimen Bank*. Available from: https://www.umweltprobenbank.de/en/documents/investigations/biometrics/specimen_types. [30 March 2016].

U.S.EPA 1994, *Motor Vehicles and the 1990 Clean Air Act*. Available from: https://www3.epa.gov/otaq/consumer/11-vehs.pdf. [10 April 2016].

U.S.EPA 2015, *Summary of the Clean Air Act.* Available from: https://www.epa.gov/laws-regulations/summary-clean-air-act .[10 April 2016].

U.S.EPA 2016, *Possible Exposure Pathways During Emergencies*. Available from: https://www.epa.gov/emergency-response/possible-exposure-pathways-during-emergencies. [10 Februry 2016].

Wang J, Christopher SA. Intercomparison between satellite-derived aerosol optical thickness and PM2.5 mass: Implications for air quality studies, Geophysical Research Letters 2003; 30 (21), 2095, doi:10.1029/2003GL018174.

Wang Z, Chen L, Tao J, Zhang Y, Su L. Satellite-based estimation of regional particulate matter (PM) in Beijing using vertical-and-RH correcting method. Remote Sensing of Environment 2010; 114: 50-63

Wiegner, M., Emeis, S., Freudenthaler, V., Heese, B., Junkermann, W., Muenkel, C., Schaefer, K., Seefeldner, M., Vogt, S., 2006. Mixing layer height over Munich, Germany: variability and comparisons of different methodologies. Journal of Geophysical Research—Atmospheres 111, D13201

Weschler, CJ & Nazaroff, WW 2010, SVOC partitioning between the gas phase and settled dust indoors, *Atmospheric Environment*, vol.44, no.30, pp. 3609–3620.

World Health Organization. Establishing A Dialogue On Risks From Electromagnetic Fields. In: Environment RAEHDOPOTH, editor, Geneva, 2002.

Ye, X, Pierik, FH, Hauser, R, Duty, S, Angerer, J, Park, MM, Burdorf, A, Hofman, A, Jaddoe, VW, Mackenbach, JP, Steegers, EA, Tiemeier, H & Longnecker, MP 2008 Urinary metabolite concentrations of organophosphorous pesticides, bisphenol A, and phthalates among pregnant women in Rotterdam, the Netherlands: the Generation R study, *Environmental Research*, vol.108, no.2, pp.260-267.

Zaffanella LE, Rauch GB, Johnson GB. Research facility for the study of power system magnetic fields, 1992, pp. 749-757. [http://www.scopus.com/inward/record.url?eid=2-s2.0-0026626057&partnerID=40&md5=4d18158de33ac7d671c85a21a7dd75b9]