

# Characterizing Environmental Inequalities Using Integrated Exposure Assessment and Spatial Approach

**Julien Caudeville** (✉ [julien.caudeville@ineris.fr](mailto:julien.caudeville@ineris.fr))

INERIS (French National Institute for Industrial Environment and Risks) <https://orcid.org/0000-0001-8939-9369>

**Corentin REGRAIN**

INERIS

**Frederic TOGNET**

INERIS

**Roseline BONNARD**

INERIS: Institut National de l'Environnement Industriel et des Risques

**Mohammed GUEDDA**

UPJV

**Celine BROCHOT**

Ineris

**Maxime BEAUCHAMP**

INERIS: Institut National de l'Environnement Industriel et des Risques

**Laurent LETINOIS**

INERIS

**Laure MALHERBE**

INERIS: Institut National de l'Environnement Industriel et des Risques

**Fabrice MARLIERE**

INERIS: Institut National de l'Environnement Industriel et des Risques

**Francois LESTREMAU**

INERIS

**Karen CHARDON**

University of Picardy Jules Verne: Universite de Picardie Jules Verne

**Veronique BACH**

University of Picardy Jules Verne: Universite de Picardie Jules Verne

**Florence Anna ZEMAN**

INERIS

**Keywords:** Inequalities, spatial, exposure, modeling, integrated

**Posted Date:** November 12th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-103756/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

At a regional or continental scale, the characterization of environmental health inequalities (EHI) expresses the idea that populations are not equal in the face of pollution. It implies the analysis in order to identifying and managing areas at risk of overexposure where increasing risk to human health is suspected. The development of methods is a prerequisite for the implementation of public health actions aimed at the protection of populations.

## Methods

This paper presents the methodological framework developed by INERIS (French national institute for industrial environment and risks) to identify a common framework for conceptualizing and operationalizing environmental exposures as an important step towards articulating a science of EHI. We develop an integrated exposure assessment approach capable to integrate the multiplicity of exposure pathways from various sources, through a series of models up to the internal exposure.

## Results

Measured data from environmental networks reflecting the actual contamination of the environment are reused to characterize the population's exposure. Sophisticated methods of spatial analysis are applied to include additional information and take benefit from spatial and inter-variable correlation to improve data representativeness and characterize associated uncertainty. Integrated approaches bring together all information necessary for assessing the source-to-human-dose continuum using Geographic Information System, multimedia exposure and toxicokinetic model.

## Conclusion

This framework could be used for many purposes, such as mapping EHI, identifying vulnerable populations and determinants of exposure to manage and plan remedial actions and assessing spatial relationships between health and environmental to identify factors that influence the variability of disease patterns.

## 1. Introduction

World Health Organization in a recent report (2012) [1] has identified environmental inequalities as a priority issue in need to be addressed by the national governments in Europe. Reducing health inequalities requires identifying and characterizing exposure in order to interpret how they accumulate across a territory and prioritize interventions. As the health status of a population is the result of complex interactions between several social, territorial and environmental factors, all related information needs to be studied in order to assess it. At a regional or continental scale, the characterization of environmental health inequalities (EHI) expresses the idea that populations are not equal in the face of pollution. It

implies the analysis in order to identifying and managing areas at risk of overexposure where increasing risk to human health is suspected. The development of methods is a prerequisite for the implementation of public health actions aimed at the protection of populations. Constructing tools to guide public action in order to reduce EHI requires the evaluation of phenomena not always simple to apprehend and the reliability and representativeness of available information that usually demands statistical processing [2].

After more than 10 years of actions aimed to the prevention of environmental health risks the third national plan for health and environment (PNSE 3, 2015-2019) proposes a new EHI approach that is not only more robust and connected to the territories but also integrating the scientific concept of exposome. The recently emerged term of exposome [3] is used to describe these complex exposures, considering all sources, routes, and - when possible - the interactions of stressors, that are likely to contribute to the health alteration of individuals. The external contribution to the human exposome is determined by environmental exposure, also termed the eco-exposome [4] and includes exposure from air, water, soil and food exposure media. A coherent exposure assessment conceptual framework is needed to tackle EHI permitting the estimation of the magnitude, frequency and duration of exposure to chemicals, along with the number and characteristics of the population exposed.

Quantitative exposure assessment for environmental inequality characterization poses specific questions that need to be addressed:

- Identification of contamination(s) source(s);
- Characterization of exposure mechanisms (pathways and relevant routes);
- Prioritization of vulnerable populations or specific susceptible groups (e.g. infants);

Contamination process is extremely complex and variable through space and time, with localized multiple sources at larger scale. At a regional scale to better evaluate exposure to large chemical emissions, fate and transport models could provide both the ability to account for relevant spatial variability (e.g., around emissions sources or highly populated areas) and temporal variability during a specific time of contamination [5].

Exposure assessment to identify and characterize the territorialized EHI depends on the availability of data. Exposure assessment is generally complex due to a lack of data and the inherent natural variability in exposure levels, leading to uncertainty in the estimates [6]. The temporal support also differs between the available data (punctual measurements, annual averages, etc.) which also requires additional treatment. Furthermore, they often lack a common spatial support and therefore preliminary spatial analysis is required in order to homogenize them or increase their resolution. The available databases are often assembled for diverse objectives, and often re-processed using statistical methods. Spatialization and crossing of these data pose several methodological difficulties and can introduce uncertainties in the cartography process carried out. For this reason, different methods and techniques are employed to

specifically treat environmental databases in order to take benefit from all available information and reduce the uncertainties (see Section 3.2).

This paper presents the methodological framework developed by INERIS to identify a common framework for conceptualizing and operationalizing environmental exposures as an important step towards articulating a science of EHI.

In order to build a calculation infrastructure able to characterize the eco-exposome at the territorial level, we had to solve several methodological issues: (1) define an integrated exposure assessment framework that first requires different scientific limitations to be overcome, such as the linkage of the global source-effect chain, (2) provide statistical methods and numerical tools that would allow spatial and temporal data processing from existing environmental and populational databases, (3) link, adapt or develop transport and transfer models.

## 2. The Integrated Exposure Assessment Framework

The characterization of the territorialized exposome implies the development of dynamic, multidimensional, longitudinal approaches, and information systems that require the adoption of transdisciplinary methods of data analysis. To respond to the general objective, it is required to integrate and combine of various levels of data from different environmental compartments and exposure media. Data and information emerging from an expanding field of exposure science can be integrated in the exposome conceptual framework that provides the necessary linkages between source and internal exposure and helps to identify and compare relationships between different levels at critical life stages, personal health outcomes, and health disparities at a population level across space, place, and time [7]. This framework could be a layered structure that describes the elements of exposure pathways (Fig. 1), the relationship between those elements, and how data describing the elements is stored and used for selected outputs, such as exposure assessment, exposure prediction, epidemiology or public health decision making [8]. Refined aggregate exposure assessment is data-intensive, requiring detailed information at every step of the source-to-dose pathway. Integrated exposure assessment requires 1) methodologies to allow calculating the aggregate exposure systematically and 2) computational research tools to estimate the exposure from the different contributing sources.

### Fig. 1 : Modeling framework example to characterize integrated EHI

For example, integrated approaches could bring together all information necessary for assessing the source-to-exposure continuum referring to the linkage, connecting the source of the exposure with the target exposure. In the context of mapping the environmental inequalities enabling the identification of vulnerable individuals and communities at risk in order to target public health interventions, additional requirements are needed in the exposure assessment processes compared with classical risk assessment methodology.

The environmental inequalities operate at different scales (global, regional, local) and could not be apprehended by the study of a single medium, but by the integration of varied contamination pathways: air, water, soil and food. The design study should be able to:

- integrate the processes that take place at the interface between the environment contaminants of interest and the organisms,
- characterize the principal exposure pathways,
- define realistic scenarios that integrate the past and present sources,
- describe the phenomena at a fine temporal and spatial resolution.

Based on the needs described above, research objectives are to bring together all available information within a coherent methodological framework for assessing the source-to-dose continuum covering an extensive chemical space. We develop an integrated exposure assessment approach able to integrate the multiplicity of exposure pathways from various sources, through a series of models up to the internal exposure. The main objective of our projects, i.e. testing the feasibility of the methodology, has been achieved. Our framework allows for substances of interest:

- identifying areas of potential overexposure by analyzing the variations of the indicators in space,
- analyzing sources and environmental compartments potentially associated with overexposure,
- explaining the variability of exposure inequalities for pollutants and study areas,
- estimating of internal exposure and link with human biomonitoring data.

This approach involved the implementation of different models, namely atmospheric dispersion modeling, spatial analysis for environmental and population data processing, a multimedia exposure model and a physiologically-based pharmacokinetic (PBPK) model. The models have been adapted and coupled to allow the integration of the output data of an upstream model as input data of the downstream model. The coupling also allows the integration of information on the sources of contamination, the quality of environmental media and resident populations on the same analysis medium, namely the reference grid. A Geographic Information Systems (GIS) thus provides the opportunity to cross the estimated exposure with biological impregnation data to provide interpretive elements of the environmental determinants of exposure. The coupling of numerical and statistical models has established a scientific and technical basis for the integration, data processing and assessment of the transfer of contaminants from the environment to the populations. In that way, it is possible to integrate all available data, despite their heterogeneity in a common spatial support. The referent grid selected, allows the reflection of local variations, and the integration of environmental monitoring databases in France. The GIS modeling platform enables the coupling and interoperability of

all spatial data via the reference grid on which the input data and the variables of interest are discretized after processing.

### **3. In Search Of Data And Representativeness**

## **1. Integrated data from existing environmental health monitoring programs**

Many different approaches can be used for quantifying environmental exposures: direct methods (measuring, monitoring or biomonitoring) or indirect methods, involving exposure estimations from measurements and existing data, like environmental monitoring, questionnaires and exposure models. The availability of data on the geographic area of interest for the pollutants evaluated is an essential prerequisite. The quality and usability of all environmental data should be assessed before employing them in the health or risk assessment processes, as many factors can bias environmental sampling results [9]. Ideally, direct measures of exposure (e.g., biomarkers or personal monitoring data) for all key stressors related health effects, throughout the critical time-period of exposure, and in the population of interest would be necessary<sup>10</sup>. However, exclusive use of biomarker data in exposure assessment to characterize EHI is currently not practicable when considering a large number of diverse chemicals due to analytical and resource limitations [11] specifically when the assessment should cover a large territory and fine resolution. Environmental quality data are often available at a fine administrative or resolution level and enable the building of environmental indicators on a regional or national scale. The processing of variables for the identification and characterization of environmental inequalities depends on the reuse of this type of data, which is very diverse by nature regarding its initial intended objectives. Determining how representative those measured levels of contamination are of other locations or time frames is not always a simple task [12].

Databases in health and environment have been developed for several years. They evolve and are in full expansion. Actions to identify and monitor the quality of the environment for soils, water and air are conducted by different agencies, institutes or observatories. The production of this type of data and advances in computer technology allow their reuse in conceptual frameworks and with objectives different from those that prevailed in their implementation. The emergence of quality data and their integration into GIS make it possible to conduct territorial analysis work. These environmental data reflect the actual contamination of the environment and therefore of the global exposure of the populations. The indicators based on these data allow to characterize the population's exposure and its evolution regarding the implementation of public prevention policies. In the context of reuse of this type of data for the purpose of expology, a database must be set up in which the variables are associated with the modes of exposure (concentrations in the environmental and exposure media are present, eating behavior, space-time budget, ...).

These variables must know several stages of process to allow the construction of indicators:

- the identification of data sources allowing the construction of the different variables,
- the acquisition of these data in view of the access modalities, the financial, legal or human aspects,
- the analysis of the quality and representativeness of the databases regarding the objective of the study (choice of a database, validity and representativeness of the data) sometimes involving the approximation or the application of simplifying assumptions,
- the preprocessing of databases: cleaning the databases, rebuilding missing data,
- the construction of ad-hoc data where the appropriate data sources are not available or exhaustive in relation to the objectives of the study,
- data transformation (homogenization, aggregation or disaggregation of data).

The estimation of exposure requires knowledge of the concentrations of environmental compartments to which an individual or a population is exposed. These concentrations can be measured or modeled. A wide range of data might potentially be mobilized for integrated assessment. The database selection or study design definition should be guided to reach the best compromise between data representativeness and method robustness, consistent with the objectives of the study.

Characteristics of air pollution (e.g., chemical components, particle properties) vary spatially [13] and may differ between areas near and far from monitors [14]. Automated monitoring networks operate in Europe providing detailed air quality information on a regular basis. The soil routes of exposure to humans are inhalation of dust and vapor coming from soil contaminants, ingestion of contaminated soil particles (mainly for children) or contaminated food, and dermal absorption through the skin. Once a site is considered as contaminated, it is necessary to provide enough accurate data to minimize lack of statistical representativeness and increase the spatial quantification. The time spent for evaluating the presence and extent of contamination can be reduced by an adequate sampling plan [15] which can at the same time, reduce the project costs [16]. A soil monitoring system could be the source of the comparable and objective data on the current state and evolution of soils. The database of the soil monitoring system allows the creation and maintenance of data for each of the monitoring sites of agricultural land as well as the preparation of data for further processing through specialized programs [17]. Position information provides a link to the GIS, and thus opens the possibilities for further spatial analysis, the identification of risk areas and their assessment. By example, in France, soil pollutant stocks and properties and most explanatory variables were derived from the French National Soil Monitoring Network (Réseau de Mesures de la Qualité des Sols or RMQS). The RMQS surveys soils and their properties on a regular 16 km grid across the French mainland territory (around 2,200 sites covering 550,000 km<sup>2</sup>) [18].

The Drinking Water Directive (80/778/EEC), and its successor (98/83/EC which comes in force in 2003), aims to ensure that water intended for human consumption is safe. In addition to microbiological and physicochemical parameters, a number of toxic substances such as pesticides, polyaromatic



hydrocarbons, cyanide compounds, and heavy metals are to be monitored. This is because the raw supply may be contaminated, for example, with pesticides from agricultural land which have leached into groundwater or from contamination within the distribution system, such as lead from piping. In France, 300,000 samples are tested each year. Indeed, tap water is one of the most strictly controlled foodstuffs. Each year, the health agencies carry out close on 12.3 million tests covering all of the country's public water and wastewater services (both publicly and privately managed). In 2013, more than 8.1 million tests were carried out on services managed by private water company.

Work has been carried out by INERIS to identify environmental and spatialized databases for the purpose of characterizing exposures by associating the main producers and data managers identified [19, 20]. It allows to propose elements for the specification of environmental health platforms and to improve the integration of data in the framework of building an environmental health tracking information system. However, spatial data used to characterize environmental exposures have not always been initially collected and collated to meet these objectives, resulting in use bias. Measuring frequencies or spatial densities of sampling are not always sufficient. To partially overcome these problems, different techniques are adopted to specifically address the different environmental, behavioral or population databases. The selection of a treatment method depends on the problem to be solved and the quality of the data available.

## **2. Statistic approaches to link and optimize data representativeness**

The data available in a region of interest characterize levels of contamination at very specific locations, over a given spatial support (i.e. the support on which the data is measured such as point, surface or volume), and for very specific time frames. In order to construct the exposure maps from spatialized databases in the context of evaluating environmental inequalities, the development of methods is required to process and harmonize the available data, with respect to their specificities (missing values, limited number of observations, etc.) in the same resolution and support.

In the mathematical field of numerical analysis, interpolation is a method of constructing new data points within the range of a discrete set of known data points. During the last years the increasing availability of spatial and spatiotemporal data pushed the developing of many spatial interpolation methods, including geostatistics [21]. Spatial interpolation includes any of the formal techniques which study entities using their topological, geometric, or geographic properties. Spatial dependence is the co-variation of properties in a geographical space: features at nearby locations seem to be correlated. The fundamental principle is Tobler's first law on geography: if the interrelation between entities increases with proximity in the real world, representation in geographical space and evaluation using spatial analysis techniques are appropriate [22]. These interactions are all stronger as the locations concerned are closer. In statistics, spatial autocorrelation measures the correlation of a georeferenced variable with itself. It makes it possible to measure the degree of similarity between neighboring observations. This spatial dependence

implies the infringement of the assumptions made in the classical statistical techniques which suppose the independence between the observations. Spatial dependence should also be considered as a source of information. To characterize the different scales of local, regional and global variability of the phenomena studied, the analysis of spatial data structures through geostatistical tools (variogram, autocorrelation analysis) is often employed [23].

Several more sophisticated methods of spatial analysis can be applied to include additional information and take benefit from spatial and inter-variable correlation to improve data representativeness and characterize associated uncertainty [24].

For air, several methods for estimating exposure to air pollutants exist, including monitor-based approaches such as proximity-based assessments and statistical interpolation, as well as land-use regression and air quality modeling [25]. Using data from existing monitoring networks remains popular, due to cost considerations, data availability, and population coverage. Such statistical methods are aimed at using multiple types of information to inform exposure estimates and allow to estimate exposure in areas far from monitors. In addition to fused data, several other approaches have been developed to estimate individual- and population-level exposures, including various interpolation methods, land use regression (LUR) models, aerosol measurements obtained from satellites, and source- and traffic-proximity analysis [26]. Stochastic methods such as kriging are preferred [27]. An issue commonly reported is the availability of data. Some databases include some limitation (as a limited number of observations by example) and therefore it is not possible to assess the population's exposure adequately. External drift kriging is then widely used in air and soil quality modeling, in order to combine different kind of information to include secondary information in the model.

Machine learning use algorithms and statistical methods to “learn” information directly from data without relying on a predetermined equation as a model. The algorithms adaptively improve their performance as the number of samples available for learning increases. Machine learning allow for example to build a metamodel from a dataset of deterministic model outputs. The fundamental concepts of machine learning and its usages in spatially distributed data are given in Kanevskij et al [28].

In order to construct the exposure maps from spatialized databases in the context of evaluating health risks, methods have been developed to process and harmonize the available data, with respect to their specificities (missing values, limited number of observations, etc) in the same resolution and support. A GIS-based modeling platform developed by INERIS for quantifying human exposure to chemical substances (PLAINE: environmental inequalities analysis platform [29]) aims to spatialize an environmental indicator related to human health using risk assessment methods and mapping environmental disparities at a fine resolution. The main aim of the PLAINE Project, developed in France, is to develop a platform of environmental and health data. This platform is developed for systematic collection, integration, and analysis of data on emission sources, environmental contamination, exposure to environmental hazards, and population and health. Ad-hoc methodologies are used to align the available data to the same pixels. Spatial analysis and statistical methods are employed to process

(georeferencing, data controlling, pre-processing, re-formatting) and assemble the databases for the purpose of the study, using R and QGIS. By example, atmospheric concentration data were collected in France in the context of regulatory surveillance for two years (2010 and 2011). Estimation of concentrations over France by classical interpolation method could lead to a misrepresentation of the spatial distribution due to the limited number of observations. To address this issue, auxiliary variables in the context of external drift kriging [30] were employed. The best auxiliary variable to define linear drifts was found to be the one that includes the atmospheric emissions as well as the population and the altitude. Measurements of PAH topsoil concentrations are available through the French Soil Monitoring Network. Qualitative data on the polluted sites localization are integrated by processing distance-to-polluted soil proxy. These, along with 14 variables about physicochemical soil properties were combined in a hybrid regression-kriging and fitted using Random Forest [31] models, were shown to outperform the traditionally used linear regression. Due to its hydrophobic nature, B[a]P is found in water in small concentrations; therefore, the exact measurement cannot always be reported. The observations under the detection limit rate is quite high, which requires careful handling. A complex multiple imputation method was developed in order to extract the maximum information from the available measurements without introducing too much bias in the results. This one permits to take advantage of the temporal aspect and correlations between substance of interest and other PAH substances. Spatial estimation of water concentrations was carried out by taking into account the multi-annual data and the network water distribution complexity using a bootstrap based expectation-maximization algorithm. The above methods permitted the construction of a representative spatial database in a 9 km<sup>2</sup> grid of reference on the whole France (550,000 km<sup>2</sup>) used to perform the integrated exposure assessment [3].

## 4. Fate And Transport Models

### 1. Outdoor air dispersion modeling

Atmospheric chemistry and dispersion modelling experienced important improvements in the last two decades. Nowadays, a large variety of modelling systems and options exist, from simpler to more complex ones, covering global or regional to urban and street level scales.

Air quality models simulate the atmospheric concentrations and deposition fluxes to the Earth's surface of air pollutants by solving the transport equations that represent the emissions, advection, diffusion, transformations and removal of those air pollutants and associated chemical species.

Contemporary air quality models can be grouped into two major categories:

- models that calculate the concentrations of air pollutants near a source (source-specific models). The Gaussian models simulate the atmospheric dispersion of non-reactive pollutants near the source (steady-state approach). Lagrangian models are also source-specific models, which treat atmospheric dispersion of reactive substances as a source-specific process;

- Eulerian models that calculate concentrations of reactive air pollutants over large areas ranging from an urban area, to a region, a continent and the globe (grid-based models).

Inputs to air quality models include the emission rates of primary air pollutants and precursors of secondary air pollutants, meteorology (three-dimensional fields of winds, turbulence, temperature, pressure, boundary layer height, relative humidity, clouds and solar radiation ...), and boundary conditions (baseline or background conditions). For grid-based models, an emission model is used to translate an emission inventory into a spatially distributed and temporally resolved grid structure.

As an example, INERIS used BaP as a tracer of the carcinogenic risk associated with PAH has been the subject of several recent studies using the CHIMERE model at European scale [36, 37]. The population exposure estimate shows that 20% of the European population is exposed to BaP background ambient concentrations above the EU target value and only 7% live in areas with concentrations under the estimated acceptable risk level of  $0.12 \text{ ng.m}^{-3}$ . Heavy metals have also been addressed using the CHIMERE model [38], modelling Pb, Cd, As, Ni, Cu, Zn, Cr and Se air background concentration in Europe. Evaluation of the model performance in order to see its capability to reproduce observed levels shows that more recent annual totals, information on snap activities for each metal, higher spatial resolution and a better knowledge of the temporal emission behavior are necessary to adequately model these air pollutants.

## 2. Multimedia exposure models

Spatially resolved multimedia fate and multi pathway exposure models facilitate the prediction of environmental concentration distributions, related levels of contaminants in different sources, and the fraction of a chemical release that will be taken in by the entire human population (the intake dose) at the regional or local scale. When spatial resolution of computations is low, usually variations in environmental characteristics tend to average out, and adoption of roughly selected representative or characteristic values allows the depiction of the correct orders of magnitude of outputs. Research has been starting to cope with spatially explicit models of fate and transport with increasing resolution, and now a few models with resolution from a few tens of km up to 1 km are available for calculations at the continental scale [39,40]. However, in any case the computational effort associated with this modeling strategy is generally quite high and limits routine applications when a large number of chemicals need to be evaluated.

A multimedia fate and exposure model called Modul'ERS [41, 42] developed by INERIS is used to estimate intakes from air inhalation and soil, tapwater, marketed food products, as well as local-produced fruits and vegetables ingestion. Local foodstuff concentrations are estimated using atmospheric deposition of particulate pollutants, air (for POP) and soil concentrations. As mechanistic and dynamic models for plants required many input data that can be difficult to define (lack of data, difficulty for estimating the magnitude of variability and uncertainty of data and even anticipating the qualitative effect of variation of input data on results), contributions of air gaseous and soil concentrations to edible organs of plants

are estimated from bioconcentration factors, which are specific to the different categories of fruit and vegetables cultivated in domestic gardens and time average concentration during culture duration. Therefore, the inputs of the model for media concentration estimates are georeferenced environmental databases (with a direct reuse of the treated data incorporated into the GIS for tapwater and marketed food products).

In the model used, attention was focused on the quality of values used to define all inputs (exposure, environmental and chemical parameters). Analyses of available data were systematically conducted. For most of the parameters, all the data collected with their contextual information, as well as the selection criteria used are described in dedicated reports. Depending on the level of knowledge, the quantity and the relevance of the data available, the parameters are finally defined with a point value, a range of values or a probabilistic distribution.

### **3. Physiologically-based pharmacokinetic (PBPK) models**

PBPK models are a specific class of biokinetic models based on the physiology and the anatomy of the individuals that are able to predict the kinetics and metabolism of substances in the body. Those models describe the body as a set of compartments corresponding to specific organs or tissues (e.g., adipose, bone, brain, gut, heart, kidney, liver, lung, muscle, skin, and spleen, etc.). Between compartments, the transport of substances is dictated by various physiological flows (blood, bile, pulmonary ventilation, etc.) or by diffusion [43, 44]. The model structure can be described by a set of differential equations, with parameters representing blood flow rates, organ volumes etc., for which information is available in the published scientific literature or may be obtained in vitro [45]. Numerical integration of that differential system computes the quantity and concentration of the drug considered in each compartment, as a function of time and exposure dose. A stochastic whole-body physiologically-based pharmacokinetic model over the human lifespan has been developed by INERIS [46] and integrated in the EHI context to predict internal concentration such as concentrations in blood but also in other tissue or biological matrices (urine) from multi-route exposure (inhalation, ingestion, dermal exposure). Those models are used to link exposure with biomarker data [47,48] and have proven to be successful in integrating and evaluating the influence of age or gender-dependent changes with respect to the pharmacokinetics of xenobiotics throughout the lifetime [49, 50].

Each model represents a different component of the continuum emission-environmental quality-exposure-internal dose and effects). These models can operate in different spatio-temporal scales, which poses a challenge when coupling them in a coherent framework and can result in structural uncertainty and deep time calculation problem.

### **5. Conclusion**

The exposome concept has been proposed as an emergent exposure science paradigm for conceptualizing the cumulative effects of environmental exposures across the whole human life. The

need for risk manager to identify population at-risk in the context of substantial data deficiencies that hinder evaluation of cumulative health risks brings the operational declination of the concept at the territorial scale in the EHI characterization context. The characterization of the territorialized exposome implies the development of dynamic, multidimensional, longitudinal approaches, and information systems that require the adoption of transdisciplinary methods of data analysis. For example, integrated approaches bring together all information necessary for assessing the source-to-human-dose continuum using GIS, multimedia exposure and toxicokinetic model.

This framework could be used for many purposes, such as:

- mapping EHI;
- identifying vulnerable populations and determinants of exposure to manage and plan remedial actions;
- highlighting hotspot areas with significantly elevated exposure indicator values to define environmental monitoring campaigns;
- assessing spatial relationships between health and environment to identify factors that influence the variability of disease patterns.

## **Declarations**

### **1. Ethics approval and consent to participate**

Not applicable.

### **2. Consent for publication**

Not applicable.

### **3. Availability of data and materials**

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

### **4. Competing interests**

The authors declare that they have no competing interests.

### **5. Funding**

This research was funded by the French minister in charge of the environment and by the French Biodiversity Agency (Agence française de la biodiversité -AFB) into the Ecophyto 2 plan context.

## 6. Authors' contributions

Conceptualization, J.C.; validation, All authors; writing—original draft preparation, F.Z., C.R., R.B., F.T.; writing—review and editing, all authors; supervision, J.C.; funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

## References

1. WHO Regional Office for Europe (2012). Environmental health inequalities in Europe: assessment report. Copenhagen: WHO Regional Office for Europe.
2. Caudeville J, Bonnard R., Boudet C., Denys S., Govaert G., Cicoletta A. Development of a spatial stochastic multimedia model to assess population exposure at a regional scale. *Journal of Total Environment* 2012 Volume 432, Pages 297–308.
3. Wild CP. Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev.* 2005;14:1847-50; doi:10.1158/1055-9965.EPI-05-0456.
4. Liroy PJ; Smith KR. A Discussion of Exposure Science in the 21st Century: A Vision and a Strategy. *Environ Health Persp* 2013;121:405–409.
5. Bulle C, Margni M, Patouillard L, et al (2019) IMPACT world+: a globally regionalized life cycle impact assessment method. *Int J life cycle assess* 1–22. doi: <https://doi.org/10.1007/s11367-019-01583-0>
6. Ciffroy P, Péry ARR, Roth N. Perspectives for integrating human and environmental exposure assessments. *Sci Total Environ.* 2016;568:512-21; doi:10.1016/j.scitotenv.2015.11.083.
7. Juarez, P. Matthews-Juarez, D. Hood, W. Im, R. Levine, B. Kilbourne, M. Langston, M. Al-Hamdan, W. Crosson, M. Estes Jr, S. Estes, V. Agboto, P. Robinson, S. Wilson, and M. Lichtveld. The public health exposome: A population-based, exposure science approach to health disparities research. (11: 12866 - 12895), 2014.
8. J. G. Teeguarden, Y. M. Tan, S. W. Edwards, J. A. Leonard, K. A. Anderson, R. A. Corley, M. L. Kile, S. M. Simonich, D. Stone, R. L. Tanguay, K. M. Waters, S. L. Harper, and D.E. Williams. Completing the link between exposure science and toxicology for improved environmental health decision making: The aggregate exposure pathway framework. *Environmental Science and Technology*, (50(9): 4579–4586), 2016.
9. U.S. EPA. 1992. Guidelines for exposure assessment. *Fed Reg* 57:22887-22938.
10. A.M. Evans, G.E. Rice, J.M. Wright, and L.K. Teuschler. Exploratory cumulative risk assessment (cra) approaches using secondary data. *Human and Ecological Risk Assessment: An International Journal*, (20(3): 704-723), 2014.

11. Agency for Toxic Substances and Disease Registry. 2005. Public Health Assessment Guidance Manual. Atlanta, GA: Department of Health and Human Services; January 2005.
12. Sarnat SE, et al. Spatial considerations in a study of ambient air pollution and cardiorespiratory emergency department visits. *Epidemiology*. 2006; 17:S242.
13. Bell ML, et al. Spatial and temporal variation in PM<sub>2.5</sub> chemical composition in the United States for health effects studies. *Environ. Health Perspect.* 2007; 115:989–995.
14. Bravo, M. A.; Fuentes, M.; Zhang, Y.; Burr, M. J.; Bell, M. L. Comparison of exposure estimation methods for air pollutants: Ambient monitoring data and regional air quality simulation. *Environ. Res.* 2012, 116, 1–10.
15. L. Malherbe. Designing a contaminated soil sampling strategy for human health risk assessment. *Accred. Qual. Assur.*, 7 (5) (2002), pp. 189-194
16. A. Horta, B. Malone, U. Stockmann, B. Minasny, T. Bishop, A. McBratney, R. Pallasser, L. Pozza. Potential of integrated field spectroscopy and spatial analysis for enhanced assessment of soil contamination: a prospective review. *Geoderma*, 241 (2015), pp. 180-209.
17. Dominique Arrouays, Anne C. Richer-de-Forges, Florence Héliès, et al. Impacts of national scale digital soil mapping programs in France, *Geoderma Regional*, Volume 23, 2020, e00337.
18. Lacarce, N.P.A. Saby, M.P. Martin, B.P. Marchant, L. Boulonne, J. Meersmans, C. Jolivet, A. Bispo, D. Arrouays. Mapping soil Pb stocks and availability in mainland France combining regression trees with robust geostatistics *Geoderma*, 170 (2012), pp. 359-368
19. Caudeville J. Caractérisation des inégalités environnementales : Inventaire des bases de données nationales environnementales et spatialisées. INERIS-DRC-15-152407-11231B, 2015.
20. Caudeville J. Rapport « Caractérisation des inégalités environnementales : Inventaire des bases de données régionales et locales, environnementales et spatialisées ». INERIS-DRC-17-164533-00415A, 2016.
21. Caudeville J, Goovaerts P, Carré F, Bonnard R, Arrouays D, Cicoella C, Jolivet C, Saby NPA, Boudet C. 2012. Spatial modeling of human exposure to soil contamination- an example of Digital Soil Assessment. Chapter in "Digital Soil Assessments and beyond". Eds. Minasny B, Malone BP, McBratney AB. Published CRC Press.
22. J. Chakraborty. Revisiting Tobler's first law of geography: spatial regression models for assessing environmental justice and health risk disparities. J.A. Maantay, S. McLafferty (Eds.), *Geospatial Analysis of Environmental Health, Geotechnologies and the Environment*, 4, 3 (2011), pp. 337-356
23. Goovaerts, P. (1997), *Geostatistics for Natural Resources Evaluation*, Oxford Univ. Press, New York.
24. Goovaerts, P. (2001), Geostatistical modelling of uncertainty in soil science, *Geoderma*, 103, 3– 26.
25. Jerrett M, A Arain, P Kanaroglou, et al. A review and evaluation of intraurban air pollution exposure models. *J. Exposure Sci. Environ. Epidemiol.* 2005a;15:185–204.
26. Bravo, M. A.; Fuentes, M.; Zhang, Y.; Burr, M. J.; Bell, M. L. Comparison of exposure estimation methods for air pollutants: Ambient monitoring data and regional air quality simulation. *Environ.*



Res. 2012, 116, 1–10.

27. J.A. Cattle, A.B. McBratney, B. Minasny. Kriging Method Evaluation for Assessing the Spatial Distribution of Urban Soil Lead Contamination. *Journal of Environmental Quality*, 31 (2002), pp. 1576-1588.
28. Kanevski M., Pozdnoukhov A. and Timonin V., 2009, *Machine Learning for Spatial Environmental Data : theory, applications and software*, EPFL Press, 2009, 377 p.
29. J. Caudeville. Développement d'une plateforme intégrée pour la cartographie de l'exposition des populations aux substances chimiques. PhD thesis, University of Technology of Compiègne, 2011.
30. Van de Kasstele, J.; Stein, A.; Dekkers, A.L.M.; Velders, G.J.M. External drift kriging of NO<sub>x</sub> concentrations with dispersion model output in a reduced air quality monitoring network. *Environ. Ecol. Stat.* 2009, 16, 321–339.
31. Breiman, L. Random Forests. *Mach. Learn.* 2001, 45, 5–32.
32. Ioannidou D, Malherbe L, Beauchamp M, Saby N, Bonnard R and Caudeville J. Characterization of environmental health inequalities due to Polyaromatic Hydrocarbons exposure in France. *Int. J. Environ. Res. Public Health* 2018, 15, 2680; doi:10.3390/ijerph15122680
33. C.B.B. Guerreiro, J. Horálek, F. de Leeuw, F. Couvidat, Benzo(a)pyrene in Europe: Ambient air concentrations, population exposure and health effects, *Environmental Pollution*, Volume 214, 2016, Pages 657-667, ISSN 0269-7491.
34. Nuno Ratola, Pedro Jiménez-Guerrero, Combined field/modelling approaches to represent the air-vegetation distribution of benzo[a]pyrene using different vegetation species, *Atmospheric Environment*, Volume 106, 2015, Pages 34-42, ISSN 1352-2310,
35. Gonzalez, M., Vivanco, M. G., Palomino, I., Garrido, J., Santiagao, ´ M., and Bessagnet, B.: Modelling some heavy metals air concentration in Europe, *Water Air Soil Pollut.*, 223, 5227–5242, doi:10.1007/s11270-012-1274-6, 2012.
36. D. W. Pennington, M. Margni, C. Ammann, and O. Jolliet. Multimedia fate and human intake modelling: spatial versus non spatial insights for chemical emissions in western europe. *Environmental Science and Technology*, (39(4): 1119–1128), 2005.
37. K. Murasawa, T. Sakurai, K. Nansai, K. Matsushashi, Y. Moriguchi, K. Tanabe, O. Nakasugi, and N. Morita M Suzuki. Geo-referenced multimedia environmental fate model (g-ciems): Model formulation and comparison to the generic model and monitoring approaches. *Environmental Science and Technology*, (38(21): 5682–5693), 2004.
38. INERIS, R. Bonnard, Sets of equations for modeling exposure linked to soil contamination or emissions from an industrial facility, 2010
39. R. Bonnard and T.E. McKone. Integration of the predictions of two models with dose measurements in a case study of children exposed to the emissions of a lead smelter. *Human and Ecological Risk Assessment*, (15(6): 1203-1226), 2010.
40. Bois, F.Y. and Paxman, D. (1992) An analysis of exposure rate effects for benzene using a physiologically based pharmacokinetic model. *Regulatory Toxicology and Pharmacology* 15, 122-

41. Gerlowski, L.E. and Jain, R.K. (1983) Physiologically based pharmacokinetic modeling: principles and applications. *J Pharm Sci* 72, 1103-1127.
42. Parrott, N., Jones, H., Paquereau, N. and Lavé, T. (2005) Application of full physiological models for pharmaceutical drug candidate selection and extrapolation of pharmacokinetics to man. *Basic and Clinical Pharmacology and Toxicology* 96, 193-196.
43. Beaudouin R, Micaleff S, Brochot C. 2010. A stochastic whole-body physiologically based pharmacokinetic model to assess the impact of inter-individual variability on tissue dosimetry over the human lifespan. *Regul Toxicol Pharmacol* 57(1):103-116 20122977
44. F.A. Zeman, C. Boudet, K. Tack, A. Floch Barneaud, C. Brochot, A.R.R. Péry, et al. Exposure assessment of phthalates in French pregnant women: results of the ELFE pilot study. *Int J Hyg Environ Health*, 216 (2013), pp. 271-279
45. P. Quindroit, R. Beaudouin, C. Brochot. Estimating the cumulative human exposures to pyrethroids by combined multi-route PBPK models: application to the French population. *Toxicol. Lett.*, 312 (2019), pp. 125-138, 10.1016/j.toxlet.2019.05.007
46. Marie Vanacker, Paul Quindroit, Karine Angeli, et al (2020). Aggregate exposure to pyrethroids: comparison of predicted and measured urine metabolite concentrations and cumulative risk assessment for adults in France. *Food and Chemical Toxicology*. Volume 143, 111519. <https://doi.org/10.1016/j.fct.2020.111519>
47. Brochot C, Casas M, Manzano-Salgado C, et al. *Toxicology and Applied Pharmacology* 379 (2019) 114640.

## Figures

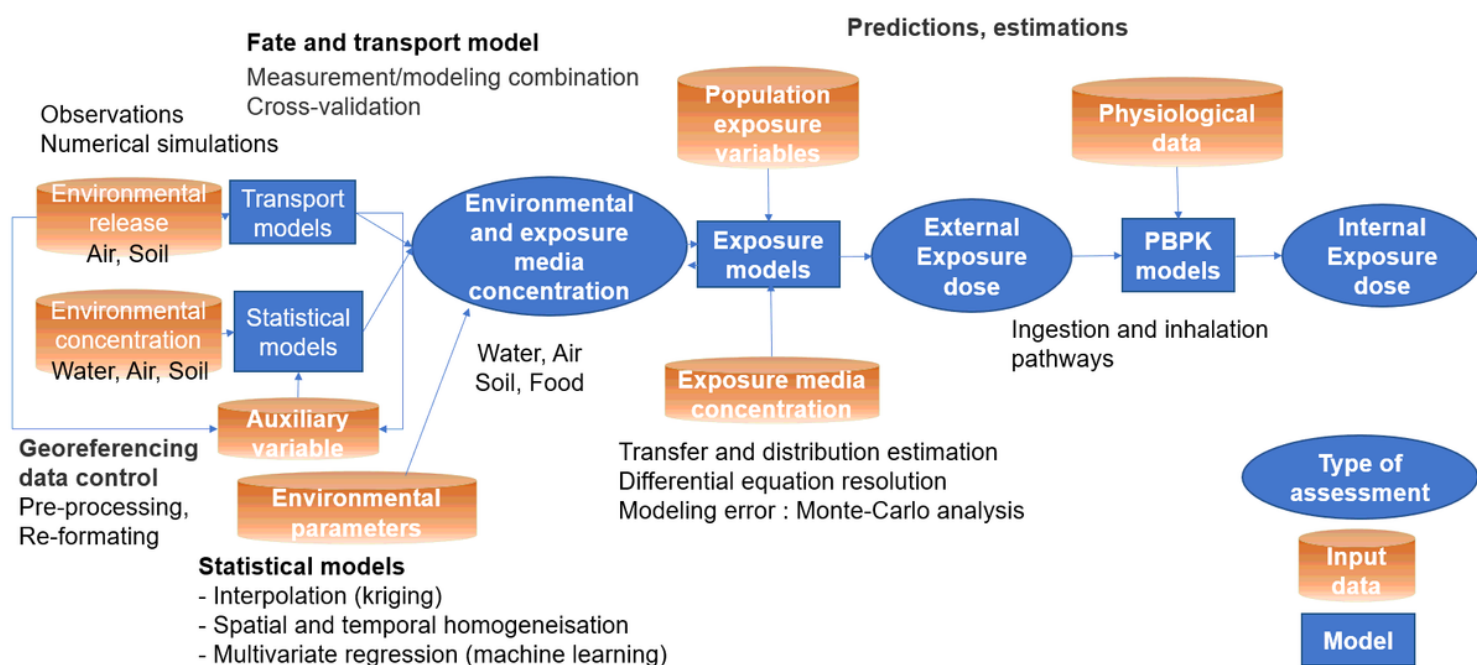


Figure 1

Modeling framework example to characterize integrated EHI

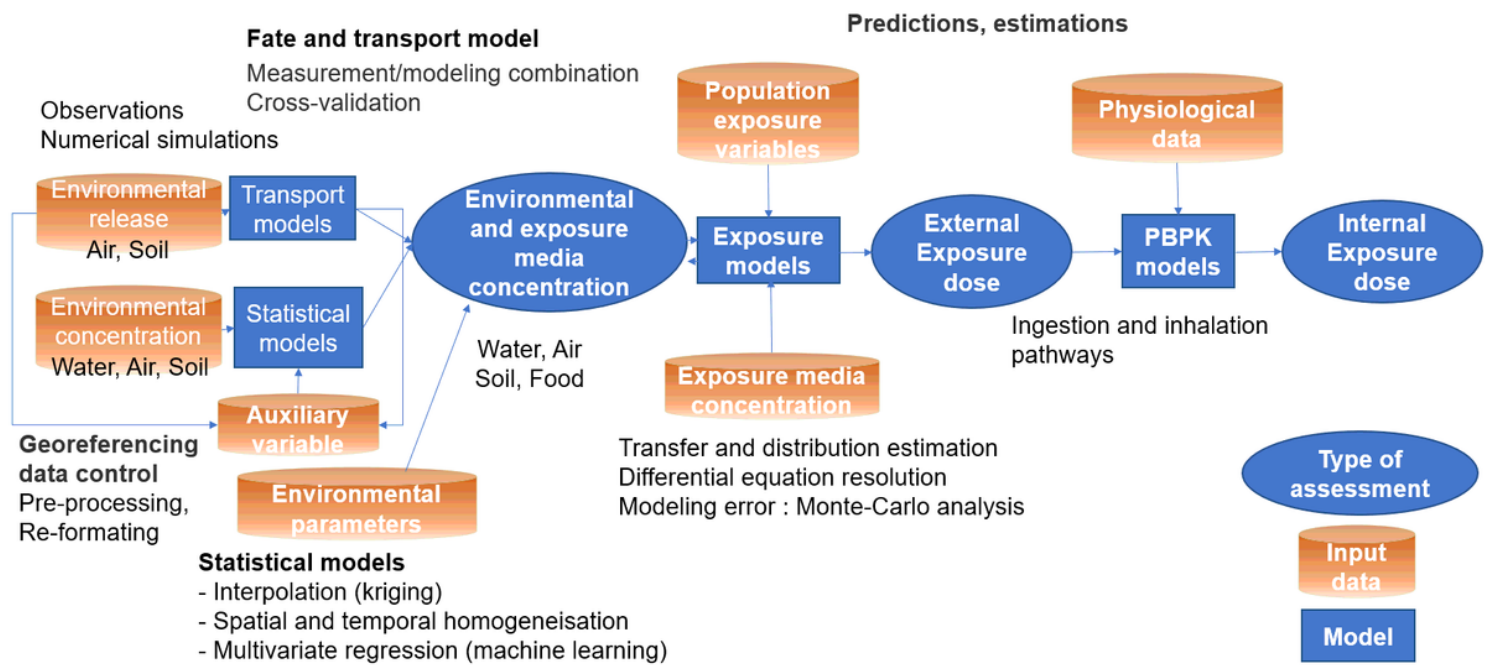


Figure 1

Modeling framework example to characterize integrated EHI