UNIVERSITA' DEGLI STUDI DI PARMA
DOTTORATO IN ECOLOGIA
XXVI CICLO

PhD Thesis

Models for exposure assessment to sources of atmospheric pollution

Supervisors:
Dott. Andrea Ranzi
Dott. Paolo Lauriola
Prof. Giulio A. De Leo

Coordinator:
Prof. Paolo Menozzi

Candidate:
Michele Cordioli

A mia moglie Veronica,
che mi ha sempre supportato (e sopportato)
nel corso di questi quattro anni di lavoro.

AKNOLEDGMENTS

This PhD work would not have been realized without the boost and constant encouragement of Professor Giulio A. De Leo, to whom I express my deepest thanks and appreciation.
My research have been founded and supported by the Regional Reference Centre "Environment & Health" of the Emilia Romagna Regional Environmental Protection Agency (ARPA). Specifically, my thanks goes to Dr. Andrea Ranzi, who constantly and professionally supervised all my work and thesis writing, and to Dr. Paolo Lauriola, who allowed me to have a great learning and professional experience in these four years. A big thank goes also to every one of the ARPA team.
I would like to thank Dr. Kees De Hoogh for his thesis review and very useful comments.
Last but not least, the warmest thanks goes to my wife Veronica, my parents Tiziano and Lucia and to all my large family.

Abstract (English)

Atmospheric pollution is almost ubiquitous and is recognized as an important cause of damage to human health, ecosystems and materials. All the disciplines that study the effect of environmental pollution on human and ecosystem health (e.g. ecotoxicology, toxicology, ecological risk assessment, human health risk assessment, etc.) share a key evaluation step: the definition of the magnitude and duration of the contact between the risk factor and the target receptor, i.e., exposure assessment.

This thesis is focused on models for human exposure assessment to atmospheric pollution. Despite recent advances in GIS, pollution modelling and environmental data handling, the accuracy of the exposure assessment process is often kept in low esteem in published studies. The overreaching goal of my thesis is to draw the attention of the environmental health researchers on the key role of exposure assessment in determining the reliability of risk results. Specifically, the aim of my research was (i) to review available exposure assessment methods, (ii) to define a quality classification framework, (iii) to evaluate the possible effects of poor exposure assessment on risk estimation and (iv) to explore the applicability of various exposure assessment methods in the field epidemiology and risk assessment.

In the absence of a practicable gold standard measure of exposure, it becomes important to evaluate different exposure assessment methods and compare their performance and fields of applicability. In the first section of this thesis I presented a case-study where I showed a fairly good agreement between self-reported and GIS-derived proxies of exposure to environmental pollution in a case-control study on lung cancer.

The heterogeneity of available methods makes it difficult to interpret the results of epidemiological studies on environmental exposures. I thus proposed a classification scheme for the quality of exposure assessment to a point-source emission based on a three-level numerical classification that consider (i) the approach used to define the intensity of exposure to the emission source, (ii) the scale at which the spatial distribution of the exposed receptors is accounted for and (iii) whether temporal variability in exposure is considered or not. This classification was then applied to 42 published studies on health effects of incinerators, highlighting the strengths and weaknesses of each method.

In the presence of exposure assessment errors that are not correlated with the disease status (i.e. non differential), it is generally assumed that the calculated health risk will be lower than "real" risk. I showed with a simulation study that when categorical exposure is poorly characterized, we cannot be confident that, because of non-differential exposure misclassification, the risk we measure is lower than the "real risk" we would measure with a better exposure assessment.

Exposure assessment methods used in published studies on health effect of incinerators generally consider only the inhalation exposure pathway. Risk assessment models can be used to compare different emission scenarios and study the importance of indirect exposure pathways, like soil or food ingestion. I carried out a case-study where I showed that (i) risk assessment can be used to identify emission compensation strategies that reduces human exposure and health effects, (ii) indirect exposure pathways plays an important role for some persistent pollutants, (iii) careful definition of the dietary habits and food origin (i.e. home-grown vs. market food) is essential to conduct adequate exposure and risk studies.

Finally, since the emission source under study is rarely the only relevant emission source on a territory, I proposed the use of Land Use Regression (LUR) models as a suitable tool to take into account intra-urban differences of exposure to diffuse air pollution and to adjust for effect confounding in studies on industrial emission sources.

In conclusion, this work highlight the need for more accurate exposure assessment in many published studies and the key role of models and spatial analysis in enhancing exposure science. Exposure information is crucial for predicting, preventing and reducing human health and ecosystem risks.

Abstract (Italiano)

L'inquinamento atmosferico è pressoché ubiquitario e rappresenta un'importante causa di danno alla salute umana, agli ecosistemi ed ai materiali. Tutte le discipline che studiano gli effetti dell'inquinamento ambientale sulla salute degli ecosistemi e dell'uomo (es. eco tossicologia, tossicologia, analisi del rischio sanitario ed ecologico) hanno un elemento comune fondamentale: la valutazione dell'esposizione, ovvero la definizione dell'entità e durata del contatto tra il fattore di rischio e il recettore di interesse. Questa tesi ha per oggetto lo studio e l'applicazione di modelli per la valutazione dell'esposizione umana all'inquinamento atmosferico. Nonostante i recenti progressi nei sistemi GIS, nei modelli per l'analisi dei fenomeni di inquinamento e la gestione dei dati ambientali, l'accuratezza del processo di valutazione dell'esposizione è spesso tenuta in scarsa considerazione negli studi pubblicati. L'obiettivo generale della mia tesi è quello di porre l'attenzione della comunità scientifica che si occupa di effetti sanitari dell'inquinamento sul ruolo chiave della valutazione dell'esposizione nel determinare la credibilità dei risultati ottenuti. Nello specifico, gli obiettivi del mio lavoro sono stati (i) la revisione dei metodi per la valutazione dell'esposizione utilizzati in letteratura, (ii) la definizione di uno schema di classificazione per la qualità dell'esposizione, (iii) l'analisi dei possibili effetti che modelli espositivi di bassa qualità possono avere sulle stime di rischio ottenibili in uno studio, (iv) l'applicazione di diverse metodiche per la valutazione dell'esposizione nel campo dell'epidemiologia ambientale e dell'analisi del rischio.

In mancanza di una tecnica standard di riferimento per la valutazione delle esposizioni ambientali, risulta importante l'utilizzo parallelo di metodologie diverse, l'analisi dell'applicabilità di ciascun metodo ed il confronto dei diversi risultati. Nel secondo capitolo della tesi ho presentato un caso-studio nel quale ho mostrato la presenza di un buon grado di concordanza tra misure di esposizione ottenute somministrando dei questionari ai soggetti in studio e attraverso elaborazioni spaziali in ambiente GIS, nell'ambito di uno studio caso-controllo sul tumore al polmone.

L'eterogeneità dei metodi di valutazione dell'esposizione rende difficile l'interpretazione coerente dei risultati di studi di epidemiologia ambientale diversi. Nel terzo capitolo ho quindi proposto uno schema di classificazione della qualità dell'esposizione per sorgenti industriali di emissioni atmosferiche basata su una classificazione numerica a tre livelli che considera (i) il metodo utilizzato per definire l'intensità della contaminazione ambientale, (ii) il livello di definizione con cui viene definita la distribuzione spaziale dei recettori di interesse, (iii) se la variabilità temporale nell'esposizione è stata considerata o meno. Questa classificazione è stata applicata a 42 studi sugli effetti sanitari degli inceneritori di rifiuti pubblicati tra il 1984 ed il 2013 ed ha consentito di sottolineare i pregi ed i limiti di ciascun metodo utilizzato.

In presenza di errori di valutazione dell'esposizione non correlati con lo stato di salute dei soggetti in studio (non-differenziali), viene generalmente assunto che le stime di rischio calcolate siano inferiori al rischio "reale" che caratterizza la popolazione studiata. Nel quarto capitolo di questa tesi ho mostrato con uno studio di simulazione che quando le variabili categoriche di esposizione sono valutate con metodi di scarsa qualità, non possiamo essere sicuri che, a causa della misclassificazione non differenziale, il rischio calcolato sia inferiore del rischio "reale" che potremmo misurare con metodi di miglior qualità.

I metodi di valutazione dell'esposizione utilizzati negli studi sugli effetti sanitari degli inceneritori considerano generalmente solo l'esposizione attraverso la via inalatoria. I modelli per la valutazione del rischio possono essere utilizzati per confrontare diversi scenari emissivi e valutare il ruolo delle vie di esposizione indirette (es. ingestione di suolo o alimenti contaminati). Nel quinto capitolo della tesi ho presentato un caso-studio nel quale ho dimostrato che (i) l'analisi del rischio può essere utile nella definizione di misure di compensazione delle emissioni di una sorgente atmosferica che riducano l'esposizione della popolazione, (ii) per alcuni inquinanti persistenti le vie di esposizione indiretta rivestono un ruolo importante e (iii) analisi di esposizione e di rischio di buona qualità devono prevedere un'accurata definizione delle abitudini alimentari delle popolazioni esposte e dell'origine geografica degli alimenti consumati.

Poiché l'emissione industriale oggetto di studio raramente rappresenta l'unica sorgente rilevante di inquinamento presente sul territorio, nel sesto capitolo ho proposto l'utilizzo di modelli *Land Use Regression* (LUR) come metodo utilizzabile per tenere in considerazione le differenze nell'esposizione ad inquinamento diffuso in aree urbanizzate e per ridurre l'effetto di confondimento in studi riguardanti specifiche sorgenti emissive industriali.

In conclusione, questo lavoro ha messo in luce la necessità di condurre analisi di esposizione più accurate in epidemiologia ambientale ed il ruolo chiave della modellistica e delle analisi spaziali nell'aumentare la qualità del processo di valutazione dell'esposizione. Una buona caratterizzazione dell'esposizione è cruciale per predire, prevenire e ridurre i rischi per la salute umana e gli ecosistemi causati dall'inquinamento ambientale.

TABLE OF CONTENTS

1
INTRODUCTION

1.1 Atmospheric pollution: a major environmental problem

Atmospheric pollution may be defined as any situation in which substances are present in the atmosphere at concentrations sufficiently high above their normal ambient levels to cause any measurable effect on humans, animals, vegetation or materials (Seinfeld and Pandis, 2006).

Since it is practically impossible in many cases to define natural background levels of chemicals in the atmosphere, in this thesis I will follow this operational definition of air pollution, which focuses principally on adverse effects.

Atmospheric pollution is caused both by natural or anthropogenic sources. Examples of natural sources are soil dust resuspension, wildfires, volcanoes, volatile organic compounds (VOCs) emission from vegetation. Examples of anthropogenic sources are road transport, industrial activities, waste disposal and treatment and use of chemicals in agriculture. For many atmospheric pollutants, human activities represents the principal source into the atmosphere (Olivier et al., 1998; Seinfeld and Pandis, 2006; Zhang and Tao, 2009).

The problem of air pollution has been recognized and studied for centuries. First examples of regulations on emission of air pollutants are found in the Middle Ages, whereas a more systematic regulatory corpus on air pollution is established starting from the end of the XIX century (Sportisse, 2010). From a scientific point of view, the study of the atmospheric chemistry can be traced back to the XIII century, while from the XIX-XX century the interest shifted from major atmospheric constituents to trace species (Seinfeld and Pandis, 2006). Searching on SCOPUS© for the words "atmospheric pollution" or "air pollution" results in 28,685 published scientific articles between 2010 and 2014, confirming the great scientific interest that the issue still raises.

In spite of the historical interest of policy makers and scientists, nowadays air pollution still represent a major environmental issue in many parts of the world. Atmospheric pollution is almost ubiquitous and is recognized as an important cause of damage to human health, ecosystems and materials.

Pollutants emitted into the atmosphere by human activities move between air, soil, water and food chains and may have significant local and global ecological impacts. Air pollution may impact ecosystems at different ecological levels (Lovett et al., 2009):

- direct effect of pollutants on biological functioning of organisms (e.g., health effects, toxicity, mortality, effects on growth, or reproduction);
- effects on species composition in communities;
- effects on abiotic ecosystem characteristics that are likely to affect the biota over the long term;
- indirect effects in which species are affected through food web or competitive interactions.

Table 1 reports some of the most relevant ecological effects of air pollution, as recently reviewed by the U.S. Environmental Protection Agency. Some effects are expected to occur shortly after exposures to high concentrations (acute), while other may emerge after an extended exposure also to low concentrations (chronic).

Table 1 - Review of Effects of Air Pollutants on Ecological Resources (adapted from Industrial Economics Inc. (2010))

Impact Class	Pollutants	Acute effects	Long-term effects
Acid deposition	Sulfuric and nitric acid Precursors: Sulfur dioxide (SO_2), nitrogen oxides (NO_x)	Direct toxic effects to plant leaves or acquatic organisms	Deterioration soil quality due to nutrient leaching, forest health decline, acidification of surface waters, reduction in pH buffering capacity, enhanced bioavailability of toxic metals, community changes
Nitrogen load	Nitrogen compounds (e.g., NO_x)		Nutrient imbalances, fertilizing effect and eutrophication, changes in global nitrogen cycle, community changes
Ozone	Ozone (O_3) Precursors: NO_x, Volatile Organic Compounds (VOCs)	Direct toxic effects to plants (e.g., leaf injuries, photosynthesis reduction)	Alteration ecosystem energy flows and nutrient cycling, community changes
Hazardous pollutants	Heavy metals, dioxins	Direct toxic effects to plants and animals	Biomagnification and accumulation in the food chain, sublethal impacts, genotoxicity

Concerning human health, air pollution exerts a wide variety of effects, ranging from nausea and difficulty in breathing to reduced lung function and lung cancer (WHO, 2013a). A recent systematic analysis of all major global health risks has found that outdoor air pollution in the form of fine particles ($PM_{2.5}$) is a much more significant public health risk than previously known: ambient air pollution represent globally the 9th health risk factor (the 4th in East Asia), contributing annually to over 3.2 million premature deaths worldwide (Lim et al., 2012). In 2013 the International Agency for Research on Cancer (IARC) classified outdoor air pollution as a cancer-causing agent (carcinogen, Group 1) (Loomis et al., 2013). According to the Clean Air For Europe (CAFE) program, the average damage per ton of $PM_{2.5}$ emitted by human activities in Europe varies between 26,000 € and 75,000 € (Holland et al., 2005).

Emerging countries are experiencing extremely high levels of air pollution in their urban centres: according to a recent analysis from the World Health Organization (WHO, 2014), the world's highest annual average concentrations of particulate matter (PM_{10}) are registered in nations like Pakistan, Afghanistan, Senegal and India (national average annual PM_{10} concentrations respectively equal to 282, 268, 179, 134 $\mu g\ m^{-3}$). On the other hand, in more developed countries the efforts to reduce emissions of atmospheric pollutants through regulation and technological innovation does not always translate into a reduction in atmospheric concentrations. Carslaw et al. (2011) reported the case of nitrogen oxides (NO_x) from vehicles in Europe: after a decrease from the '80s to late '90s, NO_x atmospheric concentrations in Europe has levelled off in recent years, with little additional reduction. This was due

to (i) the discrepancy between emissions over the regulatory test cycle and those in real driving conditions, (ii) the increase in the percentage of diesels in new car sales, (iii) the continuous increase in the power of diesel cars. As an example, the Po Valley in northern Italy remains one of the most polluted areas in the World (Figure 1), due to a high density of emissions and unfavourable meteorological conditions that promote pollution accumulation.

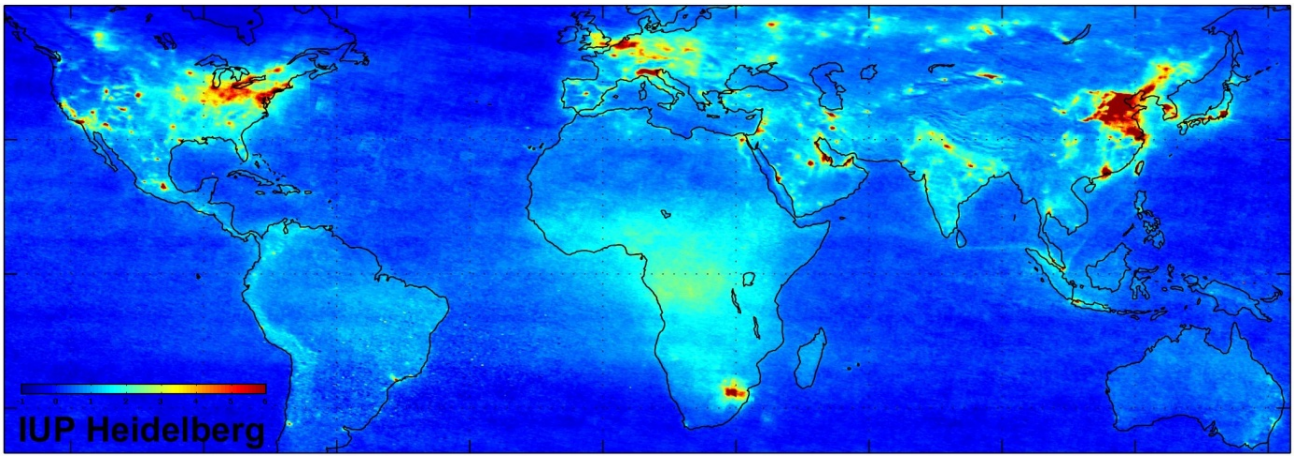

Figure 1 - Global air pollution map: the image shows the global mean tropospheric nitrogen dioxide (NO_2) vertical column density (VCD) between January 2003 and June 2004, as measured by the SCIAMACHY instrument on ESA's Envisat. The scale is in 10^{15} molecules cm^{-2}. Available from: http://www.esa.int/ (accessed on 04/06/2014)

For all the reasons mentioned above, atmospheric pollution represents an important field of scientific research, both in ecology and in public health. Due to its physical properties, the atmosphere is a continuously and rapidly evolving environment, thus the study of atmospheric pollution and its effects on ecosystems and human health poses interesting challenges in the definition of spatially and temporally representative data. Monitoring and mathematical modelling are two interconnected aspects of air pollution analysis that could help research in producing more accurate estimates of concentrations and quantification of effects.

Moreover, atmospheric pollution represents a relevant concern also for public opinion and policy makers. Especially in winter months, local and national news frequently claim that health protection limits for concentrations of fine particles, ozone and other pollutants has been exceeded. Civic and political movements have made the fight against pollution and, in some cases, specific industrial facilities, their warhorse. But what is the real danger resulting from exposure to air pollutants caused by release of chemical plants, or combustion? What are the routes of exposure and what causes the greatest risk? How do we define a good monitoring plan for assessing the possible health risks? The purpose of this thesis is to find some answers to these and other similar questions.

1.2 Ecological and health risk assessment, environmental epidemiology

Different disciplines study the effects of anthropogenic atmospheric emissions on the ecosystem.

Ecological risk assessment (ERA) is defined as the process that evaluates the likelihood that adverse effects may occur or are occurring as a result of exposure to one or more stressors, i.e. any chemical, physical, or biological entity that can induce adverse effects on individuals, populations, communities, or ecosystems (US-EPA, 1992). A risk does not exist until (i) the stressor has the capability of

determining an adverse effect on the receptor and (ii) there is an effective contact between the stressor and the target component of an ecosystem (i.e., exposure).

The process of risk assessment is conventionally divided into at least four phases (Bartell, 2008; US-EPA, 1998a):

- Problem formulation
- Characterization of potential effects
- Exposure assessment
- Risk characterization

The problem formulation phase establishes the goals of the assessment, describes the system under analysis and develops the so called *conceptual model*, i.e. the description of sources, stressors, receptors, exposure routes and expected effects (Figure 2).

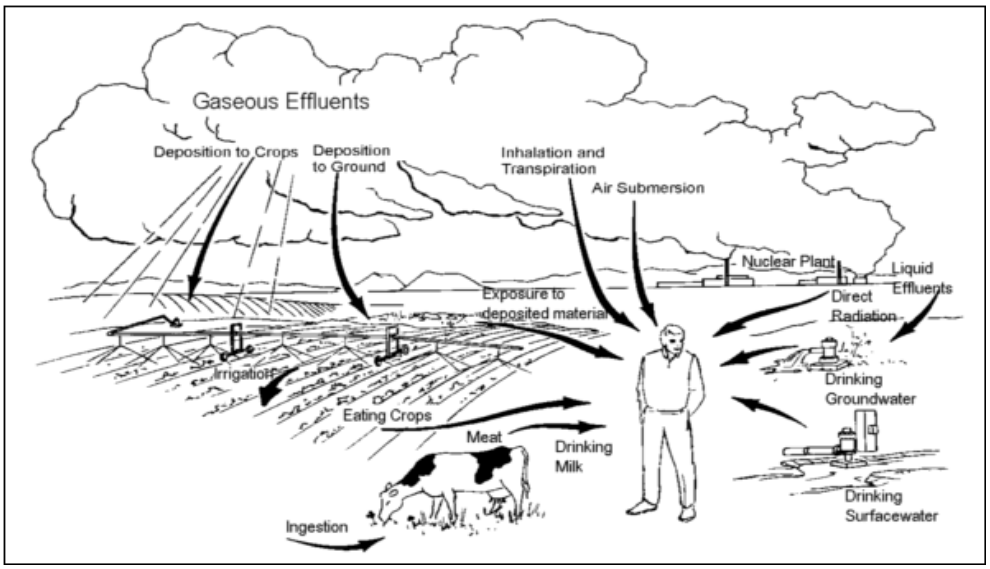

Figure 2 - Example of a generic conceptual model representing potential routes of human exposure to environmental pollutants (source: http://www.atsdr.cdc.gov)

Characterization of potential effect encompasses the analysis of available literature and data that support the evidence of effects caused by the stressor to the target receptor. The output of this phase should be the definition of a quantitative exposure-response relationship that relates the magnitude, duration, frequency, and timing of exposure in the study setting to the magnitude of effects. Possible effects that are evaluated in ERA are alterations in population dynamics (e.g., decrease in population size, impaired reproduction, alterations in genetics, likelihood of extinction), community structure (e.g., number and abundance of species), ecosystem processes (e.g., primary production, respiration, nutrient cycling) or lethal and sub-lethal effects on individual organisms.

In the exposure assessment phase, the temporal and spatial distribution of the stressor and the receptors is described and both are combined to quantify the level of exposure. Through the use of monitoring data and models, an exposure profile is defined for each receptor of interest.

The last phase, risk characterization, include the evaluation of the likelihood of adverse effects occurring as a result of exposure and the quantification of effects on target organisms or populations. One common approach is to compare estimated exposure concentrations with toxicity reference values (e.g., LOEC, LC_{50}, EC_{50}).

Human health risk assessment (HRA) can be seen as a particular form of ERA, where the target receptor of environmental contamination is a particular specie of the community, i.e. humans.

Moreover, the large number and different kinds of ecological effects that are of potential concern in ERA is reduced to direct biological effect on individuals in HRA. In spite of this, HRA frameworks were developed earlier than ERA (Suter, 2008): HRA in the environmental context developed in the United States in the 1970s, and the first comprehensive guidebook on HRA is considered the report *Risk Assessment in the Federal Government: Managing the Process* from the National Research Council (NRC, 1983), while the earliest definition of the ERA framework appears in the 1990s (US-EPA, 1998a, 1992).

From the late 1990s the scientific community started to work on the integration of HRA and ERA in a unique process (WHO 2001; Suter II et al. 2003), although in current practice HRA and ERA are still typically conducted independently. Many reasons support the vision of HRA and ERA as an integrated process: the inherent interdependence of risks to humans and nonhuman species that results from commonalities in the sources and routes of exposure, the similarity across species of many toxic mechanisms, and the fact that quality of human and that of the environment are strictly interdependent (Suter et al., 2005; Vermeire et al., 2007). The major benefits of this integration comes from sharing information, methodologies, data and from the possibility for environmental managers to make appropriate decisions on the basis of a comprehensive evaluation of effects on both humans and ecosystems.

Epidemiology is the science that studies the causes, patterns and occurrences of disease conditions in defined populations. Unlike risk assessment, epidemiology is an observational science: disease risks are estimated from the analysis of disease incidence in controlled/experimental or pure observational studies, thus epidemiology can study only the effects of past exposures. *Environmental epidemiology* is the study of the effects induced by environmental pollution on health and disease in populations (Baker and Nieuwenhuijsen, 2008). Like risk assessment, environmental epidemiology has been developed in two separate fields: the study of effects caused by environmental stressors on animal populations (sometimes called *wildlife epidemiology* or *ecoepidemiology*) and the study of effects caused on human health (*human epidemiology*).

The study of diseases in animal populations caused by environmental pollution can be considered a part of a broader discipline called *ecotoxicology* (Newman, 2008). Ecotoxicology spans a wide range of biological levels: from the study of effects on biomolecular mechanisms to the study of effects on ecosystems and landscape. Many studies about environmental pollution and animal population health are aimed at inferring the possible risks for human health: animals are thus considered as *sentinels* for human environmental health (Reif, 2011). The famous book *Silent Spring* by Rachel Carson (Carson, 1962) relied on reports of bird population die-offs to infer human health threats from pesticides in the environment. A review by Rabinowitz et al. (2005) identified 338 studies published between 1969 and 2002 that assessed both health outcomes and exposures in animal populations in order to investigate possible associations between chemical/physical hazards and health effects. Fish, birds and mammals were the most common animals studied. The majority of the studies were observational, although a relevant number of studies were experimental in nature (i.e., laboratory toxicological experiments, replicate ecosystem and field enclosure study). Observational studies of animal disease induced by pollution in natural populations can provide additional insights not available from laboratory-based studies of experimental animals.

Human environmental epidemiology has developed separately from ecotoxicology. The works of John Snow (1813-1858) on the linkage between water contamination and cholera in London is seen by many as the starting point of modern environmental epidemiology (Baker and Nieuwenhuijsen, 2008). Early epidemiology focuses largely on occupational exposures and extreme pollution events or disasters, while modern epidemiology focuses more on the effects of low exposures spread over large populations. Outdoor air pollution epidemiology is a typical example of the study of the effect caused by exposures that are far lower than those found in toxicological experiments or occupational settings, but affect

almost the entire human population. Since it is considered unethical to conduct experimental studies on humans, human environmental epidemiology is essentially an observational science: the assignment of subjects into an exposed group versus a control group is outside the control of the investigator. Thus, correct reconstruction of population exposure represent one of the more challenging fields of research in human environmental epidemiology (NRC, 2012).

Apart from the choice of the animal species and effect of interest, ecoepidemiology and human environmental epidemiology have their fundamental common element in exposure assessment to environmental contamination

1.3 Merging the gap: the concept of exposure

As mentioned in the previous chapter, HRA and ERA, as ecoepidemiology and human epidemiology, differ principally for the definition of target receptors and possible effects. Nevertheless, the two methodologies share a substantial step, i.e. exposure assessment: exposures of human and nonhuman organisms to chemicals in the environment result from the same sources that release pollutants to the same media, where they are transported and transformed through the same processes. Thus, quantification of stressor emission from sources is common to both, the same environmental physics and chemistry data are applicable to both, and the same transport and fate models and environmental monitoring results can be used to estimate concentrations in exposure media (Suter et al., 2005).

I generically defined exposure in the previous chapter. More precisely, *exposure* is defined as the contact between an agent (i.e., chemical, physical or biological entity) and a target (i.e., a physical, biological or ecological object) in a specific zone of the space during a specific time interval (Zartarian et al., 1997). Thus, the concept of exposure has three dimensions: (i) the amount of a specific agent in the environment, which is usually represented as a concentration level of contaminants in the exposure medium (e.g., air, water, soil), evaluated (ii) in a precise position in space that allow the contact with the surface of the target, (iii) during a specific time interval relevant for the arise of the effect of interest.

It is important to distinguish between environmental concentration, exposure concentration, and dose. *Environmental concentration* refers to the presence of a contaminant in the environment, while *exposure concentration* refers to the concentration of a contaminant in a carrier medium at the point of contact with the target organism. Frequently, atmospheric concentrations are assumed as proxies of the exposure concentrations, under the assumption that the atmosphere is uniformly mixed and the target is surrounded by contaminated air. The *dose* is instead the amount of a contaminant that is absorbed or deposited in the body of an exposed organism for an increment of time. *Potential dose* is the exposure multiplied by a contact rate (e.g., rates of inhalation, ingestion) and assumes total absorption of the contaminant. *Internal dose* refers to the amount of the environmental contaminant absorbed in body tissue, while *biologically effective dose* is the amount of the deposited or absorbed contaminant that reaches the cells or target site where an adverse effect occurs (NRC, 1991). Figure 3 schematically represents the relationship between environmental concentrations, exposure, dose and effect.

All my thesis work is based on the concept of exposure. The focus of my work is human exposure evaluation in environmental epidemiology, as a great part of my research was conducted in collaboration with the Regional Reference Center "Environment & Health" of the Regional Environmental Protection Agency (ARPA). As previously discussed, there is a strict interconnection between exposure assessment in the evaluation of ecological impacts and in the quantification of human health effects (NRC, 2012). The conceptual framework that will be developed for evaluation of exposure assessment quality and the methods used to estimate environmental concentrations remain largely

valid also for ecological risk and impact assessment. The essential difference is in the choice of the final receptors (i.e., human beings vs. nonhuman species, populations, communities or ecosystems) and measures of effect (i.e., health effect vs. impact on population dynamics or ecosystem processes). When possible, references to possible ecological applications of the exposed methods will be highlighted in the various chapters of the thesis.

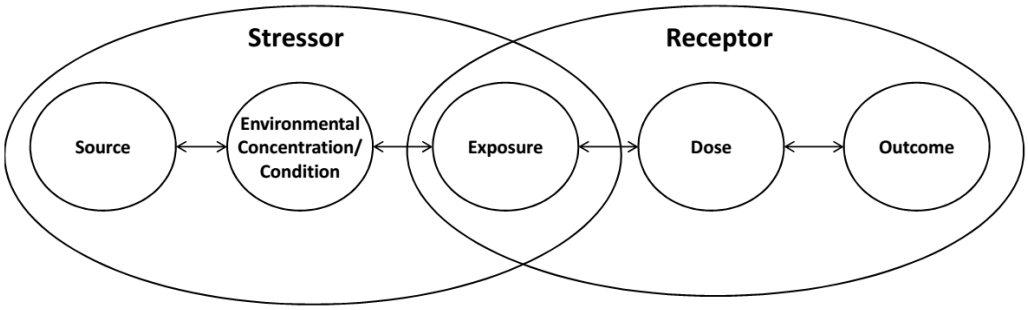

Figure 3 - Schematic representation of environmental-health continuum. Exposure occurs at the point of contact between the receptor and the environment. Dose is the amount of substance entering the receptor organism (from: (NRC, 2012)**).**

1.4 Objectives of the thesis, structure and research questions

The overreaching goal of my thesis is to draw the attention of the environmental health researchers on the key role of exposure assessment in determining the reliability of risk results

Specifically, the aim of my research is a) to review available methods for exposure assessment to atmospheric pollution, b) to define a classification framework for exposure assessment quality; c) to identify the most advanced methods available; d) to evaluate the possible effects of poor exposure assessment on risk estimation and e) to explore the applicability of exposure assessment methods in the field of epidemiology, risk assessment and environmental monitoring.

In many chapter of the thesis I will present an application of the methods described to a real case-study. In almost all the cases, the reference system for my application will be the Province of Parma and, in particular, the waste incinerator that was recently built near the city centre.

The thesis is organized as follows.

In Chapter 2, I present a brief overview of exposure assessment methods, focusing on the fundamental shift from classical epidemiological methods based on self-reported information to modern spatial methods based on the use of Geographical Information Systems (GIS). Objective GIS techniques are expected to be more reliable, nevertheless lack of digital data and historical records still leads to the use of self reported data. A case-study from an epidemiological analysis on lung cancer is presented, where self reported proxies of exposure to atmospheric pollution are compared to objective GIS evaluation, to verify the degree of association between the results of the two methods and emphasize the advantages of combining multiple exposure assessment processes. Part of this chapter has been published in a peer-review journal: **Cordioli M, Ranzi A, Freni Sterrantino A., et al.. *A comparison between self-reported and GIS-based proxies of residential exposure to environmental pollution in a case–control study on lung cancer*. Spatial and Spatio-temporal Epidemiology. 9:37–45 (2014), DOI: 10.1016/j.sste.2014.04.004**.

In Chapter 3, I report the results of a literature review on exposure assessment to industrial sources of atmospheric pollution, taking waste incinerators as a reference type of emission source. What is the state of the art for exposure assessment? Studies based on very simplified assumptions on exposure are still published nowadays, claiming for the necessity of reinforcement of the role of exposure assessment. The aim of the review is (i) to define a classification framework for the quality of exposure assessment to point emission sources, applying it to available published studies on human health effects and (ii) to identify what are the most advanced methods for exposure assessment to industrial point sources. Rarely in the published literature epidemiological reviews have been focused on the quality of exposure assessment. Part of this chapter has been published on a peer-review journal: **Cordioli M, Ranzi A, De Leo GA, Lauriola P. *A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators*. Journal of Environmental and Public Health 2013, Vol.2013, Article ID 129470, 12 pp, DOI: 10.1155/2013/129470.**

In Chapter 4, I examine the effect that the choice of different exposure assessment methods may exert on health risk estimation. The general hypothesis in most of epidemiological studies is that, in presence of exposure assessment errors that are not correlated with the disease status (i.e. non differential), the estimated health risk will be lower than "real" risk. Is this assumption justified in every case? After a brief introduction on exposure error and misclassification, I develop a simulation case-study based on data from the incinerator of Parma and evaluate the degree of error that the application of the different methodologies identified in Chapter 3 may determine in health risk estimation. Part of this chapter has been published on a peer-review journal: **Cordioli M, Ranzi A, De Leo GA, Lauriola P. *A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators*. Journal of Environmental and Public Health 2013, Vol.2013, Article ID 129470, 12 pp, DOI: 10.1155/2013/129470.**

Risk assessment studies were excluded from the review conducted in Chapter 3. Nevertheless, exposure assessment plays a fundamental role also in health risk assessment procedures. In Chapter 5 I more deeply explore how exposure is characterized in published risk assessment studies on atmospheric emissions from industrial sources. Moreover, I report the results of a risk assessment model developed to predict the possible health effects caused by the activation of the incinerator in Parma, trying to answer the following questions: (i) What is the expected health effect from a well-managed modern incinerator? (ii) What is the role of food ingestion pathway in determining the risks and how assumptions about food provenience and consumptions influence risk estimations? (iii) Can the health risk be compensated by reduction in emission from other sources? Part of this chapter has been published on a peer-review journal: **Cordioli M, Vincenzi S, De Leo GA. *Effects of heat recovery for district heating on waste incinerator health impact: a simulation study in Northern Italy*. Science of the total Environment 2013; 444:369–380, DOI: 10.1016/j.scitotenv.2012.11.079.**

Finally, Chapter 6 is focused on exposure assessment to diffuse atmospheric pollution in urbanized contexts. How can we model intra-urban differences in exposure? How does diffuse urban pollution bias the observed relationship between exposure to a specific industrial source and observed health effects? The aim of this chapter is to present a cost-effective method to obtain high resolution estimates of population exposure through the method of *regression mapping*, better known as *Land Use Regression* (LUR) modelling. I apply the method to the case-study of nitrogen dioxide (NO_2) pollution in the Province of Parma, and discuss the issue of model transferability in space and the role of diffuse pollution as a confounder for the effect of emissions from a single point emission source.

2

A BRIEF OVERVIEW OF EXPOSURE SCIENCE: FROM QUESTIONNAIRE TO SPATIAL MODELLING

Part of this chapter has been published in *Spatial and Spatio-temporal Epidemiology*. Full reference: Cordioli M, Ranzi A, Freni Sterrantino A., et al. *A comparison between self-reported and GIS-based proxies of residential exposure to environmental pollution in a case–control study on lung cancer*. Spatial and Spatio-temporal Epidemiology 9 (2014) 37–45.

2.1 Overview of methods for exposure assessment to air pollution

Methods for exposure assessment to atmospheric pollution can be classified based on different criteria (Baker and Nieuwenhuijsen, 2008; Nieuwenhuijsen et al., 2006; NRC, 1991; Zou et al., 2009). A first generally accepted classification divides exposure assessment methods in two broad classes: direct methods and indirect methods (Figure 4).

Direct approaches include:

- Personal exposure monitoring: levels of air pollution are recorded directly by the study subjects wearing active/passive lightweight measuring devices. Questions arise about sensitivity and accuracy of these measuring instruments.
- Biological markers of exposure: chemicals of concern, or a metabolite, are measured in a suitable biological matrix (e.g., blood, urine or tissues). Biomonitoring measures the internal dose of a chemical (*biomarkers of exposure*) or early biological changes (*biomarkers of effect*). Many biomarkers only reflect recent exposure and not cumulative or chronic exposures.

Indirect approaches include:

- Environmental monitoring: levels of air pollution are measured somewhere in the environment where subjects spend their time. Typical examples are fixed ambient air monitors within a city, passive samplers positioned at the address of residence or monitoring devices for indoor air quality. Concentration in the environment is assumed as proxy of inhaled concentration.
- Modelling: this class comprise a vast group of methods, which will be further discussed in the following chapters.
- Questionnaires and diaries: subjects are directly interviewed to collect information about their exposure. Questions may be asked to simply categorize subject as exposed/non exposed (e.g., existence of sources, source use, characteristics of each microenvironment), obtain information on the activity patterns of individuals, residential and work history and provide retrospective information on past exposures. Doubt may arise about the validity of the questionnaire, since subjects may not correctly remember past situation, may not understand the questions, the interviewer may influence the answers (*interviewer bias*) and those with disease may be more likely to report an exposure (*recall bias*).

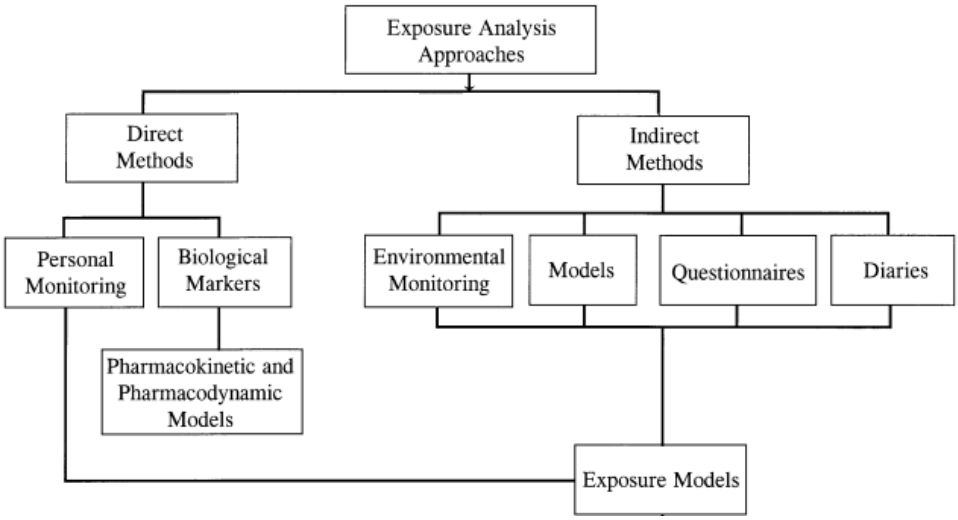

Figure 4 - Overview of exposure assessment methods. Methods are classified as direct and indirect and may be integrated to obtain the final exposure model (source: NRC, 1991)

An overview of available methods for exposure assessment to atmospheric pollution is reported in Table 2, together with some indication about each method's strengths and weaknesses. The following chapters of the thesis will be focused on the use of indirect methods (i.e. exposure modelling).

Exposure models can be divided into two categories: deterministic models and statistical models.

Deterministic models (i.e. physical models) are those describing the mathematical relationship between variables on the basis on knowledge of the physical, chemical and biological mechanisms governing the processes. Examples of this kind of models are *atmospheric dispersion models*: models that solve the mathematical equations and algorithms which simulate the pollutant dispersion into the atmosphere. They process data on emission sources (e.g., emission fluxes, geometric characteristics, emission temperature, etc.) and meteorological data (e.g., wind speed and direction, temperature, atmospheric stability) to yield concentrations of pollutants over space and time. Another group of deterministic models are *multimedia fate and transport models*: models that calculate the distribution of a chemical among different environmental compartments (e.g., air, soil, water, vegetables, dairy products) usually treating environmental media as uniformly mixed, steady state sub-systems. The transport processes are described by simple equations based on measured or estimated parameters to describe transport rates between the different compartments. Human pollutant intake may be estimated using estimated chemical concentrations and data on the contact rate (e.g., inhalation volume, food ingestion rate) and the exposure duration. These types of models are largely applied in risk assessment, while their use in epidemiological research is rare.

Statistical (stochastic) models calculate exposure on the basis of statistical relationship between environmental variables, not requiring the knowledge of the underlying physical and chemical mechanisms. *Interpolation methods* (e.g., inverse distance weighting, kriging, splines) are useful to derive continuous surfaces of pollution concentration from sparse monitoring data. *Kriging*, whose methods were developed in the mining industry during the '50s and '60s, is generally regarded as one of the most robust approach. The majority of the interpolation algorithms estimate the value at a given location using a weighted sum of values at surrounding locations, assigning weights according to *a priori* defined functions. *Kriging* assigns weights according to a data-driven weighting function based on spatial autocorrelation of measured data, and generate an estimation of the spatial distribution of errors. Although *kriging* is one of the more useful interpolation method, its performance in reproducing air pollution surfaces in presence of complex emission patterns (e.g. urban areas) may be poor (Jerrett et al., 2005). *Regression mapping*, also known as *land use regression modelling*, reconstruct the

statistical relationship between observed concentrations and geographical predictors (e.g., road distance, land use, population density, altitude) to predict concentrations in unmeasured sites. Another important modelling technique is *remote sensing*: exposure maps are obtained from the analysis of satellite images coupled to ground-level measurements from stationary monitoring sites.

One of the lead technology for exposure assessment to atmospheric pollution are *Geographical Information Systems* (GIS) (Briggs, 2006; Nuckols et al., 2004). As we described in the previous chapter, the concept of exposure has an intrinsic spatial dimension, thus air pollution exposure assessment relies heavily on a spatial context. GIS have the capability of integrating data into coherent databases that connect object properties (e.g., exposure and health attributes) on the basis of their geographic location and spatial relationships. Figure 5 represents an example of integration of different spatial and non-spatial information in a GIS environment: all data that have an underlying spatial nature (e.g., atmospheric pollution, location of sources, population distribution, socio-economic and health data, etc.) can be handled in a unique environment. The use of GIS has greatly enhanced epidemiological research in terms of definition of study populations, identification of sources and routes of exposure, and estimation of environmental concentrations in the exposure assessment process. By automating spatial analyses, GIS enable calculations to be extended across much larger populations and study areas, leading to more unified analytical approaches, supported by more accurate and detailed exposure data (Briggs, 2006).

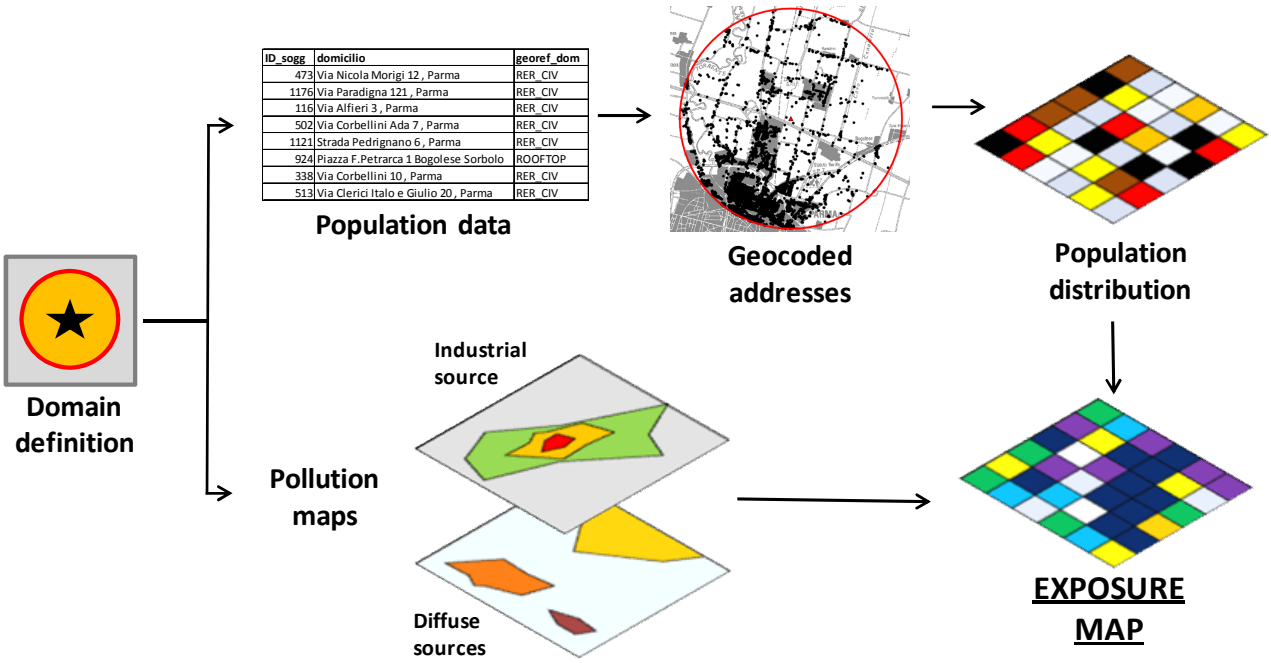

Figure 5 - Schematic representation of data integration in a GIS environment for exposure assessment. Data on receptors (population distribution) and stressors (environmental pollution) are overlaid to calculate the spatial distribution of exposure.

As a closing remark, it is important to remember that "true" exposure is almost impossible to measure. Exposure variables used in environmental epidemiology are generally regarded as approximations (*proxies*) of real exposure. Moreover, there is no a "gold standard" measure of exposure valid in each situation (Forastiere and Galassi, 2005). Firstly, the exposure measure needed may change substantially depending on the study design (e.g., individual vs. group analysis). Secondly, the size of the study population determines how refined the exposure assessment could be: studies conducted on large population spread over large distances impose the use of models, since personal

monitoring is not practicable. In addition, many epidemiological studies are based on routinely collected data or reconstruction of past exposures, without direct contact with the subjects under study: in these cases the use of questionnaire, biomonitoring or personal measurement is not possible. Last but not least, the health outcome and the time window of interest determines the type of exposure assessment needed: some measures of exposure are representative of short-term acute exposures while others are more suitable for determining long-term chronic exposures.

It is thus fundamental to understand the strengths and limitations of each method and, when possible, use multiple exposure assessment methods, linking modelling and monitoring data.

Table 2 - Overview of possible methods for exposure assessment to atmospheric pollution (modified from: Baker & Nieuwenhuijsen 2008)

Method	Input	Strength	Limitations
Questionnaire	Contact with subjects Compilation of a questionnaire	Relatively easy to obtain Information on duration of exposure Information about many health determinants	Lack of objectivity Recall Bias Interviewer bias
Monitoring station	Data from fixed air pollution monitoring station(s)	Easy to obtain from standard air quality monitoring networks	Not measure within-community variability in exposure Station placement may not be representative Provide only concentration estimates
Distance from the source (proximity models)	GIS data (e.g., point source location, road network)	Easy to calculate Applicable to vast populations Reflect exposure from "all" pollutants	Not giving concentrations, only a proxy Spatial anisotropy not considered Assume linear decay of concentrations
GIS-based interpolation (e.g., kriging)	Locations and concentrations from monitoring sites	A further refinement of the distance measure Relatively easy to calculate Provides concentrations Validation studies show good correlation with actual measurements	Requires a dense monitoring network Not adequate in urban settings with complex spatial distribution of emission sources (spatial autocorrelation may be absent)
GIS-based regression modeling (land use regression models)	Locations and concentrations from monitoring sites Geographical predictors (e.g., land use, road distance)	Provides concentrations In complex urban setting refines the representation of spatial gradients	Requires good input data on land characteristics Requires a specific model for each location Quite time-consuming
GIS and air dispersion modeling	Location , emission rates, physical characteristics of sources Meteorological data Mathematical models	Consider the effect of meteorology on air pollution Can reproduce temporal variability and past exposures Provides concentration estimates	Suitability of the dispersion model must be evaluated Many parameters needed Quite time-consuming
Personal monitoring	Contact with subjects Personal monitors	Reflect personal exposure and account for time-activity patterns Incorporate exposure from different sources Useful for validation studies	Very time-consuming and expensive Only applicable for a limited number of subjects, not suitable for large studies Need consent of individuals Does not distinguish sources
Biological monitoring	Contact with subjects Biomonitoring methods	It reflect actual uptake and internal dose Useful for validation studies	Very few good biomarkers Reflect recent exposures Expensive Need consent of individuals Hard to differentiate between exposure pathways and chemicals
Multimedia fate and transport models	Characteristics of sources Meteorological data Chemical-physical properties database Mathematical fate models	Model the transfer of pollutants between compartments (air, soil, water) Estimate pollutant intake and dose	Many parameters needed Time-consuming Many approximations needed

2.2 Case study: a comparison between self-reported and GIS-derived proxies of exposure

During the 1990s many epidemiological studies assessed the effects of atmospheric pollution using proxies of exposure derived from questionnaires (Ciccone et al., 1998; Duhme et al., 1996; Lercher et al., 1995; Rotko et al., 2002; Weiland et al., 1994). In recent years, the use of GIS has emerged as a useful method for exposure assessment to environmental pollution (Nuckols et al., 2004). In this section, I evaluated the degree of association between self-reported exposure to a variety of environmental risk factors and GIS-derived proxies of exposure. Data for the analysis comes from an epidemiological study on lung cancer, conducted in the Province of Modena (Northern Italy).

To date, few studies have investigated the relationship between self-reported and modelled exposure to environmental pollution (Cesaroni et al., 2008; Gunier et al., 2006; Heinrich et al., 2005; Migliore et al., 2009). Published studies have focused on exposure to traffic-related pollution with regard to respiratory symptoms like asthma, cough, bronchitis or rhinitis, while none of these studies have referred to lung cancer.

The aims of this analysis were to evaluate the potential of GIS data in exposure assessment compared with self-reported information, and to highlight the importance of using multiple exposure assessment methods in epidemiological studies where "true" measures of personal exposure are not available. GIS offers many new opportunities to improve the exposure assessment process, nevertheless it is important to provide some evidence of the validity of these new methods.

2.2.1 Study design and population

Lung cancer has been associated with exposure to various environmental risk factors (Alberg and Samet, 2003; Field and Withers, 2012). Although less significant than smoke and other environmental agents (e.g. occupational exposure to carcinogens, radon, asbestos, etc.) outdoor air pollution is considered a possible risk factor in the aetiology of this pathology (Pope III et al., 2011; Raaschou-nielsen et al., 2011; Turner et al., 2008). Many studies suggest relative risks up to 1.5 for high versus low estimates of exposure to air pollution (Boffetta and Nyberg, 2003).

Various methodologies have been applied in studies on lung cancer to assess human exposure to possible environmental risk factors. These include: (i) self reported exposure (Chan-Yeung et al., 2003; Hosgood III et al., 2010; Hosseini et al., 2009); (ii) comparison between subjects living in urban versus rural areas (Curwen et al., 1954; Katsouyanni and Pershagen, 1997); (iii) proximity analysis (Garcia-Perez et al., 2009; Simonsen et al., 2010); and (iv) more sophisticated estimates of air-pollution levels (Hystad et al., 2012; Nyberg et al., 2000; Pope III et al., 2011; Raaschou-nielsen et al., 2011; Turner et al., 2008).

Data on cancer incidence for the Province of Modena (Northern Italy) in the years 2000-2005 showed a possible cluster for lung cancer in the District of Mirandola, where the Standardized Incidence Ratio for males (i.e., the ratio between the observed incidence in the area and the expected incidence in the general population) reached the value of 1.26 (CI95%: 1.13-1.40) (Pirani et al., 2007).

A prospective case-control study (the *IDEALE* project) was initiated in 2010 to investigate the association between environmental risk factors and lung cancer in an area comprising the 9 municipalities belonging to the Mirandola Health District. The case-control study design involved the identification of individuals with (cases) and without (controls) the disease. The level of exposure to air pollution was then evaluated in each group. If the exposure among cases and controls is different, it is

possible to infer that the exposure may be associated with an increased or decreased occurrence of the outcome of interest (Porta, 2008).

A total of 649 subjects were enrolled in this study (case:control ratio of 1:4). Cases were defined as incident events of lung cancer in the period 2009-2010 and controls were coupled on the basis of sex and age. A summary of the main characteristics of the population enrolled is given in Table 3

Table 3 – Summary of the main characteristics of the enrolled population

		n	%
Number of subjects	total	649	100%
	cases	130	20%
	controls	519	80%
Sex	males	504	78%
	females	145	22%
Age (years)	<50	33	5%
	50-70	262	40%
	>70	354	55%
Smoking habits	Smoker	110	17%
	Ex-smoker	351	54%
	Non-smoker	188	29%
Education	None/Primary school	362	56%
	Junior high school	148	23%
	High school	119	18%
	Degree	20	3%

Subjects were interviewed face to face, using a questionnaire designed to collect personal data and information about lifestyle, active and passive smoking habits, food and alcohol consumption, health status, residential and occupational history.

In particular, participants were asked to give information about exposure to environmental pollution at each address of residence since year 1980. The questions, which derived from standardized questionnaire formats (Erspamer et al., 2007; Goldoni et al., 2003; Migliore et al., 2009; SIDRIA, 1997), were the following:

1) *The zone of residence is predominantly: rural/residential/industrial*
2) *The street of residence is: busy/quiet*
3) *Do the majority of windows look out directly onto busy roads: yes/no*
4) *Are there crossroads or traffic lights within 100m of the house: yes/no*
5) *Do you find dust on windowsills: always or frequently/sometimes/never*

2.2.2 Exposure assessment

I evaluated exposure to atmospheric pollution using residence location of each subject (Huang and Batterman, 2000). I geocoded each address reported in the questionnaire (coordinate system: UTM32, datum ED50) through record-linkage by street name and street number to the Regional Database (RDB) of the Emilia Romagna Region. Since some of the municipalities in the study area were not included in this database, I directly geocoded some addresses using a global positioning system (GPS); I geocoded those remaining using free web services (Google Maps and Microsoft Bing).

In case-control studies, exposure history must be reconstructed for each enrolled subject. In the entire study area there was only one fixed air-pollution monitoring stations, so these data were not usable to differentiate the geographic variability of individual exposure. Thus, for each residence I evaluated exposure to possible environmental risk factors by adopting a geographical approach based on GIS (ArcGIS® v.9.3, post-elaboration with Stata SE® v.12 and R v.2.13.1).

I focused on three possible risk factors: land use in the neighbourhood of the residence (Eberhard and Pamuk, 2004; Mitchell and Popham, 2007), road traffic (Cook et al., 2011; Kim et al., 2008; Nuvolone et al., 2011) and industrial pollution (Hendryx and Fedorko, 2011; Luo et al., 2011).

I also collected available information about land use, roads, and industrial emissions for the municipalities within 5 km of the study area, in order to correctly characterize exposure for residences located near the study area boundaries. Past residences of enrolled subjects located outside the study area and over 5 km from it were excluded from the exposure assessment.

Road traffic

Exposure to traffic was defined through road proximity analysis (Bayer-Oglesby et al., 2006; Cook et al., 2011). Since no homogeneous and sufficiently detailed information on traffic flows were available for the study area, I used the Regional road network cartography (RER, 2011), defining the importance of each road based on the functional classification used by the administration. According to this classification, roads were defined as:

- "major" (i.e., highways and important roads used by national or regional traffic);
- "secondary" (i.e., streets connecting city centres or used by cross-city traffic);
- "minor" (i.e., small roads used to reach a specific address or location).

Since there was only one major road in the entire study area, I finally grouped roads into two classes, i.e. (i) "major and secondary roads" (*majsec*) and (ii) "minor roads" (*minor*).

For each residence I computed the following exposure variables (Figure 6-A): typology of the nearest road; minimum distance from a *majsec* road (variable *mindist_majsec*); sum of the length of all streets and *majsec* streets inside a buffer within a radius of 100 m and 200 m around the residence (*all_100/200* and *majsec_100/200*). I used buffer radius slightly larger than typical buffers used in epidemiological studies on traffic (e.g. 25, 50 m) because the study area was predominantly sub-urban/rural, and the majority of the address would have no roads within a short distance.

To test the association between the two exposure assessment methods, I compared the distributions of GIS-derived exposure indices between the groups created on the basis of self-reported proxies of exposure. Since GIS variables showed non-normal distributions, I used non parametric statistical tests: the Wilcoxon rank-sum test (one tailed, $\alpha=0.05$) was used to test the difference between two groups (questions *2*, *3 and 4*) while the Cuzik's test for trend ($\alpha=0.05$) was used to test the presence of a trend in exposure within three groups (question *5*).

Land use in the neighborhood

I characterized the prevalent land use inside three concentric buffers within a radius of 250, 500 and 1000 m around each residence by means of an intersection with the Regional land use cartography (RER, 2011) for the years 1976, 1994, 2003 and 2008 (Figure 6-B). Each map was considered as representative respectively of the periods: <1985, 1986-1999, 2000-2006, 2007-2011.

The original Corine Land Cover categories (EEA, 2000) were divided into 4 new groups: agricultural/natural green areas (*green*), dump/construction sites (*dump*), industrial/commercial (*indu*) and urban fabric/urban green areas (*urban*). The coverage by each land use group was then computed as a percentage (%) of the total area of each buffer so as to define the prevalent land use (i.e., at least 20% more area than all the other three groups) within each buffer for each period.

I then assessed the degree of association between answers to question *1)* and the prevalent land use within each buffer by means of the Chi-square test or Fisher's exact test (α=0.05).

To compare data with the questionnaire responses, I chose the land use map which corresponded to the last year of each period of residence. As a sensitivity analysis I also considered the central year of each residence.

Industrial pollution

I characterized exposure to industrial pollution by obtaining data on the amount of pollutants emitted into the atmosphere by factories and commercial activities around the place of residence (Hendryx & Fedorko 2011; Luo et al. 2011; Willis et al. 2010; Agarwal et al. 2010) (Figure 6-C).

I collected information on authorized emission fluxes (i.e., based on Italian regulations DPR 203/88 and D.lgs 152/2006) of total suspended particles (TSP) and volatile organic compounds (VOC) from the Regional Environmental Protection Agency. Additional information about the start and end date for each activity was obtained from the local Chamber of Commerce. The addresses of industrial plants were geocoded using the methodology described above for residences.

For each residence, and for each year, I computed the total flux of TSP and VOC (tons year^{-1}) emitted into the atmosphere inside three buffers within a radius of 100, 500 and 1000 m. The average height of the emission sources in the area was 10 m, thus atmospheric dispersion over 1 km was considered to be of minor importance. I considered emission fluxes for each industrial plant as constant over time, but I used available information on the start and end date of both residences and industrial activities to modulate the inter-annual variability inside the buffers. For each buffer I then computed the average emission rate over the period of residence (*ptsov100/500/1000).*

I evaluated the association between the answers to question *1)* and *5)* and the GIS-derived exposure (categorized as "presence" or "absence" of emissions in the buffer) using the Chi-square test or Fisher's exact test (α=0.05). I evaluated the presence of a trend in total emission fluxes within the groups defined by answer to question *5)* using Cuzik's test for trend (α=0.05).

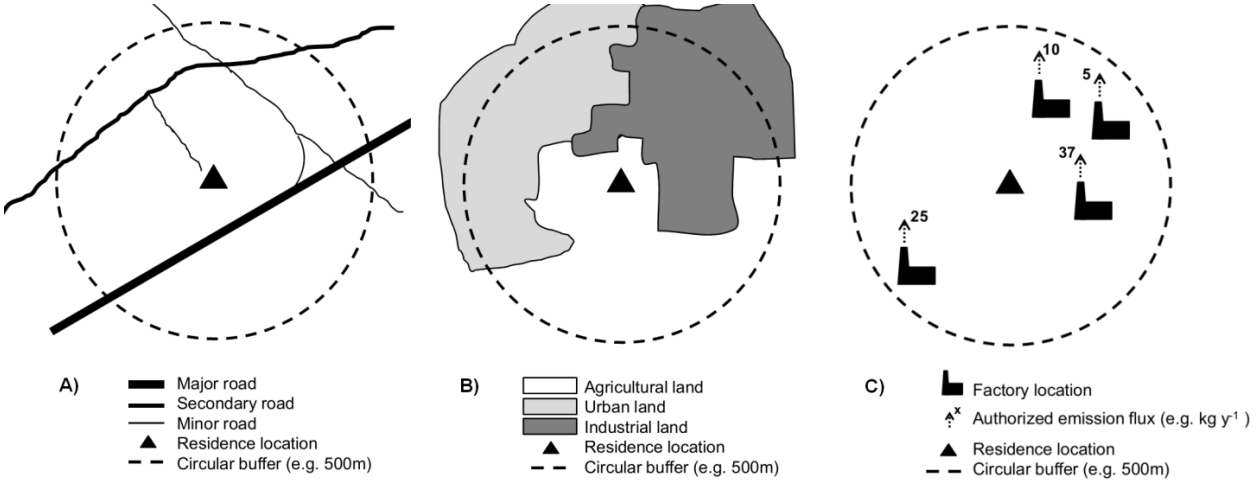

Figure 6 – Schematic representation of GIS exposure assessment: A) length of roads by typology inside a buffer around the residence, B) percentage of buffer area covered by different land uses, C) tons of pollution emitted inside a buffer.

2.2.3 Results

The 649 subjects interviewed reported a total of 838 residential addresses. The population can be considered quite stable: 73% of subjects have never changed residence since 1980, 25% moved once, and 3% had more than 2 residences. 37 residences fell outside the study area and the 5km buffer zone.

A total of 801 residential addresses inside the study area were geocoded, 60% using the Regional database (RDB), 9% using GPS and 30% using free web services. Figure 7 shows the study area and the location of the residences.

To assess the degree of comparability between the results of the three geocoding methods, for one subset of residences I computed the Euclidean distance between the coordinates derived from Google and those derived from the RDB database (n=129) or the GPS (n=50). The median error for the Google-RDB and Google-GPS comparison was respectively 28 m and 35 m, while the 90th percentile was respectively 559 m and 196 m. In some limited cases different geocoding methods yielded very different coordinates (up to 2 683 m distance).

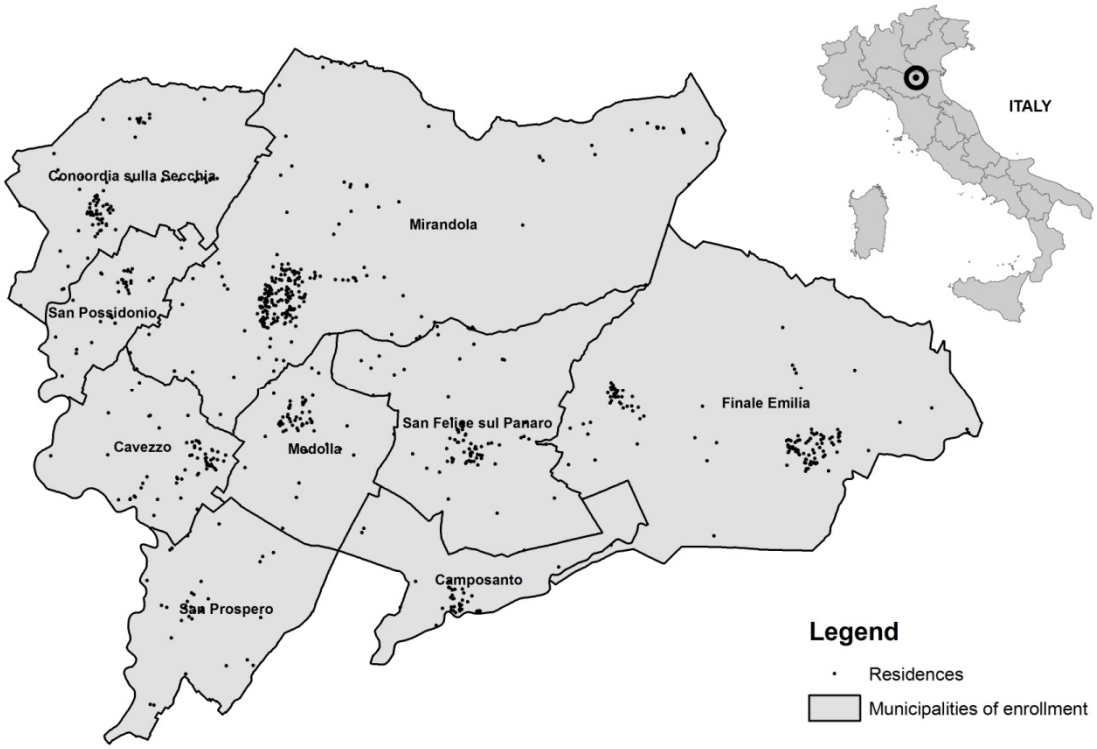

Figure 7 – Study area and distribution of residences in the 9 municipalities of enrolment (dots). Residences are more concentrated around the urban centres of the study area.

Comparison between questionnaire and GIS exposure assessment

Table 4 reports frequencies for the answers to the questionnaire, grouped by the educational level and disease status. Questionnaire answers were not differently distributed between cases and control subjects (Piro et al., 2008), while answers to question *1)* and *4)* were differently distributed within educational levels (i.e., people with higher education more frequently reported residences in "urban" and "industrial" areas and the presence of crossroads).

Table 4 - Frequencies for the answers to the questionnaire, grouped by the educational level and disease status. The p-values refers to the Chi-square test of association.

		Zone (1)			Street type (2)		Windows (3)		Crossings (4)		Dust (5)		
		rural	**urban**	**indu**	**buisy**	**quiet**	**yes**	**no**	**yes**	**no**	**alw/freq**	**somet**	**never**
All		28%	70%	2%	38%	62%	33%	67%	48%	52%	14%	16%	70%
Education	None/Primary sc.	30%	68%	3%	38%	62%	33%	67%	50%	50%	12%	16%	72%
	Junior high sc.	32%	68%	0%	39%	61%	35%	65%	42%	58%	16%	18%	66%
	High school	19%	77%	4%	37%	63%	34%	66%	50%	50%	16%	17%	67%
	Degree	23%	73%	4%	42%	58%	32%	68%	69%	31%	13%	4%	83%
	p-value Chi.		*p=0.03*		*p=0.95*		*p=0.95*		*p=0.04*		*p=0.4*		
Health status	Controls	27%	71%	3%	39%	61%	34%	66%	49%	51%	13%	17%	70%
	Cases	31%	68%	1%	37%	63%	31%	69%	47%	53%	16%	16%	68%
	p-value Chi.		*p=0.41*		*p=0.67*		*p=0.55*		*p=0.65*		*p=0.55*		

Table 5 shows frequencies of GIS-derived prevalent land use inside three buffers grouped by answer to question *1)*. Almost all the residences were classified as *green* or *urban*, regardless of the buffer chosen.

For all considered buffers, frequencies of prevalent land use show very different distributions within the groups defined by answer to question *1* (Chi-square test highly significant). Bold values in Table 2 represent the percentage of matching between self reported and GIS-derived classification of land use. The maximum level of matching for "residential" and "industrial" residences, i.e. 86% and 67% respectively, was observed for the 250 m buffer. An increase in the buffer radius yielded a decrease in the percentage of matching for these land use classes. On the other hand, when the largest buffer was used nearly all the residences (i.e. 99%) declared in the questionnaire as "rural" were correctly classified in the *green* group.

The same calculations were performed using land-use maps for the intermediate year of each residence (Table 5). Overall, the level of matching decreased, especially for the "residential" and "urban" respondents. Residences of more educated subjects yielded slightly higher frequencies of *urban* and *indu* prevalent land use compared with less educated subjects (e.g., for *urban* 81% and 63% respectively for degree and none/primary school), in accordance with self-reported data (Table 4).

Table 6 shows the results for exposure assessment to traffic. The residences were at an average distance of 182 m from *majsec* roads in the area, with an average length of 97 m and 328 m of *majsec* roads respectively within the 100 m and 200 m buffer.

All exposure variables showed highly significant ($p<0.01$) or significant ($p<0.05$) differences between the groups created on the basis of self-reported proxies of exposure, the sole exception being the total length of all streets inside buffers (i.e., *all_100* and *all_200*) for question *2)*. Only for 26% of the residences reported to be on a "busy road" was the nearest street identified with GIS a *majsec* road, whereas 92% of residences reported to be on a "quiet road" were closer to a *minor* road.

For question *5)*, about the presence of dust on windowsills, all variables showed a highly significant trend (Cuzick's test, $p<0.01$) between groups. Compared with residences in which dust is frequently found, residences without reported dust are on average at twice the distance from a *majsec* road. For these residences also the total length of *majsec* roads inside the 100 m buffer is on average 75% lower.

Table 5 – Results for the characterization of the area of residence, considering land-use maps for the last and intermediate year of each residence. For each buffer, the number represents the fraction of residences in each prevalent land-use category (i.e., row percentages). Bold values represent the land-use categories we consider matching with each response to the questionnaire. P-values refer to the Chi-squared test of association.

	Prevalent land use inside buffer - last year of each residence								
	250m			**500m**			**1000m**		
Zone of residence [a]	***green*** [b]	***urban*** [c]	***indu*** [d]	***green***	***urban***	***indu***	***green***	***urban***	***indu***
All (n=801)	31%	65%	3%	44%	54%	2%	74%	26%	0%
Rural (n=222)	**82%**	17%	1%	**91%**	9%	0%	**99%**	1%	0%
Residential (n=561)	12%	**86%**	2%	25%	**74%**	1%	63%	**37%**	0%
Industrial (n=18)	17%	17%	**67%**	33%	22%	**44%**	100%	0%	**0%**
p-value Chi.		p<0.01			p<0.01			p<0.01	
	Prevalent land use inside buffer - intermediate year of each residence								
	250m			**500m**			**1000m**		
Zone of residence [a]	***green*** [b]	***urban*** [c]	***indu*** [d]	***green***	***urban***	***indu***	***green***	***urban***	***indu***
All (n=801)	40%	58%	2%	52%	47%	1%	83%	17%	0%
Rural (n=222)	**87%**	12%	1%	**95%**	4%	0%	**100%**	0%	0%
Residential (n=561)	21%	**77%**	2%	34%	**65%**	1%	75%	**25%**	0%
Industrial (n=18)	28%	22%	**50%**	61%	11%	**28%**	100%	0%	**0%**
p-value Chi.		p<0.01			p<0.01			p<0.01	

[a] responses to question *1)*. Total number of subjects may differ due to to incomplete responses
[b] agricultural/natural green areas.
[c] urban fabric/urban green areas
[d] industrial/commercial areas

Table 6 – Indexes of exposure to road traffic. Mean (standard deviation) are shown for each group. P-values refer to the one-tailed Wilcoxon test for the difference between groups (questions 2,3,4) or to the Cuzick's test for trend (question 5).

			Traffic exposure indexes				
Question	**Answer**	**n** [a]	**all_100** [b] **(m)**	**majsec_100** [c] **(m)**	**all_200** [d] **(m)**	**majsec_200** [e] **(m)**	**mindist_majsec** [f] **(m)**
None		801	588 (234)	97 (106)	1936 (806)	328 (262)	182 (241)
Street type (2)	busy	308	586 (224)	133 (111)	1933 (770)	401 (254)	124 (186)
	quiet	493	589 (240)	75 (97)	1938 (828)	282 (257)	219 (264)
		p-value Wilc.	p=0.54	p<0.01	p=0.52	p<0.01	p<0.01
Windows (3)	yes	264	616 (220)	132 (109)	2050 (780)	410 (256)	124 (186)
	no	526	573 (241)	80 (100)	1874 (816)	287 (256)	211 (261)
		p-value Wilc.	p<0.01	p<0.01	p<0.01	p<0.01	p<0.01
Crossings (4)	yes	387	650 (222)	107 (112)	2189 (741)	373 (275)	141 (158)
	no	411	529 (231)	88 (100)	1698 (793)	285 (242)	222 (295)
		p-value Wilc.	p<0.01	p<0.05	p<0.01	p<0.01	p<0.05
Dust (5)	always/frequently	106	615 (196)	153 (120)	2061 (684)	450 (282)	103 (145)
	sometimes	127	631 (212)	93 (111)	2199 (750)	359 (288)	173 (258)
	never	540	571 (241)	87 (100)	1856 (822)	298 (246)	201 (252)
		p-value trend	p<0.01	p<0.01	p<0.01	p<0.01	p<0.01

[a] Total number of subjects may differ due to to incomplete responses
[b] length of all streets inside the 100m buffer
[c] length of major and secondary streets inside the 100m buffer
[d] length of all streets inside the 200m buffer
[e] length of major and secondary streets inside the 200m buffer
[f] distance to the nearest major or secondary street

I identified a total of 897 industrial and commercial activities with authorized TSP and VOC emissions. Pollution sources were geocoded using the RDB (78%) and Google (22%). Evaluation of exposure to industrial pollution showed that only 5% of residences had authorized TSP and VOC emission sources within a distance of 100 m, while this percentage increased to 34% and 65% respectively for the 500 m and 1000 m buffer (Table 7, classified as "yes"). For the exposed residences, the median levels of TSP plus VOC emissions inside the three buffers was respectively 0.15, 0.17 and 2.4 t year^{-1} (Table 7). All residences with *indu* prevalent land use inside the 500 m buffer had emission sources within the same distance, while this was true only for 39% of *urban* and 24% of *green* residences.

The proportion of residences with TSP and VOC emissions is lower for houses declared as "rural" and higher for houses declared as "residential" and "industrial". The Pearson Chi-squared test confirmed a different distribution of the exposure indices in the three groups defined by answers to question *1)*, with a stronger concordance on "industrial" area when buffers increased. Consistently, the highest median level of emission for each buffer radius was found for residences reported as "industrial" (Table 7). No relation was found between the presence of emission sources and the presence of dust on windowsills (question *5*), except for the 1000 m buffer, where the proportion of exposed residences was slightly higher in the "always/frequently" group. On the other hand, the level of emission around exposed residences showed a statistically significant positive trend (Cuzik's test, $p<0.05$) towards increasing levels of reported dust for the 500 m and 1000 m buffer.

Table 7 – Results for exposure to industrial pollution. Numbers represent row percentages of residences for each group. For exposed residence, the median flux of TSP and VOC in t year^{-1} is shown in parenthesis. P-values refer to the Chi-squared test of association or to the Cuzick's test for trend for pollution fluxes.

				industrial activities exposure indices					
				ptsov100 [b]		ptsov500 [c]		ptsov1000 [d]	
Question		**Answer**	**n** [a]	**no**	**yes**	**no**	**yes**	**no**	**yes**
None			801	95%	5% (0.2)	66%	34% (0.2)	35%	65% (2.4)
Zone	(1)	Rural	222	95%	5% (0.1)	80%	20% (0.2)	59%	41% (0.9)
		Residential	561	97%	3% (0.0)	63%	37% (0.1)	26%	74% (2.0)
		Industrial	18	50%	50% (5.3)	6%	94% (12.1)	0%	100% (24.9)
		p-value Chi/F		p<0.01		p<0.01		p<0.01	
Dust	(5)	always/frequently	106	92%	8% (0.2)	67%	33% (0.5)	25%	75% (7.9)
		sometimes	127	94%	6% (4.1)	62%	38% (0.1)	30%	70% (1.6)
		Never	540	96%	4% (0.1)	67%	33% (0.2)	37%	63% (1.5)
		p-value Chi/F		p=0.29		p=0.58		p<0.05	
		p-value Trend		p=0.49		p=0.04		p=0.01	

[a] Total number of subjects may differ due to to incomplete responses
[b] PTS and SOV emission inside the 100m buffer
[c] PTS and SOV emission inside the 500m buffer
[d] PTS and SOV emission inside the 1 000m buffer

2.2.4 Discussion

This work presented a comparison between self-reported and GIS-derived proxies of exposure to environmental pollution from a case-control study on lung cancer. To our knowledge, this is one of the first studies on lung cancer to have used both questionnaire and GIS to derive proxies of exposure (Erspamer et al., 2007). Previous studies applied questionnaires to derive residential history and GIS to

assess exposure (Bellander et al., 2001; Hystad et al., 2012; Nyberg et al., 2000) but no data on self-perceived exposure to environmental risks was collected (environmental tobacco smoke and occupational exposure excluded).

Some previously published works have addressed the problem of validation of self-reported exposure, but their focus was normally on traffic-related pollution only. Heinrich et al. (2005) reported a relatively low level of agreement between self-reported exposure to traffic and a combination of air-pollution measurement and GIS modelling, with subjective assessment overestimating the modelled exposure, especially in the symptomatic group. Gunier et al. (2006) reported a good correspondence between self-reported traffic exposure and GIS-based traffic-density measurements, while there was no evidence of association between self-reported exposure and levels of Polycyclic Aromatic Hydrocarbons (PAH) in the urine. Cesaroni et al. (2008) analyzed the association between various indices of exposure to traffic and respiratory health in Rome, highlighting a statistically significant association between self-reported exposure and modelled proxies of exposure. Migliore et al. (2009) found no evidence of bias in a validation analysis where they compared self-reported traffic intensity with measured traffic fluxes. Recently, Birk et al. (2011) reported a fair agreement between self-reported noise annoyance and GIS-modelled noise levels in Munich.

Overall, my results showed moderate to good agreements between the information from the questionnaire and exposure indices computed using spatial analyses.

The comparison between GIS-derived prevalent land use around the residence and subjective characterization of the neighbourhood gave good results. In particular, residences reported as "rural" were surrounded by a large area of prevalently agricultural land (i.e., the 1000 m buffer), while the perception of living in a "residential" or "industrial" context was better related to the characteristics of the neighbourhood (i.e., the 250 m buffer). This was expected, since the entire study area was prevalently rural without extensive urbanized zones.

I replicated the analysis which refers to prevalent land use in the last year of each residence, also using land-use maps for the intermediate year between each residence (Table 5). Interestingly, the matching quota was reduced for "residential" and "industrial" responses (i.e., respectively 77% and 50% for the 250 m buffer). In fact, there was a remarkable modification in land use inside the study area during the last years, with an increase in *urban* and *indu* land-use types and a corresponding decrease in the *green* category. For example, between 1976 and 2008, 96% of residences experienced a reduction in the *green* land use inside the 500 m buffer, 91% an increase in the *urban* land use, while 60% experienced an increase in the *indu* land use. The higher level of matching obtained using more recent maps may be due to the fact that people more easily remember the characteristics of the residence in the latter years.

Almost all the GIS-derived indices of traffic exposure were coherently associated with self-reported proxies. Residences located on "busy" roads, with windows exposed to traffic or where crossroads were reported in the neighbourhood, had more *majsec* streets inside the buffers and a minor distance to *majsec* roads. Interestingly, all indices of exposure showed a trend towards the categories of self-reported dust on windowsills. Some evidence of a correlation between the characteristics of indoor dust and outdoor road dust from traffic resuspension has been previously reported in the literature (Kuo et al., 2012; Tong et al., 2000). Nevertheless, my assessment was based on self-reported qualitative data and was insufficient to draw conclusions in this sense.

Finally, residences self-reported as "industrial" proved to be more exposed to industrial TSP and VOC emissions, while no association was found between the presence of dust on windowsills and the presence of industrial sources in the surroundings. Nevertheless, the analyses of quantitative data on emission fluxes suggest a possible association between increasing levels of dust on windowsills and the quantity of TSP and VOC emitted from industries inside the 500 m and 1000 m buffer.

It is difficult to identify a "gold standard" between self-reported information and GIS-derived models (Forastiere and Galassi, 2005). Both methodologies suffer from possible errors. The collection of both exposure and disease data from questionnaires may be prone to so- called recall bias (Coughlin, 1990) and interviewer bias (Wynder, 1994), and this could lead to errors in risk estimations (Kuehni et al., 2006; Rugbjerg et al., 2011). In this sense, the use of GIS is a promising development since it yields more objective exposure readings.

However, GIS modelling is also subject to errors and limitations (Krieger, 2003). In stis study I evaluated the performance of different geocoding systems, with results that are in agreement with data from other published studies on geocoding errors (Duncan et al., 2011; Whitsel et al., 2006). Shifting the position of residences by some hundred meters may lead to errors in the definition of proxies of exposure (Zandbergen, 2007). Moreover, some streets classified as *minor* may have been characterized by intense traffic due to the presence of specific poles of attraction or traffic lights, while incomplete data on some industrial emissions and their evolution in time may have led to overestimating or underestimating exposure. Last but not least, some errors may arise because the GIS data does not accurately represent the real situation, e.g. land use maps have a minimum resolution below which we cannot distinguish land features.

Another shortcoming in GIS use may derive from the absence of information about past exposure. In this study, I had good quality information about past land use in the area and relatively complete records about industrial emission. Using such data, I was able to model the evolution of exposure in time. This would not have been possible if we had used only self-reported information. On the other hand, no data were available on the evolution of the road network over recent years.

The direct comparison between subjective and objective assessment presents certain limitations. To compute the prevalent land use inside a buffer, I used an *a priori* cut-off (i.e. 20% more surface then all the other categories), and it is unlikely that this procedure would have been used in self-assessment. Furthermore, the nearest street identified through GIS was not necessarily the street to which people referred in answering question *2)*, since the actual entrance to the house may have been in another street.

Finally, one shortcoming of my analysis may be represented by the fact that both self-reported and GIS-derived exposure variables may exhibit a certain degree of spatial autocorrelation. Spatial autocorrelation occurs when characteristics at proximal locations appear to be correlated, so that near things are more similar than distant things. The spatial nature of our data may thus bias the results of the statistical tests used (Dale and Fortin, 2009).

2.2.5 Conclusions

The use of GIS in epidemiological research is an interesting challenge for improvement in the quality of analyses and reliability of findings. Nevertheless, it is important to provide evidence of the validity of these new tools with respect to "classical" epidemiological methods. The results of my work showed moderate to good agreements between GIS-derived proxies of exposure to environmental risks and self-reported evaluation from a questionnaire. I was able to highlight some peculiarity of self-reported exposure, e.g. a tendency in interviewed people to refer to more recent land use characteristics and to consider the presence of industrial areas in the immediate vicinity of the home address. Moreover, I identified some agreements between self-reported presence of dust on home window sills and proxies of exposure to road traffic.

Both the methods employed in this study have their specific weaknesses and I could not define which one performed best. Nevertheless, finding a good agreement between different methodologies of exposure assessment is essential to strengthen epidemiological evidence. Even if the information about environmental risk factors in the area were in certain cases incomplete or approximate, the use of GIS represented a powerful tool in characterizing the exposure of residents. When good quality data are available, the use of objective measures of exposure instead of self-reported data is encouraged. GIS and spatial analysis enable considerations about the temporal and spatial dimension of environmental exposure, which subjective evaluation cannot manage/handle. In particular, the use of these spatial proxies of exposure is essential in cases of non-availability of self-reported exposure, not uncommon in epidemiological studies involving a larger number of enrolled subjects.

3

EXPOSURE ASSESSMENT METHODS FOR INDUSTRIAL EMISSION SOURCES: A REVIEW

Part of this chapter has been published on *Journal of Environmental and Public Health*. Full reference: Cordioli M, Ranzi A, De Leo GA, Lauriola P. *A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators*. Journal of Environmental and Public Health 2013, Vol.2013, Article ID 129470, 12 pp.

3.1 Development of a classification framework for exposure assessment quality

Primary pollutants are released directly into the atmosphere from a variety of anthropogenic activities. Among others, industrial emissions raise concern for the possible health effects on populations living in the neighbourhood. Releases from industrial sources can include complex mixtures of substances, depending on processes and input materials. *Point sources* of air pollution can be defined as industrial locations or fixed facilities from which pollutants are released into the atmosphere from a defined process stream, such as a stack (Kibble and Harrison, 2005). Examples of point emission sources are power plants, incinerators, foundries, refineries, etc.

People living near point sources can be exposed through a number of pathways depending on the point source and the type of release. An example of a conceptual model for the emission-exposure pathways is sketched in Figure 8. After the emission from the stack, pollutants dispersion in the atmosphere depends upon a number of physical and environmental variables such as stack height, wind speed and direction, temperature, atmospheric stability, etc. Some gases may undergo various chemical transformations and part of the contaminants may eventually settle down on a variety of surfaces such as soil, vegetation, water, etc. Concentrations in the atmosphere and in soil may be either directly inhaled, ingested or absorbed through dermal contacts or they can enter the agricultural food-chain (Schuhmacher et al., 2006). The actual exposure to potentially hazardous contaminants is thus determined by the time spent by various sectors of the population in different environments (outdoor, indoor at home or at work) and could be due to inhalation, ingestion of contaminated water or food and dermal contact with contaminated vectors (e.g., soil, water) (US-EPA, 2005a). When an industrial emission source releases persistent pollutants in the environment (e.g., dioxins, heavy metals, etc.) ingestion can represent a relevant exposure pathway.

In Chapter 1 I described exposure as a phenomenon with three intrinsic dimensions: (1) the amount of contaminant in the environment, (2) the spatial relationship with the receptor, that determines the contact, and (3) the temporal evolution of this contact. The green-coloured section of the conceptual model depicted in Figure 8 is related to the definition of environmental concentrations. The orange section is more related to the definition of the effective contact with the receptor and the contaminant, determined on the basis of receptor movements in space and time.

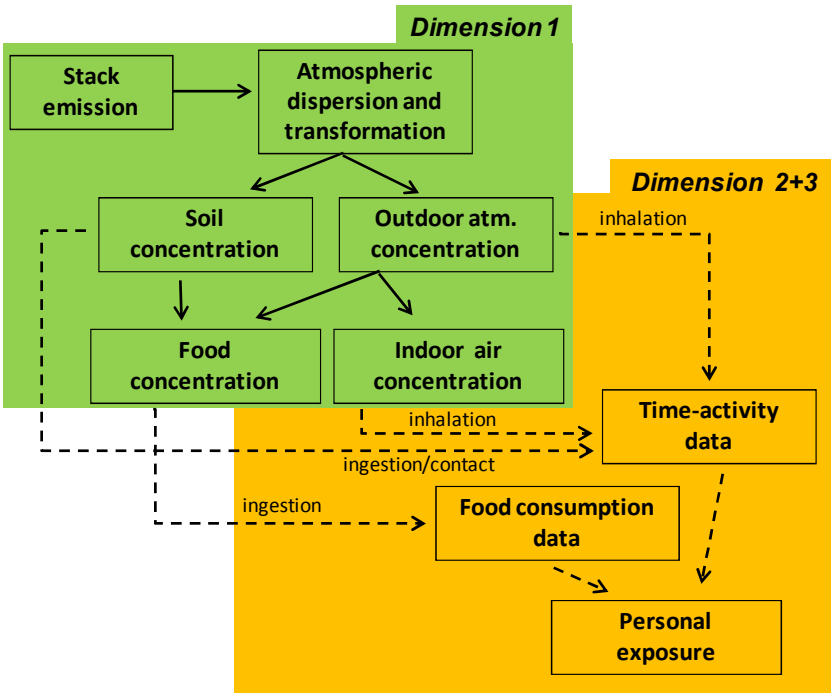

Figure 8 – Conceptual mode representing the principal impact pathways that determine human exposure to atmospheric emissions from a point emission source. The green area regards the first dimension of exposure (contamination intensity), the orange area regards the other two dimensions (receptor spatial distribution and temporal variability). Contamination of drinking water is not represented.

This framework helps me defining a classification scheme useful for determining the quality of the exposure assessment process in an epidemiological study about an industrial emission source, based on three criteria:

1) the approach used to define the intensity of exposure to the emission source;
2) the scale at which the spatial distribution of the exposed receptors is accounted for;
3) whether temporal variability in exposure is considered or not.

For each criteria, it is possible to define a ranked classification of available methods, ordered from the less to the more accurate one.

In the following chapter I will apply this classification scheme to available published studies about the health effects attributable to exposure to waste incinerators. The aims are (i) to investigate what methods and approaches are commonly used in the published literature to characterize exposure levels from waste incinerators and (ii) to test the applicability of the developed exposure classification framework. Although the focus of the review is on waste incinerators, the results of our analysis can be extended to any point source of atmospheric pollution or more generally to industrial sites, where the presence of multiple sources has to be taken into proper account.

3.2 Literature review on incinerators

Incineration is one of the most common technologies for waste disposal (Eurostat, 2010). The number of incineration plants in Europe has been constantly rising in the last years, in an effort to manage and treat an ever–increasing waste production according to the EU directives and minimizing landfill disposal (Saner et al., 2011). As waste incineration releases chemicals that are potentially toxic in the atmosphere (Reis, 2011), there is increasing public concern about the possible adverse effects on human health caused by this waste management technology (Allsopp et al., 2001; BSEM, 2008).

The literature on health effects of waste incinerators is extensive and can be essentially classified into two groups: observational studies (i.e., epidemiological analyses) and simulation studies (i.e., health risk assessment). The first group includes studies that make use of a variety of statistical techniques to describe the potential relationship between the *observed* health status of the population and the exposure level from incinerators. The second group includes studies aimed at estimating the *expected* impact, in terms of health risk and/or number of sanitary cases, of a measured or simulated exposure to environmental contaminants (Cordioli et al., 2013; Forastiere et al., 2011; Roberts and Chen, 2006). Available epidemiological studies have been well reviewed in many published papers (Franchini et al., 2004; S. S. Hu and Shy, 2001; Porta et al., 2009) and reports published by international agencies (DEFRA, 2004; WHO Europe, 2007). However, the lack of a common framework for study designs make the results of the different investigations on the health impacts hard to compare and thus inconclusive. Poor exposure assessment is claimed as one of the main reasons of inconsistent results in published studies (Franchini et al., 2004; Porta et al., 2009; Reis, 2011; WHO Europe, 2007).

The studies reviewed in this chapter, rather than defining a relationship between a single atmospheric pollutant and human health, aimed at evaluating the possible association with a specific industrial source of pollution (i.e., incinerators). Although direct measures of exposure (i.e., personal monitoring or biomonitoring) can be considered the best measures for assessing the effect of a specific mix of substances on the target population, indirect measures of exposure (e.g., simulations of atmospheric dispersion) have greater utility for source emission assessment and control, since they are capable of linking population health to specific pollution emission sources (NRC, 2012). Moreover, if we excluded emission from incidents or disasters (e.g., Bhopal, Seveso, Chernobyl of Fukushima), environmental pollution from single industrial sources respecting emission restrictions is generally low and hardly mesurable at receptors with direct methods.

3.2.1 Materials and methods

In this chapter I analyzed papers referenced in previously published reviews on incinerator health effects (DEFRA, 2004; Franchini et al., 2004; Giusti, 2009; S. S. Hu and Shy, 2001; Porta et al., 2009; WHO Europe, 2007) and, additionally, searched for further references on *MEDLINE*, *PubMed*, *Scopus* and *Google Scholar* by using a number of keywords combinations (e.g., "epidemiology", "incinerator", "adverse effect", etc.). I focused my analysis only on observational epidemiological studies. Human biomonitoring (Chen et al., 2004; Sampaio et al., 2007) and risk assessment studies (Cordioli et al., 2013; Forastiere et al., 2011; Roberts and Chen, 2006) were not considered here. Human health risk assessment will be the focus of the following Chapter 5. I excluded also studies on incinerator's workers (Porta et al., 2009) as the exposure pathway and levels can be completely different from those experienced by the population living around the incinerator plants.

I reviewed the selected literature focusing only on the approaches used to define the exposure level and classified them on the basis of three criteria described in paragraph 3.1. After a first analysis of the methods employed in the studies, I defined a scoring scheme with 18 possible methods of exposure assessment (Table 8), hereafter referred to as "*x.y.z*", where *x* represents the method used to estimate expected intensity, *y* the method used to estimate population distribution and *z* whether the exposure was variable in time or not. For example, a published study classified as "2.3.1" means that the exposure level was evaluated as a function of the distance from the source, population distribution in the territory was assessed by using exact residential address location and exposure was fixed in time.

Exposure assessment methods were categorized only on the basis of the exposure variables actually used in the epidemiological model. As discussed afterward, some studies reported additional

information (such as measured concentrations of pollutants in various media) useful to interpret or support exposure model outcomes even though this information was not used in statistical calculations.

Another important element of the exposure assessment process is the control of confounding factors, i.e., variables that may hide or enhance the measure of effect (Blair et al., 2007; Sheppard et al., 2012). These factors can be socio-economic (e.g., people living in industrial areas near incinerators may be more deprived) or environmental (e.g., frequently incinerators are located in areas with high pollution from other industrial sources and traffic). For each reviewed study I analysed also whether and how confounding factors were accounted for. Since evaluation of confounding factors can follow a variety of approaches, I decided not to include this aspect as a fourth criterion in our classification scheme. Nevertheless I thoroughly comment the role of confounding factors as well as their importance in epidemiological studies in the discussion.

Table 8 – Proposed classification of exposure assessment methods. The three numerical categories are joined in a unique index in the form "x.y.z".

Criterion 1: Definition of exposure intensity	
Category	**Description**
1.	Qualitative (e.g., presence/absence of the source/contamination in an area)
2.	Distance from the source (e.g., linear distance)
3.	Dispersion models (e.g., average annual atmospheric concentration)
Criterion 2: Definition of population distribution	
Category	**Description**
.1	Municipality/Community/Postcode sector
.2	Census unit/Full postcode
.3	Exact residential address location
Criterion 3: Temporal variability	
Category	**Description**
.1	Time-invariable (i.e. fixed) exposure (e.g. at enrolment)
.2	Time-variable exposure (e.g., residential history and/or variability in emissions from the source)

3.2.2 Results of the review

A total of 42 studies published between 1984 and 2013 were identified by the literature search. Table S1 in APPENDIX A reports the resulting categorization of exposure methods and other relevant information for each study. The column "covariates" lists the confounding factors that were evaluated in each study.

Figure 9 represents the evolution of methodologies in time, based on the year of publication. Methods on the y-axis are sorted from the less precise to the best one.

With reference to first classification criterion, i.e. method used to assess exposure intensity, 19 studies (45%) used a measure of distance, both on a continuous scale or more commonly by defining concentric areas with arbitrary radius. In some cases (Barbone et al., 1995; S. Hu and Shy, 2001; Mohan et al., 2000; Shy et al., 1995) also wind direction was used to introduce some spatial anisotropy in exposure. Lee et al. (Lee and Shy, 1999) used distance to define exposed communities, but developed

also a longitudinal study using daily PM_{10} measurement from fixed air monitors. One study (Viel et al., 2000) analysed spatial clustering of disease cases: since the analysis was based on the position of the community of residence, I classified this method as 2.1.1. One study (S. Hu and Shy, 2001) presented multiple assessment methods: presence/absence of the incinerator, distance from the plant and an exposure index based on distance, wind direction and time spent outdoor by people.

12 studies (28%) used atmospheric dispersion models to define population exposure. Generally models were used to estimate long-term average atmospheric concentrations at ground level, although one study used cumulated depositions (Goria et al., 2009). Two studies (Cordier et al., 2004; Ranzi et al., 2011) also used heavy metals as indicator of exposure, one study used PM_{10} (Candela et al., 2013), while all the other used dioxins.

The remaining 11 studies (26%) used a qualitative definition of exposure to contrast the health status of communities/municipalities with and without incinerators. One study (Fukuda et al., 2003) developed quantitative indicators to classify municipalities, using emission inventories for dioxin from incinerators.

All the published studies used the residence as the place where exposure to atmospheric pollution occur (criterion #2). Nevertheless, different levels of detail were used in defining residence location. The majority of the papers (n=19, 45%) considered the municipality or community of residence (e.g., postcode sector, school, hospital, etc.), 13 studies (31%) used the exact geographic coordinates of the address of residence and 10 (24%) used the full postcode or census unit.

Finally, all the published literature, with two exceptions (Candela et al., 2013; Zambon et al., 2007), defined exposure proxies that did not account for temporal variability in population spatial distribution and incinerators' emissions (criterion 3), i.e. they considered the residence at the time of diagnosis, at enrolment or the longest residence of the subject. Residential histories and changes in exposure intensity (e.g., as a consequence of changes in combustion and gas depuration technologies) were not accounted for in the other examined studies.

Overall, Figure 9 shows a trend of improvement in the quality of exposure assessment during the examined years, although three studies published after 2010 still used linear distance as the exposure proxy.

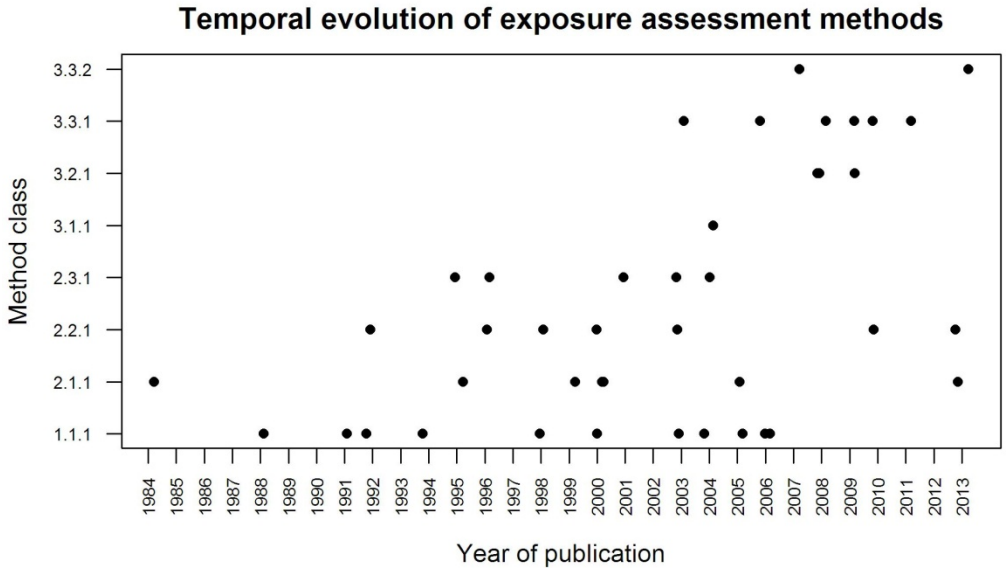

Figure 9 – Temporal evolution of exposure assessment methods. Methods are classified according to Table 1 and sorted in the y axis from the less precise to the best one.

3.2.3 Discussion

Evaluation of exposure intensity (Criterion #1)

The majority of the papers reviewed in the present study appear to suffer from poor exposure characterization. A relevant part of these papers (26%) used qualitative definitions of exposure (e.g. presence/absence of the source or anecdotic presence of pollution). These methods cannot account for the complexity of impact pathways described in Figure 8 nor for the heterogeneity in the exposure level that is normally expected as a consequence of the uneven distribution of the resident population and of the anisotropic dispersal of pollutants in the atmosphere.

Epidemiological analyses carried out on a significant number of municipalities still represent a valuable instrument for public health tracking, since they can identify possible disease clusters in some regions that must be studied further. Even though any departure of disease incidence in large communities from background levels has to be taken very seriously, it is very difficult to use this type of evidence to infer about the role of specific emission sources (i.e., an incinerator), as many other potential confounding factors might exert a significant effect, particularly in highly urbanized areas. Moreover, the risk of false positive and, to a greater extent, false negative results, common to all exposure assessment methods, can be exacerbated when epidemiological data are averaged out on a vast territory with large internal differences in the exposure levels, as in method 1.1.1.

Almost half of the studies used distance to measure exposure. This is certainly a substantial improvement with respect to just an absence/presence evaluation, as contamination from an atmospheric emission source (e.g. air, soil, locally produced food) is generally expected to decrease with distance. However, the assumption of isotropy in atmospheric dispersion of contaminants could lead to errors in exposure assessment. Many features of the emission source (e.g., stack height, gas flow temperature and velocity, pollutant concentration) and of the local environment (e.g., local meteorology, topography, land use) determine where and how far stack emissions disperse how ultimately enter different environmental compartments. In many circumstances, ground level concentrations from a stack may first increase with distance (because stack height prevent the pollutants to fall in the immediate surroundings) and then decrease again, leading to errors when using distance as a measure of exposure.

The use of well-tuned atmospheric dispersion models allows a substantial improvement in the estimation of exposure level, especially if carried out along with a fine scale estimation of the spatial distribution of the vulnerable population. Anyway, atmospheric pollution models are themselves affected by a considerable level of uncertainty (Rao, 2005) depending upon assumptions on actual atmospheric conditions, reconstruction of wind fields, type of dispersion processes, including the possibility of simulating chemical transformation which are known to be highly relevant for the formation of tropospheric ozone and secondary fine particulate matter.

A significant number of the published papers analysed in the present study provided only a limited information on the atmospheric model used: generally there was no discussion about the type of model used, the type and source of meteorological data, model adequacy to represent complex morphological natural or urban landscape and/or wind calms, the assumptions made about pollutant's emission rates and physical-chemical properties.

Only few studies explicitly acknowledged limitations in the modelling approach used. For example, instead of adopting a different dispersion model as suggested by the same authors in a previous study (Floret et al., 2006), in Viel et al. (2008) a part of the study area was excluded because dispersion model results were judged unreliable in that area. Another study (Tessari et al., 2006) used maps of ground level concentrations estimated on the basis of emissions and meteorological data, but no dispersion model was cited. Almost all the studies used dioxins as an impact indicator: dioxins represent a family

of 210 congeners, each one with different physical-chemical characteristics: no study clearly explained how these chemicals were treated in the model (e.g. using 2,3,7,8-TCDD congener properties). Moreover, some studies did not report a clear definition even of the most basic variables used to measure exposure, i.e. averaging time for concentrations (Cordier et al., 2004; Floret et al., 2003; Viel et al., 2008b) or distinction between concentrations and depositions to ground (Viel et al., 2008b). Ground level atmospheric PM_{10} concentrations and depositions from a point source have similar patterns with some significant departure, nevertheless the choice of one or the other measure of exposure should be at least discussed, related to the main route of exposure considered. All this information are important in order to judge the quality of the exposure assessment process, its uncertainties, and to allow comparability and reproducibility of methods.

Regardless of how detailed, accurate and advanced the model is to simulate atmospheric dispersion, it is still only a part of the impact pathways described in Figure 9. All the studies implicitly assumed that inhalation represent the principal exposure pathway, while no published literature measured or modelled the possible exposure through ingestion of contaminated food or contact with contaminated soil.

No study used measured levels of pollution in different media (e.g., atmosphere, soil, food) as the exposure variable in the epidemiological model, except for one work (Lee and Shy, 1999) that used also measured 24h average PM_{10} concentrations in each community as a predictor for pulmonary function, although there were no differences in average levels between communities defined *a priori* as exposed and not exposed. Many studies presented information on measured levels of pollution (Gray et al., 1994; Parodi et al., 2004; Tango et al., 2004; Viel et al., 2000), but these data were not included in the statistical model. This is not surprising, as it is very difficult to discriminate between the contribution of single point sources to the observed concentrations levels. The latter, in fact, invariably depends on the contribution of several other confounding emission sources (Caserini et al., 2004; Nadal et al., 2002), especially if they are located in urbanized areas with intense traffic or industrial activities. Thus, indirect measures of exposure obtained through modelling represent a valid alternative useful to identify the possible role of a specific emission source.

Evaluation of receptor's exposure (criterion #2 and #3)

The actual exposure of an individual to the pollutants emitted by an incinerator may occur in different environments and last a variable amount of time. All published studies used the residence as the place where exposure to atmospheric pollution occur (criterion #2). Notably, one study (Vinceti et al., 2008) considered also the location of workplace of studied subjects.

Residence location can be determined with various degrees of precision. The majority of revised studies (45%) used community level to determine residence location (i.e., town, municipality, postcode sector, school). In this way the same exposure level is assigned to large groups of population, but this assumption was rarely discussed and no measures of exposure variability inside groups was reported. Thus, it was impossible to evaluate the degree of ecological bias (Shaddick et al., 2013), i.e. how well the variation in risk between groups with different average exposure apply to the variation in risk between individuals.

Some studies used census block or full postcode for determining residence position. The dimension of these blocks may vary greatly depending on the location: normally these blocks are smaller in populated areas but may become very large in other rural zones. Moreover, no information were generally given about blocks extension, and it was difficult to compare very different blocks type like the French census blocks (Goria et al., 2009; Viel et al., 2008b), UK census postcode system (Williams and Lawson, 1992) or UK Lower Layer Super Output Areas (LSOA) (Reeve et al., 2013).

The most precise way to locate residences is address geocoding: this procedures assign a couple of geographic coordinates to each address. Errors in address positioning depend on the quality of the

database used, but is generally in the order of tens to hundreds meters (Duncan et al., 2011; Zandbergen, 2009), thus small in comparison with the use of census blocks or full postcode.

In future studies maximum disaggregation of data, to maximise information and minimize potentially differential ecological biases (Diggle and Elliott, 1995), is thus recommended.

The use of residence as exposure location is a very common assumption in environmental epidemiology, since it is easily derived and there is evidence that people normally spend a great part of their time inside their residences, e.g. on average 69% (Klepeis et al., 2001) and 80% (Wu et al., 2011). Nevertheless, home location may not well represent total exposure because people may experience shorter but more intense exposures outside home and residence is a proxy only for inhalation exposure and does not account for indirect pathways (Huang and Batterman, 2000) (Figure 8). Although this technique has well-known limitations, it is often the only method available, particularly for large populations or for reconstructing historical exposures.

Temporal variability in exposure is an issue rarely explored in the reviewed studies. Temporal variability may results both from changes in source emissions over time or from residential mobility of the population and may be a cause of incorrect exposure assignment (Canfield et al., 2006; Meliker et al., 2005). Zambon et al. (2007) explicitly account for historical exposure variability by reconstructing residential histories and evolution of dioxin emissions from the sources considered. However the exposure indicator chosen (i.e., the average exposure over time) may introduce some bias: since emissions were drastically reduced during the '90, the longer a subject was exposed the lower resulted the average exposure. A better indicator could have been cumulative exposure. Candela et al. (2013) studied adverse pregnancy outcomes using a weighted average of monthly values of PM_{10} from the incinerator during the 9 month of pregnancy as exposure indicator. Mother residence at birth was used as exposure location and population mobility was not considered. One study (Ranzi et al., 2011) considered the modification of incinerator emissions over time indirectly, without considering changes in the final statistical model, but evaluating how the morphology of fallout maps was similar in time.

Although difficult to achieve because of data unavailability, especially for studies on old incinerators, in future studies efforts should be developed in reconstructing residential histories and variability in incinerator's emission over time, at least as a sensitivity analysis for the model.

Confounding factors

Another issue that is only partially dealt with in reviewed literature is confounding. Confounding occurs when a risk factor different from the exposure variable under study causes bias in the estimation of association between exposure and disease, due to its differential distribution in exposed and non exposed groups (McNamee, 2003).Various confounding factors may affect a study on incinerators' health effects, i.e. socio-economic differences (e.g., poverty, occupation), personal lifestyles (e.g., alcohol, smoke) and presence of other sources of pollution.

Many reviewed studies did not account for any confounder in the epidemiological model (Bianchi and Minichilli, 2006; Biggeri and Catelan, 2006, 2005; Comba et al., 2003; Rydhstroem, 1998; Tusscher et al., 2000; Williams and Lawson, 1992; Zambon et al., 2007). Some studies collected information about personal lifestyles or socio-economic status directly through questionnaires (Barbone et al., 1995; Biggeri et al., 1996; Floret et al., 2003; Hsiue et al., 1991; Mohan et al., 2000; Shy et al., 1995). Unfortunately the use of questionnaires and surveys is unfeasible for large populations, thus a large part of the studies did not consider personal lifestyles but included socio-economic indicators (e.g., deprivation indexes) evaluated at municipality/census block of residence (Candela et al., 2013; Cordier et al., 2010, 2004; Elliott et al., 1996, 1992; Federico et al., 2010; Goria et al., 2009; Ranzi et al., 2011; Viel et al., 2008a). These indexes are generally constructed based on census statistics.

Of particular concern is the general lack of information about environmental confounding. Many of the pathologies under study have been associated with various atmospheric pollutants (e.g., PM_{10}, NO_x, etc.) or specific anthropogenic sources (e.g., road traffic, industrial emissions). Often, waste incinerators are located inside industrial areas or near other major sources of pollution. It will be difficult to correctly identify the possible health effect of this incinerator, unless we have some information about the difference in population exposure to other sources between the exposed and non-exposed groups. Only few studies included information about environmental confounders. Biggeri et al. (1996) used measured particulate depositions from the nearest monitoring station, Cordier et al. (2004) used proxies for the presence of industrial activities and road traffic at community level, Two studies (Goria et al., 2009; Viel et al., 2008b) used proxies for traffic and industrial pollution at census block level. Notably, two recent studies (Candela et al., 2013; Ranzi et al., 2011) used atmospheric dispersion models to estimate pollution concentrations at the address of residence from other local sources of atmospheric pollution (road traffic, industrial plants, heating). This represent a notable improvement, since the confounding factor was evaluated with the same spatial resolution as exposure to the incinerator.

As the quantitative contribution of well managed modern incinerators to total pollution levels in a study area and to baseline health risks is expected to be low, we suggest to put a careful attention to other local sources of pollution and to implement multi-site studies on large populations where feasible.

3.2.4 Conclusions

In this chapter I reviewed 42 articles from literature on incinerators with the aim to define and test a classification framework for the quality of exposure assessment and to identify what are the most advanced methods for exposure assessment to industrial pollution.

Overall, my analysis showed a trend of improvement in exposure assessment quality over time, with a massive use of dispersion models in exposure assessment after year 2003. Nevertheless, the lack of a common framework for exposure assessment is demonstrated by the use of a variety of methods, also in recent papers, with different quality of epidemiological findings and difficulties in comparisons of results.

In most of the selected studies the characterization of exposure can be significantly improved by using more detailed data for population residency and better simulation models. Recent development of informative systems and high availability of environmental and demographic data suggests the use of dispersion models of pollutants emitted from a source, combined with precise methods of geographic localizations of people under study, as the state of the art method to assess exposure of population in epidemiological studies. Considerations about residential mobility, home-work movements, temporal variations in pollution emissions, latency period of investigated diseases and treatment of environmental and socio-demographic confounders can improve exposure assessment accuracy.

All these aspects of exposure assessment are particularly relevant as most of environmental conflicts usually arise from the evaluation of the contribution of the various pollution sources to the overall contamination.

4
EXPOSURE MISCLASSIFICATION: A SIMULATION STUDY

Part of this chapter has been published on *Journal of Environmental and Public Health*. Full reference: Cordioli M, Ranzi A, De Leo GA, Lauriola P. *A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators*. Journal of Environmental and Public Health 2013, Vol.2013, Article ID 129470, 12 pp.

4.1 Exposure misclassification

Almost all papers revised in Chapter 3 used categorical definitions of exposure: exposure variables originally measured on a continuous scale (e.g. distance, concentrations) are categorized into defined classes (e.g. low-medium-high exposure). Although this practice have been discouraged (Altman and Royston, 2006; Bennette and Vickers, 2012; Frøslie et al., 2010; Walraven and Hart, 2008) it is still widely used in epidemiology, mainly because it leads to a more simple interpretation and presentation of results. Criteria for classification varies: classes may be arbitrary chosen or based on distribution percentiles (e.g. quintiles, quartiles) to guarantee classes of comparable size and thus a more robust statistical analysis.

When categorical exposure variables are measured with error, they are said to be *misclassified* (i.e. the subject is assigned to a wrong exposure category). Misclassification can be differential or non-differential with respect to disease status of an individual person (Blair et al., 2007): differential misclassification occurs when the probability of misclassification is not the same among diseased and non-diseased. Non-differential misclassification is more probable in those studies where exposure assessment is conducted independently from the diagnosis of disease and using "objective" methodologies (e.g., GIS, mathematical models), while differential misclassification may affect studies where exposure assessment is conducted after diagnosis and using "subjective" methods (e.g., questionnaires or expert judgment).

Non-differential misclassification of a dichotomous exposure (i.e. yes/no) is generally considered to cause a bias towards the null or no-effect value in risk estimations, thus resulting in the underestimation of real health risks (Baker and Nieuwenhuijsen, 2008; Rothman et al., 2008). Nevertheless, non-differentiality alone does not guarantee bias toward the null when categorization is applied (Dosemeci et al., 1990; Jurek et al., 2008, 2005; Sorahan and Gilthorpe, 1994; Weinberg and Umbach, 1994), e.g. in presence of more than two exposure categories, non-differential misclassification can move estimates of risk away from null and disrupt exposure-response trends.

4.2 Case study: the incinerator of Parma

The literature review presented in Chapter 3 highlight the great variability of exposure measures used in studying the health effects of industrial atmospheric emission sources. None of the studies reviewed reported multiple measures of exposure nor analyzed the possible effect of exposure misclassification on estimated health risks. Misclassification sometimes received mention in discussion section only as an argument that study's results are conservative (i.e. calculated health risks are lower

than the real ones), although this is not a generalizable rule (Rothman et al., 2008). Two opposite questions may thus arise:

1. Given a study that identifies a significant health effect of the source using a good-quality measurement of exposure (eg. atmospheric dispersion modelling coupled with precise identification of residence location), what would be the health effect detectable with an exposure of poorer quality?
2. Given a study that find a significant health effect of the source using a poor-quality measure of exposure (e.g. distance to the source from the census block centroid), what would be the health risk detectable with the best available methods?

Some simulation studies that analyzed the issue of exposure misclassification in a quantitative way exists (Flegal et al., 1991; Höfler, 2005; Jurek et al., 2008, 2005), but they are generally based on theoretical consideration about sensitivity (i.e., the probability of correctly classifying an exposed individual) and specificity (i.e., the probability of correctly classifying a non-exposed individual) and are based on simple 2×2 tables with binary exposures (yes/no). Frequently, the sensitivity and specificity of the exposure method used is unknown, and the degree of precision in exposure assessment is conditional on availability of site-specific data.

Here I developed a simulation study to analyze how the use of the different exposure assessment methodologies described in Table 8 may ultimately affect exposure classification and health risk estimation in a real case study on the health impacts of a municipal waste incinerator.

The aims of this analysis were:

a. To quantify the degree of misclassification introduced by the use of different methods of exposure assessment in a real situation
b. To determine the effect that this misclassification can have on the strength of the association between exposure and disease

I restricted the analysis to those exposure assessment methods with high spatial resolution applicable on the local scale, thus excluding qualitative methods (1.y.1 in Table 8) and methods based on comparison between large communities (x.1.1 in Table 8).

4.2.1 Materials and methods

For the simulation study I used the following data from an epidemiological surveillance program for the incinerator of the city of Parma, located in Northern Italy (Figure 10):

- Location of the stack of the incinerator;
- Exact location of the address of residence for 31,019 people living within 4km from the incinerator. Geographic coordinates of addresses were provided by the local registry office;
- Boundaries of the 2001 Italian census blocks for the area, as defined by the Italian National Institute of Statistics (ISTAT);
- Concentration and deposition of PM_{10} emitted from the incinerator simulated with an atmospheric dispersion model;

I simulated atmospheric dispersion of pollutants emitted by the incinerator using the ADMS Urban model (Cambridge Environmental Research Consultants Ltd., 2001), a second generation quasi-Gaussian model already employed in other studies on health effects of incinerators (Candela et al., 2013; Cordier et al., 2010; Ranzi et al., 2011; Viel et al., 2008b). Since the study area is located in a flat plane, this model was judged suitable to compute long-term average concentrations and depositions (Floret et al., 2006).

I used five years of hourly meteorological data (2005-2010) from the nearest meteorological station (about 4km from the plant) and source characteristics from the authorized project (i.e., stack height: 70m; gas temperature: 150°C; PM_{10} emission flux: 231 mg s^{-1}) to calculate average hourly concentrations at ground level (ng m^{-3}) and average hourly deposition (ng m^{-2} h^{-1}) of PM_{10} over the period 2005-2010. Concentrations and depositions were calculated on a regular 200 m receptor grid and then interpolated (using quadratic inverse distance weighting, IDW) to obtain continuous raster maps with a resolution of 20m (Figure 10). Although the model could calculate concentrations and deposition at receptors with specific coordinates (e.g., home addresses), the use of a regular interpolated grid makes it suitable also for other applications, without the need to run the model again.

For each individual, I evaluated residential time-invariant exposure to the incinerator using the six variables described in Table 9 , developed in a GIS environment (*Q-GIS* v. 2.2).

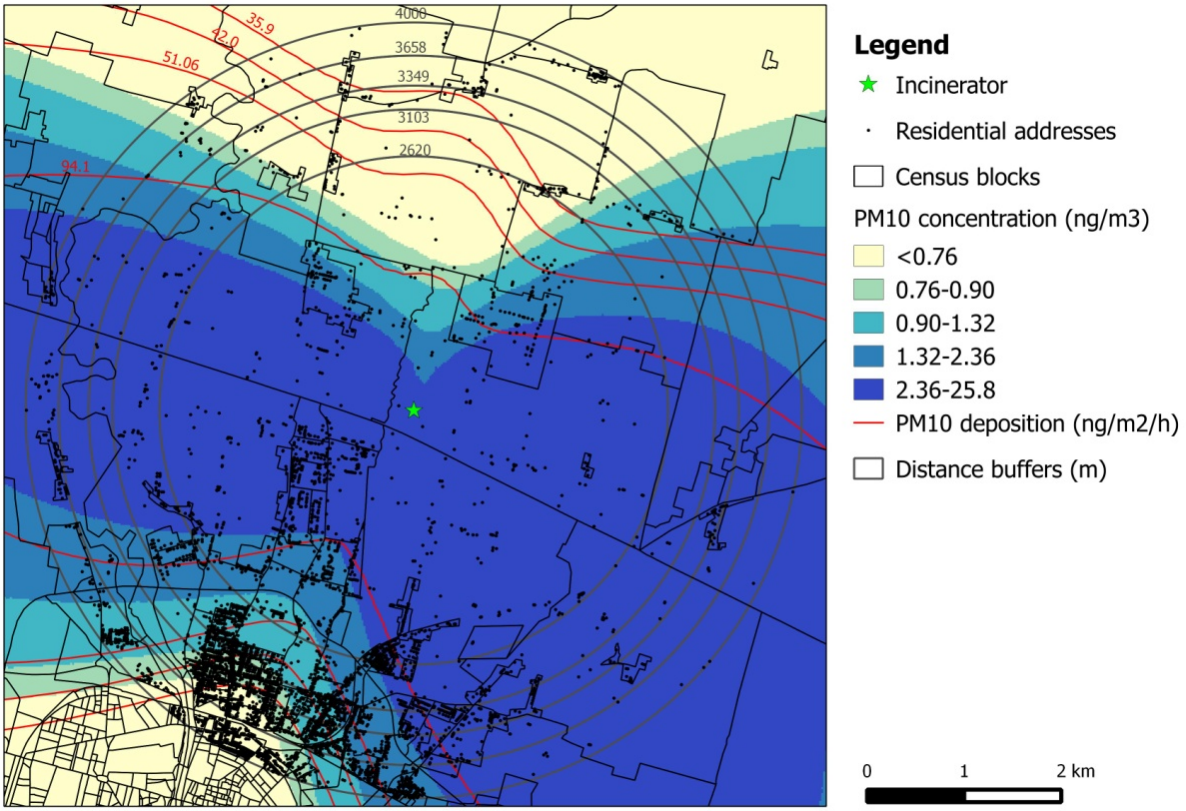

Figure 10 – Representation of the data used to develop the simulation study. The star is the stack of the incinerator. Black dots represent residential addresses within 4km from the incinerator. The irregular polygons represents census blocks. The coloured areas represent average PM_{10} atmospheric concentrations calculated with the atmospheric dispersion model. The red lines represent areas with different level of atmospheric depositions. The cut-offs chosen are the distribution quintiles used in the simulation study.

Table 9 – Description of exposure variables that define population exposure to the incinerator

Acronym	Description	GIS procedure	Method from Chp.3 - Table 8
ADCO (ADdres Concentration)	Concentration at the exact address location	Extraction of raster value at the exact coordinates of the address from concentration map	3.3.1
ADDE (ADdres Deposition)	Deposition at the exact address location	Extraction of raster value at the exact coordinates of the address from deposition map	3.3.1
CBCO (Census Block Concentration)	Average concentration inside census block of residence	Average of raster values from concentration map within the polygon representing census block (*zonal statistics*)	3.2.1
CBDE (Census Block Deposition)	Average deposition inside census block of residence	Average of raster values from deposition map within the polygon representing census block (*zonal statistics*)	3.2.1
ADDI (ADdress Distance)	Distance from exact address location to the incinerator	Euclidean distance between points representing address and incinerator	2.2.1
CBDI (Census Block Distance)	Distance from the census block to the incinerator	Euclidean distance between points representing census block centroid and incinerator	2.2.1

I here assumed that the use of atmospheric concentrations evaluated with the dispersion model coupled with exact geocoding of residential addresses (ADCO method) represent the closest estimate to the actual exposure (i.e., my "gold standard" for exposure assessment to industrial atmospheric emissions). Nevertheless, concentration and deposition estimates based on dispersion models are affected by their own degree of error and uncertainty (Chang and Hanna, 2004; Colvile et al., 2002) and should be possibly validated with field measurements and/or experiments (Floret et al., 2006). Distance of the census block centroid (CBDI) was indeed the less precise exposure measure.

All exposure variables were categorized in 5 classes (i.e.,0: lowest exposure, 4: highest exposure) using quintiles of each variable distribution (Figure 10). Thus, each exposure class contains approximately the same number of subjects.

To evaluated the degree of exposure misclassification, I compared the ADCO exposure categorization with all other methods using two-way tables. For each comparison I calculated the percentage of subjects misclassified by one and more than one exposure classes. I evaluated the level of agreement between two different exposure classifications using Cohen's kappa test of agreement (Eriksson et al., 2012; Peters et al., 2011) applying quadratic weighting to assign less importance to misclassification between adjacent classes and higher importance to other misclassifications.

To quantify the effect of exposure misclassification on the strength of association between exposure and disease, I compared the health risk derived with a reference exposure method with those obtained using all other available methods.

In presence of a binary health outcome (i.e. 0=healthy, 1=sick), the relationship between exposure and outcome can be analyzed through *logistic regression* (Baker and Nieuwenhuijsen, 2008; Rothman et al., 2008). Logistic regression model the *logit-transformed* probability of being sick (p) as a linear relationship with a series of predictor variables (x_i).

$$logit(p) = \ln\left(\frac{p}{1-p}\right) = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + c$$

The term $\frac{p}{1-p}$ represent the ratio between the probability of having the disease (p) and the probability of not having the disease (1-p), and is named *odds*. It can be shown that exponentiating the coefficients of the logistic regression (β_n) gives the value of the ratio between the *odds* (i.e. *odds ratio*, OR) for a unitary increase in the predictor variable x_n. The OR is used as a measure of risk and strength of association between the predictors and the disease When the exposure variable is categorical, the OR represent a comparison between the risk in the different categories. If OR=1 there is no association between the predictor under study and the occurrence of the disease. If OR>1 then an increase in the predictor increase the risk of disease. If OR<1 then the predictor can be seen as a protective factor for the disease. Confidence intervals can be calculated for the ORs: an OR is statistically significant if its confidence interval does not include the value of no-effect (i.e., OR_{CI95-} >1 for a positive association, OR_{CI95+} <1 for a protective association).

Since real data on health status of the enrolled subjects was not available at the moment of this analysis, I developed a simulation that randomly assign the health status to each subject under study. The procedure was the following:

1. I calculated the number of healthy and sick subjects in each category of the reference exposure metric needed to obtain a series of predefined ORs between the low exposure class (i.e., 0) and each class with increased exposure. The predefined ORs were: 1.20, 1.50, 1.80, 2.01, that are typical order of magnitude of OR values in environmental epidemiology studies. I calculated OR between class n and the class 0 using the following equation:

$$OR_{0-n} = \frac{sick_n \times healty_0}{sick_0 \times healty_n}$$

where:
$sick_n$= number of sick subjects in class n
$healty_0$= number of healthy subjects in the reference class 0
$sick_0$= number of sick subjects in the reference class 0
$healty_n$= number of healthy subjects in class n

2. I randomly set the status of illness (healthy/sick) of each subject, keeping fixed the totals of sick and healthy subjects within each exposure class of the reference method. Thus, although the subjects that are sick and healthy vary randomly, the OR between different classes of the reference exposure remain constant.
3. I ran a series of logistic regression using all the available exposure measures (i.e., ADCO, ADDE, CBCO, CBDE, ADDI, CBDI) and calculate the OR for each exposure class. While the ORs for the reference exposure method remains fixed by construction, the ORs for the other exposure variables will be different because a certain number of subjects will be misclassified (i.e. change exposure class with respect to the reference exposure). Misclassification will randomly affect sick and healthy subjects in each iteration of the simulation, thus affecting the OR value.

The steps 2. and 3. of the simulation were repeated 10000 times.

Two answer the two questions described in the introduction, I defined two simulation scenarios (Table 10):

a. *Real risk scenario*: the reference exposure method was the best available one, i.e. ADCO. I here assumed that there was a *real* health risk associated with the incinerator, since the ADCO exposure represented well the *real* exposure of the population.
b. *Apparent risk scenario*: the reference exposure was the less precise, i.e. CBDI. I here assumed that there was an *apparent* health risk associated with the incinerator, since the CBDI exposure did not represent a good measure of the *real* exposure to the incinerator.

Table 10 – Number of sick and healthy subjects in each ADCO (*real risk* scenario) and CBDI (*apparent risk* scenario) exposure classes needed to obtain predefined odds ratios (OR)

Simulation scenario	Exposure class	Predefined OR	Total population	Sick	Healty
Real risk ref. exposure: ADCO	0 (ref.)	-	6176	2470	3706
	1	1.20	6217	2763	3454
	2	1.50	6191	3096	3095
	3	1.80	6215	3393	2822
	4	2.01	6220	3564	2656
Apparent risk ref. exposure: CBDI	0 (ref.)	-	6654	2662	3992
	1	1.20	5797	2577	3220
	2	1.50	6280	3140	3140
	3	1.80	6177	3373	2804
	4	2.01	6111	3502	2609

The output of the analysis is a distribution of ORs for each class of each exposure measurement method. This simulation try to answers the following question: *given an health risk equal to X associated with the reference exposure Y categorized in quintiles, what would be de risk detectable using the other exposure classification methods*? Thus, I calculated the number of times the misclassified ORs are below or above the originally fixed ORs based on the reference exposure measure, the quote of non statistically significant ORs, and represented the ORs distributions graphically.

4.2.2 Concordance between different exposure metrics and exposure misclassification

Figure 11 contrasts the results of alternative approaches to assess exposure level of the cohort in terms of intensity (simulated concentration vs distance from the emission sources) and accuracy in residence location (exact address vs. census block), using continuous variables.

As expected, PM_{10} concentrations from the incinerator decrease with distance from the stack (Figure 11-a). Nevertheless the graph shows that (i) this relationship is not linear and (ii) exposure to atmospheric PM_{10} may vary consistently within subjects residing at the same distance from the plant, due to the anisotropy of atmospheric dispersion (Figure 10). The relationship between ground concentration and deposition at address location is almost linear, with some outliers that have very high depositions (Figure 11-b). The ADMS-Urban model calculates depositions from atmospheric concentrations, on the basis of deposition velocities and rain washout (Cambridge Environmental Research Consultants Ltd., 2001): a good relationship between these measures of exposure is thus

expected. The choice of either methods should be based on the specific interest in considering exposure through inhalation and/or through other routes (e.g. soil ingestion/contact).

The effect of geocoding residential addresses with different precision is analysed in Figure 11-c and Figure 11-d. In this case study census blocks had an average area of 0.4 km^2 (min: 968.4 m^2; max: 6.3 km^2) and contained on average 26 addresses (min:1; max:130), thus both address distances, depositions and concentrations varies widely inside some census blocks. This was true especially for more exposed areas, since the incinerator is located in a less densely populated area with large census blocks. This aspect could lead different degree of errors in exposure assignment, that increase with level of pollutant or proximity to the incinerator. Moreover, the use of the average exposure inside the census area or exposure at its centroid may not correctly represent exposure variability in the studied population: in the Parma study area in some cases within-block exposure variability is greater than between-block variability.

Table 11 reports the results of the comparisons between different categorizations of exposure. The table reports the share of individuals over the 31,019 sample assigned to the same class of exposure, the share of individuals classified in an adjacent exposure class and that of individuals classified into two or more class apart. Table 11 shows also Cohen's Kappa indices of agreement between concentration maps and other exposure assessment methods.

The use of the census block instead of the exact address location in evaluating atmospheric PM_{10} concentrations (ADCO vs. CBCO) leads to a misclassification for 10.5% of subjects. Almost all misclassified subjects move to an adjacent exposure category, and the weighted Kappa is very high (0.97). The use of atmospheric depositions (ADCO vs. ADDE) introduces misclassification for about a third of the cohort. The concordance is slightly higher for the comparison with average depositions inside census block (ADCO vs. CBDE), mainly because the spatial averaging process smooth the differences between concentrations and deposition estimated by the ADMS Urban model. The share of subjects experiencing a misclassification by more than one category remains low when using depositions. When using distance, 45.1% and 44.5% of the subjects are mislcassified respectively for ADDI and CBDI, and weighted Kappa values fall to 0.6. Interestingly, the use of distance cause misclassification by more than one category for up to 16% of the subjects. Misclassification in non-adjacent categories is one of the causes of possible change in the slope of the exposure-response trend (Dosemeci et al., 1990).

Table 11 – Evaluation of the agreement between different classifications of exposure. The table shows the percentages of subjects classified in the same exposure class or in different classes and quadratic weighted Cohen's kappa of agreement.

Comparison exposure	Matching subjects	Misclassified 1 category	Misclassified > 1 category	Weighted Kappa [a]
ADCO vs CBCO	89.2%	10.5%	0.3%	0.97
ADCO vs ADDE	69.6%	29.3%	1.1%	0.91
ADCO vs CBDE	70.0%	27.8%	2.2%	0.90
ADCO vs ADDI	38.9%	45.1%	16.0%	0.61
ADCO vs CBDI	40.2%	44.5%	15.3%	0.60

ADCO= address concentration (quintiles), ADDE=address deposition (quintiles), CBCO=census block concentration (quintiles), CBDE=census block deposition (quintiles), ADDI=address distance (quintiles),, CBDI= distance between census block centroid and incinerator. [a]all kappa with $p<0.001$.

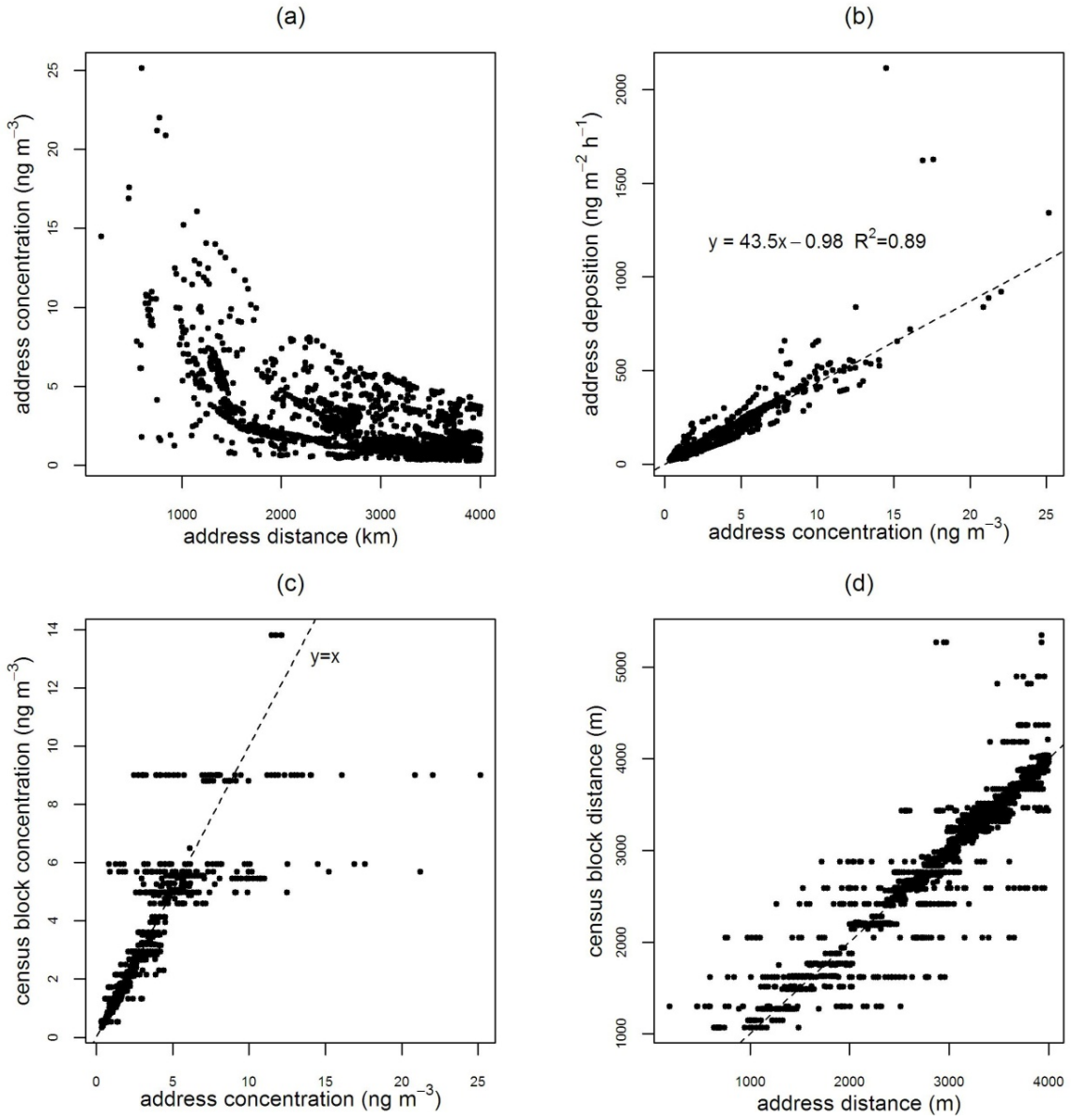

Figure 11 – Results of exposure assessment by using different methodologies.
(a) Residential address concentration (ADCO) versus address distance (ADCO). (b) Relationship between ground concentration (ADCO) and deposition (ADDE) at addresses location. The line represent the linear regression model. (c) Relationship between simulated concentrations evaluated at exact address (ADCO) and at census block level (CBCO). The line represent the 1:1 relationship. (d) Relationship between distance of the exact address ($ADDI_1$) and distance of the census block centroid (CBDI). The line represent the 1:1 relationship.

4.2.3 Effect of exposure misclassification on estimated health risks

The degree of exposure misclassification described in the previous paragraph depends solely on the method used to evaluate exposure to the incinerator. The fact that a subject is misclassified with respect to a reference exposure measure depends only on geographical location of its residence. This type of misclassification is by design non-differential, since there is no relationship between the health status of the subject (actually not defined) and the probability of being misclassified. It is thus expected that this type of misclassification will cause underestimation of the health risk.

The distribution of ORs for the *real risk* and *apparent risk* scenarios are depicted in Figure 12 and Figure 13 as box-plots, together with the minimum values of the lower bound confidence interval and the maximum values of the upper bound confidence interval for the ORs. Table 12 reports the percentages of simulated ORs that are below and above the reference values for ADCO and CBDI, together with the percentage of non statistically significant and significantly protective (i.e. $OR_{CI95+} < 1$) ORs.

Real risk scenario

By construction, in this scenario the ORs for ADCO remain constant at the values of A) 1.2, B) 1.5, C) 1.8 and D) 2.0. The use of the concentration at census block level (CBCO) introduces a certain variability in the results: the medians of the ORs are slightly lower, but a certain quote of the simulated ORs exceeds the reference values (i.e., 34%, 10%, 8% and 14% respectively for $OR_{0\text{-}1}$, $OR_{0\text{-}2}$, $OR_{0\text{-}3}$ and $OR_{0\text{-}4}$). The use of depositions instead of concentrations (ADDE, CBDE) introduces a further bias toward the null: a great proportion of simulated ORs falls below the reference values, although in some cases the simulated ORs are higher. Moreover, Figure 12-A shows that in a few cases simulated $OR_{0\text{-}1}$ for depositions are not statistically significant (i.e., the lower bound of the confidence interval falls below the null value of 1). Distance to the incinerator (ADDI, CBDI) gave the lowest ORs. Simulated ORs are always below the reference value, and a relevant quote of the simulated ORs are non-significant (for CBDI up to 100% of $OR_{0\text{-}1}$). Additionally, a certain quote of $OR_{0\text{-}1}$ calculated on distance were significantly below the null value of 1, i.e. identify a protective effect of exposure to the incinerator in class 1.

Apparent risk scenario

By construction, in this scenario the ORs for CBDI remain constant at the values of A) 1.2, B) 1.5, C) 1.8 and D) 2.0. For the $OR_{0\text{-}1}$ values similar or higher than 1.2 are obtained using all other exposure measures. The situation reverses for all the other exposure classes: the use of more precise measures of exposure gives ORs that are lower than the reference ones. Refining the distance measure using exact address location (ADDI) generates slightly lower ORs. The use of atmospheric concentrations (ADCO, CBCO) gives the lowest ORs for exposure classes 3 and 4. Simulated ORs remain always statistically significant for all the exposure classes (except for one single case for $OR_{0\text{-}1}$ using the CBCO exposure) and no significant protective ORs are detected.

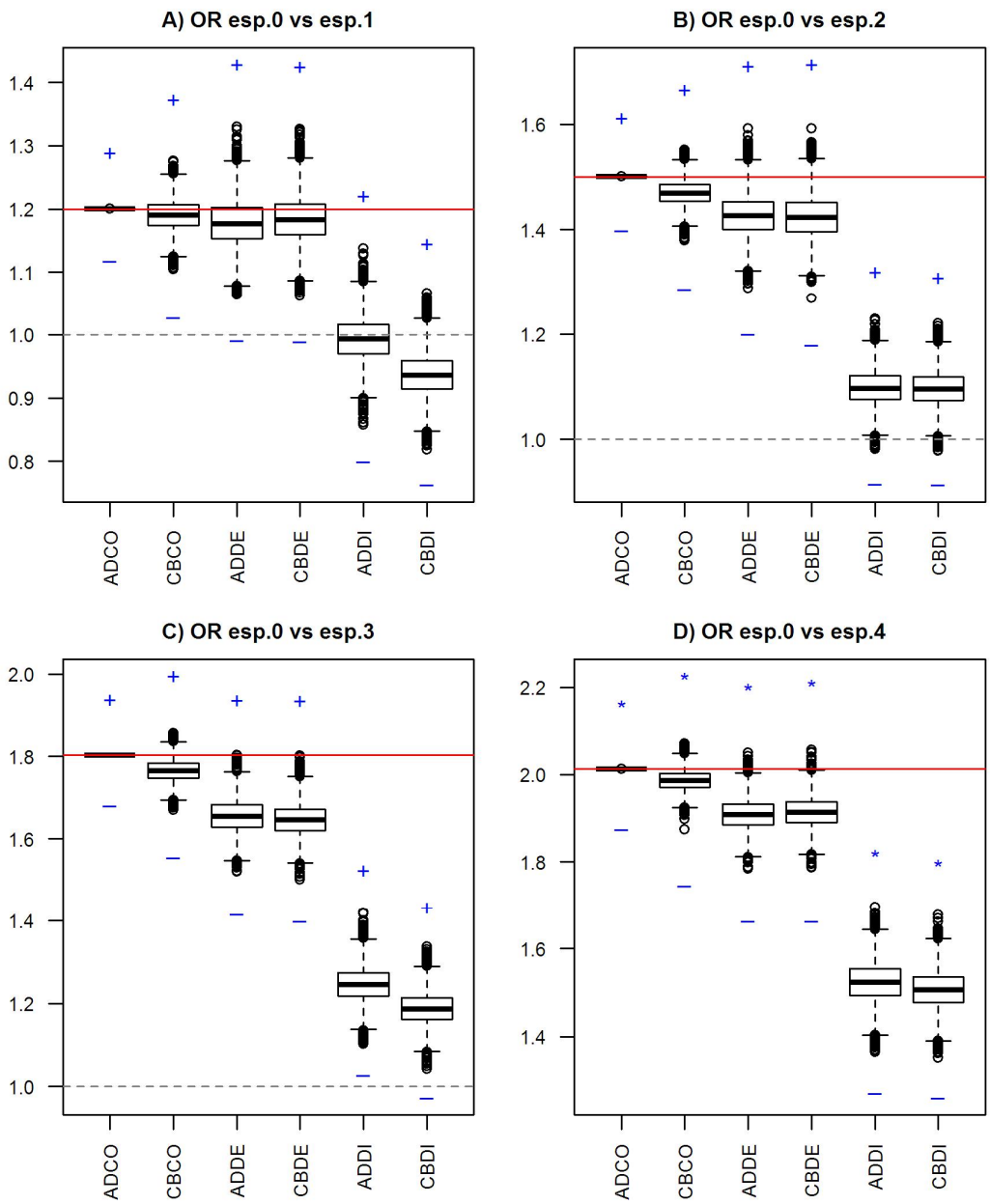

Figure 12 – Distributions of simulated Odds Ratios (OR) using different exposure assessment methods in the *real risk* scenario. The box represent the inter-quartile range (IQR) of the distribution, the horizontal line inside the box is the median value, the whiskers extend to 1.5 times the IQR from the box. The red line represent the predefined OR based on ADCO exposure. The dashed line represent the "no effect" value (i.e. OR=1). The blue markers represent the minimum value of the lower bound confidence interval (-) and the maximum value of the upper bound confidence interval (+) for the calculated ORs.

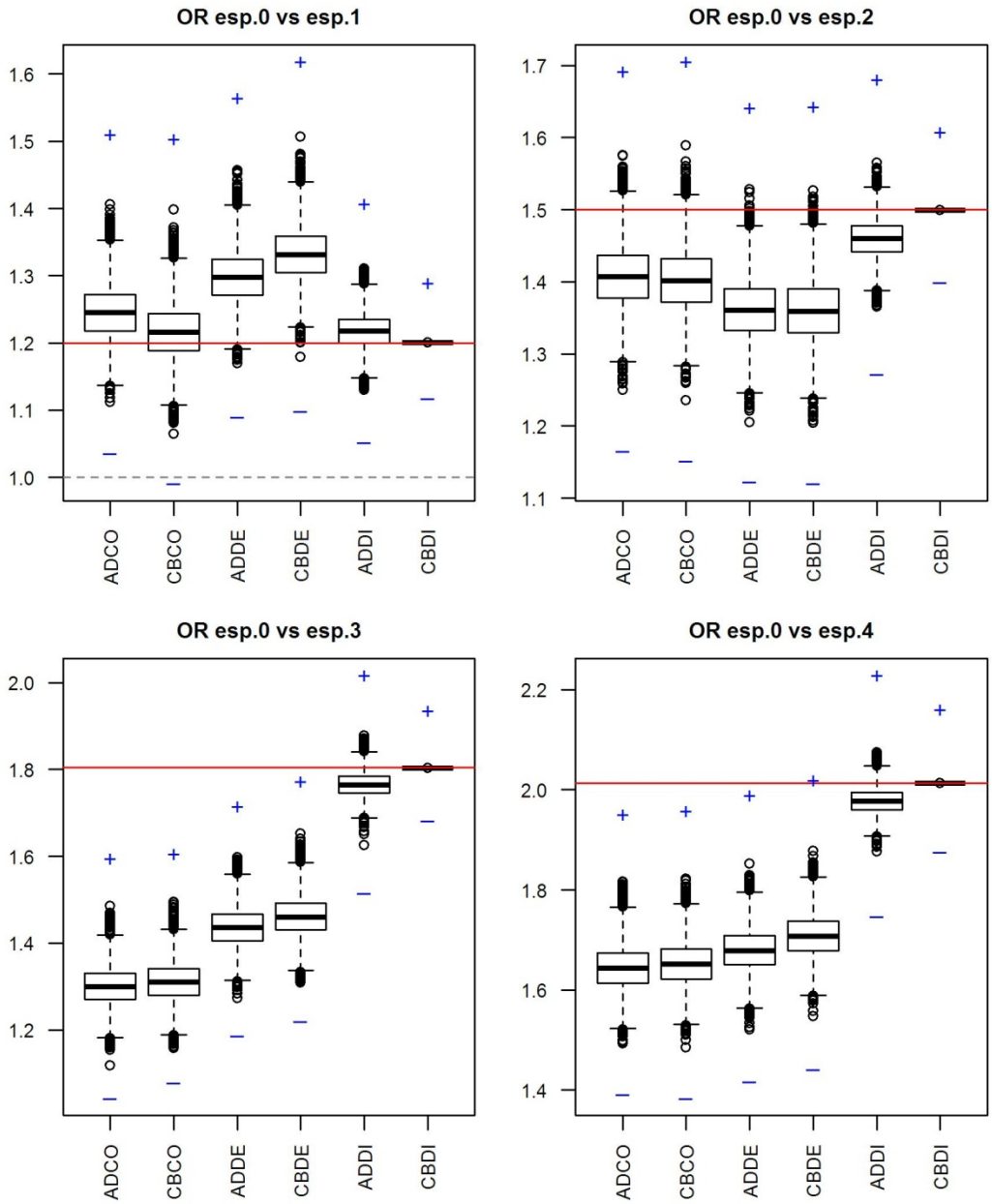

Figure 13 – Distributions of simulated Odds Ratios (OR) using different exposure assessment methods in the *apparent risk* scenario. The box represent the inter-quartile range (IQR) of the distribution, the horizontal line inside the box is the median value, the whiskers extend to 1.5 times the IQR from the box. The red line represent the predefined OR based on CBDI exposure. The dashed line represent the "no effect" value (i.e. OR=1). The blue markers represent the minimum value of the lower bound confidence interval (-) and the maximum value of the upper bound confidence interval (+) for the calculated ORs.

Table 12 – Result for the random simulation. Data represent the fraction of simulated ORs below and above the reference values, the fraction of non statistically significant (n.s.) ORs and the fraction of ORs that are significantly below the null value (protective).

		Real risk				*Apparent risk*			
Exposure	**OR**	**< OR_{ADCO} %**	**> OR_{ADCO} %**	**n.s. %**	**protect. %**	**< OR_{CBDI} %**	**> OR_{CBDI} %**	**n.s. %**	**protect. %**
ADCO	**$OR_{0\text{-}1}$**	-	-	-	-	12.7	87.3		
	$OR_{0\text{-}2}$	-	-	-	-	97.9	2.1		
	$OR_{0\text{-}3}$	-	-	-	-	100.0	0.0		
	$OR_{0\text{-}4}$	-	-	-	-	100.0	0.0		
CBCO	**$OR_{0\text{-}1}$**	66.0	34.0			34.7	65.3	0.01	
	$OR_{0\text{-}2}$	90.4	9.6			98.5	1.5		
	$OR_{0\text{-}3}$	92.1	7.9			100.0	0.0		
	$OR_{0\text{-}4}$	87.3	13.7			100.0	0.0		
ADDE	**$OR_{0\text{-}1}$**	73.1	26.9	0.1		0.4	99.6		
	$OR_{0\text{-}2}$	96.7	3.3			99.9	0.1		
	$OR_{0\text{-}3}$	100.0	0.0			100.0	0.0		
	$OR_{0\text{-}4}$	99.7	0.3			100.0	0.0		
CBDE	**$OR_{0\text{-}1}$**	67.6	32.4	0.1		0.1	99.9		
	$OR_{0\text{-}2}$	97.1	2.9			99.9	0.1		
	$OR_{0\text{-}3}$	100.0	0.0			100.0	0.0		
	$OR_{0\text{-}4}$	99.7	0.3			100.0	0.0		
ADDI	**$OR_{0\text{-}1}$**	100.0	0.0	98.8	3.5	25.4	74.6		
	$OR_{0\text{-}2}$	100.0	0.0	22.7		93.0	7.0		
	$OR_{0\text{-}3}$	100.0	0.0			91.2	8.8		
	$OR_{0\text{-}4}$	100.0	0.0			91.2	8.8		
CBDI	**$OR_{0\text{-}1}$**	100.0	0.0	100.0	43.9	-	-	-	-
	$OR_{0\text{-}2}$	100.0	0.0	22.9		-	-	-	-
	$OR_{0\text{-}3}$	100.0	0.0	0.1		-	-	-	-
	$OR_{0\text{-}4}$	100.0	0.0			-	-	-	-

4.2.4 Discussion

For exposure measures based on distance to the source (ADDI, CBDI) a relevant part of the population is classified in the *wrong* exposure category (assuming that dispersion model better represent *real* exposure), with relevant percentages of subjects moving by more than one category. The classification based on atmospheric depositions (ADDE, CBDE) introduces a lower degree of exposure misclassification.

The simulation study highlight that, under the specific conditions tested, non-differential exposure misclassification tends to give health risks that are biased toward the null in both *real risk* and *apparent risk* scenarios. Nevertheless, there are some cases (e.g. CBCO, ADDE, CBDE in the *real risk* scenario) where misclassified results are slightly higher than reference risks.

The misclassification caused by the use of distance as a proxy of exposure leads always to an underestimation of the *real* risk associated with the exposure to the atmospheric concentrations of

pollutants emitted by the stack (ADCO). On average, the ORs for the CBDI exposure are 22%, 27%, 44% and 25% lower than the *real* imposed risk, with many cases of non statistically significant ORs. In some extreme cases, using the CBDI exposure we would observe a non statistically significant $OR_{0\text{-}3}$ in place of a *real* $OR_{0\text{-}3}$ of 1.8, which represent a relatively high risk in environmental epidemiology. The answer to the first question I asked is that the probability of not being able to detect the *real* effect of the incinerator in Parma using distance metrics is high. Of the 10000 hypothetical situations giving an $OR_{0\text{-}4}$ of 2.01 associated with the ADCO exposure, no one gives $OR_{0\text{-}4}$ greater than 1.7 with the ADDI and CBDI exposures.

On the other hand, when an *apparent* risk associated with the worst exposure distance metric (i.e. ADCO) is present, the probability that this risk will be conservative and that the use of more accurate measures of exposure will identify a risk equal to or higher than the observed one is very low. Of the 10000 random simulations giving an $OR_{0\text{-}4}$ of 2.01 associated with the CBDI measure, no one gives $OR_{0\text{-}4}$ greater than 1.82 with the ADCO exposure.

In this simulation I defined two *a priori* scenarios, one that impose an health risk associated with the best available exposure measure, the other that impose a risk associated with the less precise one. The degree of exposure misclassification remain the same in the two scenarios, while the distribution of healthy and sick subjects in the study area varies. In the *real risk* scenario, a study conducted using exposure proxies based on distance will give conservative risk estimations that are biased toward the null. On the contrary, in the *apparent risk* scenario a study based on distance proxies will give risk estimations that are overestimated with respect to those obtainable with a better representation of exposure. I called this risk *apparent* since there is no reason to believe that the linear distance from the census block of residence to the incinerator can represent people's exposure better than an estimation of average pollution concentration at address location, that accounts for meteorology and characteristics of the emission source. In both scenarios, the use of distance as a measure of exposure is thus discouraged. Although this may seems an obvious finding, many studies that uses distance as a proxy of exposure are continuously published in the literature, e.g. some of the most recent studies found on incinerators (García-pérez et al., 2013; Reeve et al., 2013).

In practice, when conducting epidemiological studies based on observed prevalence of the disease, we will never know if we are in a *real risk* or *apparent risk* situation. If we are aware that we are using a measure of exposure that inaccurately represent the real exposure, we cannot be confident that the health risks we are measuring are underestimated because of non-differential misclassification with respect to those obtainable with a better exposure characterization. In some cases, the risk we will measure with a badly characterized exposure would be only *apparent*.

It must be specified that while the use of atmospheric dispersion models certainly represent an improvement with respect to the use of linear distance, the choice of atmospheric concentrations or depositions is more controversial. The deposition of pollutants to the ground is of interest to investigate the role of indirect exposure pathways, like soil contact or ingestion (Abrahams, 2002) and food contamination (Pandey et al., 2012; Schiavon et al., 2013). Nevertheless, while direct exposure through air breathing occurs in every place people spend time (e.g. residential address), soil deposition of pollutants may influence people's exposure only in specific places, e.g. recreational areas for dermal contact with soil (Kissel et al., 1996) or cultivated areas for ingestion through vegetables. Thus, in many cases atmospheric depositions at residence location may not be a good exposure indicator. This issue will be partially explored in Chapter 5, where an health risk assessment model for the incinerator of Parma is presented and the role of soil ingestion and food contamination in different areas of the territory is discussed.

Of course, better (e.g. microenvironmental exposures models (Mölter et al., 2012)) and worse (e.g. qualitative methods reviewed in Chapter 3) measures of exposure exist with respect to ADCO and CBDI

used in this simulation. Whether the trends found in this simulation extent to other exposure measures remain to be explored.

One limitation of my simulation may be the use of categorized exposure variables. Categorization of a continuous variable introduces by itself some degree of misclassification (Bennette and Vickers, 2012). I decided to use categories because (i) the great majority of the studies reviewed in Chapter 3 used categorical exposures and (ii) it becomes an hard task to compare measures of association (i.e. ORs) evaluated using exposures with very different scales and units of measurement. While the OR for categories represent the increase in risk in a category compared to the reference one, and is thus comparable among different measures of exposure, the use of OR for continuous variables would represent the increase in risk for a unitary increase in exposure, i.e. 1 ng m^{-3} for concentrations, 1 ng m^2 h^{-1} for deposition and 1 m for distance. My simulation is based on categories defined on quintiles of exposure. This guarantee a certain stability in the simulation results, since the classes compared in the logistic regression model have the same population. To test the validity of the results with other types of categories I developed a simulation using also *a priori* defined cut-offs for ADCO, ADDE and ADDI (i.e. 0.3-0.8-1.6-3.2-6.4-25.2 ng m^{-3}, 21-40-60-150-300-2114 ng m^2 h^{-1}, 0.8-1.6-2.4-3.2-4.0 and 1.5-2.5-3.0-3.5-4 km respectively), as was done in some of the published studies reviewed in Chapter 3 (Comba et al., 2003; Cordier et al., 2004; Elliott et al., 1996; Floret et al., 2003; Michelozzi et al., 1998; Viel et al., 2008a; Vinceti et al., 2009). The results of this simulation are reported in APPENDIX C. The results of the sensitivity analysis confirmed the trends that emerged from the main analysis.

4.2.5 Conclusions

The analysis of published studies on incinerators presented in Chapter 3 revealed that in many cases a poor measure of exposure was used. Many of these studies describe their results as conservative, because of non-differential misclassification of exposure. In this chapter I conducted a simulation study to examine the degree of exposure misclassification related to the use of different methods for exposure assessment to a point emission source of atmospheric pollution and the effect that this misclassification have on estimated health risks.

With this simulation I demonstrated that (i) the use of distance and - to a limited extent - atmospheric deposition introduces a certain degree of exposure misclassification with respect to exposure to atmospheric concentrations, with many subjects being misclassified by more than one category; (ii) non-differential exposure misclassification based on categorical exposure is more likely to introduce a bias toward the null or no effect value, although in some simulated situations risk may be higher; (iii) the use of linear distance must be avoided in studies on atmospheric emission sources, unless it is demonstrated that atmospheric dispersion in the area is exactly homogeneous in all directions; (iv) when exposure is poorly characterized, we cannot be confident that the risk we measure is lower than the real risk we would measure with a better exposure assessment because of non-differential exposure misclassification.

Although the quantitative results remain significant only in the specific context under study, the analysis underlined some interesting issues that may be common in many other epidemiological studies. Sometimes, the degree of error in exposure assessment can be evaluated with a validation study, i.e. comparing modelled exposure with “gold-standard” measurement of exposure collected for a random subsample of the population, such as direct measurement of individual exposure. In practice, since no such gold-standard is generally available in exposure analyses, I recommend researchers to conduct sensitivity analyses on exposure assessment and discuss the magnitude of error that may be present in their data.

5
HEALTH RISK ASSESSMENT: ESTIMATING POTENTIAL HEALTH RISKS

Part of this chapter has been published on a peer-review journal: Cordioli M, Vincenzi S, De Leo GA. *Effects of heat recovery for district heating on waste incinerator health impact: a simulation study in Northern Italy*. Science of the total Environment 2013; 444:369–380, DOI: 10.1016/j.scitotenv.2012.11.079.

5.1 Environmental health risk assessment: general principles and exposure assessment methods

Environmental health risk can be defined as the likelihood of the occurrence of adverse health effects to an individual, a population or an ecosystem due to the exposure to hazardous agents or chemicals in the environment (Aral, 2010). Environmental health risk assessment (HRA) is the scientific process used to evaluate environmental risks that uses toxicology data collected from animal studies or human epidemiology, combined with information about the degree of exposure, to quantitatively predict the likelihood of a particular health effect in a specific human population (Simeonov and Hassanien, 2009). The ultimate goal of HRA is not to eliminate all risks, but rather to identify acceptable levels of exposure and compare different technological scenarios (Mckone, 1996).

The risk assessment process is generally divided into four main stages:

1. *Hazard identification*
2. *Exposure/dose-response assessment*
3. *Exposure/dose assessment*
4. *Risk characterization*

In the *hazard identification* phase it is established if exposure to the agent under study can determine, at some intensity, an adverse health effect in exposed subjects. This stage is based on a weight-of-evidence classification of hazardous agents and chemicals. The *exposure/dose-response assessment* stage defines the relationship between the exposure or dose of an agent and the probability of a specific health effect. Usually this relationship is defined on the basis of evidences from laboratory studies on animals but for some substances also epidemiological data on humans may be available. The *exposure/dose assessment* process evaluates the contact and uptake of a chemical from the environment into human body through different exposure pathways (e.g., inhalation, ingestion, dermal contact). In this stage, models and analytical techniques are used to quantitatively evaluate (i) the transformation and transfer of pollutants in the environmental media (e.g. air, water, soil, food) and (ii) the magnitude of contact with the human receptor, incorporating knowledge of lifestyles and time-activity patterns. The *risk characterization* process links estimated exposure with the exposure-response relationship and leads to the quantitative estimation of health risks and associated uncertainties. Normally, risk estimates are evaluated separately for *stochastic* and *non-stochastic* health effects (Mckone, 1996). The first group refers to those effects for which the probability of occurrence, rather that the magnitude, is proportional to the dose (e.g. carcinogenesis and many genetic effect). The second group include those effects for which the magnitude and severity is a function of the dose (e.g. neurotoxins).

Exposure assessment has been the focus of this PhD thesis. In the previous chapters exposure methodologies were applied to the context of environmental epidemiology. The role of exposure assessment is crucial also in HRA, and it is one of the biggest sources of uncertainty (Palma-Oliveira et al., 2012). One of the peculiar characteristics of environmental HRA is the necessity of estimating the dose or intake of a pollutant, and to couple it with a dose-response function. This call for quantitative estimates implies that (i) qualitative methods or proxies like questionnaires or distance to the source are unusable and (ii) models or measures of environmental concentrations must be coupled with models or measures that reconstruct human contact with the agent and the intake process through multiple pathways (Fryer et al., 2006).

The quantitative estimation of exposure and dose can be approached following three ways (Simeonov and Hassanien, 2009):

1. Point-of-contact measurement: exposure is measured at the point of contact with the organism through time (e.g. personal monitors);
2. Scenario evaluation: exposure concentrations and contact duration are modeled separately and then combined in different exposure scenarios
3. Reconstruction: exposure is estimated from measure dose in the organism (e.g. biomarkers)

The use of models to define different risk scenarios is of particular interest because can be used (i) to predict potential future exposures, (ii) to compare different emission options and technological alternatives, (iii) to analyze what aspect of human behavior is more determinant for health risk, (iv) to reduce the need of intensive monitoring programs (Fryer et al., 2006). Since environmental pollution is a multimedia problem, models that simultaneously consider multiple exposure pathways are particularly useful in HRA. One of the most used framework for multimedia risk assessment is the *Multiple Pathway Exposure Methodology* (MPE) developed by the United States Environmental Protection Agency (US-EPA) (Reisman and Brady-Roberts, 1999). This framework considers all the possible way a contaminant emitted into the atmosphere by a source can reach a human receptor (Table 13) and has been implemented in many US-EPA's risk assessment guidelines (US-EPA, 2005a, 2002). As discussed in Chapter 4, atmospheric deposition may represent a relevant pathway for the entry of pollutants into soil, water and the food chain.

Table 13 – Direct and indirect routes of human exposure to pollutants emitted into the atmosphere considered in the Multiple Pathway Exposure Methodology

Compartment	Cause of contamination	Exposure routes
Air	Atmospheric concentration	Air inhalation
Soil	Atmospheric deposition to ground	Soil ingestion
		Dermal contact with soil
		Inhalation of suspended soil particles
Water	Atmospheric deposition to water surfaces	Groundwater consumption and contact
		Surface water consumption and contact
		Fish consumption
Food	Atmospheric concentration and deposition to cultivated surfaces	Vegetable consumption
		Meat consumption
		Dairy products consumption
		Breast milk consumption

5.2 Health risk assessment for incinerators: state of the art and open issues

The common thread of this thesis is exposure assessment to industrial sources of atmospheric emission, with particular focus on waste incinerators. Several HRA studies have recently estimated the health effects of pollutants from municipal solid waste incinerators (MSWI) and other industrial plants considering multiple exposure pathways (Cangialosi et al., 2008; Domingo, 2004; Karademir, 2004; Lonati and Zanoni, 2012; Morra et al., 2009; Nouwen et al., 2001; Ollson et al., 2014; Roberts and Chen, 2006; Rovira et al., 2010; Schuhmacher et al., 2004). Almost all the published studies followed the *Human Health Risk Assessment Protocol (HHRAP) for hazardous waste combustion facilities* (US-EPA, 2005a). This guidance, based on the previously outlined MPE framework, provides methodologies and equations for estimating the transfer of pollutants between different environmental compartment and for evaluating human exposure.

A few more studies (Forastiere et al., 2011; Kim et al., 2011; Phillips et al., 2010) have estimated health impacts of incinerators using a different methodology, based on epidemiological methods that quantify the public health burden of disease attributable to a specific risk factor (Steenland and Armstrong, 2006). In these studies only inhalation exposure is normally considered, and the impact of the incinerators is estimated considering the increase in concentrations of some pollutants for which exposure-response functions are available from the epidemiological literature (e.g. PM_{10}, NO_2, O_3, SO_2).

While HRAs on incinerators have been routinely performed in the last twenty years, there are still a number of issues that have not been thoroughly investigated in the literature.

First, it is not clear whether the potential health impacts caused by incinerator stack emissions can be partially compensated for by the reduction in atmospheric emissions achievable through district heating and the switch-off of a substantial number of domestic boilers (Rezaie and Rosen, 2012). It is critical to determine whether the choice to locate a waste incineration facility near a densely populated area may be justified by the benefits derived from heat recovery from waste incineration and the activation of extensive district heating.

Second, in standard HRAs it is often assumed that only food produced at the exposure location (i.e., home-grown at the residential address) is potentially contaminated by stack emissions (US-EPA, 2005a), while food of animal origin consumed by residents in urban areas is commonly assumed to be produced elsewhere and, as such, considered to be uncontaminated. Although this might be true in large urban settings, in many semi-urbanized areas “farmers markets” selling local food products are becoming increasingly popular. As a consequence, it is important to assess whether people living in small- and medium-sized urban settings could be potentially affected by stack emissions by regularly eating food produced in the nearby countryside that may be potentially contaminated by waste incineration.

Third, previous risk assessment studies performed in Italy used the typical diet of North American (Cangialosi et al., 2008; Morra et al., 2006) or Spanish citizens (Lonati and Zanoni, 2012). However, the typical Italian diet is substantially different from the North American one (da Silva et al., 2009), and therefore it is relevant to assess whether assuming the consumption of one diet or the other may change the outcome of HRA studies.

5.3 Case-study: expected impact of the incinerator of Parma

In this section, I present the results of a multi-compartment model to estimate the potential long-term consequences on human health of the operations of the MSW incinerator of Parma (Italy). The

incinerator of Parma has started its activity in 2013. The Environmental Impact Assessment (EIA) procedure was carried out in 2008 according to the European legislation (EU Directives 85/337/CEE and 97/11/CE), but no quantitative estimation of expected health risks was ever performed. The emission data presented in the Environmental Impact Statement were used in our study to conduct the HRA.

The analysis was carried out specifically to investigate the effect on HRA of: (i) the activation of district heating fuelled by the MSW incinerator and the switch-off of domestic boilers, and (ii) different assumptions on dietary habits and geographical origin of production of food consumed by the resident population.

5.3.1 Materials and methods

HRA studies of incinerators' emissions typically focus on two COPC classes (Ollson et al., 2014; Roberts and Chen, 2006; Schuhmacher et al., 2004): (i) criteria pollutants (i.e., PM_{10}, NO_2, etc.) and (ii) micropollutants (i.e., dioxins, heavy metals, etc).

Different approaches have been developed to analyze these two categories of pollutants, as described below. Here, I focused on health effects due to chronic exposure under the assumption that the waste incinerator plant under study will be well managed (and thus emissions will never exceed legal limits) and, consequently, the probability of acute short-term exposure to high level of pollutants will be negligible. Figure 14 shows the conceptual model I adopted to estimate human health risks.

Risk assessment for criteria pollutants

Criteria pollutants have short residence times in the atmosphere due to degradation reactions, and do not usually show bioaccumulation properties. Direct inhalation through contaminated air is usually considered the main pathway of exposure (WHO Europe, 2006, 2000).

The number E of new cases per unit time (case year^{-1}) caused by a projected increase in ground concentration of atmospheric pollutants (or the number of cases avoided due to a reduction in ground concentration of pollutants) was estimated by means of Exposure-Response Functions (ERFs), as described in previous studies (Forastiere et al., 2011; Künzli et al., 2000; Phillips et al., 2010). Appendix B reports in detail the methodology used in the present study.

There is no general consensus on ERFs for chronic health effect of pollutants other than PM_{10}, e.g. NO_x and SO_x. In fact, PM_{10} concentrations generally exhibit a very high correlation with these criteria pollutants. Therefore, in epidemiological studies it may be particularly problematic to disentangle the health effect of SO_x and NO_x from that of PM_{10} (Torfs et al., 2007; WHO Europe, 2003). As consequence, I performed HRA for criteria pollutants only with reference to primary PM_{10} emissions so as to avoid double counting. The evidence for an independent effect of tropospheric ozone (O_3) is stronger. However, O_3 modelling requires the use of more complex photochemical models, and this was beyond the scope of the present work.

Appendix B reports the ERFs for PM_{10}. I noted that the relative risks for some health outcomes (i.e. stroke, acute bronchitis and asthma) are not statistically significant. However, since the central estimate suggests that an effect of PM_{10} on those outcomes is indeed possible, I decided to use these ERFs in our RHA consistently with a precautionary approach. For ERFs relative to $PM_{2.5}$, a ratio $PM_{2.5}/PM_{10}$ of 0.7 was assumed according to Medina et al. (2005).

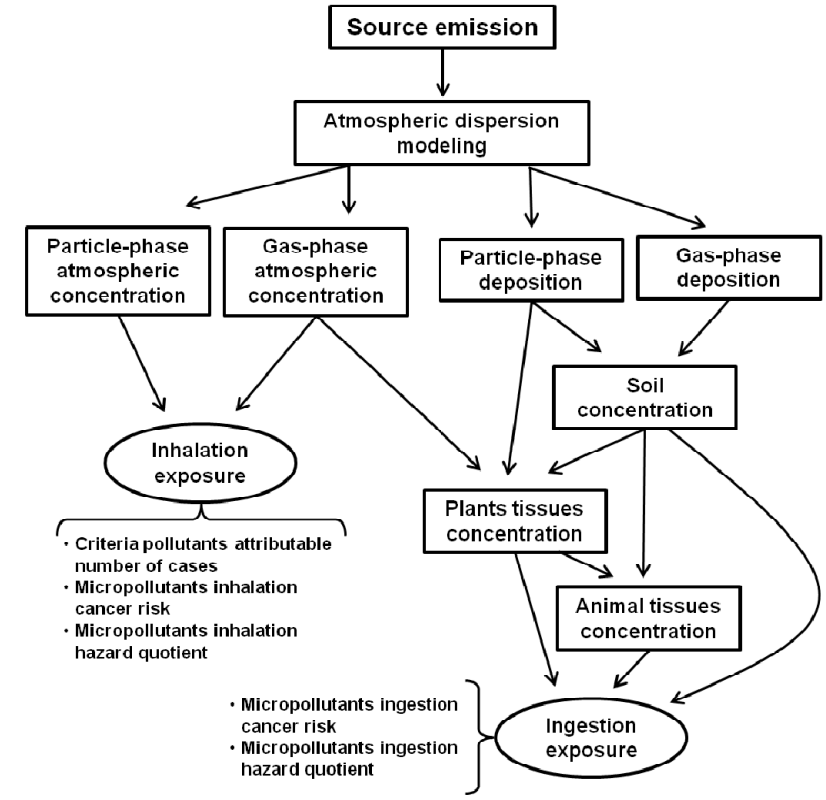

Figure 14 - Conceptual model for the diffusion of Chemicals Of Potential Concern (COPC) emitted from a point source through different environmental media and exposure pathways

Risk assessment for micropollutants

Micropollutants generally have high persistence in the environment and may exhibit bioaccumulation properties. The most relevant exposure pathway for humans is ingestion through diet (Fries, 1995; Linares et al., 2010; Llobet et al., 2003; Sweetman et al., 2000), even if inhalation and other ways of exposure, like ingestion of contaminated soil and water or dermal contact, may be of interest in particular situations (US-EPA, 2005a; WHO Europe, 2000). For this group of pollutants, I applied the model proposed by the United States Environmental Protection Agency (US-EPA) in the *Human Health Risk Assessment Protocol* (HHRAP) (US-EPA, 2005a). More details about the equations used for health risk calculation are reported in Appendix B.

For carcinogen pollutants, such as dioxins (PCDD/F) polycyclic aromatic hydrocarbons (PAH), polychlorinated biphenyl (PCB) and some heavy metals (HMs), the US-EPA model (US-EPA, 2005a) assumes there is no safe threshold dose below which there is no health risk. The health risk, measured as the probability of developing cancer during the entire lifetime (assumed to be 70 years), is estimated by multiplying the exposure dose by a Cancer Slope Factor (CSF), i.e. the estimate of the carcinogenic potency of the chemical. As a screening procedure, risks caused by different exposure pathways and carcinogen pollutants can be summed up to obtain the Total Cancer Risk (TCR).

For non-carcinogenic pollutants, a threshold dose is assumed to exist below which no appreciable health effects are expected. For each contaminant, the risk is computed in terms of Hazard Quotient (HQ), i.e. the ratio between the estimated daily dose and the reference dose (RfD). Accordingly, HQ>1 implies that the reference dose is exceeded for a specific contaminant. HQs due to exposure to different toxic pollutants can be then combined to determine the overall Hazard Index (HI).

As for micropollutants, I modeled PCDD/F (as equivalent 2,3,7,8-tetrachloro-dibenzo-p-dioxin), PAH (as benzo[a]pyrene), mercury (as Hg^0 , Hg^{2+} and methyl-Hg) and cadmium (as the sum of Cd+Tl authorized emissions). The CSFs and RfDs values reported in APPENDIX B were derived from the Integrated Risk Information System (IRIS) (US-EPA, 2008) and the Risk Assessment Information System (RAIS) (Oak Ridge National Laboratory, 2008).

Study area

The study area was chosen as a 16.75 x 20 km^2 rectangle centred on the city of Parma, divided in a 250 x 250 m^2 regular cell grid, for a total of 5360 cells. For each cell, I determined the prevalent land use (i.e., agricultural, urban and surface water), the farming (ha $cell^{-1}$) and breeding (tons $cell^{-1}$) intensity and the number of residents using available information from the Cartographic Database of the Emilia Romagna Region (RER, 2011). For the very few cells in which the prevalent land use was "surface water", I did not calculate the indirect risk for ingestion of contaminated soil or home-grown foods, nor did I assume exposure scenarios through contaminated fish as this is not relevant in this region.

The total population in the study area was 191,330 people. Information on the geographical distribution and the general age structure of the population was retrieved from the local registry office. Baseline incidences (I_0) for all causes of mortality, lung cancer, heart attack and stroke were respectively 1,102.48, 64.79, 90.12 and 108.99 case $year^{-1}$ per every 100,000 dwellers (ASR-ER, 2007).

Simulation of pollution dispersion from the incinerator and the district heating network

The MSW incinerator is located in a northern area of the city of Parma, about 4 km away from the city center. It has two grid furnaces, each one with a treating capacity of about 190 t day^{-1} of MSW and a maximum of 130,000 t $year^{-1}$. The plant is authorized to burn municipal solid wastes, sewage sludge, sanitary wastes and non hazardous special wastes. A description of stack emissions is reported in Table 14 along with pollutant emission rates derived from expected pollutant concentrations and gas flow rate reported in the Environmental Impact Statement.

Heat recovered by waste incineration will supply a district heating network of about 20,000 equivalent inhabitants, distributed in ten residential districts of the city (Figure 15). The estimated total amount of energy that will be provided in these districts by the heating network is 6.22x107 kWh $year^{-1}$.

I used the software WINDIMULA3® (Cirillo and Manzi, 1991; MAIND S.r.l., 2006) to model the atmospheric dispersion of pollutants. WINDIMULA3® is a multi-source Gaussian model that calculates deposition fluxes and allows the simulation of calm winds (i.e., wind speed < 1 m s^{-1}), a condition frequently encountered in the study area, namely ca. 20% of annual data on hourly basis at the stack height. I used meteorological data relative to the annual period 15/10/2005 – 14/10/2006 provided by the Regional Environmental Protection Agency.

I followed EPA (2005a) guidelines for the partitioning of different micropollutants between the gas and particle phase (APPENDIX B). Mercury was assumed to be released as Hg^0 and $HgCl_2$ and 2% of deposited mercury was assumed to speciate to Methyl-Hg in soil (US-EPA, 2005a).

Table 14 - Characteristics of point emission sources used in the atmospheric dispersion model.

Parameter	Value
Stack height (m)	70
Diameter (m)	2.15
Gas exit velocity (m s^{-1})	17
Gas flow (Nm^3 s^{-1})	40.0
Gas temperature (°C)	150
Functioning (hours $year^{-1}$)	8760
PM emission rate (mg s^{-1})	120.0
PCDD/F emission rate (ng s^{-1})	1.9
PAH emission rate (mg s^{-1})	0.6
Cd emission rate (mg s^{-1})	3.1
Hg emission rate (mg s^{-1})	3.1
Latitude (UTM WGS84, m N)	4,966,055
Longitude (UTM WGS84, m E)	607,155

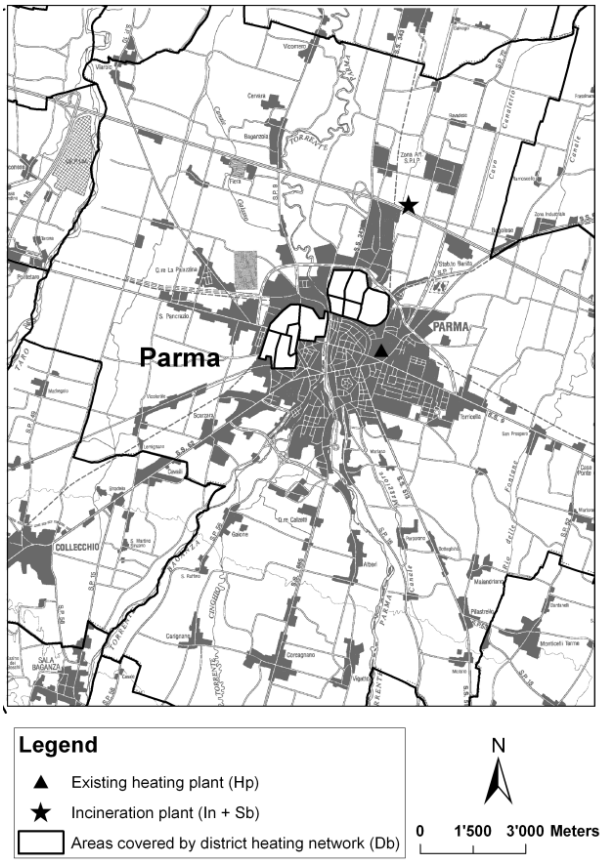

Figure 15 - Area considered in the study and location of the principal emission sources

I analyzed two emission scenarios for criteria pollutants. In the first scenario (ES1), I considered only the new MSW incinerator (In) modeled as a point source. In the second scenario (ES2), I also considered the activation of the district heating network, the concurrent switch-off of domestic boilers (Db) as well as the reduction of the activity of a pre-existing district heating plant (Hp). In addition, I also considered the emissions of a supplementary gas boiler (Sb) to be constructed within the incineration plant.

Emissions under ES2 scenario were analyzed on a seasonable basis, i.e. during the cold season (15 October – 15 April) and during the warm season (16 April – 14 October).

The emission balances (EB, mg year^{-1}) for the cold and warm seasons were:

$$EB_{cold} = +In+Sb-Db_{cold}-Hp \quad (1)$$

$$EB_{warm} = +In-Db_{warm} \quad (2)$$

I quantified emissions from domestic boilers on the basis of the estimated energy consumption in the residential areas that will be supplied by district heating and assumed that this energy is produced through methane combustion (Energy Agency of Parma, personal communication). I used an emission factor of 24.1 mg kWh^{-1} for PM_{10} as reported for methane combustion in residential boilers (SNAP code 020202) in the national guidelines for emission inventories (ISPRA, 2012).

To account for seasonal changes in heat and water uses, I assumed that 20% of the total energy consumption is attributable to hot water production throughout the entire year and 80% to house heating during the cold season only (ENEA, 2005)

The warm season emission rate (EF_{warm}, mg s^{-1}) for hot water production was assumed to be constant over the year and computed as follows:

$$EF_{warm} = Ehw_{tot}/(8760 \cdot 3600) \quad (3)$$

where Ehw_{tot} is the total emission due to hot water production (mg year^{-1}) and 8760·3600 is the number of seconds in an year.

The cold season emission rate (EF_{cold}, mg s-1) was computed as:

$$EF_{cold} = Erh_{tot}/(4392 \cdot 3600) + Ehw_{tot}/(8760 \cdot 3600) \quad (4)$$

where Erh_{tot} is the total emission as a result of house heating (mg year-1), and 4392·3600 is the number of seconds in the cold season.

The ten residential areas that will be connected to the new heating network (Figure 15) were aggregated for modeling purposes into two circular areas of equivalent surface and treated in the second emission scenario (ES2) as distributed pollution sources: the pollutant concentration from these two sources was subtracted from that derived from the incinerator. Their emission height was assumed to be 15 m (Table 15).

Table 15 – Characteristics of diffuse emission sources, treated in the atmospheric dispersion model as area sources. The ten residential areas that will be connected to the new heating network were aggregated into two circular areas of equivalent surface

Parameter	Source	
	Domestic boilers (Db) Zone A	Domestic boilers (Db) Zone B
Stack height (m)	15	15
Area diameter (m)	809	798
Functioning (hours year-1)	4392	4392
PM flux (mg s-1)	41.2	44.1
Latitude (UTM WGS84, m N)	4 963 401	4 962 363
Longitude (UTM WGS84, m E)	605 674	603 824

Dietary habits

It is well known that food consumption may change substantially depending upon regional habits and local culinary traditions, and this might significantly affect the potential intake of environmental pollutants from food (Undeman et al., 2010). The North American diet in particular is known to be quite different from the typical Italian diet (da Silva et al., 2009). I thus investigated whether alternative assumptions on diet composition may significantly affect health risk assessment.

Since detailed dietary data were not available for the province of Parma, I derived the average Italian diet composition from Turrini et al. (1991) on the basis of nation-wide food consumption data. According to the inclusion criteria presented in EPA (1997), I grouped food items in the eight food categories used in the EPA model (US-EPA, 2005a). When needed, I derived dry weight consumption values by correcting Turrini et al.'s data (1991) for water content on the basis of food composition tables developed by the Italian Institute for Research on Foods and Nutrition (INRAN, 2012). Average body weight for adults was set to 70 kg as in Walpole et al. (2012). No preparing and cooking losses were considered.

Health risk was first estimated for the Italian diet under the following assumptions (Table 16):

- all food consumed by residents people was produced in the study area and, thus, potentially contaminated (i.e., $Floc + Favg = 1$);
- in "rural" cells 100% of vegetables in the residents' diet were home-grown ($FlocVEG$=1) while only 50% of animal products was home-grown ($FlocANI$=0.5) and the other 50% was produced elsewhere within the study area ($FavgANI$=0.5);
- in "urban" cells 50% of vegetables was home-grown ($FlocVEG$=0.5, $FavgVEG$=0.5) and 100% of animal products came from the study area ($FavgANI$=1)
- 50% of livestock's diet came from the cell, the other 50% from the study area.

We named this scenario "Partial Home-grown Italian Diet" (PHItD) and used it as our reference exposure scenario.

In order to test the sensitivity of model results to alternative assumptions of food origin and diet type, we also computed health risk for the following exposure scenarios (APPENDIX B):

- "Full Homegrown North American Diet" (FHNAD): US-EPA (2005a) standard North American diet, with the dietary consumption for "rural" cells exclusively supported by food produced in the same cell of residence and no consumption of contaminated animal products in "urban" cells;
- "Full Homegrown Italian Diet" (FHItD): as above but with the Italian diet;
- "Full Mixed Italian Diet" (FMItD): Italian diet under the extreme assumption that the dietary consumption is entirely supported by a perfect mix of food produced within the overall study area. Under this assumption, the same food contamination level applies to all the cells of the study area

Table 16 - Assumptions for the four food consumption scenarios. Numbers represents fraction of food from the cell of residence (*Floc*) and from the study area (*Favg*).

	PHItD [b]		FHItD [b]		FMItD [b]		FHNAD [b]	
	farm	urban	farm	urban	farm	urban	farm	urban
Consumption profile [a]	ITA	ITA	ITA	ITA	ITA	ITA	USA	USA
FlocVEG for humans	1.0	0.5	1.0	1.0	0.0	0.0	0.0	0.0
FavgVEG for humans	0.0	0.5	0.0	0.0	1.0	1.0	1.0	1.0
FlocANI for humans	0.5	0.0	1.0	0.0	0.0	0.0	0.0	0.0
FavgANI for humans	0.5	1.0	0.0	0.0	1.0	1.0	1.0	0.0
Floc for animal diet	0.5		1.0		0.0		1.0	
Favg for animal diet	0.5		0.0		1.0		0.0	

a USA = from EPA (2005a) , ITA = calculated from Turrini et al. (1991)
b **PHItD** = Partial-Homegrown Italian Diet; **FHItD** = Full-Homegrown Italian Diet; **FMItD** = Full-Mixing Italian Diet; FHNAD = Full-Homegrown Nord American Diet.

Sensitivity analysis through Monte Carlo sampling

The results of a health risk assessment depend upon a large number of assumptions on processes and model parameterization. Therefore, I carried out a Monte Carlo analysis to assess what model parameters health risk is most sensitive to (Hwong-Wen, 2002; Lonati and Zanoni, 2012; Schuhmacher et al., 2001).

I implemented the "Tier 2" methodology presented in US-EPA (2001) as follows: first, for each model parameter I defined an uninformative uniform probability distribution over a range of ± 50% the mean value reported in US-EPA (2005a). I then drew a parameter value from each respective probability distribution and estimated the corresponding health risk in each of the 67x80 cells in which the study area was discretized. I replicated this process 10,000 times. For each cell I derived six percentiles of the distribution of risk (i.e., 2.5, 25, 50, 75, 97.5 and 99), and for each percentile I reported the average and maximum value over the entire area.

To determine which parameters mostly affected health risk, I computed for each model parameter j the Spearman rank correlation coefficients ρ_j between the 10,000 casually extracted values and the corresponding average health risk over the studied domain. ρ_j were then squared and normalized so as to sum to 1 and then ranked from the largest to the smallest. Thus, each coefficient represents the relative contribution of each input parameter to the total variance of the average risk of the area (Hwong-Wen, 2002; US-EPA, 2001).

To assess whether the resulting ranking was strongly affected by the shape of the probability distribution for model parameters, I re-ran the sensitivity analysis also by using beta distributions (shape parameters a=4, b=5) instead of uniform ones.

The Monte Carlo analysis was carried out only for the indirect risk of ingestion of micropollutants, as preliminary sensitivity analyses showed that inhalation caused only a very small incremental risk.

5.3.2 Results

Air quality simulations

Air quality simulations (Table 18, Figure 16) show that the emission sources analyzed in the present study provide a very small contribution to the observed annual mean concentration of PM_{10} in the study area, i.e. about 40 µg m^{-3} (APAT, 2008). The incinerator (I) and the domestic boilers (Db) are the most important emission sources, with maximum values of the annual mean of modelled 1-hour

concentrations - equal to 0.02 and 0.40 μg m^{-3}, respectively - expected to occur close to the emission sources. In terms of mass balance (Table 17), the activation of the new incinerator increases PM_{10} input in the atmosphere (+ 0.71 and + 1.74 tons $year^{-1}$ during the cold and warm season, respectively), even when considering the switching-off of the domestic gas boilers and the activation of the district heating network (Scenario ES2). Nevertheless, in the cold season the contribution of domestic boilers to the mean annual atmospheric concentration at ground level is expected to be an order of magnitude higher than PM_{10} concentration due to the other emission sources, including stack emissions from incinerators (Table 18). As a consequence, the full activation of district heating powered through heat recovery from the incinerator and the switch-off of domestic boilers results in a general reduction in atmospheric concentrations at ground level (max reduction over the study area: -0.76 μg m^{-3}) during the cold season. In the warm season, a small increase in concentrations is expected in a large part of the study area (max increase over the study area: + 0.01 μg m^{-3}) except for the city centre of Parma, where a reduction of atmospheric concentrations is still expected (max reduction over the study area: -0.04 μg m^{-3}) due to the switching-off of domestic boilers for the production of hot water (Table 18).

Average annual concentration at ground level and cumulative annual deposition for micropollutants are reported in Table 19. Maximum fall-out concentrations are expected in less populated areas close to the waste incinerator plant, which is the only source of micropollutants analyzed in the present analysis. Maximum concentrations calculated by the model are one to four order of magnitude below international guideline values for air quality (Table 19).

Table 17 - Tons of PM_{10} emitted by all sources in different periods. Negative signs correspond to an emission reduction according to Scenario ES2.

Emission Source	Total PM10 emission (t)		
	Cold period	Hot period	Year
Incinerator (In)	+ 1.90	+ 1.89	+ 3.78
Supplementary gas boilers (Sb)	+ 0.19	n.a.	+ 0.19
Domestic boilers (Db)	- 1.35	- 0.15	- 1.50
District heating plant (Hp)	- 0.03	n.a.	- 0.03
Net balance	+ 0.71	+ 1.74	+ 2.45

n.a. = not active

Table 18 - Results of atmospheric dispersion modelling for PM_{10}. Minimum, median and maximum over the study area for mean atmospheric concentrations of each period.

Averaging period	Statistic on the area	Mean PM_{10} concentration at ground level (μg m^{-3}) for each source [a]			
		In	Sb	Db	Hp
Cold period	minimum	$4.2x10^{-4}$	$4.3x10^{-5}$	$1.7x10^{-3}$	$3.4x10^{-5}$
	median	$2.1x10^{-3}$	$2.6x10^{-4}$	$1.1x10^{-2}$	$1.4x10^{-4}$
	maximum	$2.3x10^{-2}$	$3.9x10^{-3}$	$7.6x10^{-1}$	$8.9x10^{-3}$
Warm period	minimum	$3.8x10^{-4}$	n.a.	$6.4x10^{-5}$	n.a.
	median	$1.9x10^{-3}$	n.a.	$6.0x10^{-4}$	n.a.
	maximum	$1.8x10^{-2}$	n.a.	$4.4x10^{-2}$	n.a.

[a] Legend: In = incinerator, Sb = supplementary boilers, Db = domestic boilers, Hp=existing heating plant. (n.a. = not active)

Table 19 - Results of atmospheric dispersion modelling for micropollutants. Minimum, median and maximum on the study area for mean annual atmospheric concentrations and cumulative annual deposition fluxes. For concentrations, available reference values for health protection are shown.

Model output	Area statistic	*Chemical of potential concern (COPC)*				
		PCDD	Cd	PAH	Hg^0 [a]	Hg^{2+}
Mean annual concentration at ground level ($\mu g\ m^{-3}$)	minimum	$6.5x10^{-12}$	$1.1x10^{-5}$	$6.4x10^{-7}$	$2.2x10^{-6}$	$8.6x10^{-6}$
	median	$3.1x10^{-11}$	$5.1x10^{-5}$	$3.0x10^{-6}$	$1.0x10^{-5}$	$4.1x10^{-5}$
	maximum	$2.8x10^{-10}$	$4.6x10^{-4}$	$2.7x10^{-5}$	$9.2x10^{-5}$	$3.7x10^{-4}$
	reference value	*$3.0x10^{-7}$* [b]	*$5.0x10^{-3}$* [c]	*$1.0x10^{-3}$* [c]	*1.0* [b]	
Cumulative annual deposition ($mg\ m^{-2}\ anno^{-1}$)	minimum	$1.1x10^{-9}$	$1.1x10^{-4}$	$1.3x10^{-4}$	$2.6x10^{-4}$	$1.4x10^{-3}$
	median	$1.0x10^{-8}$	$1.5x10^{-3}$	$9.5x10^{-4}$	$1.6x10^{-3}$	$1.0x10^{-2}$
	maximum	$1.4x10^{-7}$	$1.8x10^{-1}$	$1.4x10^{-2}$	$3.0x10^{-2}$	$3.8x10^{-1}$

[a] *for Hg^0 only the gas phase is considered.* [b] *WHO Europe, 2000.* [c] *EU Directive 2004/107/CE*

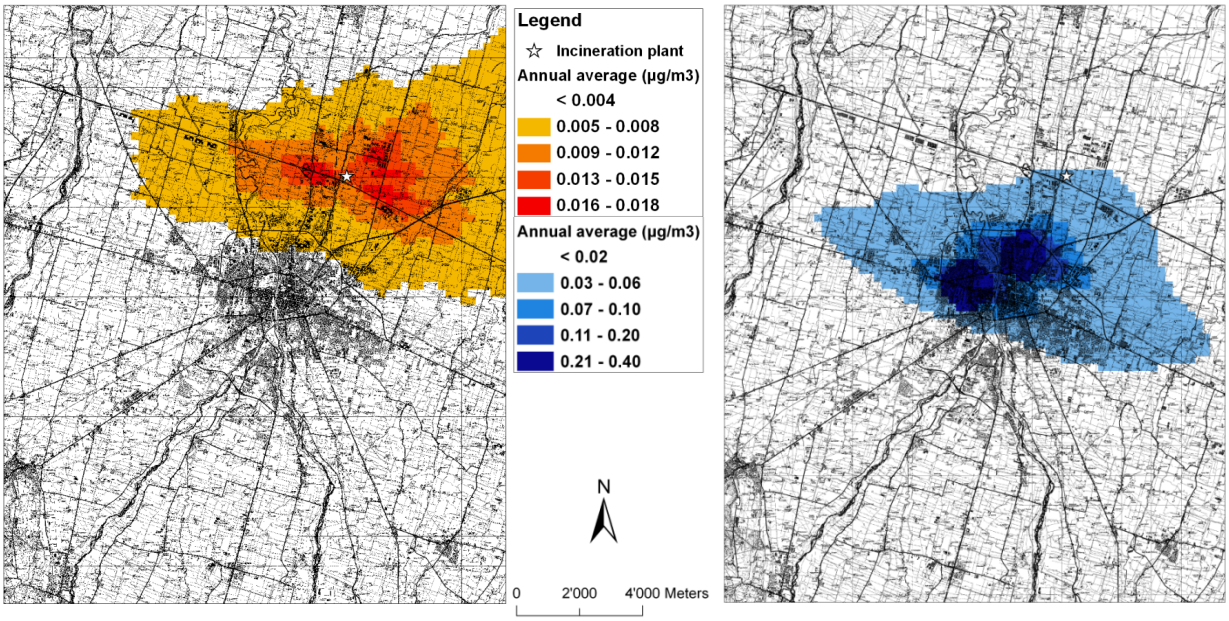

Figure 16 – Results for atmospheric dispersion of PM_{10} from the incinerator (sx) and domestic boilers (dx) (annual average of 1-hour concentrations, $\mu g\ m^{-3}$)

Health effect due to PM_{10}

The activation of the new incinerator alone (emission scenario ES1) is expected to cause a marginal increase in mortality and morbidity in the exposed population due to the increase in PM_{10} chronic exposure (Table 20). For general mortality, $1.6x10^{-2}$ additional cases $year^{-1}$ (95%CI: $5.3x10^{-3} \div 2.9x10^{-2}$) on a population of 191,330 exposed residents are expected, i.e. a +0.001% increase in the expected annual number of deaths. The activation of the district heating network powered by the incinerator (emission scenario ES2) is expected to reduce atmospheric PM_{10} concentrations in some populated areas in the centre of the city, resulting in an overall reduction, albeit small, in mortality and morbidity (Table 20, Figure 17). For general mortality, the model showed a reduction of $5.1x10^{-1}$ cases $year^{-1}$ (CI 95%: $-1.7x10^{-1} \div -9.3x10^{-1}$), namely ca. 10 cases less over 20 years, on the entire population (-0.024% in annual number of deaths).

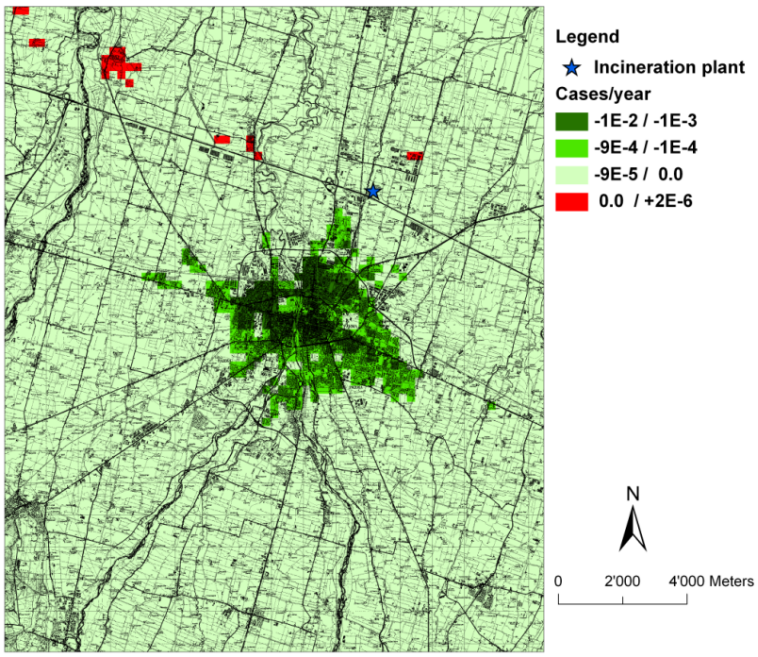

Figure 17 - Increase and decrease in mortality in different areas (case year^{-1} for each cell) for PM_{10} based on annual ground level concentration balance for emission scenario ES2

Table 20 - Health effects from exposure to PM_{10} for emission scenarios ES1 (incinerator only) and ES2 (incinerator and district heating), computed as a sum over the entire area (95% CI in parenthesis). For ES2, the resulting number represent the net balance between the expected reduction of cases in the areas of reduced exposure and the expected increase in the areas of increased exposure.

Health outcome [units]	Scenario ES1	Scenario ES2
All causes mortality [case year^{-1}]	$1.6x10^{-2}$ ($5.3x10^{-3}$; $2.9x10^{-2}$)	$-5.1x10^{-1}$ ($-1.7x10^{-1}$; $-9.3x10^{-1}$)
Lung cancer [case year^{-1}]	$1.2x10^{-3}$ ($1.5x10^{-4}$; $2.5x10^{-3}$)	$-4.0x10^{-2}$ ($-5.0x10^{-3}$; $-8.0x10^{-2}$)
Infraction [case year^{-1}]	$3.9x10^{-3}$ (3.0x10-3 ; $4.9x10^{-3}$)	$-1.2x10^{-1}$ ($-9.7x10^{-2}$; $^{-1}.6x10^{-1}$)
Stroke [case year^{-1}]	$5.2x10^{-4}$ (-1.3x10-3 ; $2.6x10^{-3}$)	$-1.7x10^{-2}$ ($4.2x10^{-2}$; $-8.4x10^{-2}$)
Acute bronchitis [case year^{-1}]	$9.8x10^{-3}$ ($-7.0x10^{-4}$; $2.0x10^{-2}$)	$-3.2x10^{-1}$ ($2.3x10^{-2}$; $-6.4x10^{-1}$)
Asthma in children, <15 years of age [extra days of bronchodilator usage year^{-1}]	$6.4x10^{-1}$ (-2.5 ; 3.8)	$-2.1x10^{1}$ ($7.9x10^{1}$; $-1.2x10^{2}$)
Asthma in adults, $\geq$ 15 years of age [extra days of bronchodilator usage year^{-1}]	$3.7x10^{1}$ ($-3.7x10^{1}$; $1.1x10^{2}$)	$-1.2x10^{3}$ ($1.2x10^{3}$; $-3.6x10^{3}$)
Restricted Activity Days , 15-64 years of age [num. year^{-1}]	$1.9x10^{1}$ ($1.7x10^{1}$; $2.2x10^{1}$)	$-6.2x10^{2}$ ($-5.4x10^{2}$; $-6.9x10^{2}$)
Work Lost Days ,15-64 years of age [num. year^{-1}]	4.7(4.0 ; 5.4)	$-1.5x10^{2}$ ($-1.3x10^{2}$; $-1.8x10^{2}$)
Minor Restricted Activity Days ,18-64 years of age [num. year^{-1}]	$1.2x10^{1}$ ($1.0x10^{1}$; $1.5x10^{1}$)	$-4.0x10^{2}$ ($-3.2x10^{2}$; $-4.7x10^{2}$)
Lower Respiratory Symptoms , 5-14 years of age [extra days year^{-1}]	6.6 (3.3 ; 9.8)	$-2.1x10^{2}$ ($-1.1x10^{2}$; $-3.2x10^{2}$)
Lower Respiratory Symptoms , >15 years of age [extra days year^{-1}]	$5.3x10^{1}$ (6.1 ; $9.9x10^{1}$)	$-1.7x10^{3}$ ($-2.0x10^{2}$; $-3.2x10^{3}$)

The analysis of the Italian reference diet shows great differences with respect to the standard US-EPA farmer consumption profile (US-EPA, 2005a). As shown in Table 21, vegetable consumption in the average Italian diet is three times greater than what US-EPA's suggest for North American diet, while animal products consumption is three times lower.

Detailed results of health risk estimation for the reference PHItD exposure scenario are reported in Table 22 and Figure 18, while risk assessment under alternative assumptions of diet composition and food origins are presented in Figure 19.

The maximum value of lifetime average daily dose (LADD) for PCDD/F over the area is $2.5x10^{-3}$ pg-TEQ kg_{bw}^{-1} day^{-1}, i.e. three order of magnitude smaller than the Tolerable Daily Intake (TDI) for this contaminant, i.e. 1-4 pg-TEQ kg_{bw}^{-1} day^{-1} (van Leeuwen et al., 2000).

Food of animal origin represents the principal exposure pathway to PAH, PCDD/F and Hg^{2+}, accounting respectively for 97%, 83% and 82% of the total dose, as a mean over the area (Figure 20). Exposure to Cd and MeHg, is primarily through the consumption of vegetables (respectively 82% and 61% of the total dose). The importance of soil ingestion for health risk is negligible. Inhalation is the only exposure pathway for Hg^0, and also plays an important role in exposure to Cd (16% of the dose).

The maximum value of total risk for carcinogens (TCR) over the entire area is $4.1x10^{-6}$. By combining the spatial distribution of resident population with that of health risk, a total of 0.29 new cases in 70 years attributable to incinerator emissions should be expected in the entire study area. Ingestion risk is strongly determined by the effect of PAH, while inhalation risk is mostly attributable to Cd.

The total Hazard Index (HI) for toxic pollutants reaches its maximum, i.e. 7.3% of the RfD, close to the incinerator, with a median over the entire study area of about 0.3% . For 94% of the population the HI value is below 1% of RfD. On average, over the entire study area, ingestion of MeHg and Hg^{2+} represents the most important contribution to total HI.

Table 21 - Comparison between the North American and Italian diet, calculated from Turrini et al. (1991). Food categories are those described in EPA (2005a). When needed, we derived dry weight consumption values by correcting Turrini et al.'s data (1991) for water content on the basis of food composition tables developed by the Italian Institute for Research on Foods and Nutrition (INRAN, 2012). Average body weight for adults was set to 70 kg as in Walpole et al. (2012). No preparing and cooking losses were considered.

Food category	ITA [a] (g kg_{BW}^{-1} day^{-1})	USA [b] (g kg_{BW}^{-1} day^{-1})	USA vs ITA (%)
Aboveground vegetables	0.57	0.47	-18%
Aboveground protected vegetables	3.36	0.64	-81%
Belowground vegetables	0.18	0.17	-5%
Beef	1.02	1.22	+20%
Pork	0.47	0.55	+16%
Poultry	0.46	0.66	+43%
Milk and dairy products	3.59	13.67	+281%
Eggs	0.34	0.75	+122%

a = Calculated from Turrini et al. (1991) , b = from EPA (2005a), farmer scenario

Table 22- Doses and health risks for micropollutants under the Partial Home-grown Italian Diet (PHItD) scenario. Median and (maximum) on the study area are shown. Number of cases is computed only for carcinogens, as a sum over the entire area.

COPC	Dose (mg kg^{-1} day^{-1})		Risk for carcinogens (adimensional)			Hazard Quotient for toxics (adimensional)	
	Ingestion	Inhalation	Ingestion	Inhalation	number of cases	Ingestion	Inhalation
PCDD	9.4x10^{-13} (2.5x10^{-12})	4.8x10^{-15} (4.3x10^{-14})	7.7x10^{-8} (2.1x10^{-7})	5.5x10^{-10} (5.0x10^{-9})	1.6x10^{-2}		
PAH	3.5x10^{-7} (8.5x10^{-7})	4.7x10^{-10} (4.2x10^{-9})	1.4x10^{-6} (3.4x10^{-6})	1.8x10^{-9} (1.6x10^{-8})	2.7x10^{-1}		
Cd	7.7x10^{-8} (1.3x10^{-5})	8.0x10^{-9} (7.2x10^{-8})		5.0x10^{-8} (4.5x10^{-7})	1.2x10^{-2}	7.4x10^{-5} (1.2x10^{-2})	
Hg0		1.6x10^{-9} (1.4x10^{-8})					3.3x10^{-5} (2.9x10^{-4})
Hg2+	4.9x10^{-7} (3.7x10^{-6})	6.3x10^{-9} (5.7x10^{-8})				1.6x10^{-3} (1.2x10^{-2})	3.5x10^{-5} (3.2x10^{-4})
MeHg	1.3x10^{-7} (5.4x10^{-6})					1.3x10^{-3} (5.2x10^{-2})	
TOTAL			1.5x10^{-6} (3.6x10^{-6})	5.3x10^{-8} (4.7x10^{-7})	2.9x10^{-1}	3.0x10^{-3} (7.3x10^{-2})	6.8x10^{-5} (6.1x10^{-4})

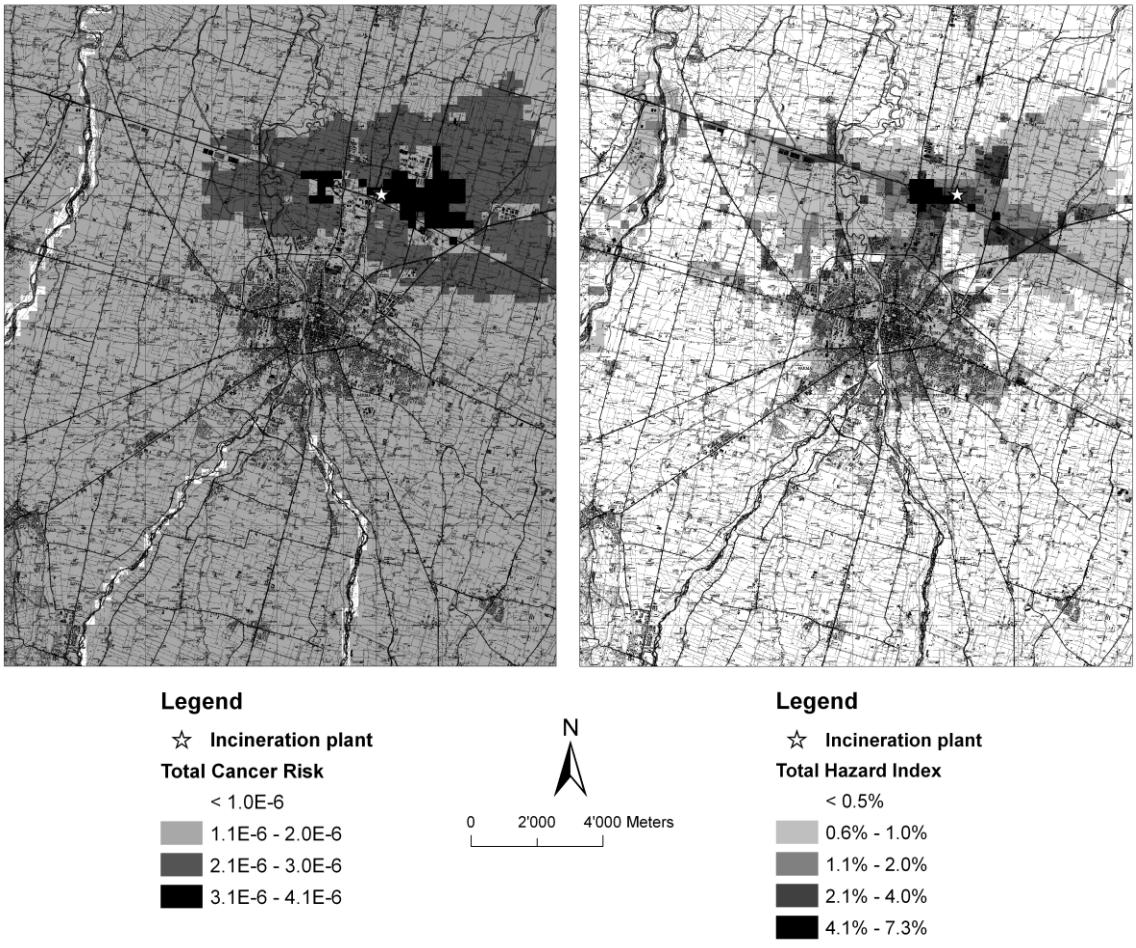

Figure 18 - Spatial distribution of Total Risk for Carcinogens (sx) and total Hazard Index (dx) over the study area.

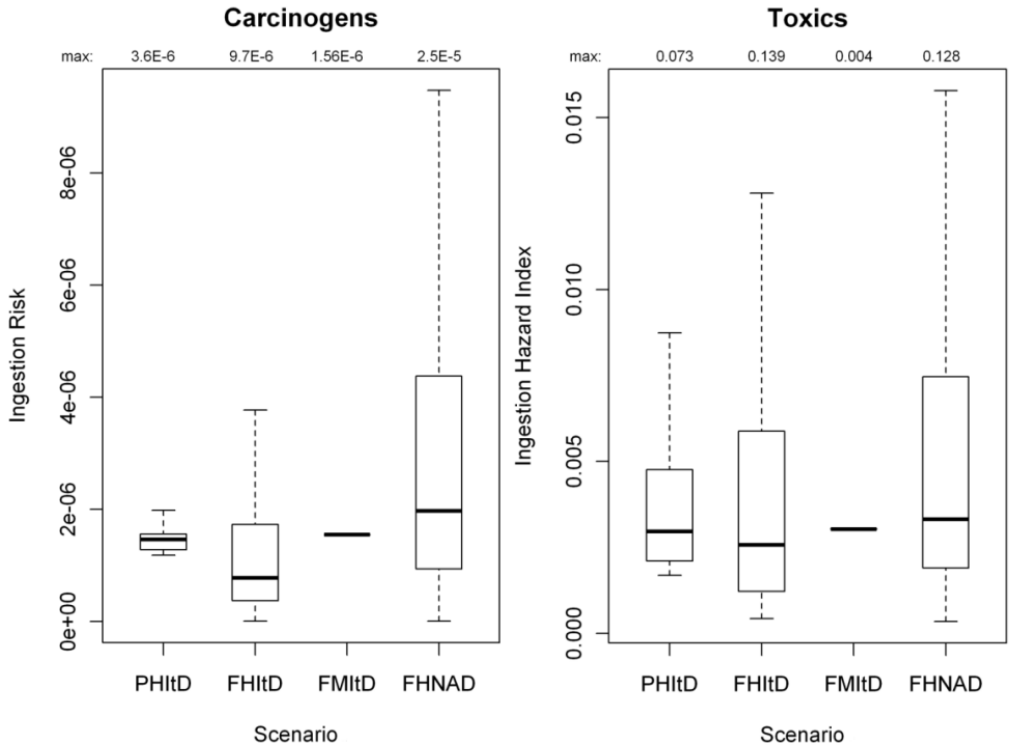

Figure 19 – Sensitivity analysis for different food consumption scenarios*. The box-plot show the distribution of ingestion risk for carcinogens (sx) and ingestion Hazard Index (dx) over the entire area (outliers not shown). The box represent the inter-quartile range (IQR), the horizontal line inside the box is the median value, the whiskers extend to 1.5 times the IQR from the box.

* **PHItD** = Partial-Homegrown Italian Diet; **FHItD** = Full-Homegrown Italian Diet; **FMItD** = Full-Mixing Italian Diet; **FHNAD** = Full-Homegrown Nord American Diet.

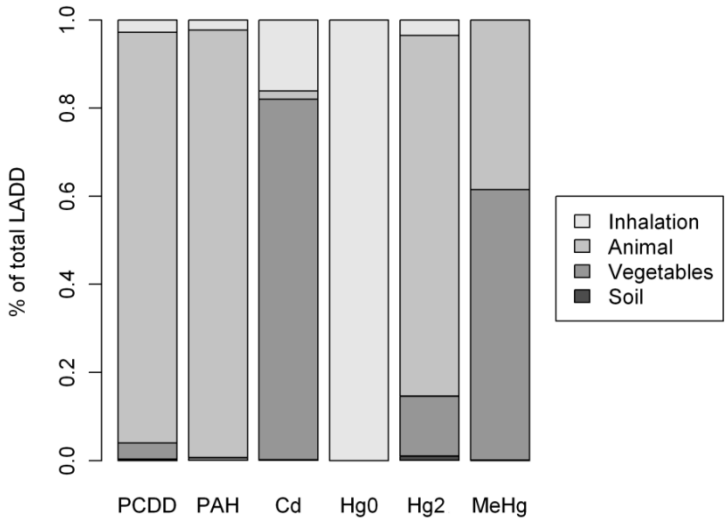

Figure 20 - Relative contribution (average over the study area) of different exposure pathways (inhalation and ingestion of food of animal origin, vegetables and resuspended soil) to the total dose for each contaminant

The Monte Carlo analysis highlights a substantial variability in risk for micropollutants. The 2.5th and 99th percentiles of the maximum TCR over the study area span over almost two order of magnitude between $3.3x10^{-7}$ and $1.1x10^{-5}$, respectively (Table 23). Furthermore, by accounting for population distribution in the study area, the resulting cumulative number of cases expected in 70 years ranges between 0.03 (2.5th percentile) and 0.65 (99th percentile).

For toxics pollutants, the maximum HI over the study area never exceeds 1 and ranges between 0.5% (2.5th percentile) and 15.2% (99th percentile). The 99th percentile of the Monte Carlo simulations for maximum HQs over the area equal to 6% for Cd, 5% for Hg^{2+} and 4% for MeHg. Simulations performed using a beta distributions of model parameters provided very similar results (data not shown).

The Monte Carlo analysis demonstrates that only a few parameters (between 7 and 12 over more than 90) explain the majority of the variability in health risk from micropollutant ingestion. Health risk is mostly sensitive to the parameters defining the timing of exposure, such as the exposure frequency (*EF*, 16-43% of variance), the averaging time for carcinogenic effects (*AT*, 18-14%) and the emission duration (*tD*, 2-10% of variance). Other influential parameters are the toxicological reference values (*CSF*: 12-18% ; *RfD*:, 17-21%), biotransfer factors between different compartments (e.g., *BvFOR*: 6-11%; *BrAG*: 13%) and parameters related to the food consumption for both humans (e.g., *CrMILKagr*: 2-8%; *CrAGPurb*: 3%; *FlocVEGurb*: 9%) and animals (e.g., *QpSILmilk*: 7-12%). Finally, parameters that determine the initial conversion from atmospheric deposition values to soil concentrations, such as the soil bulk density (*BD*, 2-14%) and soil mixing depth (*ZsURB*, 6%), play an important role.

Table 23 - Lifetime cancer risks and Hazard Indexes for micropollutants computed with Monte Carlo simulation. Median and (maximum) values on the study area are shown for six cut-off of the output risk distribution. Risk values are cumulated by type of effect and exposure pathway.

COPC category	Exposure pathway	Percentile on risk distribution					
		2.5	25	50	75	97.5	99
Carcinogens (risk)	Ingestion	$1.3x10^{-7}$ ($2.7x10^{-7}$)	$3.4x10^{-7}$ ($8.1x10^{-7}$)	$5.7x10^{-7}$ ($1.5x10^{-6}$)	$9.6x10^{-7}$ ($2.6x10^{-6}$)	$2.4x10^{-6}$ ($7.7x10^{-6}$)	$3.1x10^{-6}$ ($1.0x10^{-5}$)
	Inhalation	$7.4x10^{-9}$ ($6.6x10^{-8}$)	$1.7x10^{-8}$ ($1.5x10^{-7}$)	$2.6x10^{-8}$ ($2.3x10^{-7}$)	$3.9x10^{-8}$ ($3.5x10^{-7}$)	$8.1x10^{-8}$ ($7.3x10^{-7}$)	$9.7x10^{-8}$ ($8.8x10^{-7}$)
	Total	$1.3x10^{-7}$ ($3.3x10^{-7}$)	$3.6x10^{-7}$ ($9.5x10^{-7}$)	$6.0x10^{-7}$ ($1.7x10^{-6}$)	$1.0x10^{-6}$ ($3.0x10^{-6}$)	$2.5x10^{-6}$ ($8.4x10^{-6}$)	$3.2x10^{-6}$ ($1.1x10^{-5}$)
Carcinogens (n. of cases)	Total	$3.0x10^{-2}$	$7.0x10^{-2}$	$1.2x10^{-1}$	$2.0x10^{-1}$	$5.1x10^{-1}$	$6.5x10^{-1}$
Toxics (hazard quotients)	Ingestion	$7.6x10^{-4}$ ($5.1x10^{-3}$)	$1.6x10^{-3}$ ($1.5x10^{-2}$)	$2.3x10^{-3}$ ($2.6x10^{-2}$)	$3.4x10^{-3}$ ($4.4x10^{-2}$)	$7.0x10^{-3}$ ($1.2x10^{-1}$)	$8.4x10^{-3}$ ($1.5x10^{-1}$)
	Inhalation	$2.0x10^{-5}$ ($1.8x10^{-4}$)	$3.5x10^{-5}$ ($3.1x10^{-4}$)	$4.7x10^{-5}$ ($4.2x10^{-4}$)	$6.3x10^{-5}$ ($5.7x10^{-4}$)	$1.1x10^{-4}$ ($9.8x10^{-4}$)	$1.2x10^{-4}$ ($1.1x10^{-3}$)
	Total	$7.9x10^{-4}$ ($5.2x10^{-3}$)	$1.6x10^{-3}$ ($1.5x10^{-2}$)	$2.4x10^{-3}$ ($2.6x10^{-2}$)	$3.5x10^{-3}$ ($4.5x10^{-2}$)	$7.1x10^{-3}$ ($1.2x10^{-1}$)	$8.5x10^{-3}$ ($1.5x10^{-1}$)

5.3.3 Discussion

Exposure assessment quality

Concerning exposure characterization, the methods I used can be considered quite advanced compared to the level of detail typically achieved in epidemiological studies. In fact, risk assessment procedures requires the characterization of both direct (i.e. inhalation) and indirect (i.e. ingestion) routes of exposure.

I used atmospheric dispersion models, coupled with steady-state models for the transfer of pollutants in the food chain, to characterize the intensity of exposure to the emission source. Population distribution was characterized through the use of both exact address location (for the Parma municipality) and spatial disaggregation of available census data for the remaining municipalities. Overall, the resolution of my exposure estimates was 250 m. I did not evaluate temporal variability of exposure, attributable to modifications in the emissions of the incinerator or population mobility, but rather I assumed people to be exposed to the same average concentrations for their entire life.

Health effect of PM_{10}

The increase in mortality due to primary PM_{10} from the incinerator represents a very small fraction of expected number of deaths in the area (< 0.01%). Similar results were reported in other simulation studies (Forastiere et al., 2011; Roberts and Chen, 2006; Schuhmacher et al., 2004). Nevertheless, as the annual average of daily concentration of PM_{10} recorded by monitoring stations in the city of Parma in the simulated year 2006 was about 40 $\mu g\ m^{-3}$ (APAT, 2008), i.e., the annual limit for the protection of human health as regulated by the Italian law, efforts should be made to reduce additional sources of exposure in the area. The ES2 scenario shows that the negative effects caused by the small increase in PM_{10} concentration due to the new incinerator can be offset by the positive effects due to the switching-off of domestic boilers and the activation of a large district heating network powered by heat recovery from the incinerator.

To my knowledge, this is the first study that provides quantitative estimates of the perspective benefits of thermal energy recovery from incinerators for district heating in terms of reduced human health risks at the local spatial scale. Previous Life Cycle Assessment (LCA) studies have assumed that energy recovery from waste combustion can compensate for the electricity produced by fossil-fuel power plants. For example, Morselli et al. (2008) estimated the health impact of the entire Regional incineration system through LCA, accounting for the health benefit deriving from energy recovery and avoided emissions. However, these LCA analyses are based purely on emissions mass balance and, therefore, cannot explicitly account for the expected ground level concentration of pollutants in conjunction with the actual population distribution. This study shows that in term of mass balance, the increase in PM_{10} emissions due to the activation of the incinerator is not compensated for by the switching-off of the existing emission sources (Table 17). Yet, when accounting for both atmospheric dispersion and population distribution, the increase in ground level concentrations of PM_{10} due to waste incineration is more than compensated by the reduction in PM_{10} concentration as a result of switching-off of domestic boilers achievable through district heating (Table 20, Figure 17). Therefore, heat recovery for district heating is a key factor to effectively curb the environmental burden of a new waste incineration facility. Accordingly, efforts should be made to recover as much energy as possible, extend the district heating network so as to switch-off the largest number of domestic and non-domestic boilers. Additional benefits in terms of electricity consumptions could be obtained by using the district heating network also for air conditioning during the hot season.

Health effects due to micropollutants

According to air quality simulations, the expected contribution of the MSW plant to long-term atmospheric concentrations of micropollutants at ground level is almost negligible when compared to international limits or guidelines (Table 19). Health risks computed through the use of the HHRAP model (US-EPA, 2005a) are also moderate (Table 22).

The maximum value for TCR estimated in the area (i.e., 4.1×10^{-6}) is within the acceptable risk range (i.e., 10^{-6} to 10^{-5}) recommended by US-EPA (1998b) for exposure to emissions from a single facility.

Furthermore, when accounting for population distribution, less than 1 new mortality case due to waste incineration is expected over a lifetime. The maximum value for the total HI for toxics (i.e.,7.3% of the RfD) is smaller than the reference value of 25% recommended by US-EPA (1998b) to account also for potential background exposures. These results derive from a set of conservative (i.e. protective) assumptions. For example, I assumed that the food consumed by residents, either purchased in the markets or home-grown, was in some way all contaminated.

The 99th percentiles derived from the Monte Carlo simulation (Table 23) for (i) maximum TCR (i.e., 1.1x10^{-5}), (ii) number of lifetime cancer cases over 70 years (i.e. 0.65) and (iii) maximum HI (i.e. 15%) confirm that, even in the worst-case scenario, the health impact of the incinerator is expected to be moderate.

These results are in general agreement with those from other recent studies on health risk assessment for point emission sources that used a similar methodology (Cangialosi et al., 2008; Lonati and Zanoni, 2012; Ollson et al., 2014; Roberts and Chen, 2006), although there are differences in the types of pollutants and exposure routes considered between our study and the ones cited above.

Dietary habits

My analysis shows that the Italian reference diet derived from Turrini et al. (1991) data is quite different from the standard "reasonable maximum exposure" consumption profile suggested by US-EPA (2005a) (Table 21). The imbalance between consumptions of foods of vegetable and animal origin is in line with data presented in other available datasets on national dietary habits (WHO, 2013b). Data from Turrini et al. (1991) used to derive the Italian diet profile were the only data sufficiently detailed to be grouped in the eight food categories used in the EPA model (US-EPA, 2005a). A comparison of aggregated consumption from Turrini et al. (1991) combined with the results of a more recent Italian survey (Leclercq et al., 2009) shows no appreciable differences. Another recent study confirmed that food consumptions did not vary significantly from 1991 to 2006 in Northern Italy (Pelucchi et al., 2010). While more detailed surveys of diet in the Parma province are encouraged, I am confident that the diet profile I derived by using data from Turrini et al. (1991) provides a fairer representation of the local diet than the one presented in US-EPA (2005a).

The sensitivity analysis shows that dietary choices might have relevant implications in terms of exposure to environmental contaminants (Figure 19). The more vegetarian profile of the Italian diet with respect to the North American one (US-EPA, 2005a) notably reduces health risk caused by carcinogens, as the exposure to PCDD/F and PAH is primarily through ingestion of animal food. As for toxic pollutants, the FHItD and FHNAD exposure scenarios provide similar results, since there is a compensation between reduced exposure to Hg^{2+} (primarily via animal food) and increased Cd and MeHg exposure (primarily via vegetables).

The assumption in the PHItD reference scenario that the fraction of food purchased in local markets has a contamination given by a weighted average of contamination over the whole study area (APPENDIX B), leads to a homogenization of health risk with respect to the two full home-grown scenarios FHItD and FHNAD. Risk is reduced in areas of maximum fallout and increased in more populated areas of low fallout. Overall, the PHItD scenario results in a higher number of expected cancer cases (i.e., 0.28 for PHItD, 0.04 for FHItD and 0.09 for FHNAD). This sensitivity analysis reinforce the indication given in Chapter 4 about the use of atmospheric concentrations or depositions in epidemiological analyses: atmospheric deposition at residential address is of interest only if a certain fraction of the diet is home-grown, whereas atmospheric deposition over cultivated land is more relevant whenever a relevant fraction of the diet is covered by foods purchased on the market.

It should be noted that in both the PHItD and FMItD cases, COPC concentration in the food purchased in local markets depends upon the extension of the study area: the larger the area the lower the average level of contamination due to waste incinerator emissions. A careful definition of the

extension of the study area is thus important to derive reliable estimates of food contamination and health risk.

Caveats and limitations

As in any modelling analysis, I provided here a simplified version of a complex system aimed at grasping the fundamental processes characterizing exposure to pollutants and their effects on human health. As such, my analysis is not exempt from limitations and additional research would be helpful to consolidate the methodology.

First, I analysed only the long term health effect of primary PM_{10}. I did not account for other criteria pollutants - such as NO_x and SO_2, or for photochemical pollutants like tropospheric O_3 and secondary particulate matter - nor for possible acute exposure to high concentrations in the case of malfunctioning of the incinerator's pollution abatement system. However, other studies have shown that actual emissions from these type of waste incinerator plants can be an order of magnitude smaller than the maximum legally authorized PM_{10} emissions rate used in the present study (Biancolini et al., 2011; Buonanno et al., 2011, 2009; Forastiere et al., 2011). Therefore, my analysis should provide a conservative (i.e. worst-case) estimation of the associated health risk.

A unique emission factor for domestic boilers was derived by ISPRA (2012) and used under the realistic hypothesis that all the domestic heat is produced in Parma through methane combustion. A more detailed analysis will be required in the future to provide a better characterization of the diffuse emission sources, since combustion for domestic heat generation is among the major sources of air pollution in urban settings along with traffic. Additionally, in this analysis I considered PM_{10} from domestic boilers and incinerator as causing the same health damage. Particulate matter produced by waste or natural gas combustion may be very different for size distribution and composition, thus its toxicity and health effect may be dissimilar and non-comparable. The exposure-effect functions I used are referred to urban atmospheric PM_{10}, thus they are not source-specific.

Moreover, several caveats characterize the assessment of health effects of micropollutants through the HHRAP model (US-EPA, 2005a), as thoroughly discussed in US-EPA (2005b). Hofelt et al. (2001) suggested that the HHRAP model largely over-predict bioaccumulation of PAH in animal foods. In my results, PAH was in fact the most important carcinogen, on average representing 89% of total carcinogenic risk. Other authors highlighted a substantial over-prediction of risks deriving from exposure to mercury (Palma-Oliveira et al., 2012). Overall, the HHRAP model could be considered conservative.

The number of micropollutants analysed in the present study was limited by the lack of data in the Environmental Impact Statement of the proposed project on expected emissions of contaminants such as HMs and PCB. As overall health risk is assumed to be additive, simply accounting for further micropollutants would ultimately increase the TCR and HI. Further analysis will be required to account for the potential impacts of other pollutants not considered in the present study. To partially compensate for this limitation, I modelled PCDD/F and PAH in terms of their most toxic congeners (i.e., 2,3,7,8-TCDD and benzo[a]pyrene), since there is no information currently available about the emission profile of the stack.

Finally, the Monte Carlo simulation was carried out assuming simple non-informative uniform probability distributions for parameter sampling. The use of uniform distributions increased the uncertainty associated to health risk estimation and yielded values of the upper percentiles of health risk higher than in the case of beta distributions (data not shown), thus providing health risk estimates consistent with a worst-case scenario. Moreover, as I wanted to focus our attention on the post-dispersion and -deposition processes, I did not carry out a sensitivity analysis for the parameters of the air quality model (e.g., particulate diameter, gas-particle partitioning, etc.). Here I refer to published

studies explicitly addressing this issue (Carlos García-Díaz and Gozalvez-Zafrilla, 2012; Lonati and Zanoni, 2012; Yegnan et al., 2002).

5.3.4 Conclusion

The present study showed that, under the specific set of assumptions on exposure scenarios, transmission pathways and on the basis of the state of the art methodology for HRA, the proposed waste incinerator plant appears to cause a negligible increment to health risk, as long as it will be properly managed and emissions will never exceed the legal limits.

My analysis also showed that the activation of district heating powered through heat recovery from waste incineration has the potential to compensate for stack emissions of particulate matter. Under these circumstances, it can be reasonable to locate a modern waste incineration plant in proximity of a densely populated urban area as long as it is possible to recover heat from waste combustion, convey it into an extensive district heating network and switch off as many domestic boilers as possible. Nevertheless, the reduction in exposure to macropollutants shall be carefully balanced against the increased exposure to micropollutants.

The caveats and limitations listed above suggest that my study should be considered a preliminary analysis and that the results should be taken cautiously. Moreover, even though my study showed a limited risk for the set of micropollutants analysed, concern of resident population for plant mismanagement or for unexpected accidents or unreported emissions has to be considered legitimate and should not be undervalued. Adequate surveillance and monitoring systems should be also implemented to guarantee that emissions never exceed the authorized limits.

Whenever a choice was possible or practical given the available information, I made assumptions that generally tended to overestimate rather than underestimate the potential risk for human health. While several improvements and refinement could be certainly implemented in future studies of health risk assessment, I am confident that this modelling framework provided a robust preliminary estimate of health impacts of waste incineration under a variety of alternative exposure scenarios.

6
EXPOSURE TO DIFFUSE AIR POLLUTION: LAND USE REGRESSION (LUR) MODELS

6.1 Exposure assessment to diffuse pollution

In the previous chapters I analyzed methodologies for exposure assessment to specific point sources of atmospheric emissions. The aim of this type of exposure assessment is to identify the possible health effect attributable to a single specific source of pollution in a territory.

In many cases epidemiological researchers are interested in studying the relationship between atmospheric pollution and health in the general population, independently from the specific source(s) that cause pollution. In this chapter I will refer to the concentration of a pollutant that cannot be easily traced back to a single or definite source as "diffuse pollution". A large body of literature has investigated the effects caused by diffuse pollution, like particulate matter (PM_{10}, $PM_{2.5}$), nitrous oxides (NO_x) and ozone (O_3), on human health (Anderson et al., 2012; Hamra et al., 2014; Hoek et al., 2013; IARC, 2013; Rückerl et al., 2011; WHO, 2013a).

Different approaches for exposure assessment to diffuse air pollution are available, depending largely on data availability and study design. In contrast to source-specific analyses, exposure assessment to diffuse pollution is mainly based on measurement of atmospheric concentrations. When the researchers are interested only in evaluating the health effect of a diffuse pollutant (e.g. health effect of urban PM_{10}), the inability to identify the contribution of each specific source to measured concentrations may not be a limitation.

A recent symposium at the International Society of Exposure Sciencee, has proposed a tiered classification (Figure 21) of air pollution exposure prediction approaches used in air pollution epidemiology studies (Baxter et al., 2013; Özkaynak et al., 2013).

The easiest method is the direct use of measured concentrations from fixed monitors as a proxy of exposure for large groups of population. The majority of the cohort studies conducted in the '90s have compared mortality rates between cities, with exposure characterized by the average concentration measured at a central site within each city (Dockery et al., 1993; Pope et al., 1995). This method uses data averaged over space and time and the attribution of the same exposure to large populations (e.g. an entire city): in many urban settings, within-city variation in air pollution may be as large as between city differences (Jerrett et al., 2005; Wilson et al., 2005).

A more complex approach uses Geographical Information System (GIS) techniques to increase the spatial resolution of exposure data. GIS-based modeling combine monitoring data with statistical models to predict spatial patterns in ambient concentrations. One example is the use of spatial interpolation algorithms to generate continuous surfaces of air pollution from sparse monitoring data (Liao et al., 2006; Wong et al., 2004). Another modeling technique widely used in recent epidemiological studies is Land Use Regression (LUR) modeling (Hoek et al., 2008). LUR modeling will be the focus of the following paragraphs.

The use of atmospheric dispersion models introduces a further refinement in the spatial and temporal resolution of exposure modeling (Ozkaynak et al., 2008; Zou et al., 2009). These models incorporates available knowledge on the physical and chemical processes that determine pollution transfer in the atmosphere, requiring detailed data on meteorology and emission sources. Moreover, model and parameter uncertainties may limit the reliability of such detailed results.

Data from different sources (e.g., measured concentration, remote sensing images, air quality models) can be blended together with a variety of statistical models to obtain more accurate estimation of human exposure (Dionisio et al., 2013; Fuentes and Raftery, 2005; Yu et al., 2009). Data blending can be used to enhance spatial or temporal resolution of available monitoring data.

At the highest tier there are models that combine ambient concentrations estimated with the methodologies described above with microenvironmental and behavioral exposure models. These models incorporates information on human activity and behaviors for particular individuals or groups, in addition to demographic characteristics and features of indoor environments (Breen et al., 2014; Burke et al., 2001; Georgopoulos et al., 2005; Mölter et al., 2012). The level of detail in the spatial and temporal characterization of exposure depends largely on the spatial and temporal resolution of the input data.

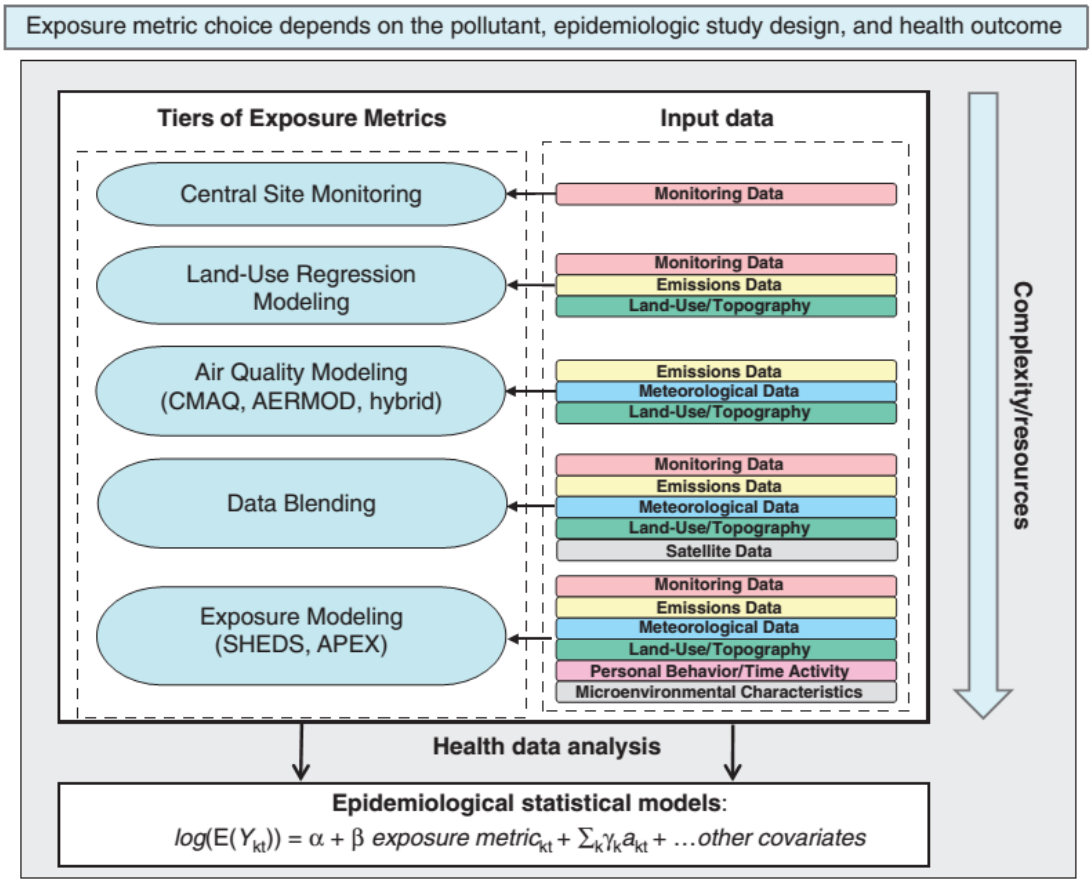

Figure 21 - Tiers of exposure metrics relevant to air pollution epidemiology studies. Source: Özkaynak et al. (2013)

The central question is thus to what extent exposure estimates generated from complex models represent an improvement over those generated from central-site monitoring data? (Özkaynak et al., 2013). The answer cannot disregard epidemiological study design, the type of health outcome considered (i.e. acute *vs.* chronic) and the specific pollutant of interest. In ecological and aggregate-level study designs (e.g. time-series analyses), where exposure assessment aims at capturing the population-level day-to-day variation in pollution levels, the use of central site monitors may be adequate as long as

pollution concentrations at different places are well correlated in time. In this case data blending techniques can be used to enhance the temporal resolution of monitoring data. In cohort studies, where exposure is estimated at individual level, both spatial and temporal contrasts in exposure are of interest and the use of sparse monitoring data must be considered inadequate, especially for pollutants that exhibit marked spatio-temporal heterogeneity (e.g. CO, NO_x, EC).

6.2 Land use regression models (LUR)

The Land Use Regression (LUR) methodology, also named regression mapping, seeks to predict pollution concentration at a given location based on surrounding land characteristics (e.g., land use, traffic intensity, proximity to emission sources, meteorology, etc.) (Hoek et al., 2008; Jerrett et al., 2005).

More formally, a statistical relationship between pollution concentrations (y) observed at a small number of sampling locations and a set of predictors (x_i) representing surrounding land features is developed, typically through least squares regression methods:

$$y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + c$$

This relationship is then used to predict concentrations at a large number of unsampled locations, given the values of the predictors x_i for those locations.

The application of land use regression (LUR) modeling for exposure assessment has been introduced in the SAVIAH study (Briggs et al., 1997). Developments in GIS have recently contributed to the increased use of LUR models in epidemiological research. Published studies developed LUR models for nitrous oxides (NO2, NOx) (Beelen et al., 2013), particulate matter (PM_{10}, $PM_{2.5}$) and its constituents (Hoogh et al., 2013), elemental and black carbon (EC, BC) (Dons et al., 2013), volatile organic compounds (VOCs) (Kheirbek et al., 2012) and polycyclic aromatic compounds (PAH) (Noth et al., 2011). Among others, NO_2 is one of the most used pollutants, largely because of the simplicity of monitoring this pollutant with passive samplers (Hoek et al., 2008).

6.2.1 The ESCAPE methodology

Recently, the European Project called ESCAPE (European Study of Cohorts for Air Pollution Effects) used LUR models to investigate the relationship between exposure to air pollution and health in 36 study areas (www.escapeproject.eu). ESCAPE developed a standardized procedure for LUR implementation that identifies common criteria for (i) selection of sampling sites, (ii) definition of GIS predictors, (iii) development of multivariate regression models and (iv) evaluation of model quality and performance (Beelen et al., 2013; ESCAPE project, 2010).

Briefly, the ESCAPE protocol requires a minimum of 20 sampling sites for $PM_{2.5}$ and 40 sampling sites for NO_2. Sites must be representative of the anticipated spatial variability of air pollution in the study area and are classified as regional background, urban background and traffic sites. Measurements are conducted for three 14-day periods in one year, representing the warm, cold and intermediate seasons. At least one sampler must be co-located with a fixed monitoring site measuring pollution year-round: this allows to correct the calculated annual average concentration of the samplers by the difference with respect to the annual average concentration at the fixed site.

The protocol also makes a series of a-priori choices on possible GIS predictors, i.e. the sources of data needed to represent specific land features (e.g. land use, traffic, altitude, population density, etc.), the expected direction of effect of each predictor (e.g. positive influence of traffic intensity in buffer, negative influence of distance from a major road, etc.) and the size of the buffers used to calculate predictor variables (e.g. for land use 100,300,500,1000,5000 m).

Model development follows a supervised forward stepwise procedure, finalized at finding the model that explains the maximum percentage of variance in the data:

1. Univariate linear regression models are developed for all available predictor variables
2. The model with the highest coefficient of determination (R^2) and a slope of the pre-specified direction is chosen as the start model
3. All the remaining variables are added consecutively to the start model and only the predictor with the highest increase in adjusted R^2 is retained if (i) the increase in adjusted R^2 is more than 1%, (ii) the coefficient conformed to the *a priori* direction of effect and (iii) the direction of effect of the predictor already in the model does not change.
4. Additional variables are added to the model until no variable add more than 1% of explained variance (adjusted R^2).
5. Finally variables with p-values larger than 0.1 are sequentially removed to obtain the final regression model.

Diagnostic tests are applied to evaluate the quality of the final model:

a. Variance Inflation Factor (VIF) is used to evaluate variable multi-colinearity: if VIF > 3 the variable is removed from the model
b. Cook's D is used to evaluate the presence of influential observations: if D > 1 for a specific observation, this can be removed from the dataset used to develop the model or the specific predictor that makes the observation too influential is not offered to the stepwise procedure.
c. Graphical analyses are used to evaluate heteroschedasticity and normality of residuals
d. Variograms of Moran's I are used to analyze possible spatial autocorrelation in model residuals.

Finally, model performance is tested using leave-one-out cross validation (LOO-CV) method: the final model is fitted to $N-1$ sites and is then used to predict the concentration of the excluded site. The procedure is repeated N times and the overall level of fit between observed and predicted values across all sites (R^2_{LOO}) and root mean square error ($RMSE_{LOO}$) are used as a measure of model performance.

6.3 Development of a LUR model for the Province of Parma

The Regional Environmental Protection Agency in Parma (ARPA-PR) maintains a network of NO_2 passive samplers that covers the territory of the entire Province of Parma, which has a surface of about 3400 km^2 (Figure 22). Multiple seasonal measurements are available for each sampling site, ranging from a minimum of 1 data per site for the newest sites introduced in 2013, to a maximum of 19 data for those sites present since 1996. Sampling is normally carried out for a 7-day period during winter and autumn, but for 1999 and 2012 data are available also for the hot season.

One of the debated aspects of LUR models is their transferability in space: since these models are purely statistical (stochastic) and are developed empirically from site-specific measurements, it is questionable whether one model can be transferred from a city or a region to another. Results from previous studies indicated caution when transferring models between different cities (Allen et al., 2011; Jerrett et al., 2005) or countries (Vienneau et al., 2010). Sometimes the bad performance of LUR models when transferred depends on the heterogeneity of available GIS predictors or the difference in pollution sampling techniques.

Moreover, in Chapter 3 I highlighted the importance of considering exposure to environmental confounding factors when analyzing the health effect attributable to a specific point emission source. The most advanced studies (Candela et al., 2013; Ranzi et al., 2011) used atmospheric dispersion models to reconstruct people's exposure to a variety of confounding emission sources around the incinerators under study. To develop a dispersion model for a large number of emission sources is highly resource and time demanding, since physical characteristics and emission rates of each single source modeled need to be known. To date, no epidemiological study on incinerators have explored the use of LUR models to control for exposure to confounding diffuse pollution in the study area.

The aim of the study presented in this chapter was twofold:

a. To develop a LUR model to estimate annual average exposure to NO_2 in the population of Parma, to be used in controlling for confounding from diffuse pollution in the analysis of the health impact of the local municipal solid waste incinerator
b. To study the difference between inner-city and outer-city NO_2 pollution, evaluating the differences in models calibrated on an urban and extra-urban context and their performance in out-of-sample validation.

I here defined an *urban area* as the area occupied by the city of Parma and an *extra-urban area* as the remaining part of the territory, although many of the extra-urban samplers were located in small towns or urbanized areas.

In the following paragraph I will describe the construction of the LUR model, while the issue of confounding of effect will be further discussed in paragraph 6.4.

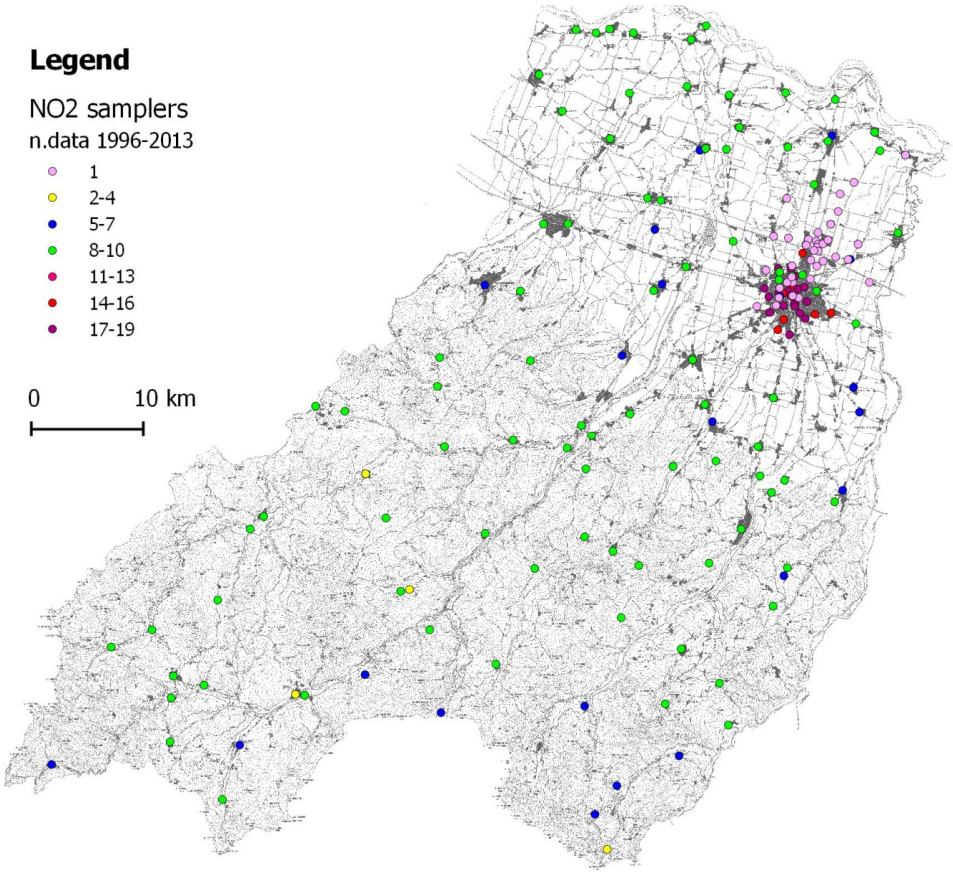

Figure 22 – Spatial distribution of NO_2 passive samplers over the Province of Parma. The colors corresponds to the number of seasonal 7-day measurement conducted between 1996 and 2013 in each sampling location.

6.3.1 Material and methods

Overview of sampling data

To develop a LUR model that can represent average annual exposure to diffuse NO_2 pollution, I used data from the three sampling campaigns of June 2012, November 2012 and February 2013 to calculate an annual average concentration in all sites. Overall, 118 sampling location with no missing data for these three campaigns were available (Figure 23). Table 24 lists the number of sampling sites available in the period 2012/2013 split by site type, as classified by ARPA-PR.

The estimated annual average NO_2 concentration varied between 3.6 and 87.7 μg m^{-3} over the Province, with the highest concentrations registered in the Parma urban area. Pollution variability and seasonal differences are higher in the urban area, while the extra-urban context is more homogeneous both in space and time (Figure 24). The variability of concentrations by site type is depicted in Figure 25: as expected, there was an increasing trend in concentrations between background, residential and traffic monitoring sites. Although the same trend is present in both urban and extra-urban samplers, concentrations at traffic sites in the extra-urban area is comparable to background concentrations in the urban area.

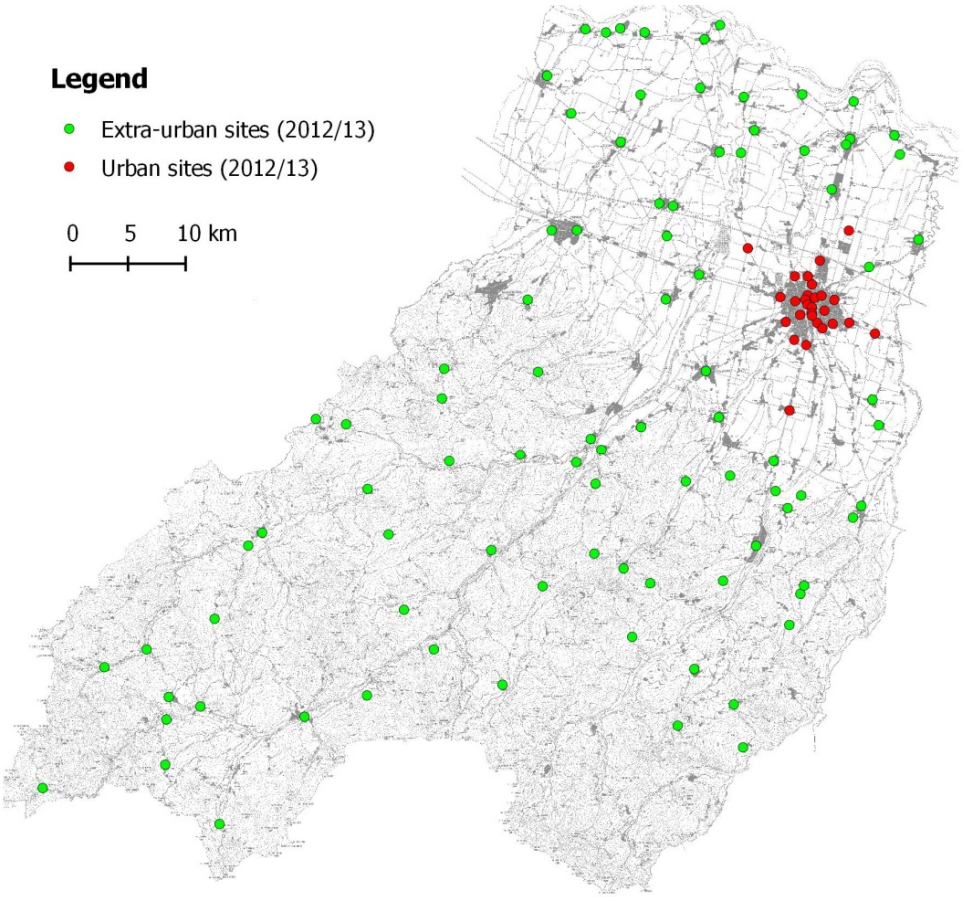

Figure 23 - Spatial distribution of NO_2 passive samplers with three available data for the period June 2012 – February 2013. The site are divided into urban (i.e. inside the Parma city area) and extra-urban.

Table 24 – Number and type of sampling site with available data for the period 2012/2013, as classified by ARPA-PR

Area	Typology	n
Parma urban area (red dots Figure 23)	Parma historical centre / traffic	2
	Urban background	2
	Residential / industrial	1
	Residential / traffic	6
	Traffic	17
Extra-urban area (green dots Figure 23)	Background	33
	Residential	46
	Residential / traffic	9
	Residential / industrial	1
	Rural background	1
	TOTAL	**118**

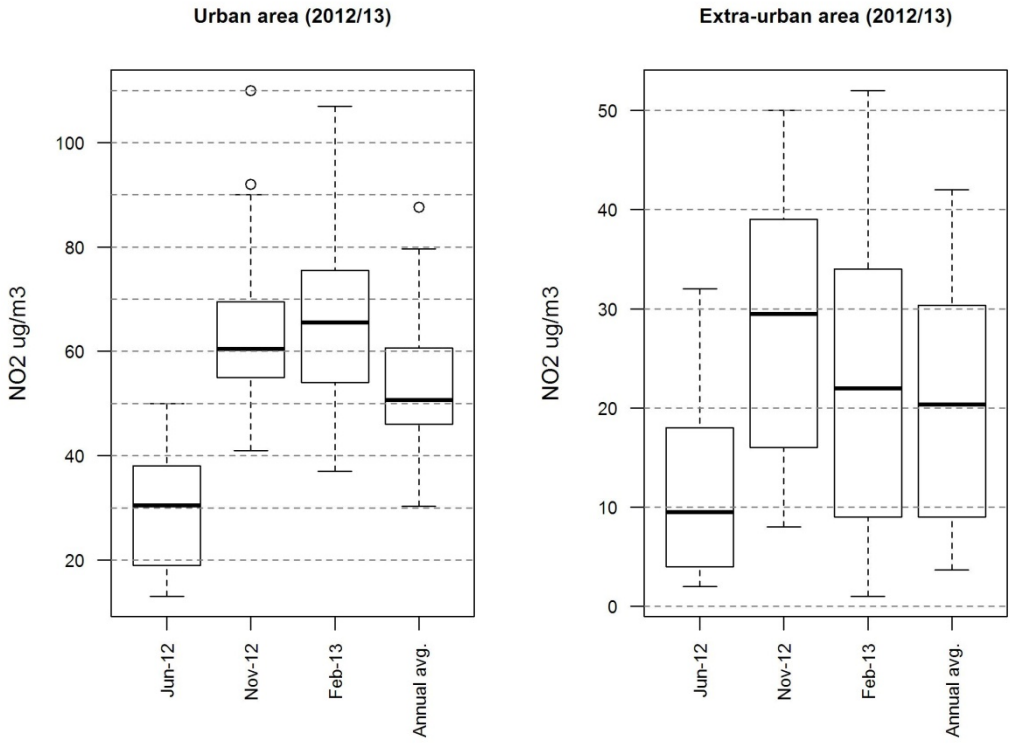

Figure 24 – Variability of NO_2 concentrations in different sampling campaigns. The box represent the inter-quartile range (IQR), the horizontal line inside the box is the median value, the whiskers extend to 1.5 times the IQR from the box. "Annual.avg" is the annual average of the three campaigns. The scales of the two graphs are different.

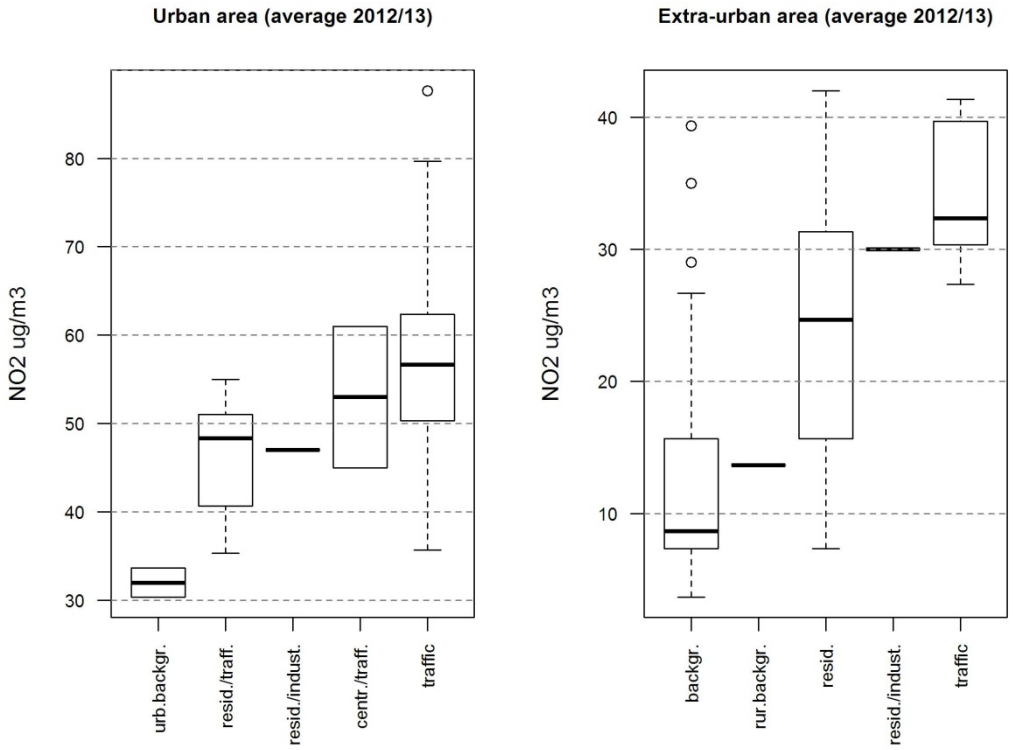

Figure 25 – Variability of annual average NO_2 concentrations by site typology. The box represent the inter-quartile range (IQR), the horizontal line inside the box is the median value, the whiskers extend to 1.5 times the IQR from the box. The scales of the two graphs are different.
"backgr." = background, "resid." = residential, "indust." = industrial, "traff." = traffic, "centr." = historical centre

Four of the samplers were co-located with a fixed air pollution monitoring station, measuring NO_2 hourly concentration all year long. This allows a comparison between the average annual concentrations estimated by averaging the three 7-day sampling campaigns (June 2012, November 2012 and February 2013) and the real annual average NO_2 concentrations (Table 25, Figure 26).

Overall, there is a good agreement between passive samplers and fixed monitors. The 7-days average concentrations from passive samplers were sometimes higher and sometimes lower then corresponding averages from fixed monitors. The 21-days average concentrations of passive samplers were lower for *Cittadella*, *Badia* and *Saragat* but higher for *Montebello* with respect to the fixed monitor. During the three sampling periods, all fixed monitor 21-days average concentrations were about 10% higher than the corresponding annual average calculated on all available data, indicating that the three sampling periods chosen were not completely representative of annual average conditions. A uniform rescaling of all concentrations from passive samplers based on fixed monitors could be done (ESCAPE project, 2010). I decided to use unadjusted concentrations to develop the LUR model since average annual concentrations estimated from passive samplers at the two background sites (*Cittadella* and *Badia*) were almost equal to real annual average concentrations at the same sites. Anyhow, a uniform rescaling would not affect model development process and would propagate linearly in the predicted values.

Table 25 – Comparison between NO_2 concentrations measured with passive samplers during the three 7-days campaigns and the concentrations measured by co-located continuous fixed monitors.

Location	Sampling campaign	P.sampler 7-day average µg m^{-3}	Fixed monitor 7-day average µg m^{-3}	P.sampler 21-day average µg m^{-3}	Fixed monitor 21-day average µg m^{-3}	Fixed monitor annual average (1/6/12 to 31/5/13) µg m^{-3}
Cittadella (urban background)	Jun-12	13.0	13.8	30.3	33.3	29.2
	Nov-12	41.0	40.2			
	Feb-13	37.0	45.9			
Montebello (urban traffic)	Jun-12	30.0	32.9	49.7	46.3	41.8
	Nov-12	60.0	48.8			
	Feb-13	59.0	57.3			
Badia (rural background)	Jun-12	6.0	9.0	13.7	15.7	15.3
	Nov-12	24.0	20.1			
	Feb-13	11.0	17.9			
Saragat (sub-urban background)	Jun-12	12.0	14.2	25.0	26.2	23.1
	Nov-12	35.0	28.7			
	Feb-13	28.0	35.6			

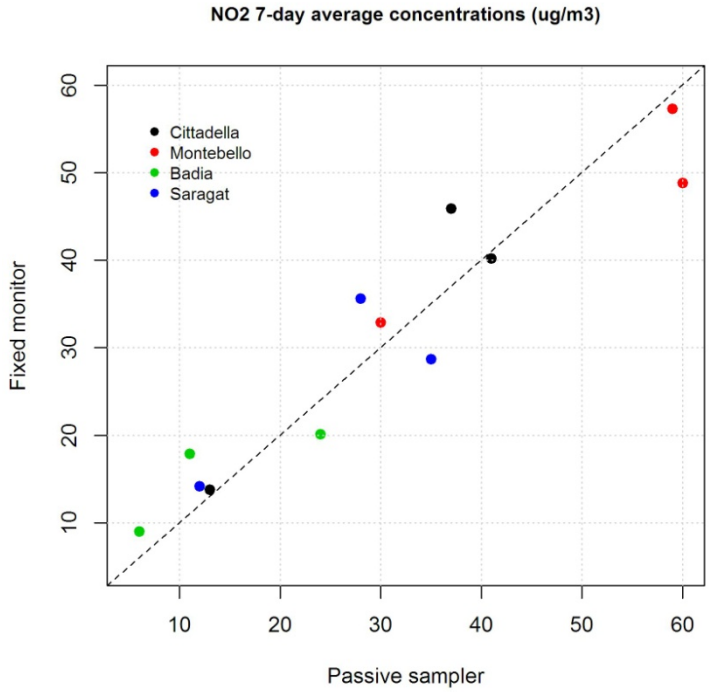

Figure 26 – Comparison between 7-day average NO_2 concentrations measured by passive samplers and by the continuous fixed monitors

Overview of spatial predictors

A variety of spatial predictors have been used in the literature for LUR development, depending on characteristic of the study area and data availability (Hoek et al., 2008). In the present study I evaluated a large set of possible spatial predictors, as described in Table 26.

Traffic is expected to exert a positive effect on NO_2 concentrations. I calculated a set of distance and length variables using the road network for year 2011 as a series of segments (Figure 27). Each road segment is classified with a functional road code (FRC) that describe the type of road (i.e. 0: Motorways; 1: 'Main road' major importance; 2: Other major roads; 3: Secondary roads; 4: Local connecting roads; 5: Local roads of high importance; 6: Local roads; 7: Local roads of minor importance; 8: Others). I identified major roads as roads with FRC $\geq$ 4 (ESCAPE project, 2010). Data on traffic counts on main roads in the rush hour (08:00-09:00 am) were available for the Province of Parma. Data were available as a projection of 2004 measured traffic based on programmed road expansions and traffic growth (Figure 28). Major roads were identified as roads with more than 400 vehicles travelling in the rush hour ($\approx$ 5000 vehicles per day) (ESCAPE project, 2010).

I evaluated land use type inside different buffers using maps by the Emilia Romagna Region for the year 2008. Different land types are represented as polygons and classified with a three digit code, following the CORINE classification scheme (EEA, 2000). I reclassified the original land use types following the ESCAPE criteria (ESCAPE project, 2010) as *high density residential*, *low density residential*, *natural*, *urban green* and *industrial land* (Figure 29). Residential and industrial land use is expected to have a positive effect on NO_2 concentrations, while the presence of natural and urban green is expected to lower NO_2 concentrations.

I evaluated total population and population density inside different buffers using most recent Census data (year: 2001), assuming uniform spatial distribution of population inside each polygon representing a census unit (Figure 30). NO_2 concentrations are expected to increase with population number and density.

I evaluated building volume and average building height inside buffers using the Emilia Romagna Region cartography (Figure 31). The presence of buildings is a proxy for the presence of resident population, but exert also a positive effect on NO_2 concentrations since it may reduce air mixing and recirculation. Recently, the use of the *skyview factor* (i.e. the portion of sky visible from a specific sampling location) has been proposed to improve air pollution model's performance in urban contexts (Eeftens et al., 2013). In this study I calculated the *skyview factor* from available data on building volume using the recently developed function *r.skyview* from GRASS GIS (v 7.0 beta).

I calculated altitude above sea level (a.s.l.) using the Shuttle Radar Topography Mission (STRM) digital elevation model (DEM) with a resolution of 90m (Figure 32). Altitude is expected to have a negative influence on NO_2 concentrations.

Table 26 – Overview of spatial predictor variables used for LUR model development. All variables are offered to the model through a supervised foreward stepwise procedure.

Group	Predictor variable	Variable name	Unit	Buffer sizes	Spatial scale	Effect direction
Traffic	Traffic load on all streets (=n.vehicles*road.length) during rush hour within buffer	*alload*	n.veh. m h^{-1}	25,50,100, 300,500,1000	Local/traffic	+
	Traffic load on all streets (=n.vehicles*road.length) during rush hour within buffer, corrected with minor road lenght	*alload.cor*	n.veh. m h^{-1}	25,50,100, 300,500,1000	Local/traffic	+
	Traffic load on major roads (n.vehicles >400/h, ≈ 5000 vehicles/d) within buffer	*majload*	veh. m h^{-1}	25,50,100, 300,500,1000	Local/traffic	+
	Traffic volume on nearest major road (n.vehicles >400/h)	*hloadnear*	veh h^{-1}	-	Local/traffic	+
	Traffic volume on nearest road	*loadnear*	veh h^{-1}	-	Local/traffic	+
	Minimum distance from a road	*distall*	m	-	Local/traffic	-
	Minimum distance from a major road (FRC<=4)	*distmaj*	m	-	Local/traffic	-
	Minimum distance from an high traffic street (n.vehicles >400/h)	*disthload*	m	-	Local/traffic	-
	Inverse minimum distance from a road	*invdistll*	m^{-1}	-	Local/traffic	+
	Inverse minimum distance from a major road (FRC<=4)	*invdistmaj*	m^{-1}	-	Local/traffic	+
	Inverse minimum distance from an high traffic street (n.vehicles >400/h)	*invdisthload*	m^{-1}	-	Local/traffic	+
	Length of major roads (FRC<=4) within buffer	*majlen*	m	25,50,100, 300,500,1000	Local/traffic	+
	Length of minor roads (FRC>4) within buffer	*minlen*	m	25,50,100, 300,500,1000	Local/traffic	+
	Length of all roads within buffer	*totlen*	m	25,50,100, 300,500,1000	Local/traffic	+
Landuse	High density residential landuse within buffer	*hdres*	m^2	100,300,500, 1000,2000, 5000	Urban	+
	Low densi residential landuse within buffer	*ldres*	m^2	100,300,500, 1000,2000, 5000	Urban	+
	Residential landuse within buffer	*res*	m^2	100,300,500, 1000,2000, 5000	Urban	+
	Industrial landuse within buffer	*indu*	m^2	100,300,500, 1000,2000, 5000	Urban	+
	Natural landuse within buffer	*natu*	m^2	100,300,500, 1000,2000, 5000	Urban	-
	Urban Green landuse within buffer	*urbgreen*	m^2	100,300,500, 1000,2000, 5000	Urban	-
Population	Average population density within buffer, 2001 ISTAT census data	*pdens*	num. km^{-2}	100,300,500, 1000,2000, 5000	Urban	+
	Total population within buffer, 2001 ISTAT census data	*pop*	num.	100,300,500, 1000,2000, 5000	Urban	+
Buildings	Skyview factor (% visible sky)	*skyview*	%	-	Urban/traffic	-
	Building total volume within buffer	*bvol*	m^3	-	Urban/traffic	+
	Average building height within buffer	*bhgt*	m	-	Urban/traffic	+
Altitude	Altitude a.s.l. DEM SRTM90	*altitude*	m (a.s.l.)	-	Regional	-
	Square root of altitude a.s.l. DEM SRTM90	*altitude.sq*	m (a.s.l.)	-	Regional	-

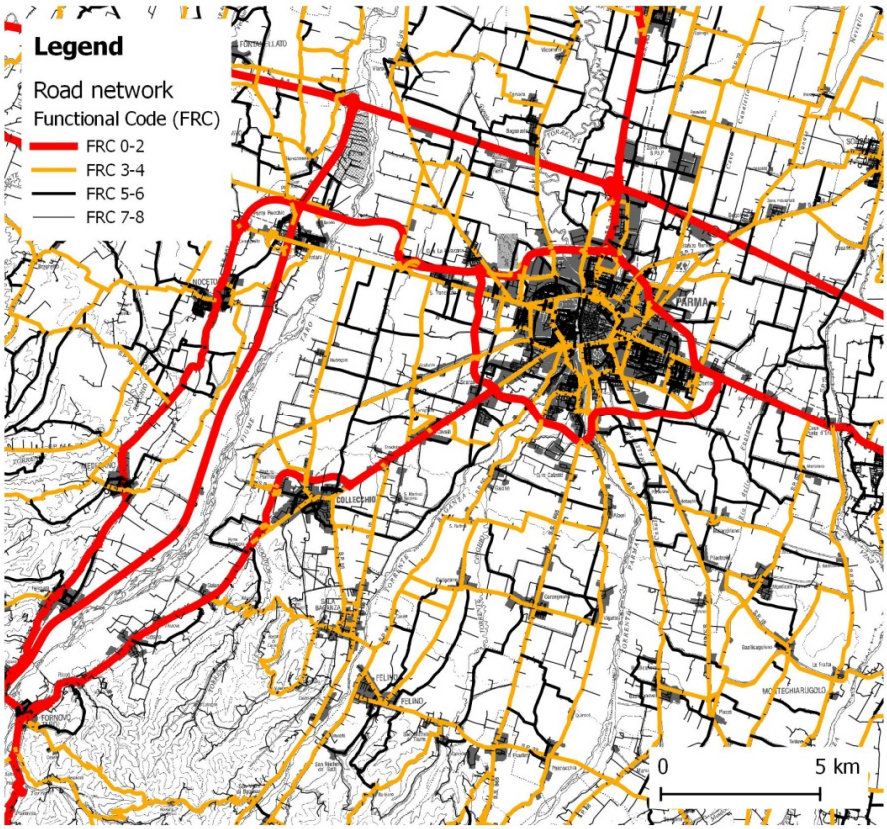

Figure 27 – Representation of the road network in a portion of the study area, classified on the basis of the functional road code (FRC). Each line represents a street.

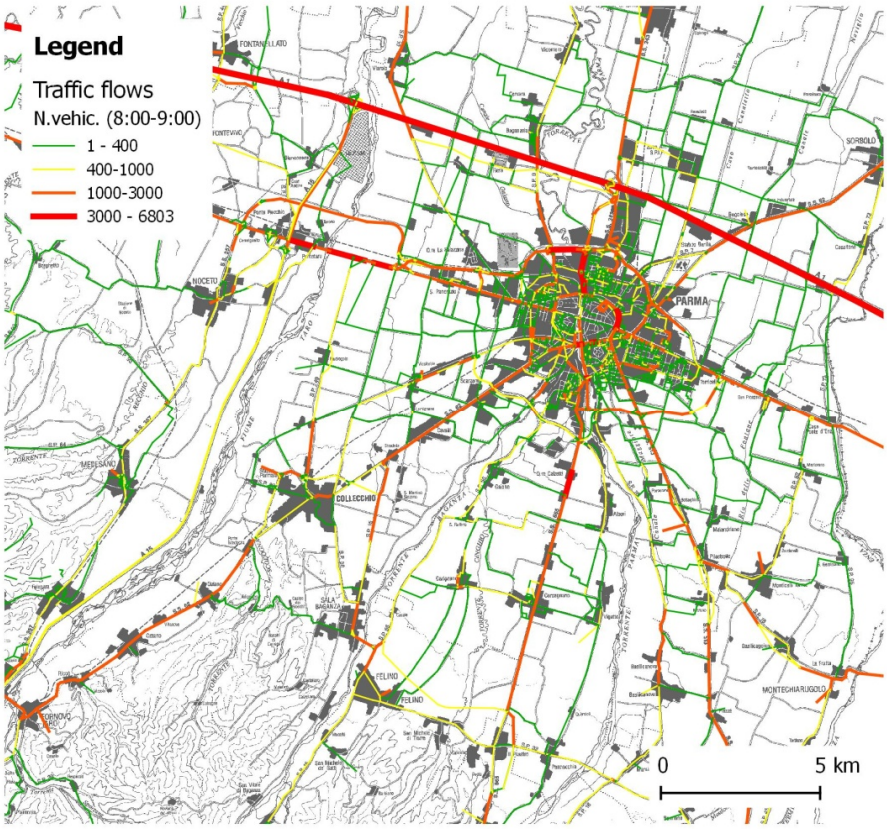

Figure 28 – Representation of the road network in a portion of the study area, classified on the basis of the traffic flux in the morning rush hour (8:00-9:00).

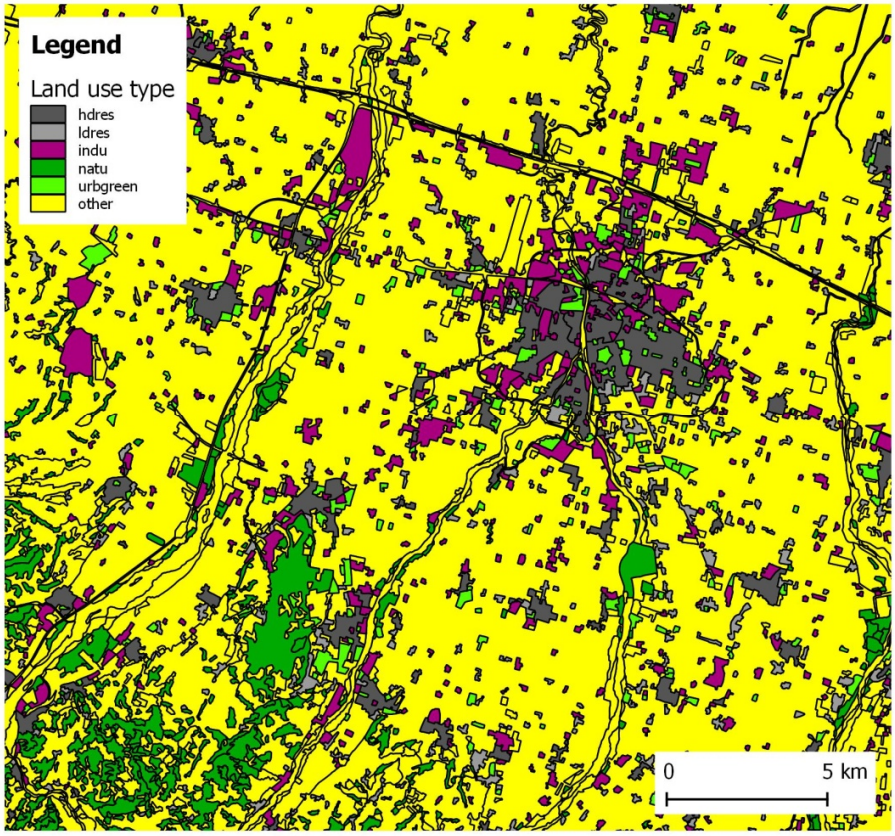

Figure 29 – Representation of land use in a portion of the study area, classified on the basis of the ESCAPE criteria. Each coloured polygon represents an area with different land use.

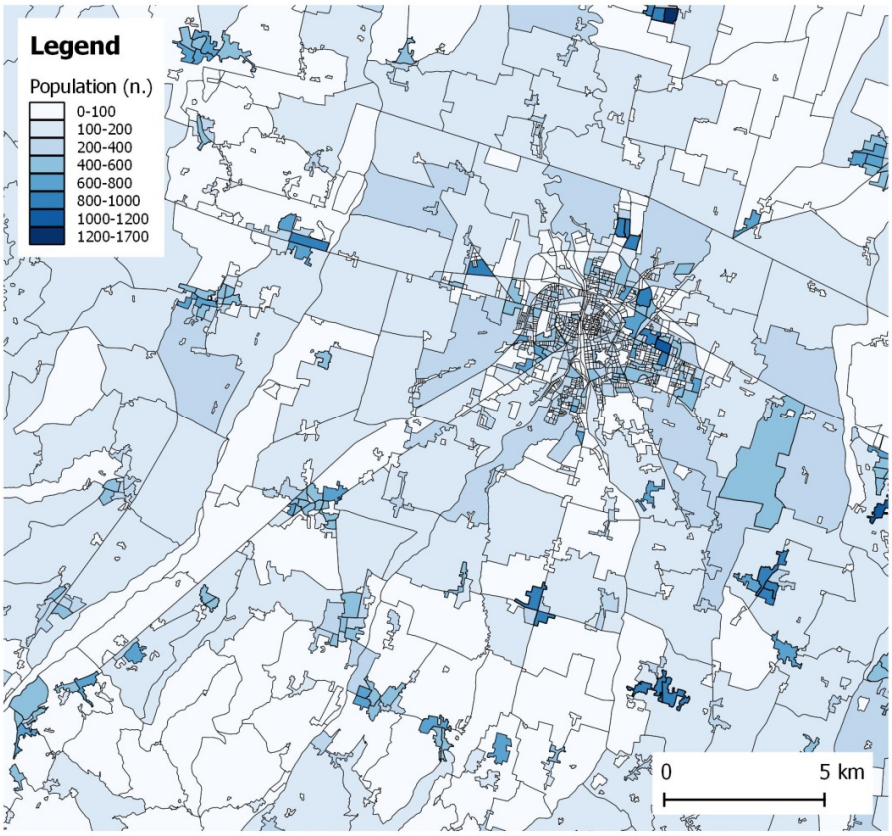

Figure 30 – Representation of 2001 census units in a portion of the study area, classified on the basis of total population

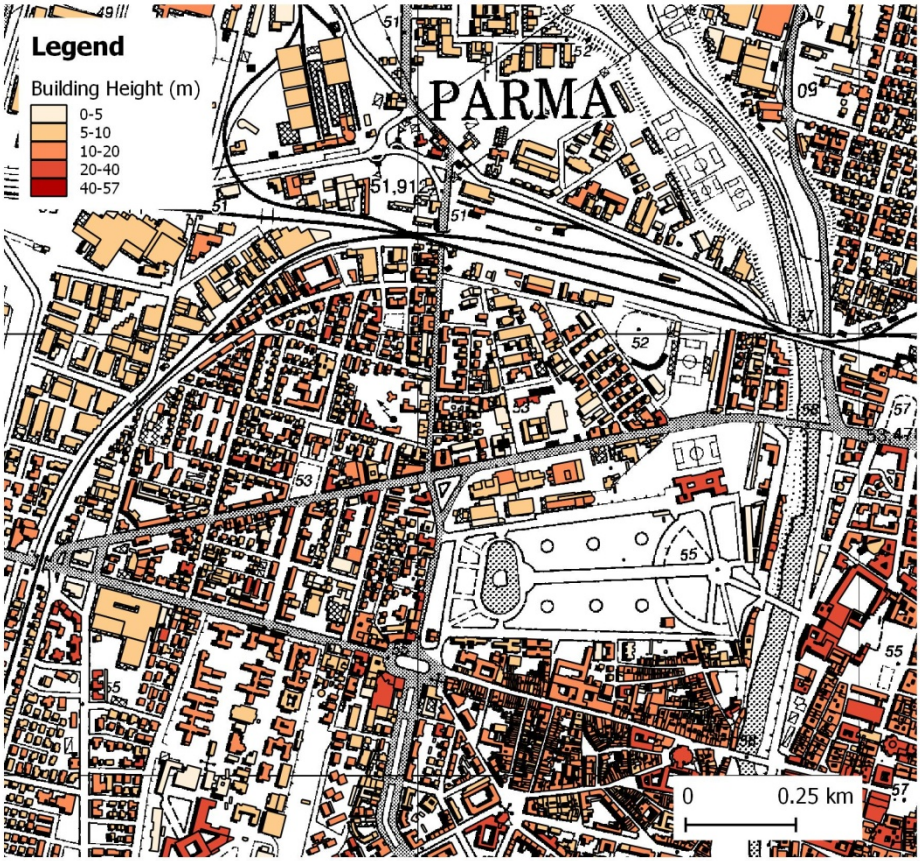

Figure 31 – Representation of buildings in a portion of the study area, classified on the basis of their height

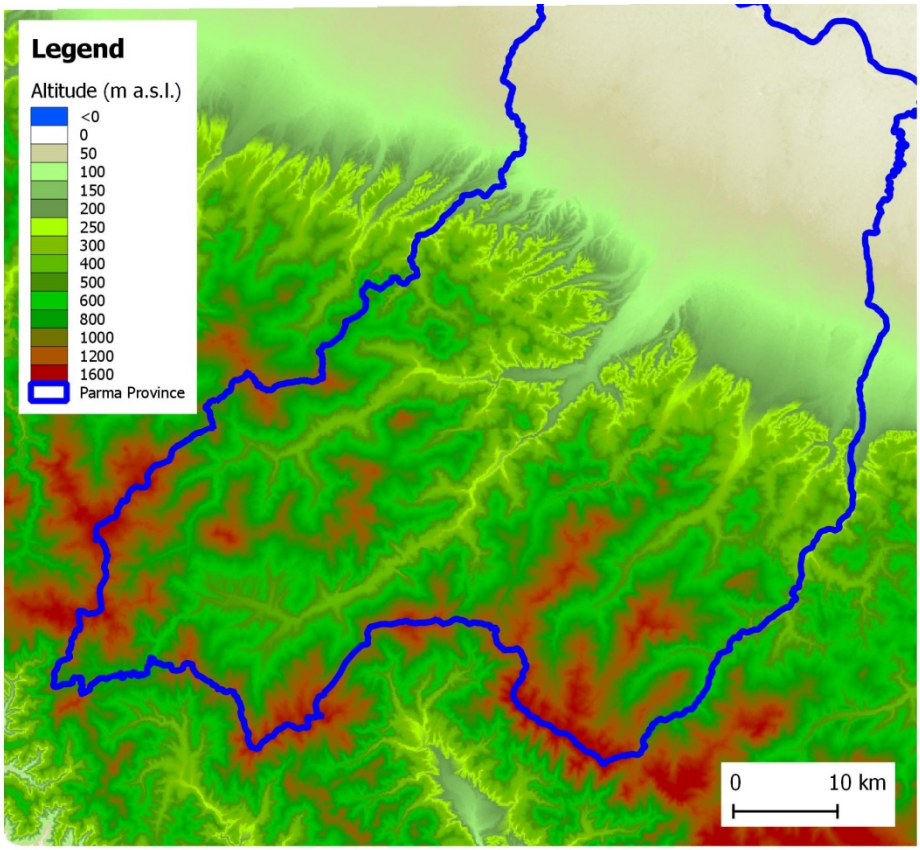

Figure 32 – Representation of the digital elevation model for the province of Parma over the entire study area

Model development

The emission sources, micrometeorology and chemical-physical phenomena that determine atmospheric pollution are different in urban and extra-urban contexts. On a regional scale, macro-meteorological phenomena or terrain characteristics may be the most important determinant of pollution spatial distribution, while at the urban scale traffic intensity, presence of buildings and population density may be more relevant. It is thus questionable if a unique linear regression model can capture the variability of NO_2 concentrations inside the city of Parma and within the territory of the entire Province (about 3400 km^2). Moreover, while the passive samplers in the Parma city centre are more representative of traffic conditions, the passive samplers of the extra-urban area are more representative of rural/suburban situations.

I thus developed three separate LUR models, using different subsets of the full dataset:

A. A model for the entire Province, using all 118 monitoring sites
B. A model for the urban area corresponding to the city of Parma, using 28 monitoring sites
C. A model for the extra-urban territory, using 94 monitoring sites

Four monitoring sites near the limits of the urban area were used in both model B (to enhance contrast in the predictors) and C (to introduce some sub-urban effect in the model).

I developed the LUR models following the ESCAPE methodology (Beelen et al., 2013), fully described in paragraph 6.2.1. I validated model performance of model A, B and C using leave-one-out cross validation (LOO-CV). Since models B and C were constructed on two different subsets of the available data, for these model an external validation was also possible. I thus evaluated the performance of the urban model (B) in predicting the extra-urban concentrations and the performance of the extra-urban model (C) in predicting the urban concentrations. I used predictors in the test dataset both as they were and truncated at the range of the values observed in the data set used for model development, to prevent unrealistic predictions based on model extrapolations (Wang et al., 2012).

I used the three LUR models to predict NO_2 annual average concentrations at residential addresses of 190510 people in the Parma municipality (Figure 33). I defined an *urban area* containing the Parma urban city centre, corresponding to the definition of the urbanized area from the National Institute of Statistics (ISTAT). This area contained all the urban passive samplers except the four samplers that were used in constructing both model B and C (Figure 33).

I tested four possible population exposure models:

1. Use of model A for the entire population
2. Use of model B for the entire population
3. Use of model C for the entire population
4. Use of model B in the predefined *urban area* and model C outside the *urban area* (named "BC" in the following)

For each residential address I calculated the values of the predictors required by each model. In each application of the models I truncated the values of the predictors for the addresses that where below the minimum or above the maximum values calculated for the passive samplers used to develop the models. This prevents the model is applied outside the ranges of observed values, which may give unreasonably low or high concentrations (Wang et al., 2012). I also interpolated the NO_2 concentration values estimated with the LURs at each address using kriging, to obtain continuous pollution surfaces over the study area.

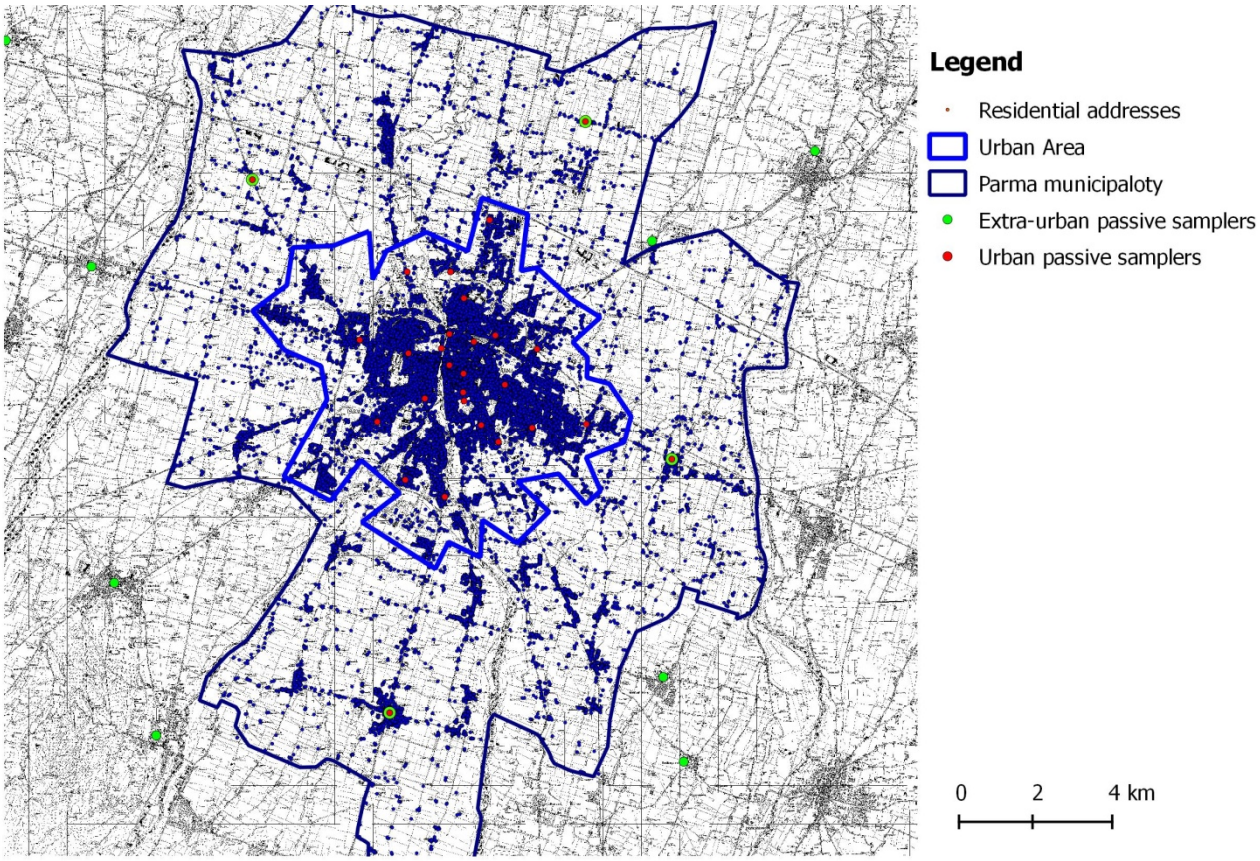

Figure 33 – Spatial distribution of residential addresses in the Parma municipality, together with passive sampler locations and definition of the *urban area* where to apply model B.

6.3.2 Description and discussion of resulting LUR models

The LUR models resulting from the supervised stepwise regression procedure are presented in Table 27. Model quality tests are briefly described in Table 28.

The model developed using all passive samplers (A) explains 90.3% of NO_2 variability in the area. The presence of industrial land use within 5000 m (that enters in the model through two separate variables: 0-500 m and 500-5000 m) explain about 74% of variability in the data. Altitude a.s.l. explains another relevant part of variability (9.5%), while traffic variables (traffic load and distance) and the building's volume have minor importance. The quality test shows the non-normality of model residuals (Table 28). The model performed quite well in the LOO-CV validation (Figure 34), with an R^2_{LOO} = 88.6%.

The model developed with the urban passive samplers (B) explains 83.1% of NO_2 variability. Variables related to traffic (traffic load and road length) explains about 44% of the variance. The presence of industrial and low density residential land uses adds about 25% of explained variance. Also buildings characteristics (volume and height) and presence of urban green spaces enters as predictors in the model. Overall, the quality of the model was good (Table 28), but the performance of the model in the LOO-CV validation was not as good as model A, with an R^2_{LOO} about 10% lower than the model R^2 (Figure 34). Nevertheless, the range of absolute errors between observed and predicted NO_2 concentrations is lower for model B cross validation [-18, +11] than for model A [-25, +18].

The model developed with the extra-urban passive samplers (C) explains 86.8% of NO_2 variability. Altitude a.s.l. is the first selected predictor and explains alone 69.2% of variance in the data, indicating a strong gradient in NO_2 concentrations between the flat and the mountain area. Traffic load on all

roads within 50 m adds 10.4% of explained variance to the model. The presence of industrial land use and buildings determine minor improvements in the model R^2. The quality tests for the model were good and the R^2_{LOO} was 83.1% (Figure 34). Model A and model C shares 80% of the observations.

Table 27 – LUR models resulting from the supervised stepwise regression algorithm. ΔR^2 is the gain in adjusted R^2 obtained adding the variable to the model.

LUR model	Adjusted R^2	Predictors	Coefficients	p-values	ΔR^2
All sites (A)	90.3%	indu500.5000	1.77E-06	<0.001	70.4%
		altitude.sq	-7.93E-01	<0.001	+9.5%
		majload25	3.97E-05	<0.001	+4.4%
		indu500	3.56E-05	<0.001	+3.2%
		bvol100	5.62E-05	<0.001	+1.3%
		distmaj	-1.67E-02	<0.001	+1.4%
		(intercept)	2.73E+01	<0.001	-
Urban (B)	83.1%	majlen25	1.48E-01	<0.01	29.3%
		indu2000	9.98E-06	<0.001	+14.6%
		majload25	2.78E-05	<0.01	+9.1%
		bvol100	6.03E-05	<0.01	+5.2%
		ldres2000	2.33E-05	<0.001	+11.5%
		majlen25.300	4.74E-03	<0.01	+5.6%
		urbgreen2000	-9.12E-06	<0.01	+4.2%
		bhgt25	6.12E-01	<0.01	+3.7%
		(intercept)	4.45E+00	0.43	-
Extra-urban (C)	86.8%	altitude.sq	-8.55E-01	<0.001	69.2%
		alload50	1.50E-05	<0.001	+10.4%
		indu2000	7.85E-06	<0.001	+4.8%
		distmaj	-1.12E-02	<0.01	+1.3%
		bvol100	4.15E-05	<0.01	+1.0%
		(intercept)	2.94E+01	<0.001	-

Table 28 – Quality tests for the three LUR models of Table 27.

Model	Multicollinearity	Residual normality	Influent observations	Residual autocorrelation	Leave-one-out cross validation
All sites (A)	all VIF< 2.3	NO (Shapiro-Wilk p<0.01)	NO (max Cook's D = 0.32)	NO (Moran I= 0.03, p=0.14)	R^2_{LOO} = 88.6% ERR_{LOO}= [-25,+18] $RMSE_{LOO}$=6.1 FAC2=96.6%
Urban (B)	all VIF< 2.6	YES (Shapiro-Wilk p=0.28)	NO (max Cook's D = 0.74)	NO (Moran I= 0.02, p=0.09)	R^2_{LOO} = 72.2% ERR_{LOO}= [-18,+11] $RMSE_{LOO}$=6.9 FAC2=100.0%
Extra-urban (C)	all VIF< 1.6	YES (Shapiro-Wilk p=0.07)	NO (max Cook's D = 0.81)	NO (Moran I= -0.01, p=0.30)	R^2_{LOO} = 83.1% ERR_{LOO}= [-14,+16] $RMSE_{LOO}$=4.9 FAC2=95.0%

VIF: variance inflativo factor, R^2_{LOO}= R^2 between predicted and measured values, ERR_{LOO}= range of errors between predicted and measured values, $RMSE_{LOO}$= root mean square error, FAC2 = fraction of predictions within a factor of 2 of observations

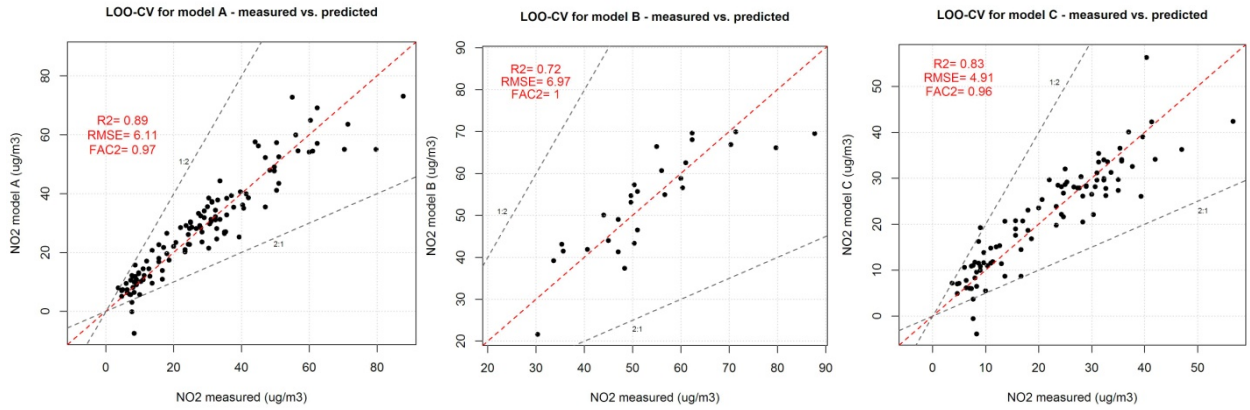

Figure 34 – Leave-one-out cross validation (LOO-CV) for models A, B and C. The scatter plots represents the comparison between concentrations measured at site *n* and concentration modeled with the *N-1* model.
R2=determination coefficient , RMSE=root mean square error, FAC2= fraction of predictions within a factor of two of observation;

The models obtained in this case study have R^2 and predictors comparable to many previously published models (Beelen et al., 2013; Dons et al., 2014; Hoek et al., 2008). Models developed in the ESCAPE project had adjusted R^2 values varying between 55% and 92% and included 2 to 7 predictors, among which traffic intensity within 100 m or less was the more common traffic variables. Also the LOO-CV performance of models in Parma was comparable to those of the ESCAPE experience, where R^2_{LOO} where generally less than 10% lower than model R^2 (Beelen et al., 2013).

The two models B and C performed badly in the external validation (Figure 35, Figure 36). This was expected, since they are calibrated on very different contexts. The urban model (B) overpredicts the majority of the concentrations of the extra-urban passive samplers (Figure 35): 61% of the predicted values are within a factor of two of observations (FAC2) and the R^2 between observed and predicted values is 0.25. Truncation of predictors worsens the performance of model B and the R^2 decreases to 0.12. Concentrations predicted with the extra-urban model (C) are more uniformly dispersed around the 1:1 line (Figure 36), with two outliers that are highly overestimated by the model. The R^2 between measure and predicted values is 0.13 for model C, but rise to 0.47 if predictors are truncated. Thus, the performance of model C in predicting urban concentrations is better than the performance of model B in predicting concentrations in the extra-urban area, especially with truncated predictors, although the performance remains weak. Overall, model performance in in-sample cross validation (LOO-CV) was better than model's performance in out-of-sample external validation.

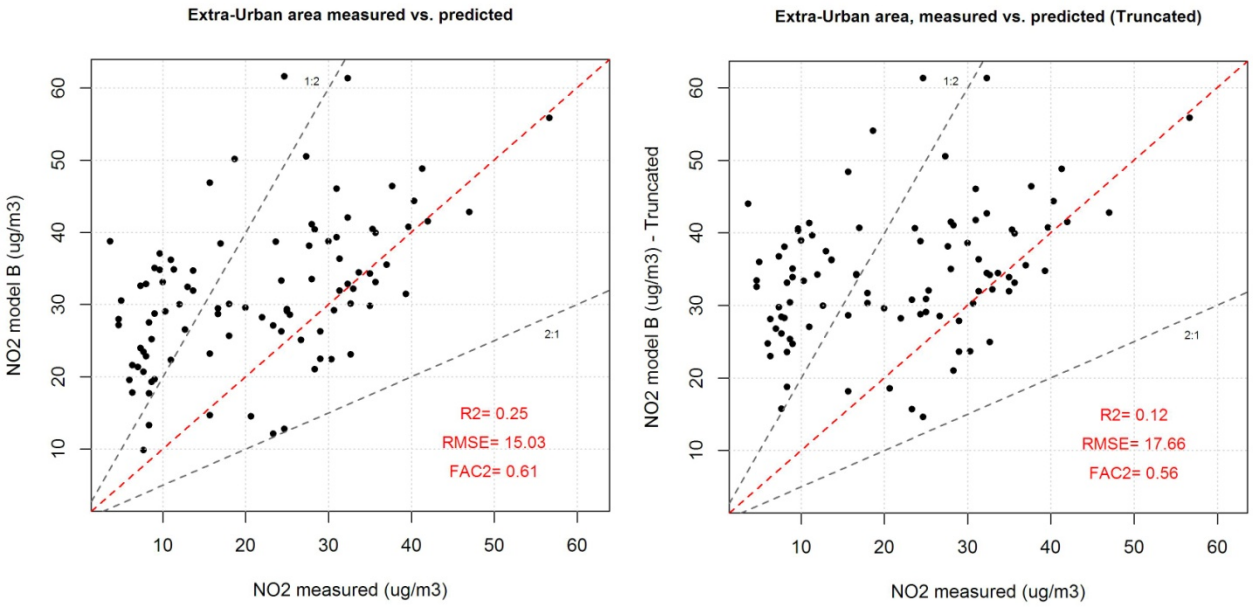

Figure 35 – Application of the urban model B to predict NO_2 concentrations at the extra-urban sampling sites. Model B is applied with original (sx) and truncated (dx) predictors.
R2=determination coefficient , RMSE=root mean square error, FAC2= fraction of predictions within a factor of two of observation;

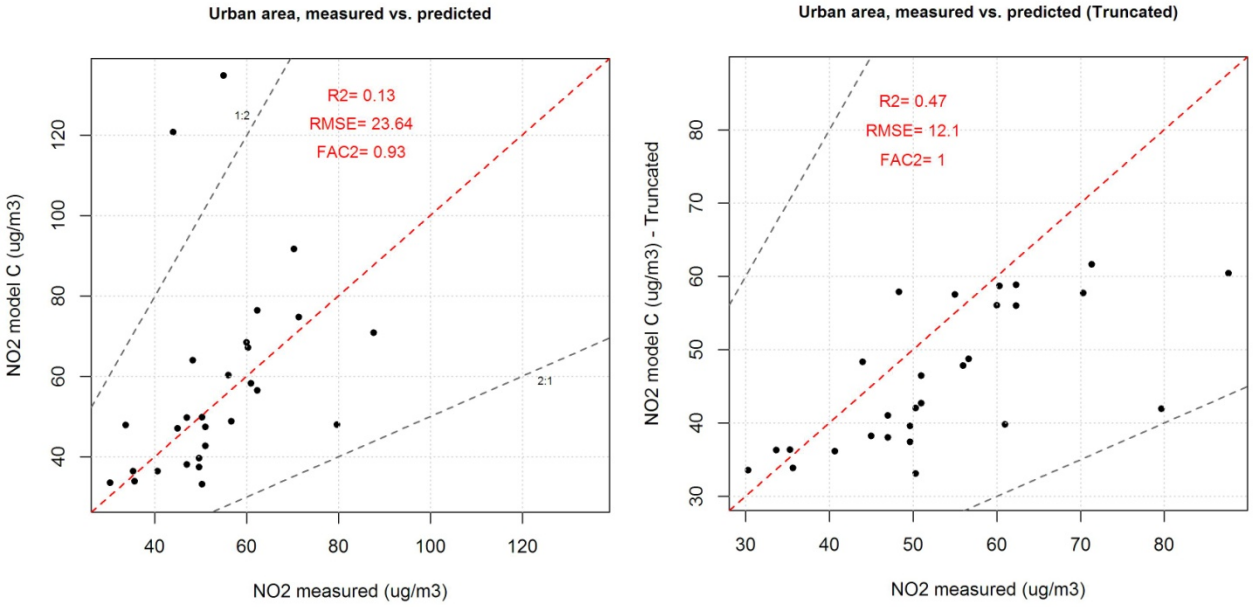

Figure 36 – Application of the extra-urban model C to predict NO_2 concentrations at the urban sampling sites. Model C is applied with original (sx) and truncated (dx) predictors.
R2=determination coefficient , RMSE=root mean square error, FAC2= fraction of predictions within a factor of two of observation;

The purpose of LUR modelling is not to study causal relationship between single predictors and NO_2 concentrations, but rather to obtain models usable to predict concentrations in unsampled locations, i.e. LUR models are not explanatory but predictive models (Sainani, 2014; Shmueli, 2010). The magnitude of association and statistical significance of each single predictor are not the focus of LUR analysis. Nevertheless, it is important that the predictors included in each model have a scientific rationale and it is interesting to analyze the differences between predictors that were selected in the three models.

The variability of NO_2 concentrations over the whole Province (model A) is almost totally characterized by the variable describing industrial land use within a buffer of 5 km and altitude above

sea level. It is unlikely that these two predictors can correctly represent the short-range variability in NO_2 concentrations in the Parma urban area. It is also improbable that the presence of industries can explain such a large fraction of NO_2 variability in the Province of Parma: more probably the predictor *indu5000* is a proxy for other pollution sources, since it is highly correlated with many other variables describing traffic (e.g., *majload1000*, *alload500*) and population (e.g. *pop5000*, *pdens2000, res5000*).

The model developed on the urban samplers (B) is based on predictors with small-scale variability (e.g. road length and traffic load within 25 m, land use within 2 km, building characteristics within 100 m and 25 m). Interestingly, the model incorporates two predictors that may explain typical urban phenomena: (i) the presence of urban green areas, which reduces local air pollution and (ii) the average building height within 25 m, which may indicate the presence of some *urban canyon effect* in the city center (Eeftens et al., 2013).

In the extra-urban model (C) the altitude a.s.l. represent the most important predictor, indicating the presence of a strong North-South gradient in NO_2 concentrations over the region. Altitude a.s.l. is correlated with many other predictors (e.g., Pearson's correlation with *hdres5000*, *indu5000*, *disthload* and *pop5000* are respectively -0.66, -0.66, 0.72, -0.57), thus it represent a general indicator for the North-South gradient in human activities. Given the altitude a.s.l., small-scale variation in traffic load determines another relevant quote of NO_2 variance.

Overall, the difference in the predictors selected for the urban (B) and extra-urban models (C), together with the poor performance of these models in the external validation, suggests caution when transferring LUR models between different spatial scales and environmental contexts. Previous studies have analyzed the issue of LUR transferability between different cities (Allen et al., 2011) or countries (Vienneau et al., 2010), suggesting some concern about the use of LURs that are not developed locally. Here I showed that also the difference between in-city and outside-city determinants in air pollution limits the spatial transferability of a LUR models. One of the issues that limits transferability is sometimes the unavailability of homogeneous GIS databases or the use of different pollution sampling methodologies. In this case studies both GIS data and NO_2 sampling method were homogeneous over the entire study area.

This case study has some limitations. First of all the passive samplers used to develop the models were not positioned for this specific purpose, but must be considered as routine monitoring data. I used the ESCAPE methodology to develop the LUR models, although this method originally requires an accurate definition of sampling locations. Thus, pollution variability inside the urban area and across the region may have not been fully characterized. On the other hand, locations of passive samplers were chosen by ARPA-PR to be representative of populated areas and this is an important aspect to consider when using LUR to estimate population exposure.

The model B obtained with the ESCAPE procedure is probably overfitted: the proportion between observation and predictors (i.e. 28:8) is very low (Babyak, 2004). Basagaña et al. (2012) highlighted the poor value of LOO-CV in-sample validation with respect to LUR model ability to predict out-of-sample concentrations. Indeed, when used to predict extra-urban concentrations (Figure 36) model B performed poorly. It is nevertheless impossible to separate the effect of overfitting from the error deriving from the application of the model to a different environmental context.

6.3.3 Population exposure in Parma

I applied the three LUR models, with truncated predictor variables, to the residential addresses of the entire population of the Parma municipality. The results for using model A, B, C and model BC are quite similar in terms of average population exposure and range of exposure variability. The extra-

urban model (C) gives lower exposure values and exposure variability (Table 29, Figure 37). Although exposure distributions were quite similar for the three models A, B and BC (Figure 37), each model predicts a quite different spatial distributions of NO_2 concentrations in the area, thus resulting in differences in exposure up to 32 µg m^{-3} for single subjects (Figure 38, Figure 39). For example, Model B predicts high concentration in a vast area in the south of the urban area, while model A predicts higher concentration in the city centre. Truncation of predictors avoids the estimation of negative (model A, Figure 40) or unreasonably high (model C and BC, Figure 40) concentrations.

Table 29 – Average NO_2 exposure and minimum-maximum range in the population of Parma, according to the four different truncated models.

Applied model	Average exposure µg m^{-3}	Exposure range µg m^{-3}
A (trunc.)	44.2	10.1 - 84.9
B (trunc.)	44.8	11.9 – 82.3
C (trunc.)	37.4	13.3 – 61.6
BC (trunc.)	44.3	13.3 – 82.3

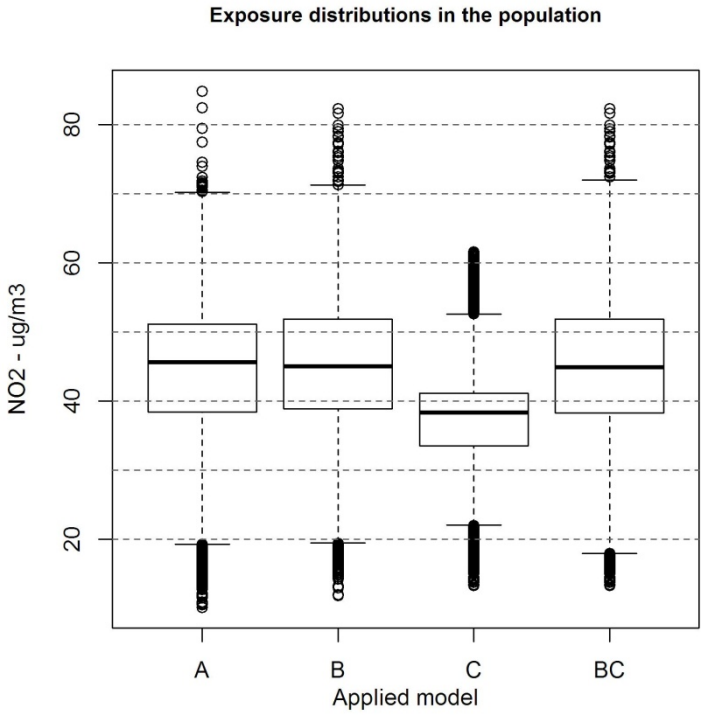

Figure 37 – Graphical representation of exposure distribution in the population of Parma, according to the four different truncated models. The box represent the inter-quartile range (IQR), the horizontal line inside the box is the median value, the whiskers extend to 1.5 times the IQR from the box.

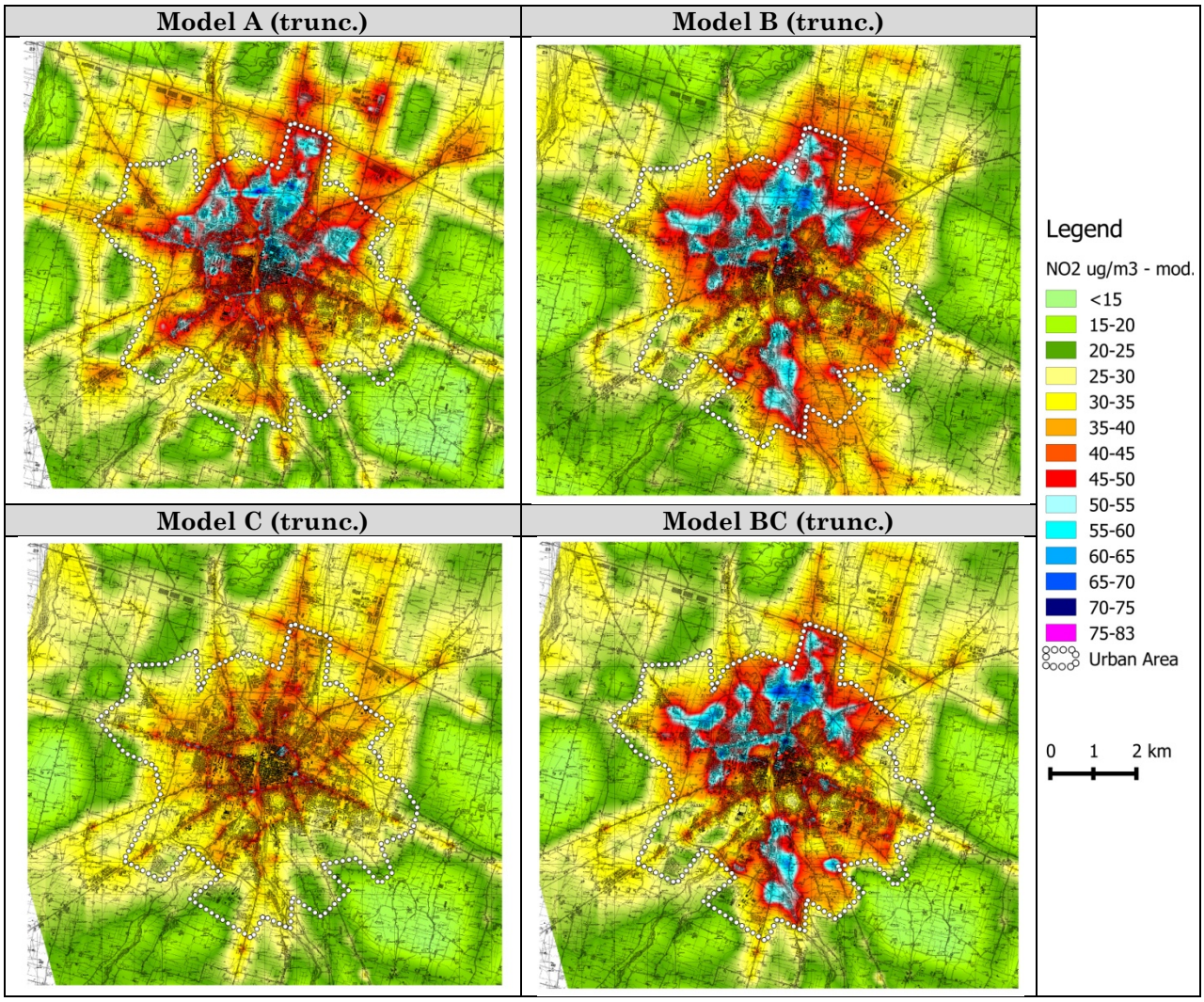

Figure 38 – NO_2 pollution maps obtained interpolating NO_2 concentrations calculated at each address with the four truncated LUR models.

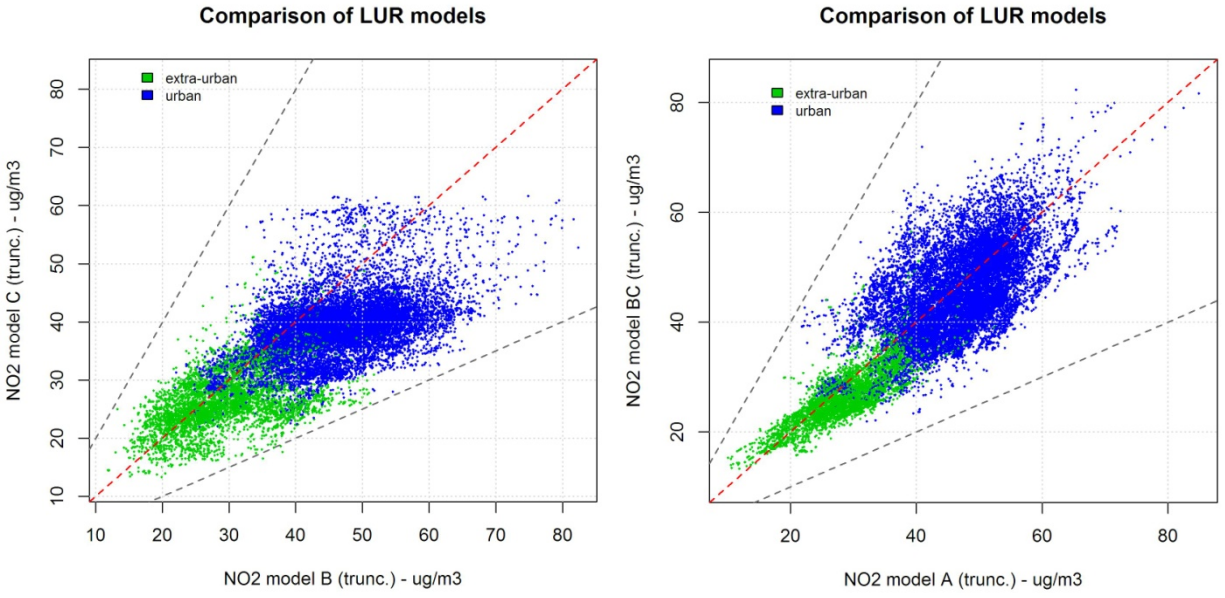

Figure 39 – Comparison of population exposure obtained using different truncated LUR models. The colors identify people living in the urban and extra-urban area

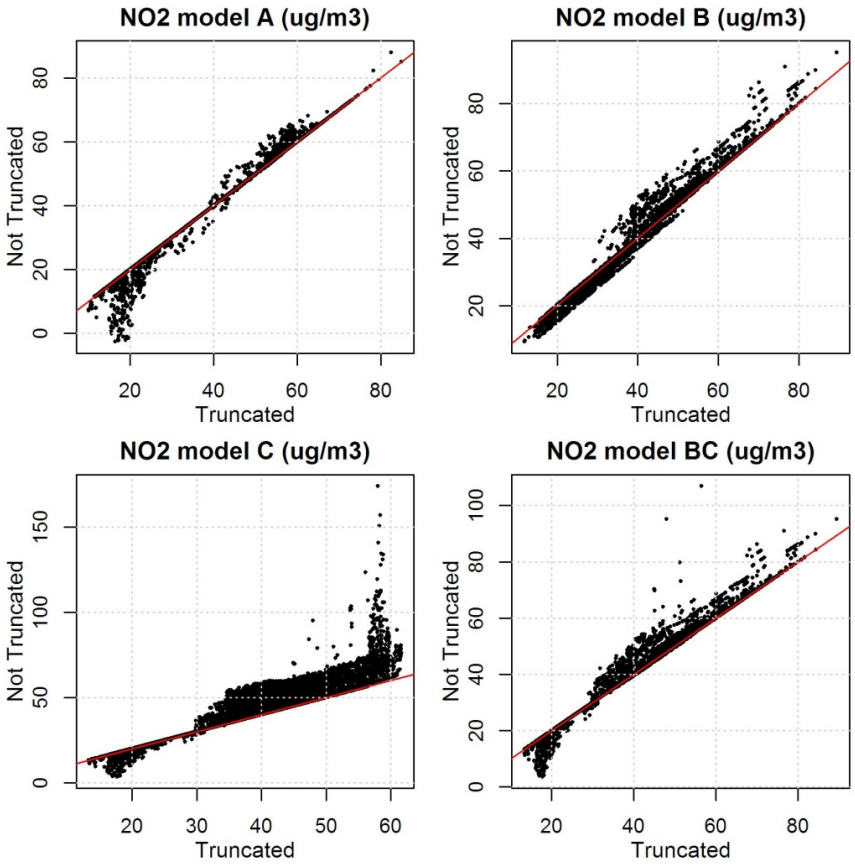

Figure 40 – Comparison between NO_2 concentrations modeled using truncated and non truncated predictors

6.4 Confounding effect of diffuse pollution in studies on point emission sources

In Chapter 3 I introduced the concept of confounding of effect in epidemiological studies on point sources of atmospheric emissions: confounding occurs when a risk factor differs from the exposure variable under study causing bias in the estimation of association between exposure and disease.

To measure the causal effect of exposure on disease is not a trivial task. Ideally, it requires the contrast between the experience of a group of exposed individuals with the same subjects "had they been unexposed" (McNamee, 2003). This is practically impossible in many environmental epidemiology studies (with the exception of some study design like the *case-crossover*), and the best observable evidence of causality is obtained comparing a group of exposed subjects with a group of different subjects that were not exposed. Thus, confounding can be seen as an issue of *comparability* between the exposed and unexposed subjects, if both has been unexposed: if the incidence of the disease in the unexposed group is not identical in the hypothetical incidence of the exposed group without exposure, the comparison will give a false representation of the causal effect and there will be confounding (Morabia, 2011).

More precisely, there are three conditions that a factor must satisfy to be a confounder (McNamee, 2003):

a. The factor must be a cause of the disease (i.e. *risk factor*) or a surrogate measure of a cause
b. The factor must be correlated, positively or negatively, with exposure (i.e. the factor has a different distribution in the exposed and unexposed group)
c. The factor must not be affected by exposure

When analyzing the health effect of a specific point source of atmospheric emissions, the difference in exposure to other pollutants that represent risk factors for the disease under study between the exposed and the unexposed to the point emission source may lead to confounding of effect. In many cases industrial emission sources are found to be aggregated in areas of high emissions from other industries and traffic. Thus, those people who are more exposed to the industrial emission source are also more exposed to other sources of atmospheric pollution.

Only few of the studies reviewed in Chapter 3 considered the issue of confounding from exposure to other pollution sources in the study areas. The most advanced studies used atmospheric dispersion models to characterize population exposure to both the incinerator and diffuse pollution (Candela et al., 2013; Ranzi et al., 2011). The use of this advanced method imply the knowledge of all the emission sources in the area and the availability of detailed information on each source (e.g., emission rate, physical characteristics of the emissions), thus it is applicable only on a small study area (i.e. a 4 km area around the incinerator in the two mentioned studies).

When the area is large with heterogeneous emission sources, the use of atmospheric dispersion models become a very resource-intensive task. In these cases, the use of stochastic models (e.g. LUR) based on measured concentration may represent a valid solution. In the following paragraph I will present an example of the use of a LUR model to characterize exposure to confounding sources of pollution in the epidemiological surveillance program for the incinerator of Parma.

6.4.1 Exposure in a cohort of asthmatics around the incinerator of Parma

The great public concern about the possible health effects determined by the activation of a new municipal solid waste incinerator in Parma led to the definition of an epidemiological surveillance plan for the exposed population. Monitoring the health status of susceptible sub-populations may help in early identifying health risks in the general population (Annesi-Maesano et al., 2003; Guarnieri and Balmes, 2014; Makri and Stilianakis, 2008). The local health authority thus decided to construct a cohort of asthmatics in Parma, characterizing their exposure to the emissions from the incinerator and other sources of atmospheric pollution in the area and monitoring their health status and pulmonary function before and after the activation of the plant.

Subjects were enrolled from a pool of 533 asthmatics under treatment at the Parma Hospital. I geocoded all residential and work addresses and for each one I evaluated exposure to (i) the emissions from the incinerator and (ii) other local sources of air pollution. I estimated exposure to the incinerator using annual average concentrations of particulate matter (PM_{10}) calculated with an atmospheric dispersion model (Figure 41). The model was previously described with more details in Chapter 4. PM_{10} was selected as a tracer for the complex mix of pollutants emitted by the stack.

I estimated exposure to other sources of air pollution using annual average nitrous dioxide (NO_2) concentrations calculated with the LUR model presented in section 6.3 (Figure 42). I decided to use the BC model, i.e. the urban model B was applied to residential and work addresses inside the *urban area*, while the model C was applied to extra-urban addresses. NO_2 was selected as a tracer for exposure to a wide range of combustion sources (i.e., traffic, home heating, industries, etc.). Exposure to NO_2 from diffuse emission sources may represent a confounder for the effect of the incinerator because (i) it represent a risk factor for the worsening of respiratory health in asthmatic subjects (Guarnieri and Balmes, 2014; Jackson et al., 2011) and (ii) exposure to the incinerator will not be affected by exposure to NO_2. Whether NO_2 will exert a confounding effect will thus depend on the difference in exposure to NO_2 between subjects exposed and unexposed to the incinerator.

For each subject I defined exposure as the weighted average concentration estimated at the place of residence and work. Assumed exposure durations were 16 h/d for the residential exposure and 8 h/d for working exposure. We identified a cohort of 62 subjects with high exposure to the incinerator (*exposed group*) which were coupled 1:1 to a *control group* on the basis of functional and clinical parameters (sex, age, body mass index, smoke habits, atopy) and exposure to diffuse pollution. We considered also exposure to diffuse pollution in the matching procedure to minimize confounding from other environmental exposures. Each subject completed a medical examination before and after the activation of the incinerator and compiled a detailed questionnaire on personal habits and environmental exposures.

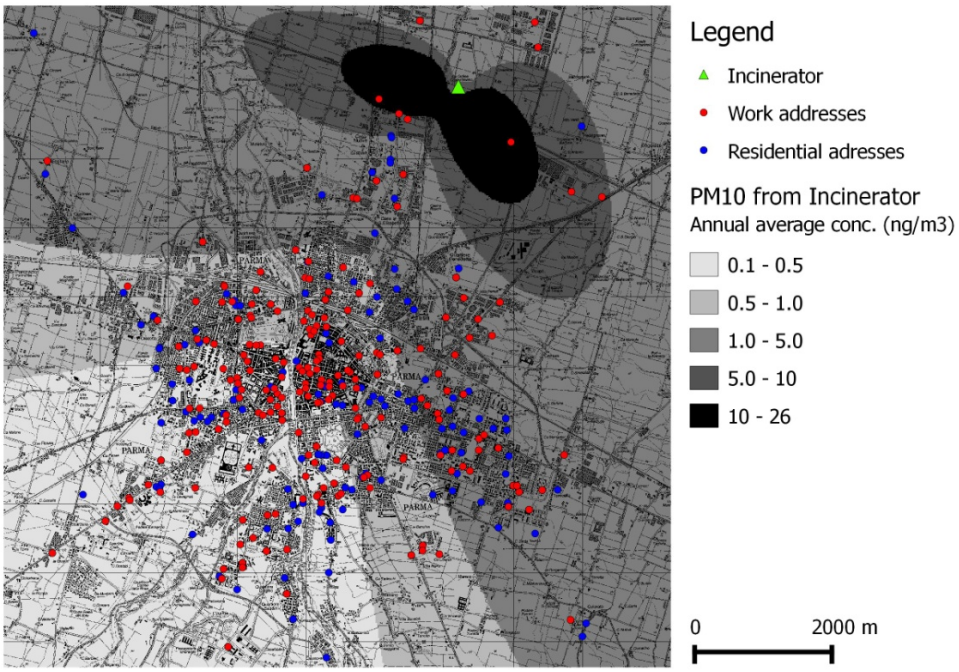

Figure 41 - Representation of home and workplace location for all asthmatic subjects in the central area of Parma and annual average exposure to the incinerator resulting from the atmospheric dispersion model for PM_{10}.

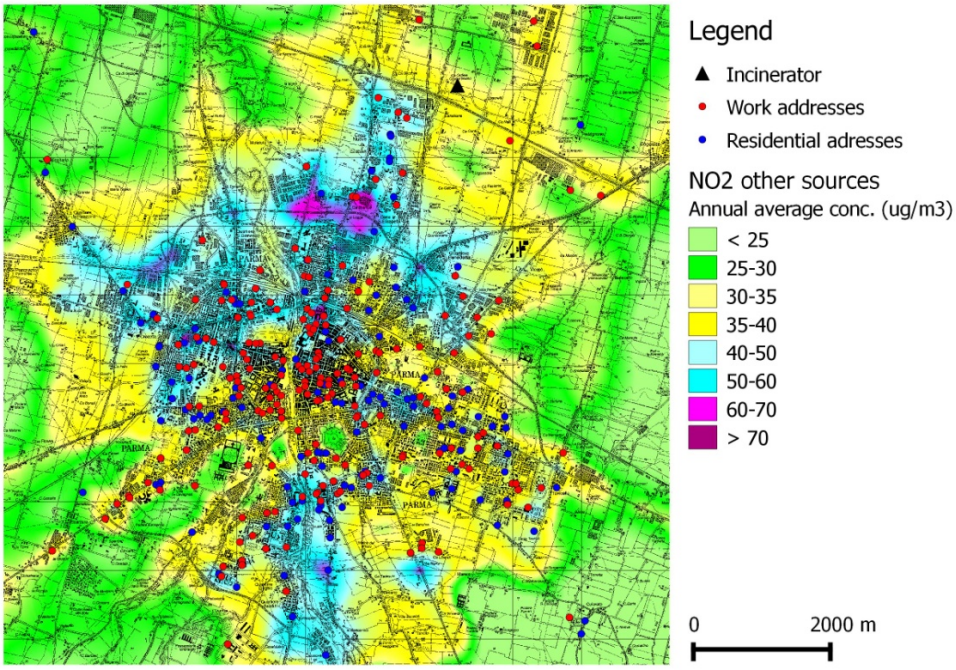

Figure 42 - Representation of home and workplace location for all asthmatic subjects in the central area of Parma and annual average levels of NO_2 resulting from the truncated LUR model BC.

Results for exposure assessment

The estimated annual average exposure to the incinerator was 2.1 [sd: 1.9] and 0.3 [sd: 0.3] ng m^{-3} (PM_{10}) for the *exposed* and *control group* respectively (Figure 43). 32% of the cohort were more exposed to the incinerator at workplace than at home. Average exposure to diffuse pollution was 46.6 [sd: 10.9] and 37.9 [sd: 9.9] µg m^{-3} (NO_2) for the *exposed* and *control group* respectively, with higher exposure values at workplaces than home for 35% of the cohort (Figure 43). Thus, data shows that for about a third of the cohort considering only residential exposure would have introduced some degree of exposure misclassification. Considering also workplace location, where people spend a considerable amount of time, will improve the quality of exposure assessment, although a full characterization of people exposure would need the definition of many microenvironmental exposures.

Even if we considered exposure to diffuse pollution as a matching criteria, the distribution of NO_2 weighted exposure in the *control group* was slightly lower than in the *exposed group*. Nevertheless, the correlation between weighted exposure to the incinerator and weighted exposure to diffuse pollution was very low in the enrolled cohort (Figure 44). Thus, when using this study design it is unlikely that exposure to diffuse NO_2 pollution will represent a confounder for the effect of the incinerator in this cohort.

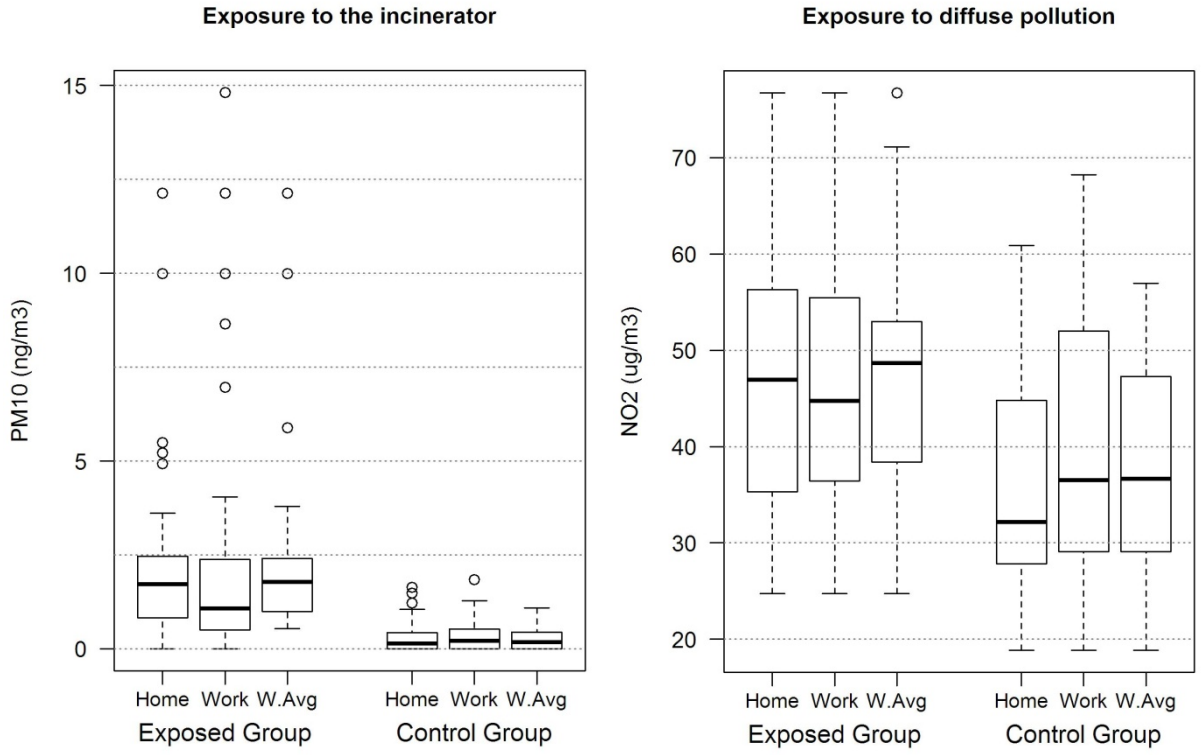

Figure 43 - Box-plots for the distribution of exposure to the incinerator (sx) and diffuse pollution (dx) in the exposed and control group. The box represent the inter-quartile range (IQR), the horizontal line inside the box is the median value, the whiskers extend to 1.5 times the IQR from the box. Residential exposure ("Home"), workplace exposure ("Work") and weighted average exposure ("W.Avg") are represented separately.

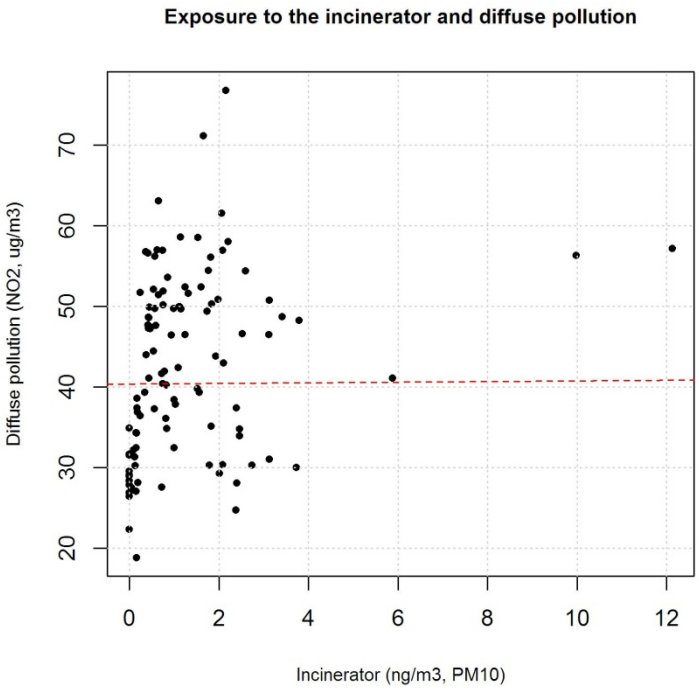

Figure 44 – Relationship between exposure to diffuse NO_2 pollution and PM_{10} from the incinerator in the cohort of asthmatics. The red line represent the linear regression line. The two variables are uncorrelated, thus it is unlikely that NO_2 will represent an environmental confounder for the effect of PM_{10} from the incinerator.

7
CONCLUSIONS

The present work focuses on the evaluation of human exposure to air pollution and to industrial sources of atmospheric pollutants, with a particular focus on waste incinerators. Exposure assessment represent a key step in many scientific disciplines that study the effect of anthropogenic pollutants on living beings, like environmental epidemiology, ecoepidemiology, ecotoxicology, human and ecological risk assessment. The common issue in all these applied sciences is the definition of the frequency and magnitude of the contact (i.e. exposure) between a polluted environmental media and a receptor.

In this thesis, I considered indirect methods of exposure assessment and human beings as the receptors of interest, although many of the methods presented here can be applied to other situations. The overreaching goal of my work was to draw the attention of the environmental health researchers on the key role of exposure assessment in determining the reliability of risk results. Specifically, the aim of my research was (i) to review available exposure assessment methods, (ii) to define a quality classification framework, (iii) to evaluate the possible effects of poor exposure assessment on risk estimation and (iv) to explore the applicability of exposure assessment methods in the field epidemiology and risk assessment.

I started my analysis with an overview of available methodologies for exposure assessment to atmospheric pollution. Classical methods used in environmental epidemiology relies on the use of questionnaires: although epidemiological theory deeply analyzed different kind of bias in this type of assessment, the most common and "safe" way to learn about a specific exposure is to ask the most informed subject, i.e. the person who is experiencing the exposure. Advances in computer science and data analysis now allow the use of more objective methods for exposure assessment, like geographical information processed in a GIS environment. The case-study I presented in Chapter 2 shows that the results of self-reported and GIS exposure assessment are to some extent comparable. Nevertheless, when good quality data are available, the use of objective measures of exposure is encouraged: GIS and spatial analysis enable considerations about the temporal and spatial dimension of environmental exposure, which subjective evaluation cannot easily handle. In particular, since most of modern epidemiological studies on environmental exposure call for the consideration of very large population, the collection of self-reported information is impractical and the use of spatial proxies of exposure and GIS analyses becomes essential.

In the third chapter I reviewed the literature about the health effects of waste incinerators and I proposed a numerical classification scheme for the quality of the exposure assessment to industrial sources of air pollution, based on (i) the approach used to define the intensity of exposure to the emission source, (ii) the scale at which the spatial distribution of the exposed receptors is accounted for and (iii) whether temporal variability in exposure is considered or not. Overall the analysis shows that many published studies suffers of poor exposure characterization. About a third of published studies used qualitative measures of exposure (e.g. presence/absence of the source). These studies are hardly usable to establish any causal exposure-effect relationship. Moreover, the heterogeneity of methods used in the literature make it difficult to compare results from different studies. I suggests the use of atmospheric dispersion models of pollutants emitted from a source, combined with precise geographic localizations of places where people spend time in the study area, as the reference method to assess exposure of population in epidemiological studies on industrial sources of atmospheric emissions.

In face of a variety of methods used for exposure assessment, it is of interest to understand what is the error arising from using different methodologies with decreasing precision and the effect of this error on measures of risk-exposure association. In Chapter 4 I showed that the use of different measures of exposure in a case study on a waste incinerator cause a substantial degree of exposure misclassification. The use of distance introduces a substantial exposure misclassification with respect to the best suggested method, i.e. the use of simulated atmospheric concentrations. The use of modelled atmospheric ground depositions is debatable, but I suggests that its role in determining exposure to a specific industrial source is generally more relevant in some specific areas (e.g. recreational areas or cultivated land) than at the residential address. Moreover, the results of the simulation conducted in Chapter 4 demonstrated that when exposure to a point industrial source is poorly characterized, we cannot be confident that because of non-differential exposure misclassification the risk we measure is lower than the "real risk" we would measure with a better exposure assessment.

In Chapter 5 I moved a step forward introducing some methods of exposure assessment that are typical of the environmental health risk assessment process. Since risk assessment models require the estimation of the dose of pollutant assumed through a variety of exposure pathways, atmospheric dispersion models need to be coupled with models that predict the transfer of pollutants through different environmental media (e.g. soil, water, food) and the magnitude and duration of the contact with human receptors. I applied a multi-pathway exposure model chain to the case study of the incinerator in Parma, showing the advantages in the use of risk assessment models, i.e. the possibility of comparing different exposure scenarios. In this case study, the activation of a district heating network powered through heat recovery from waste incineration and the switch-off of domestic boilers has the potential to compensate the health risks for PM_{10} emitted from the stack. Yet, the reduction in exposure to macropollutants shall be carefully balanced against the increased exposure to micropollutants. Moreover, the analysis highlighted that indirect exposure through consuming contaminated food represents a relevant exposure pathway and the careful definition of the dietary habits and food origin (i.e. home-grown vs. market food) is essential to conduct adequate risk assessment studies for anthropogenic sources of persistent pollutants. The routine use of residence as the exposure location in most epidemiological studies may not correctly account for the contribution of indirect exposure pathways for some pollutants, that may enter in the food chain in other locations of the impacted area.

The industrial emission source under study is rarely the only relevant emission source on a territory. In Chapter 6 I thus propose the use of Land Use Regression (LUR) models as a cost-effective method to model diffuse atmospheric pollution and control for confounding of health effect. With a few effort in sampling atmospheric concentration and the availability of digitalized information on land characteristics it is possible to estimate diffuse pollution concentration at unsampled locations. The case study I presented shows that care must be taken when extrapolating model result in environmental contexts that are different from those in which observation used to develop the models were collected. Moreover, I showed how the use of LUR to model exposure to diffuse pollution when designing epidemiological studies for specific point emission sources may help in reducing the risk of effect confounding.

Overall, my work showed that exposure information is crucial for predicting, preventing and reducing risks for human health and ecosystems. The availability of methods and technologies have historically limited exposure science but recent advances in GIS, pollution modelling and environmental data handling represent an important opportunity to enhance the quality of exposure assessment. Although this issue is generally recognized by the literature, yet many studies are published where expose is poorly characterized due to a lack of expertise in the use of exposure models or the unavailability of good quality environmental data.

BIBLIOGRAPHY

Abrahams, P., 2002. Soils: their implications to human health. Sci. Total Environ. 291, 1–32. doi:10.1016/s0048-9697(01)01102-0

Agarwal, N., Banternghansa, C., Bui, L.T.M., 2010. Toxic exposure in America: Estimating fetal and infant health outcomes from 14 years of TRI reporting. J. Health Econ. 29, 557–574. doi:10.1016/j.jhealeco.2010.04.002

Alberg, A., Samet, J., 2003. Epidemiology of lung cancer. Chest J. 123, 21S–49S. doi:10.1378/chest.123.1

Allen, R.W., Amram, O., Wheeler, A.J., Brauer, M., 2011. The transferability of NO and NO2 land use regression models between cities and pollutants. Atmos. Environ. 45, 369–378. doi:10.1016/j.atmosenv.2010.10.002

Allsopp, M., Costner, P., Johnston, P., 2001. Incineration and human health. State of Knowledge of the Impacts of Waste Incinerators on Human Health. Exeter, UK.

Altman, D., Royston, P., 2006. The cost of dichotomising continuous variables. BMJ 332, 1080.

Anderson, J.O., Thundiyil, J.G., Stolbach, A., 2012. Clearing the air: a review of the effects of particulate matter air pollution on human health. J. Med. Toxicol. 8, 166–75. doi:10.1007/s13181-011-0203-1

Annesi-Maesano, I., Agabiti, N., Pistelli, R., Couilliot, M.-F., Forastiere, F., 2003. Subpopulations at increased risk of adverse health outcomes from air pollution. Eur. Respir. J. 21, 57S–63s. doi:10.1183/09031936.03.00402103

APAT, 2008. Annuario dei dati ambientali 2007. APAT.

Aral, M., 2010. Environmental Modeling and Health Risk Analysis (Acts/Risk). Springer Science + Buisness Media, New York, NY. doi:10.1007/978-90-481-8608-2

ASR-ER, 2007. Atlante della mortalità in Emilia Romagna 1998–2004 [Italian]. Bologna.

Babyak, M. a., 2004. What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. Psychosom. Med. 66, 411–421. doi:10.1097/00006842-200405000-00021

Baker, D., Nieuwenhuijsen, M.J., 2008. Environmental Epidemiology. Study methods and application. Oxford University Press, Oxford.

Barbone, F., Bovenzi, M., Biggeri, A., Lagazio, C., Cavallieri, F., Stanta, G., 1995. Comparison of epidemiologic methods in a case-control study of lung cancer and air pollution in Trieste, Italy. Epidemiol. Prev. 19, 193–205.

Bartell, S.M., 2008. Ecological Risk Assessment, in: Jorgensen, S.E., Brian, F. (Eds.), Encyclopedia of Ecology, Five-Volume Set. Elsevier Science.

Basagaña, X., Rivera, M., Aguilera, I., Agis, D., Bouso, L., Elosua, R., Foraster, M., de Nazelle, A., Nieuwenhuijsen, M., Vila, J., Künzli, N., 2012. Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmos. Environ. 54, 634–642. doi:10.1016/j.atmosenv.2012.01.064

Baxter, L.K., Dionisio, K.L., Burke, J., Ebelt Sarnat, S., Sarnat, J.A., Hodas, N., Rich, D.Q., Turpin, B.J., Jones, R.R., Mannshardt, E., Kumar, N., Beevers, S.D., Özkaynak, H., Sarnat, S.E., 2013. Exposure prediction approaches used in air pollution epidemiology studies: key findings and future recommendations. J. Expo. Sci. Environ. Epidemiol. 23, 654–9. doi:10.1038/jes.2013.62

Bayer-Oglesby, L., Schindler, C., Hazenkamp-Von Arx, M., Braun-Fahrlander, C., Keidel, D., Rapp, R., Kunzli, N., Braendli, O., Burdet, L., Liu, S., Lauenberger, P., Ackermann-Liebrich, U., 2006. Original Contribution Living near Main Streets and

Respiratory Symptoms in Adults The Swiss Cohort Study on Air Pollution and Lung Diseases in Adults. Am. J. Epidemiol. 164, 1190–1198. doi:10.1093/aje/kwj338

Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M.-Y., Künzli, N., Schikowski, T., Marcon, A., Eriksen, K.T., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Tuomi, T., Declercq, C., Falq, G., Stempfelet, M., Birk, M., Cyrys, J., von Klot, S., Nádor, G., Varró, M.J., Dėdelė, A., Gražulevičienė, R., Mölter, A., Lindley, S., Madsen, C., Cesaroni, G., Ranzi, A., Badaloni, C., Hoffmann, B., Nonnemacher, M., Krämer, U., Kuhlbusch, T., Cirach, M., de Nazelle, A., Nieuwenhuijsen, M., Bellander, T., Korek, M., Olsson, D., Strömgren, M., Dons, E., Jerrett, M., Fischer, P., Wang, M., Brunekreef, B., de Hoogh, K., 2013. Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. Atmos. Environ. 72, 10–23. doi:10.1016/j.atmosenv.2013.02.037

Bellander, T., Berglind, N., Gustavsson, P., Jonson, T., Nyberg, F., Pershagen, G., Järup, L., 2001. Exposure to Air Pollution from Traffic and House Heating in Stockholm. Environ. Health Perspect. 109, 633–639. doi:10.1289/ehp.01109633

Bennette, C., Vickers, A., 2012. Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents. BMC Med. Res. Methodol. 12, 21. doi:10.1186/1471-2288-12-21

Bianchi, F., Minichilli, F., 2006. Mortalità per linfomi non Hodgkin nel periodo 1981-2001 in comuni italiani con inceneritori di rifiuti solidi urbani [In Italian]. Epidemiol. Prev. 30, 80–81.

Biancolini, V., Canè, M., Fornaciari, S., Forti, S., 2011. Quaderni di MONITER. Le emissioni degli inceneritori di ultima generazione. Analisi dell'impianto del Frullo di Bologna. [in Italian]. Bologna.

Biggeri, A., Barbone, F., Lagazio, C., Bovenzi, M., Stanta, G., 1996. Air Pollution and Lung Cancer in Trieste, Italy: Spatial Analysis of Risk. Environ. Health Perspect. 104, 750–754. doi:10.1016/0169-5002(95)90521-9

Biggeri, A., Catelan, D., 2005. Mortalità per linfoma non Hodgkin e sarcomi dei tessuti molli nel territorio circostante un impianto di incenerimento di rifiuti solidi urbani. Campi Bisenzio (Toscana, Italia) 1981-2001. [in Italian]. Epidemiol. Prev. 29, 156–159.

Biggeri, A., Catelan, D., 2006. Mortalità per linfomi non Hodgkin nei comuni della Regione Toscana dove sono stati attivi inceneritori di rifiuti solidi urbani nel periodo 1970-1989 [In Italian]. Epidemiol. Prev. 30, 14–15.

Birk, M., Ivina, O., Klot, S. Von, Heinrich, J., 2011. Road traffic noise: self-reported noise annoyance versus GIS modelled road traffic noise exposure. J. Environ. Monit. 13, 3237–3245. doi:10.1039/c1em10347d

Blair, A., Stewart, P., Lubin, J.H., Forastiere, F., 2007. Methodological issues regarding confounding and exposure misclassification in epidemiological studies of occupational exposures. Am. J. Ind. Med. 50, 199–207. doi:10.1002/ajim.20281

Boffetta, P., Nyberg, F., 2003. Contribution of environmental factors to cancer risk. Br. Med. Bullettin 68, 71–94. doi:10.1093/bmb/ldg023

Breen, M.S., Schultz, B.D., Sohn, M.D., Long, T., Langstaff, J., Williams, R., Isaacs, K., Meng, Q.Y., Stallings, C., Smith, L., 2014. A review of air exchange rate models for air pollution exposure assessments. J. Expo. Sci. Environ. Epidemiol. 24, 555–563. doi:10.1038/jes.2013.30

Briggs, D., 2006. The role of GIS: coping with space (and time) in air pollution exposure assessment. J. Toxicol. Environ. Health. A 68, 1243–61. doi:10.1080/15287390590936094

Briggs, D.J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., Pryl, K., Van Reeuwijk, H., Smallbone, K., Van Der Veen, A., 1997. Mapping urban air pollution using GIS: a regression-based approach. Int. J. Geogr. Inf. Sci. 11, 699–718. doi:10.1080/136588197242158

BSEM, 2008. The Health Effects of Waste Incinerators. 4th Report of the British Society for Ecological Medicine. Second Edition.

Buonanno, G., Ficco, G., Stabile, L., 2009. Size distribution and number concentration of particles at the stack of a municipal waste incinerator. Waste Manag. 29, 749–755. doi:10.1016/j.wasman.2008.06.029

Buonanno, G., Stabile, L., Avino, P., Belluso, E., 2011. Chemical , dimensional and morphological ultrafine particle characterization from a waste-to-energy plant. Waste Manag. 31, 2253–2262. doi:10.1016/j.wasman.2011.06.017

Burke, J., Zufall, M., Ozkaynak, H., 2001. A population exposure model for particulate matter: case study results for PM2. 5 in Philadelphia, PA. J. Expo. Anal. Environ. Epidemiol. 11, 470–489.

Cambridge Environmental Research Consultants Ltd., 2001. ADMS URban technical specification.

Candela, S., Ranzi, A., Bonvicini, L., Baldacchini, F., Marzaroli, P., Evangelista, A., Luberto, F., Carretta, E., Angelini, P., Sterrantino, A.F., Broccoli, S., Cordioli, M., Ancona, C., Forastiere, F., 2013. Air pollution from incinerators and reproductive outcomes: a multisite study. Epidemiology 24, 863–70. doi:10.1097/EDE.0b013e3182a712f1

Canfield, M., Ramadhani, T., Langolois, P., Waller, K., 2006. Residential mobility patterns and exposure misclassification in epidemiologic studies of birth defects. J. Expo. Sci. Environ. Epidemiol. 16, 538–543. doi:10.1038/sj.jes.7500501

Cangialosi, F., Intini, G., Liberti, L., 2008. Health risk assessment of air emissions from a municipal solid waste incineration plant–A case study. Waste Manag. 28, 885–895. doi:10.1016/j.wasman.2007.05.006

Carlos García-Díaz, J., Gozalvez-Zafrilla, J.M., 2012. Uncertainty and sensitive analysis of environmental model for risk assessments: An industrial case study. Reliab. Eng. Syst. Saf. 107, 16–22. doi:10.1016/j.ress.2011.04.004

Carslaw, D.C., Beevers, S.D., Tate, J.E., Westmoreland, E.J., Williams, M.L., 2011. Recent evidence concerning higher NOx emissions from passenger cars and light duty vehicles. Atmos. Environ. 45, 7053–7063. doi:10.1016/j.atmosenv.2011.09.063

Carson, R., 1962. Silent Spring. Houghton Mifflin, Boston.

Caserini, S., Cernuschi, S., Giugliano, M., Grosso, M., Lonati, G., Mattaini, P., 2004. Air and soil dioxin levels at three sites in Italy in proximity to MSW incineration plants. Chemosphere 54, 1279–1287. doi:10.1016/S0045-6535(03)00250-9

Cesaroni, G., Badaloni, C., Porta, D., Forastiere, F., Perucci, C.A., 2008. Comparison between various indices of exposure to traffic-related air pollution and their impact on respiratory health in adults. Occup. Environ. Med. 65, 683–690. doi:10.1136/oem.2007.037846

Chang, J., Hanna, S., 2004. Air quality model performance evaluation. Meteorol. Atmos. Phys. 87, 167–196. doi:10.1007/s00703-003-0070-7

Chan-Yeung, M., Koo, L.C., Ho, J.C., Tsang, K.W., Chau, W., Chiu, S., Ip, M.S., Lam, W., 2003. Risk factors associated with lung cancer in Hong Kong. Lung Cancer 40, 131–140. doi:10.1016/S0169-5002(03)00036-9

Chen, H.L., Su, H.J., Liao, P.C., Chen, C.H., Lee, C.C., 2004. Serum PCDD / F concentration distribution in residents living in the vicinity of an incinerator and its association with predicted ambient dioxin exposure. Chemosphere 54, 1421–1429. doi:10.1016/j.chemosphere.2003.10.033

Choi, H.S., Shim, Y.K., Kaye, W.E., Ryan, P.B., 2006. Potential Residential Exposure to Toxics Release Inventory Chemicals during Pregnancy and Childhood Brain Cancer. Environ. Health Perspect. 114, 1113–1118. doi:10.1289/ehp.9145

Ciccone, G., Forastiere, F., Agabiti, N., Biggeri, A., Bisanti, L., Chellini, E., Corbo, G., Orco, V.D., Dalmasso, P., Volante, T.F., Galassi, C., V, S.P., Renzoni, E., Rusconi, F., 1998. Road traffic and adverse respiratory effects in children. Occup. Environ. Med. 5, 771–778. doi:10.1136/oem.55.11.771

Cirillo, M.C., Manzi, D., 1991. PC DIMULA 2.0: An atmospheric multisource dispersion model of air pollutants on local scale. Environ. Softw. 6, 43–48. doi:10.1016/0266-9838(91)90018-L

Colvile, R.N., Woodfield, N.K., Carruthers, D.J., Fisher, B.E. a., Rickard, A., Neville, S., Hughes, A., 2002. Uncertainty in dispersion modelling and urban air quality mapping. Environ. Sci. Policy 5, 207–220. doi:10.1016/S1462-9011(02)00039-4

Comba, P., Ascoli, V., Belli, S., 2003. Risk of soft tissue sarcomas and residence in the neighbourhood of an incinerator of industrial wastes. Occup. Environ. Med. 60, 680–683. doi:10.1136/oem.60.9.680

Cook, A.G., Jbm, A., Pereira, G., Jardine, A., Weinstein, P., 2011. Use of a total traffic count metric to investigate the impact of roadways on asthma severity□: a case-control study. Environ. Heal. 10, 52. doi:10.1186/1476-069X-10-52

Cordier, S., Chevrier, C., Robert-Gnansia, E., Lorente, C., Brula, P., Hours, M., 2004. Risk of congenital anomalies in the vicinity of municipal solid waste incinerators. Occup. Environ. Med. 61, 8–15.

Cordier, S., Lehébel, A., Amar, E., Anzivino-viricel, L., Hours, M., Monfort, C., Chevrier, C., Chiron, M., Robert-gnansia, E., Lehe, A., 2010. Maternal residence near municipal waste incinerators and the risk of urinary tract birth defects. Occup. Environ. Med. 67, 493–499. doi:10.1136/oem.2009.052456

Cordioli, M., Vincenzi, S., Leo, G.A. De, 2013. Effects of heat recovery for district heating on waste incineration health impact□: A simulation study in Northern Italy. Sci. Total Environ. 444, 369–380. doi:10.1016/j.scitotenv.2012.11.079

Coughlin, S.S., 1990. Recall bias in epidemiologic studies. J. Clin. Epidemiol. 43, 87–91.

Curwen, M.P., Kennaway, E.L., Kennaway, N.M., 1954. The incidence of cancer of the lung and larynx in urban and rural districts. Br. J. Cancer 8, 181–198.

Da Silva, R., Bach-Faig, A., Raidó Quintana, B., Buckland, G., Vaz de Almeida, M.D., Serra-Majem, L., 2009. Worldwide variation of adherence to the Mediterranean diet, in 1961-1965 and 2000-2003. Public Health Nutr. 12, 1676–84. doi:10.1017/S1368980009990541

Dale, M.R.T., Fortin, M.-J., 2009. Spatial autocorrelation and statistical tests: Some solutions. J. Agric. Biol. Environ. Stat. 14, 188–206. doi:10.1198/jabes.2009.0012

DEFRA, 2004. Review of Environmental and Health Effects of Waste Management□: Municipal Solid Waste and Similar Wastes. London.

Diggle, P., Elliott, P., 1995. Disease risk near point sources□: statistical issues for analyses using individual or spatially aggregated data. J. Epidemiol. Community Health 49, 20–27.

Dionisio, K.L., Isakov, V., Baxter, L.K., Sarnat, J. a, Sarnat, S.E., Burke, J., Rosenbaum, A., Graham, S.E., Cook, R., Mulholland, J., Özkaynak, H., 2013. Development and evaluation of alternative approaches for exposure assessment of multiple air pollutants in Atlanta, Georgia. J. Expo. Sci. Environ. Epidemiol. 23, 581–92. doi:10.1038/jes.2013.59

Dockery, D., Pope, C., Xu, X., Spengler, J., Ware, J., Fay, M., Ferris, B., Speizer, F., 1993. An association between air pollution and mortality in six US cities. N. Engl. J. Med. 329, 1753–1759.

Domingo, L., 2004. Health risk assessment of emissions of dioxins and furans from a municipal waste incinerator□: comparison with other emission sources. Environ. Int. 30, 481–489. doi:10.1016/j.envint.2003.10.001

Dons, E., Van Poppel, M., Int Panis, L., De Prins, S., Berghmans, P., Koppen, G., Matheeussen, C., 2014. Land use regression models as a tool for short, medium and long term exposure to traffic related air pollution. Sci. Total Environ. 476-477, 378–86. doi:10.1016/j.scitotenv.2014.01.025

Dons, E., Van Poppel, M., Kochan, B., Wets, G., Int Panis, L., 2013. Modeling temporal and spatial variability of traffic-related air pollution: Hourly land use regression models for black carbon. Atmos. Environ. 74, 237–246. doi:10.1016/j.atmosenv.2013.03.050

Dosemeci, M., Wacholder, S., Lubin, J.H., 1990. Does nondifferential misclassification of exposure always bias a true effect toward the null value? Am. J. Epidemiol. 132, 746–748.

Duhme, H., Weiland, S., Keil, U., Kramer, B., Schmidt, M., Stender, M., Chambless, L., 1996. The association between self-reported symptoms of asthma and allergic rhinitis and self-reported traffic density on street of residence in adolescents. Epidemiology 7, 578–582.

Dummer, T.J.B., Dickinson, H.O., Parker, L., 2003. Adverse pregnancy outcomes around incinerators and crematoriums in Cumbria, north west England, 1956–93. J. Epidemiol. Community Health 57, 456–461.

Duncan, D.T., Castro, M.C., Blossom, J.C., Bennett, G.G., Steven, L., 2011. Evaluation of the positional difference between two common geocoding methods. Geospat. Health 5, 265–273.

Eberhard, M., Pamuk, E., 2004. The Importance of Place of Residence□: Examining Health in Rural and Nonrural Areas. Am. J. Public Health 94, 1682–1686.

EEA, 2000. Corine Land Cover technical guide - Addendum 2000. Copenhagen.

Eeftens, M., Beekhuizen, J., Beelen, R., Wang, M., Vermeulen, R., Brunekreef, B., Huss, A., Hoek, G., 2013. Quantifying urban street configuration for improvements in air pollution models. Atmos. Environ. 72, 1–9. doi:10.1016/j.atmosenv.2013.02.007

Elliott, P., Hills, M., Beresford, J., Kleinschmidt, I., Jolley, D., Pattenden, S., Rodrigues, L., Westlake, A., Rose, G., 1992. Incidence of cancers of the larynx and lung near incinerators of waste solvents and oils in Great Britain. Lancet 339, 854–858. doi:http://dx.doi.org/10.1016/0140-6736(92)90290-J

Elliott, P., Shaddick, G., Kleinschmidt, I., Jolley, D., Walls, P., Beresford, J., Grundy, C., 1996. Cancer incidence near municipal solid waste incinerators in Great Britain. Br. J. Cancer 73, 702–710.

ENEA, 2005. Rapporto Energia e Ambiente 2005 - L'analisi. Roma.

Eriksson, C., Nilsson, M.E., Stenkvist, D., Bellander, T., Pershagen, G., 2012. Residential traffic noise exposure assessment: application and evaluation of European Environmental Noise Directive maps. J Expo. Sci Env. Epidemiol.

Erspamer, L., Ranzi, A., Silvi, G., Rubbi, L., Lauriola, P., 2007. Case-control study on the long-term effects of urban-industrial pollution: the Ravenna experience. [In Italian]. Epidemiol. Prev. 21, 28–33.

ESCAPE project, 2010. European Study of Cohorts for Air Pollution Effects (ESCAPE project). Exposure assessment manual.

Eurostat, 2010. Eurostat statistical books. Environmental statistics and accounts in Europe 2010, 2010 Editi. ed. European Union, 2010.

Federico, M., Pirani, M., Rashid, I., Caranci, N., Cirilli, C., 2010. Cancer incidence in people with residential exposure to a municipal waste incinerator□: An ecological study in Modena (Italy), 1991 – 2005. Waste Manag. 30, 1362–1370. doi:10.1016/j.wasman.2009.06.032

Field, R.W., Withers, B.L., 2012. Occupational and Environmental Causes of Lung Cancer. Clin. Chest Med. 33, 681–703. doi:10.1016/j.ccm.2012.07.001

Flegal, K.M., Keyl, P.M., Nieto, F.J., 1991. Differential Misclassification Arising from Nondifferential Errors in Exposure Measurement. Am. J. Epidemiol. 134, 1233–1244.

Floret, N., Mauny, F., Challier, B., Arveux, P., Cahn, J., 2003. Dioxin Emissions from a Solid Waste Incinerator and Risk of Non-Hodgkin Lymphoma. Epidemiology 14, 392–398. doi:10.1097/01.ede.0000072107.90304.01

Floret, N., Viel, J.J.-F., Lucot, E., Dudermel, P.-M., Cahn, J.-Y.J.-Y., Badot, P.-M., Mauny, F., 2006. Dispersion modeling as a dioxin exposure indicator in the vicinity of a municipal solid waste incinerator: a validation study. Environ. Sci. Technol. 40, 2149–2155. doi:10.1021/es052309u

Forastiere, F., Badaloni, C., de Hoogh, K., von Kraus, M.K., Martuzzi, M., Mitis, F., Palkovicova, L., Porta, D., Preiss, P., Ranzi, A., Perucci, C. a, Briggs, D., 2011. Health impact assessment of waste management facilities in three European countries. Environ. Health 10, 53. doi:10.1186/1476-069X-10-53

Forastiere, F., Galassi, C., 2005. Self report and GIS based modelling as indicators of air pollution exposure: is there a gold standard? Occup. Environ. Med. 62, 508–509. doi:10.1136/oem.2005.019901

Franchini, M., Rial, M., Buiatti, E., Bianchi, F., 2004. Health effects of exposure to waste incinerator emissions: a review of epidemiological studies. Ann Ist Super Sanità 40, 101–115.

Fries, G.F., 1995. A Review of the Significance of Animal Food Products as Potential Pathways of Human Exposures to Dioxins. J. Anim. Sci. 73, 1639–1650.

Frøslie, K.F., Røislien, J., Laake, P., Henriksen, T., Qvigstad, E., Veierød, M.B., 2010. Categorisation of continuous exposure variables revisited. A response to the Hyperglycaemia and Adverse Pregnancy Outcome (HAPO) Study. BMC Med. Res. Methodol. 10, 103. doi:10.1186/1471-2288-10-103

Fryer, M., Collins, C.D., Ferrier, H., Colvile, R.N., Nieuwenhuijsen, M.J., 2006. Human exposure modelling for chemical risk assessment□: a review of current approaches and research and policy implications. Environ. Sci. Policy 9, 261–274. doi:10.1016/j.envsci.2005.11.011

Fuentes, M., Raftery, A.E., 2005. Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. Biometrics 61, 36–45. doi:10.1111/j.0006-341X.2005.030821.x

Fukuda, Y., Nakamura, K., Takano, T., 2003. Original Article Dioxins Released from Incineration Plants and Mortality from Major Diseases: an Analysis of Statistical Data by Municipalities. J Med Dent Sci 50, 249–255.

García-pérez, J., Fernández-navarro, P., Castelló, A., López-Cima, M.F., Ramis, R., Boldo, E., López-Abente, G., 2013. Cancer mortality in towns in the vicinity of incinerators and installations for the recovery or disposal of hazardous waste. Environ. Int. 51, 31–44. doi:10.1016/j.envint.2012.10.003

Garcia-Perez, J., Pollàn, M., Boldo, E., Perèz-Gòmez, B., Arragonès, N., Lope, V., Vidal, E., Lopez-Abente, G., 2009. Mortality due to lung, laryngeal and bladder cancer in towns lying in the vicinity of combustion installations. Sci. Total Environ. 407, 2593–2602.

Georgopoulos, P.G., Wang, S.-W., Vyas, V.M., Sun, Q., Burke, J., Vedantham, R., McCurdy, T., Ozkaynak, H., 2005. A source-to-dose assessment of population exposures to fine PM and ozone in Philadelphia, PA, during a summer 1999 episode. J. Expo. Anal. Environ. Epidemiol. 15, 439–457. doi:10.1038/sj.jea.7500422

Giusti, L., 2009. A review of waste management practices and their impact on human health. Waste Manag. 29, 2227–2239. doi:10.1016/j.wasman.2009.03.028

Goldoni, C., Gambini, M., Vaccari, C., Soli, A., Ranzi, A., De Girolamo, G., Valerio, L., Borella, P., Lauriola, P., 2003. The role of epidemiology applied to a local area based on a study caried out in the community of Fiorano Modenese (ceramic district). [In Italian]. Epidemiol. Prev. 27, 310–5.

Goria, S., Daniau, C., Crouy-chanel, P. De, Empereur-, P., Fabre, P., Colonna, M., Duboudin, C., Viel, F., Richardson, S., 2009. Risk of cancer in the vicinity of municipal solid waste incinerators: importance of using a flexible modelling strategy. Int. J. Health Geogr. 8. doi:10.1186/1476-072X-8-31

Gray, E.J., Peat, J.K., Mellis, C.M., Harrington, J., Woolcock, A.J., 1994. Asthma severity and morbidity in a population sample of Sydney school children: Part I – Prevalence and effect of air pollutants in coastal regions. Aust. N. Z. J. Med. 24, 168–174. doi:10.1111/j.1445-5994.1994.tb00553.x

Guarnieri, M., Balmes, J.R., 2014. Outdoor air pollution and asthma. Lancet 383, 1581–92. doi:10.1016/S0140-6736(14)60617-6

Gunier, R.B., Reynolds, P., Hurley, S.E., Yerabati, S., Hertz, A., Strickland, P., Horn-ross, P.L., 2006. Estimating Exposure to Polycyclic Aromatic Hydrocarbons: A Comparison of Survey , Biological Monitoring , and Geographic Information System − Based Methods. Cancer Epidemiol. Biomarkers Prev. 15, 1376–1381. doi:10.1158/1055-9965.EPI-05-0799

Hamra, G., Guha, N., Cohen, A., Laden, F., Raaschou-Nielsen, O., Samet, J.M., Vineis, P., Forastiere, F., Saldiva, P., Yorifuji, T., Loomis, D., 2014. Outdoor Particulate Matter Exposure and Lung Cancer: A Systematic Review. Environ. Health Perspect. 122, 906–912.

Heinrich, J., Gehring, U., Cyrys, J., 2005. Exposure to traffic related air pollutants: self reported traffic intensity versus GIS modelled exposure. Occup. Environ. Med. 62, 517–524. doi:10.1136/oem.2004.016766

Hendryx, M., Fedorko, E., 2011. The Relationship Between Toxics Release Inventory Discharges and Mortality Rates in Rural and Urban Areas of the United States. J. Rural Heal. Rural Heal. 27, 358–366. doi:10.1111/j.1748-0361.2011.00367.x

Hoek, G., Beelen, R., Hoogh, K. De, Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., de Hoogh, K., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos. Environ. 42, 7561–7578. doi:10.1016/j.atmosenv.2008.05.057

Hoek, G., Krishnan, R.M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., Kaufman, J.D., 2013. Long-term air pollution exposure and cardio- respiratory mortality: a review. Environ. Heal. 12, 43. doi:10.1186/1476-069X-12-43

Hofelt, C.S., Honeycutt, M., Mccoy, J.T., Haws, L.C., 2001. Development of a Metabolism Factor for Polycyclic Aromatic Hydrocarbons for Use in Multipathway Risk Assessments of Hazardous Waste Combustion Facilities. Regul. Toxicol. Pharmacol. 33, 60–65. doi:10.1006/rtph.2000.1438

Höfler, M., 2005. The effect of misclassification on the estimation of association: a review. Int. J. Methods Psychiatr. Res. 14, 92–101.

Holland, M., Pye, S., Watkiss, P., Droste-Franke, B., Bickel, P., 2005. Damages per tonne emission of EU25 Member State (excluding Cyprus) and surrounding seas.

Hoogh, K. De, Wang, M., Adam, M., Badaloni, C., Beelen, R., Birk, M., Cesaroni, G., Cirach, M., Nieuwenhuijsen, M., Pershagen, G.G., Peters, A., Phuleria, H., Probst-hensch, N., Raaschou-nielsen, O., de Hoogh, K., Declercq, C., Dėdelė, A., Dons, E., de Nazelle, A., Eeftens, M., Eriksen, K., Eriksson, C., Fischer, P., Gražulevičienė, R., Gryparis, A., Hoffmann, B., Jerrett, M., Katsouyanni, K., Iakovides, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Mosler, G., Nádor, G., Quass, U., Ranzi, A., Stephanou, E., Sugiri, D., Schwarze, P., Tsai, M.-Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Weinmayr, G., Brunekreef, B., Hoek, G., 2013. Development of land use regression models for particle composition in twenty study areas in Europe. Environ. Sci. Technol. 47, 5778–86. doi:10.1021/es400156t

Hosgood III, D., Boffetta, P., Greenland, S., Amy Lee, Y., McLaughlin, J., 2010. In-home coal and wood use and lung cancer risk: a pooled analysis of the International Lung Cancer Consortium. Environ. Health Perspect. 118, 1743–1747. doi:10.1289/ehp.1002217

Hosseini, M., Naghan, P.A., Karimi, S., Seyedalinaghi, S., Bahadori, M., Khodadad, K., Mohammadi, F., Keynama, K., 2009. Environmental risk factors for lung cancer in Iran□: a case-control study. Int. J. Epidemiol. 38, 989–996. doi:10.1093/ije/dyp218

Hsiue, T.R., Lee, S.S., Chen, H.I., 1991. Effects of air pollution resulting from wire reclamation incineration on pulmonary function in children. CHEST J. 100, 698. doi:10.1378/chest.100.3.698

Hu, S., Shy, C.M., 2001. Waste Incineration and Pulmonary Function: An Epidemiologic Study of Six Communities. J. Air Waste Manag. Assoc. 51, 1185–1194.

Hu, S.S., Shy, C.M.C., 2001. Health effects of waste incineration: a review of epidemiologic studies. J. Air Waste Manag. Assoc. 51, 1100–1109.

Huang, Y.U., Batterman, S., 2000. Residence location as a measure of environmental exposure□: a review of air pollution epidemiology studies. J. Expo. Anal. Environ. Epidemiol. 10, 66–85.

Hwong-Wen, M., 2002. Using stochastic risk assessment in setting information priorities for managing dioxin impact from a municipal waste incinerator. Chemosphere 48, 1035–1040.

Hystad, P., Demers, P.A., Johnson, K.C., Brook, J., Donkelaar, A. Van, Lamsal, L., Martin, R., Brauer, M., 2012. Spatiotemporal air pollution exposure assessment for a Canadian population-based lung cancer case-control study. Environ. Heal. 11, 22. doi:10.1186/1476-069X-11-22

IARC, 2013. Air Pollution and Cancer. IARC scientific report n. 161. Geneva, Switzerland.

Industrial Economics Inc, 2010. Effects of Air Pollutants on Ecological Resources□: Literature Review and Case Studies - Draft Report. Cambridge, MA.

INRAN, 2012. Tabelle di composizione degli alimenti [In Italian] [WWW Document]. URL http://nut.entecra.it/646/tabelle_di_composizione_degli_alimenti.html

ISPRA, 2012. Manuale dei fattori di emissione [In Italian].

Jackson, D.J., Sykes, A., Mallia, P., Johnston, S.L., 2011. Asthma exacerbations: origin, effect, and prevention. J. Allergy Clin. Immunol. 128, 1165–74. doi:10.1016/j.jaci.2011.10.024

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., 2005. A review and evaluation of intraurban air pollution exposure models. J. Expo. Anal. Environ. Epidemiol. 15, 185–204. doi:10.1038/sj.jea.7500388

Jurek, A.M., Greenland, S., Maldonado, G., 2008. How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null□? Int. J. Epidemiol. 37, 382–385. doi:10.1093/ije/dym291

Jurek, A.M., Greenland, S., Maldonado, G., Church, T.R., 2005. Proper interpretation of non-differential misclassification effects: expectations vs observations. Int. J. Epidemiol. 34, 680–7. doi:10.1093/ije/dyi060

Karademir, A., 2004. Health risk assessment of PCDD/F emissions from a hazardous and medical waste incinerator in Turkey. Environ. Int. 30, 1027–1038. doi:10.1016/j.envint.2004.05.008

Katsouyanni, K., Pershagen, G., 1997. Ambient air pollution exposure and cancer. Cancer Causes Control 8, 284–291.

Kheirbek, I., Johnson, S., Ross, Z., Pezeshki, G., Ito, K., Eisl, H., Matte, T., 2012. Spatial variability in levels of benzene, formaldehyde, and total benzene, toluene, ethylbenzene and xylenes in New York City: a land-use regression study. Environ. Heal. 11, 51. doi:10.1186/1476-069X-11-51

Kibble, A., Harrison, R., 2005. Point sources of air pollution. Occup. Med. (Chic. Ill). 55, 425–31. doi:10.1093/occmed/kqi138

Kim, J.J., Huen, K., Adams, S., Smorodinsky, S., Hoats, A., Malig, B., Lipsett, M., Ostro, B., 2008. Residential Traffic and Children ' s Respiratory Health. Environ. Health Perspect. 116, 1274–1279. doi:10.1289/ehp.10735

Kim, Y., Kim, J., Lee, H., 2011. Burden of disease attributable to air pollutants from municipal solid waste incinerators in Seoul , Korea□: A source-specific approach for environmental burden of disease. Sci. Total Environ. 409, 2019–2028. doi:10.1016/j.scitotenv.2011.02.032

Kissel, J.C., Richter, K.Y., Fenske, R.A., 1996. Field measurement of dermal soil loading attributable to various activities: Implications for exposure assessment. Risk Anal. 16, 115–125. doi:10.1111/j.1539-6924.1995.tb00769.x

Klepeis, N.E., Nelson, W.C., Ott, W.R., Robinson, J.P., Tsang, A.M., Switzer, P., Behar, J. V, Hern, C., Engelmann, W.H., 2001. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. J. Expo. Anal. Environ. Epidemiol. 11, 231–252.

Knox, E.G., 2000. Childhood cancers, birthplaces, incinerators and landfill sites. Int. J. Epidemiol. 29, 391–397.

Krieger, N., 2003. Place, Space, and Health: GIS and Epidemiology. Epidemiology 14, 384–385. doi:10.1097/01.ede.0000071473.69307.8a

Kuehni, C.E., Strippoli, M.F., Zwahlen, M., Silverman, M., 2006. Association between reported exposure to road traffic and respiratory symptoms in children: evidence of bias. Int. J. Epidemiol. 35, 779–786. doi:10.1093/ije/dyl022

Künzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., Herry, M., Jr, F.H., Quénel, P., 2000. Public-health impact of outdoor and traffic-related air pollution□: a European assessment. Lancet 356, 795–801.

Kuo, C.-Y., Wang, J.-Y., Liu, W.-T., Lin, P.-Y., Tsai, C.-T., Cheng, M.-T., 2012. Evaluation of the vehicle contributions of metals to indoor environments. J. Expo. Sci. Environ. Epidemiol. 22, 489–495.

Leclercq, C., Arcella, D., Piccinelli, R., Sette, S., Le Donne, C., Turrini, A., 2009. The Italian National Food Consumption Survey INRAN-SCAI 2005-06: main results in terms of food consumption. Public Health Nutr. 12, 2504–32. doi:10.1017/S1368980009005035

Lee, J.J.-T., Shy, C.M., 1999. Respiratory function as measured by peak expiratory flow rate and PM10: Six communities study. J. Expo. Anal. Environ. Epidemiol. 9, 293–299.

Lercher, P., Schmitzberger, R., Kofler, W., 1995. Perceived traffic air pollution, associated behavior and health in an alpine area. Sci. Total Environ. 169, 71–74.

Liao, D., Peuquet, D.J., Duan, Y., Whitsel, E. a., Dou, J., Smith, R.L., Lin, H., Chen, J., Heiss, G., 2006. GIS Approaches for the Estimation of Residential-Level Ambient PM Concentrations. Environ. Health Perspect. 114, 1374–1380. doi:10.1289/ehp.9169

Lim, S.S., Vos, T., Flaxman, A.D., Danaei, G., Shibuya, K., Adair-Rohani, H., Amann, M., et al., 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380, 2224–60. doi:10.1016/S0140-6736(12)61766-8

Linares, V., Perellò, G., Nadal, M., Gòmez-Catalàn, J., Llobet, J.M., Domingo, J.L., Perello, G., Gomez-Catalan, J., 2010. Environmental versus dietary exposure to POPs and metals□: A probabilistic assessment of human health risks. J. Environ. Monit. 12, 681–688. doi:10.1039/b914962g

Llobet, J.M., Domingo, J.L., Bocio, A., Casas, C., Lutz, M., Teixid, A., Muller, L., 2003. Human exposure to dioxins through the diet in Catalonia, Spain: carcinogenic and non-carcinogenic risk. Chemosphere 50, 1193–1200.

Lloyd, L., Lloyd, M.M., Williams, F.L.R., Lawson, A., 1988. Twinning in human populations and in cattle exposed to air pollution from incinerators. Br. J. Ind. Med. 45, 556–560.

Lonati, G., Zanoni, F., 2012. Probabilistic health risk assessment of carcinogenic emissions from a MSW gasi fi cation plant. Environ. Int. 44, 80–91. doi:10.1016/j.envint.2012.01.013

Loomis, D., Grosse, Y., Lauby-Secretan, B., Ghissassi, F. El, Bouvard, V., Benbrahim-Tallaa, L., Guha, N., Baan, R., Mattock, H., Straif, K., 2013. The carcinogenicity of outdoor air pollution. Lancet Oncol. 14, 1262–1263. doi:10.1016/S1470-2045(13)70487-X

Lovett, G.M., Tear, T.H., Evers, D.C., Findlay, S.E.G., Cosby, B.J., Dunscomb, J.K., Driscoll, C.T., Weathers, K.C., 2009. Effects of air pollution on ecosystems and biological diversity in the eastern United States. Ann. N. Y. Acad. Sci. 1162, 99–135. doi:10.1111/j.1749-6632.2009.04153.x

Luo, J., Hendryx, M., Ducatman, A., 2011. Association between Six Environmental Chemicals and Lung Cancer Incidence in the United States. J. Environ. Public Health 2011, 463701,. doi:10.1155/2011/463701

MAIND S.r.l., 2006. WinDimula 3.0: descrizione delle equazioni utilizzate dal modello WD3 [In Italian].

Makri, A., Stilianakis, N.I., 2008. Vulnerability to air pollution health effects. Int. J. Hyg. Environ. Health 211, 326–36. doi:10.1016/j.ijheh.2007.06.005

Mckone, T.E., 1996. Overview of the risk analysis approach and terminology: the merging of science , judgment and values. Food Control 7, 69–76.

McNamee, R., 2003. Confounding and confounders. Occup. Environ. Med. 60, 227–234. doi:10.1136/oem.60.3.227

Medina, S., Boldo, E., Saklad, M., Niciu, E., Krzyzanowski, M., Frank, F., 2005. Health Impact Assessment of Air Pollution and Communication Strategy. Air Pollution and Health: A European Information System (APHEIS). Saint-Maurice.

Meliker, J.R., Slotnick, M.J., Avruskin, G.A., Kaufmann, A., Jacquez, G.M., Nriagu, J.O., 2005. Improving exposure assessment in environmental epidemiology: Application of spatio-temporal visualization tools. J. Geogr. Syst. 7, 49–66. doi:10.1007/s10109-005-0149-4

Michelozzi, P., Fusco, D., Forastiere, F., Ancona, C., Orco, V.D., Perucci, C.A., 1998. Small area study of mortality among people living near multiple sources of air pollution. Occup. Environ. Med. 55, 611–615.

Migliore, E., Berti, G., Galassi, C., Pearce, N., Forastiere, F., Calabrese, R., Armenio, L., Biggeri, A., Bisanti, L., Bugiani, M., Cadum, E., Chellini, E., Orco, V.D., Giannella, G., Sestini, P., Corbo, G., Pistelli, R., Viegi, G., Ciccone, G., Group, S.-C., 2009. Respiratory symptoms in children living near busy roads and their relationship to vehicular traffic: results of an Italian multicenter study (SIDRIA 2). Environ. Heal. 8, 27. doi:10.1186/1476-069X-8-27

Mitchell, R., Popham, F., 2007. Greenspace, urbanity and health: relationships in England. J. Epidemiol. Community Health 61, 681–683. doi:10.1136/jech.2006.053553

Miyake, Y., Yura, A., Misaki, H., Ikeda, Y., Usui, T., Iki, M., Shimizu, T., 2005. Relationship between distance of schools from the nearest municipal waste incineration plant and child health in Japan. Eur. J. Epidemiol. 20, 1023–1029. doi:10.1007/s10654-005-4116-7

Mohan, A.K., Degnan, D., Charles, E., Feigley, C., Shy, C.M., Hornung, C.A., Mustafa, T., Macera, C.A., 2000. Comparison of respiratory symptoms among community residents near waste disposal incinerators. Int. J. Environ. Health Res. 10, 63–75.

Mölter, A., Lindley, S., de Vocht, F., Agius, R., Kerry, G., Johnson, K., Ashmore, M., Terry, A., Dimitroulopoulou, S., Simpson, A., 2012. Performance of a microenviromental model for estimating personal NO2 exposure in children. Atmos. Environ. 51, 225–233. doi:10.1016/j.atmosenv.2012.01.030

Morabia, A., 2011. History of the modern epidemiological concept of confounding. J. Epidemiol. Community Health 65, 297–300. doi:10.1136/jech.2010.112565

Morra, P., Bagli, S., Spadoni, G., 2006. The analysis of human health risk with a detailed procedure operating in a GIS environment. Environ. Int. 32, 444–454. doi:10.1016/j.envint.2005.10.003

Morra, P., Lisi, R., Spadoni, G., Maschio, G., 2009. The assessment of human health impact caused by industrial and civil activities in the Pace Valley of Messina. Sci. Total Environ. 407, 3712–3720. doi:10.1016/j.scitotenv.2009.03.005

Morselli, L., Robertis, C. De, Luzi, J., Passarini, F., Vassura, I., 2008. Environmental impacts of waste incineration in a regional system (Emilia Romagna , Italy) evaluated from a life cycle perspective. J. Hazard. Mater. 159, 505–511. doi:10.1016/j.jhazmat.2008.02.047

Nadal, M., Agramunt, M.C., Schuhmacher, M., Domingo, J.L., 2002. PCDD/PCDF congener profiles in soil and herbage samples collected in the vicinity of a municipal waste incinerator before and after pronounced reductions of PCDD/PCDF emissions from the facility. Chemosphere 49, 153–159.

Newman, M.C., 2008. Ecotoxicology: The History and Present Directions, in: Encyclopedia of Ecology, Five-Volume Set. pp. 1195–1200.

Nieuwenhuijsen, M., Paustenbach, D., Duarte-davidson, R., 2006. New developments in exposure assessment: the impact on the practice of health risk assessment and epidemiological studies. Environ. Int. 32, 996–1009. doi:10.1016/j.envint.2006.06.015

Noth, E.M., Hammond, S.K., Biging, G.S., Tager, I.B., 2011. A spatial-temporal regression model to predict daily outdoor residential PAH concentrations in an epidemiologic study in Fresno, CA. Atmos. Environ. 45, 2394–2403. doi:10.1016/j.atmosenv.2011.02.014

Nouwen, J., Cornelis, C., Fr, R. De, Wevers, M., Viaene, P., Mensink, C., Patyn, J., Verschaeve, L., Hooghe, R., Maes, A., Collier, M., Schoeters, G., Cleuvenbergen, R. Van, Geuzens, P., 2001. Health risk assessment of dioxin emissions from municipal waste incinerators: the Neerlandquarter (Wilrijk , Belgium). Chemosphere 43, 909–923.

NRC, 1983. Risk assessment in the Federal Government: Managing the process., Managing.

NRC, 1991. Human Exposure Assessment for Airborne Pollutants: Advances and Oppotunities. The National Academies Press, Washington DC.

NRC, 2012. Exposure Science in the 21st Century. A vision and a strategy. The National Academies Press, Washington DC.

Nuckols, J.R., Ward, M.H., Jarup, L., 2004. Using Geographic Information Systems for Exposure Assessment in Environmental Epidemiology Studies. Environ. Health Perspect. 112, 1007–1015. doi:10.1289/ehp.6738

Nuvolone, D., Maggiore, R., Maio, S., Fresco, R., Baldacci, S., Carrozzi, L., Pistelli, F., Viegi, G., 2011. Geographical information system and environmental epidemiology: a cross-sectional spatial analysis of the effects of traffic-related air pollution on population respiratory health. Environ. Heal. 10, 12. doi:10.1186/1476-069X-10-12

Nyberg, F., Gustavsson, P., Järup, L., Bellander, T., Berglind, N., Jakobsson, R., Pershagen, G., Ja, L., 2000. Urban air pollution and lung cancer in Stockholm. Epidemiology 11, 487–95.

Oak Ridge National Laboratory, 2008. Risk Assessment Information System (RAIS) [WWW Document]. URL http://rais.ornl.gov/

Olivier, J.G.J., Bouwman, A.F., Van Der Hoek, K.W., Berdowski, J.J.M., 1998. Global air emission inventories for anthropogenic sources of NO(x), NH3 and N2O in 1990, in: Environmental Pollution. pp. 135–148. doi:10.1016/S0269-7491(98)80026-2

Ollson, C. a, Knopper, L.D., Whitfield Aslund, M.L., Jayasinghe, R., 2014. Site specific risk assessment of an energy-from-waste thermal treatment facility in Durham Region, Ontario, Canada. Part A: Human health risk assessment. Sci. Total Environ. 466-467, 345–56. doi:10.1016/j.scitotenv.2013.07.019

Özkaynak, H., Baxter, L.K., Dionisio, K.L., Burke, J., 2013. Air pollution exposure prediction approaches used in air pollution epidemiology studies. J. Expo. Sci. Environ. Epidemiol. 23, 566–72. doi:10.1038/jes.2013.15

Ozkaynak, H., Palma, T.E.D., Touma, J.S., Thurman, J., 2008. Modeling population exposures to outdoor sources of hazardous air pollutants. J. Expo. Sci. Environ. Epidemiol. 18, 45–58. doi:10.1038/sj.jes.7500612

Palma-Oliveira, J., Zemba, S., Ames, M., Green, L., Linkov, I., 2012. Uncertainty in multi-pathway risk assessment for combustion facilities. Hum. Ecol. Risk Assess. An Int. J. 18, 501–516. doi:10.1080/10807039.2012.672881

Pandey, R., Shubhashish, K., Pandey, J., 2012. Dietary intake of pollutant aerosols via vegetables influenced by atmospheric deposition and wastewater irrigation. Ecotoxicol. Environ. Saf. 76, 200–208. doi:10.1016/j.ecoenv.2011.10.004

Parodi, S., Baldi, R., Benco, C., Franchini, M., Garrone, E., Vercelli, M., Pensa, F., Puntoni, R., Fontana, V., 2004. Lung cancer mortality in a district of La Spezia (Italy) exposed to air pollution from industrial plants. Tumori 90, 181–185.

Pelucchi, C., Galeone, C., Negri, E., La Vecchia, C., 2010. Trends in adherence to the Mediterranean diet in an Italian population between 1991 and 2006. Eur. J. Clin. Nutr. 64, 1052–6. doi:10.1038/ejcn.2010.158

Peters, S., Vermeulen, R., Cassidy, A., Mannetje, A. 't, van Tongeren, M., Boffetta, P., Straif, K., Kromhout, H., Group, on behalf of the I., 2011. Comparison of exposure assessment methods for occupational carcinogens in a multi-centre lung cancer case–control study. Occup. Environ. Med. 68, 148–153.

Phillips, C., McCarthy, M., Barrowcliffe, R., 2010. Methods for quantitative health impact assessment of an airport and waste incinerator: two case studies. Impact Assess. Proj. Apprais. 28, 69–75. doi:10.3152/146155110X488808

Pirani, M., Rashid, I., Cirilli, C., Federico, M., 2007. Analisi geografica dei tumori diagnosticati in provincia di Modena negli anni 2000-2005 [In Italian]. Modena.

Pope, C.A., Thun, M.J., Namboodiri, M.M., Dockery, D.W., Evans, J.S., Speizer, F.E., Heath, C.W., 1995. Particulate air pollution as a predictor of mortality in a prospective study of U.S. Adults. Am. J. Respir. Crit. Care Med. 151, 669–674. doi:10.1164/ajrccm/151.3_Pt_1.669

Pope III, C., Burnett, R., Turner, M., Cohen, A., Krewski, D., Jerrett, M., Gapstur, S.M., Thun, M.J., 2011. Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure–response relationships. Environ. Health Perspect. 1616, 1616–1621.

Porta, D., Milani, S., Lazzarino, A.I., Perucci, C.A., Forastiere, F., 2009. Systematic review of epidemiological studies on health effects associated with management of solid waste. Environ. Heal. 8, 1–14. doi:10.1186/1476-069X-8-60

Porta, M., 2008. A dictionary of epidemiology, 5th ed. Oxford University Press, New York, NY.

Raaschou-nielsen, O., Andersen, Z.J., Hvidberg, M., Jensen, S.S., Ketzel, M., Sørensen, M., Loft, S., Overvad, K., Tjønneland, A., 2011. Lung Cancer Incidence and Long-Term Exposure to Air Pollution from Traffic. Environ. Health Perspect. 119, 860–865. doi:10.1289/ehp.1002353

Rabinowitz, P.M., Gordon, Z., Holmes, R., Taylor, B., Wilcox, M., Chudnov, D., Nadkarni, P., Dein, F.J., 2005. Animals as Sentinels of Human Environmental Health Hazards: An Evidence-Based Analysis. Ecohealth 2, 26–37. doi:10.1007/s10393-004-0151-1

Ranzi, A., Fano, V., Erspamer, L., Lauriola, P., Perucci, C.A., Forastiere, F., 2011. Mortality and morbidity among people living close to incinerators: a cohort study based on dispersion modeling for exposure assessment. Environ. Heal. 10, 22. doi:10.1186/1476-069X-10-22

Rao, K.S., 2005. Uncertainty Analysis in Atmospheric Dispersion Modeling. Pure Appl. Geophys. 162, 1893–1917. doi:10.1007/s00024-005-2697-4

Reeve, N.F., Fanshawe, T.R., Keegan, T.J., Stewart, A.G., Diggle, P.J., 2013. Spatial analysis of health effects of large industrial incinerators in England , 1998 – 2008: a study using matched case – control areas. BMJ Open 3, e001847. doi:10.1136/bmjopen-2012-001847

Reif, J.S., 2011. Animal sentinels for environmental and public health. Public Health Rep. 126, 50–7.

Reis, M.F., 2011. Solid Waste Incinerators: Health Impacts, in: Encyclopaedia of Environmental. Health. Elsevier B.V., p. 56.

Reisman, D.J., Brady-Roberts, E., 1999. An overview of the US EPA's multiple pathway exposure methodology. Waste Manag. 18, 371–376.

RER, 2011. Regione Emilia Romagna. Servizio Sviluppo dell'Amministrazione Digitale e Sistemi Informativi Geografici.

Rezaie, B., Rosen, M.A., 2012. District heating and cooling□: Review of technology and potential enhancements. Appl. Energy 93, 2–10. doi:10.1016/j.apenergy.2011.04.020

Roberts, R.J., Chen, M., 2006. Waste incineration - how big is the health risk? A quantitative method to allow comparison with other health risks. J. Public Health (Bangkok). 28, 261–266. doi:10.1093/pubmed/fdl037

Rothman, K., Greenland, S., Lash, T., 2008. Modern Epidemiology. Third Edition. Lippincott Williams & Wilkins.

Rotko, T., Oglesby, L., Kunzli, N., Carrer, P., Nieuwenhuijsen, M.J., Jantunen, M., 2002. Determinants of perceived air pollution annoyance and association between annoyance scores and air pollution (PM2.5 , NO2) concentrations in the European EXPOLIS study. Atmos. Environ. 36, 4593–4602.

Rovira, J., Mari, M., Nadal, M., Schuhmacher, M., Domingo, J.L., 2010. Partial replacement of fossil fuel in a cement plant□: Risk assessment for the population living in the neighborhood. Sci. Total Environ. 408, 5372–5380. doi:10.1016/j.scitotenv.2010.07.060

Rückerl, R., Schneider, A., Breitner, S., Cyrys, J., Peters, A., 2011. Health effects of particulate air pollution: A review of epidemiological evidence. Inhal. Toxicol. 23, 555–92. doi:10.3109/08958378.2011.593587

Rugbjerg, K., Harris, M.A., Shen, H., Marion, S.A., Tsui, J.K.C., Teschke, K., Joseph, K.C., 2011. Pesticide exposure and risk of Parkinson ' s disease – a population-based case – control study evaluating the potential for recall bias. Scand J Work Env. Heal. 37, 427–36.

Rydhstroem, H., 1998. No Obvious Spatial Clustering of Twin Births in Sweden between 1973 and 1990. Environ. Res. 76, 27–31.

Sainani, K.L., 2014. Explanatory versus predictive modeling. Phys. Med. Rehabil. 6, 841–4. doi:10.1016/j.pmrj.2014.08.941

Sampaio, C., Aguiar, P., Maurı, J., Miguel, J.P., Pa, O., 2007. Biomonitoring of PCDD/Fs in populations living near portuguese solid waste incinerators: Levels in human milk. Chemosphere 67, 231–237. doi:10.1016/j.chemosphere.2006.05.103

Saner, D., Blumer, Y.B., Lang, D.J., Koehler, A., 2011. Resources, Conservation and Recycling Scenarios for the implementation of EU waste legislation at national level and their consequences for emissions from municipal waste incineration. Resour. Conserv. Recycl. 57, 67–77. doi:10.1016/j.resconrec.2011.09.002

Schiavon, M., Ragazzi, M., Rada, E.C., 2013. A proposal for a diet-based local PCDD/F deposition limit. Chemosphere 93, 1639–45. doi:10.1016/j.chemosphere.2013.08.041

Schuhmacher, M., Domingo, J.L., Garreta, J., 2004. Pollutants emitted by a cement plant: health risks for the population living in the neighborhood. Environ. Res. 95, 198–206. doi:10.1016/j.envres.2003.08.011

Schuhmacher, M., Jones, K.C., Domingo, J.L., 2006. Air-vegetation transfer of PCDD/PCDFs: An assessment of field data and implications for modeling. Environ. Pollut. 142, 143–150. doi:10.1016/j.envpol.2005.09.017

Schuhmacher, M., Meneses, M., Xifrò, A., Domingo, L., 2001. The use of Monte-Carlo simulation techniques for risk assessment: study of a municipal waste incinerator. Chemosphere 43, 787–799.

Seinfeld, J.H., Pandis, S.N., 2006. Atmospheric Chemistry and Physics. From Air Pollution to Climate Change, 2nd ed. John Wiley & Sons., Hoboken, New Jersey.

Shaddick, G., Lee, D., Wakefield, J., 2013. Ecological bias in studies of the short-term effects of air pollution on health. Int. J. Appl. Earth Obs. Geoinf. 22, 65–74. doi:http://dx.doi.org/10.1016/j.jag.2012.03.011

Sheppard, L., Burnett, R.T., Szpiro, A. a, Kim, S.-Y., Jerrett, M., Iii, C.A.P., Brunekreef, B., Pope, C.A., 2012. Confounding and exposure measurement error in air pollution epidemiology. Air Qual. Atmos. Health 5, 203–216. doi:10.1007/s11869-011-0140-9

Shmueli, G., 2010. To Explain or to Predict? Stat. Sci. 25, 289–310. doi:10.1214/10-STS330

Shy, C.M., Degnan, D., Fox, D.L., Mukerjee, S., Milan, J., Boehlecke, B.A., Rothenbacher, D., Briggs, P.M., Devlin, R.B., Wallace, D.D., Stevens, R.K., Philip, A., 1995. Do Waste Incinerators Induce Adverse Respiratory Effects? An Air Quality and Epidemiological Study of Six Communities. Environ. Health Perspect. 103, 714–724.

SIDRIA, 1997. Asthma and respiratory symptoms in 6-7 yr old Italian children: gender, latitude, urbanization and socioeconomin factors. Eur. Respirol. J. 10, 1780–1786. doi:10.1183/09031936.97.10081780

Simeonov, L., Hassanien, M., 2009. Exposure and Risk Assessment of Chemical Pollution: Contemporary Methodology. Springer Netherlands, Dordrecht.

Simonsen, N., Scribner, R., Su, L.J., Williams, D., Luckett, B., Yang, T., Fontham, E.T.H., 2010. Environmental Exposure to Emissions from Petrochemical Sites and Lung Cancer: The Lower Mississippi Interagency Cancer Study. J. Environ. Public Health 2010, 759645. doi:10.1155/2010/759645

Sorahan, T., Gilthorpe, M.S., 1994. Non-differential misclassification of exposure always leads to an underestimate of risk: an incorrect conclusion. Occup. Environ. Med. 51, 839–840. doi:10.1136/oem.51.12.839

Sportisse, B., 2010. Fundamentals in air pollution. From processes to modelling. Springer Science + Buisness Media.

Steenland, K., Armstrong, B., 2006. An overview of methods for calculating the burden of disease due to specific risk factors. Epidemiology 17, 512–9. doi:10.1097/01.ede.0000229155.05644.43

Suter, G.W., 2008. Ecological risk assessment in the United States environmental protection agency: a historical overview. Integr. Environ. Assess. Manag. 4, 285–9. doi:10.1897/IEAM_2007-062

Suter, G.W., Vermeire, T., Munns, W.R., Sekizawa, J., 2005. An integrated framework for health and ecological risk assessment. Toxicol. Appl. Pharmacol. 207, 611–6. doi:10.1016/j.taap.2005.01.051

Suter II, G.W., Vermeire, T., Munns, W.R., Sekizawa, J., 2003. Framework for the Integration of Health and Ecological Risk Assessment. Hum. Ecol. Risk Assess. An Int. J. 9, 281–301. doi:10.1080/713609865

Sweetman, A.J., Alcock, R.E., Wittsiepe, J., Jones, K.C., 2000. Human exposure to PCDD/Fs in the UK: the development of a modelling approach to give historical and future perspectives. Environ. Int. 26, 37–47.

Tango, T., Fujita, T., Tanihata, T., Minowa, M., Doi, Y., Kato, N., Kunikane, S., Uchiyama, I., Tanaka, M., Uehata, T., 2004. Risk of adverse reproductive outcomes associated with proximity to municipal solid waste incinerators with high dioxin emission levels in Japan. J. Epidemiol. 14, 83–93.

Tessari, R., Canova, C., Canal, F., Lafisca, S., Inio, A., Murer, B., Stracca, V., Tollot, M., Simonato, L., 2006. Indagine su inquinamento ambientale da diossine e sarcomi dei tessuti molli nella popolazione di Venezia e Mestre: un esempio di utilizzo di fonti informative elettroniche correnti [In Italian]. Epidemiol. Prev. 30, 191–198.

Tong, S.T.Y., Che, K., Chinese, T., Kong, H., 2000. Home sweet home? A case study of household dust contamination in Hong Kong. Sci. Total Environ. 256, 115–123.

Torfs, R., Hurley, F., Miller, B., Rabl, A., 2007. A set of concentration-response functions. New Energy Externalities Developments for Sustainability (NEEDS). Deliverable 3.7 - RS1b/WP3.

Turner, M.C., Krewski, D., Pope, C.A., Chen, Y., Gapstur, S.M., Thun, M.J., 2008. Long-term Ambient Fine Particulate Matter Air Pollution and Lung Cancer in a Large Cohort of Never-Smokers. Am. J. Respir. Crit. care Med. 184, 1374–1381. doi:10.1164/rccm.201106-1011OC

Turrini, A., Saba, A., Lintas, C., 1991. Study of the italian reference diet for monitoring food constituents and contaminants. Nutr. Res. 11, 861–873. doi:10.1016/S0271-5317(05)80614-1

Tusscher, G.W., Stam, G.A., Koppe, J.G., 2000. Open chemical combustions resulting in a local increased incidence of orofacial clefts. Chemosphere 40, 1263–1270.

Undeman, E., Brown, T., Wania, F., Mclachlan, M.S., 2010. Susceptibility of Human Populations to Environmental Exposure to Organic Contaminants. Environ. Sci. Technol. 44, 6249–6255.

US-EPA, 1992. Framework for ecological risk assessment. EPA/630/R-92/001.

US-EPA, 1997. Exposure Factors Handbook. Washington DC.

US-EPA, 1998a. Guidelines for Ecological Risk Assessment. EPA/630/R-95/002F.

US-EPA, 1998b. Region 6 risk management addendum - draft human health risk assessment protocol for hazardous waste combustion facilities. Report No.:EPA-R6-98-002.

US-EPA, 2001. Risk Assessment Guidance for Superfund□: Volume III - Part A , Process for Conducting Probabilistic Risk Assessment. Washington DC.

US-EPA, 2002. Guidance on Cumulative Risk Assessment of Pesticide Chemicals That Have a Common Mechanism of Toxicity. Washington DC.

US-EPA, 2005a. Human Health Risk Assessment Protocol for Hazardous Waste Combustion Facilities.EPA530-R-05-006. Washington DC.

US-EPA, 2005b. U . S . EPA Response to Comments on the Human Health Risk Assessment Protocol for Hazardous Waste Combustion Facilities. EPA530-R-05-020. Washington DC.

US-EPA, 2008. Integrated Risk Information System (IRIS) [WWW Document]. URL http://www.epa.gov/iris/

Van Leeuwen, F.X.R., Feeley, M., Schrenk, D., Larsen, J.C., Farland, W., Younes, M., 2000. Dioxins: WHO's tolerable daily intake (TDI) revisited. Chemosphere 40, 1095–1101.

Vermeire, T., Munns, W.R., Sekizawa, J., Suter, G., Van der Kraak, G., 2007. An Assessment of Integrated Risk Assessment. Hum. Ecol. Risk Assess. An Int. J. 13, 339–354. doi:10.1080/10807030701226848

Viel, J., Clément, M., Hägi, M., Grandjean, S., Challier, B., Danzon, A., 2008a. Dioxin emissions from a municipal solid waste incinerator and risk of invasive breast cancer□: a population-based case-control study with GIS-derived exposure. Int. J. Health Geogr. 7, 4. doi:10.1186/1476-072X-7-4

Viel, J., Daniau, C., Goria, S., Fabre, P., De, P., Sauleau, E., Empereur-bissonnet, P., 2008b. Risk for non Hodgkin's lymphoma in the vicinity of French municipal solid waste incinerators. Environ. Heal. 7, 51. doi:10.1186/1476-069X-7-51

Viel, J.-F., Arveux, P., Baverel, J., Cahn, J.-Y., 2000. Soft-tissue sarcoma and non-Hodgkin's lymphoma clusters around a municipal solid waste incinerator with high dioxin emission levels. Am. J. Epidemiol. 152, 13–19.

Vienneau, D., de Hoogh, K., Beelen, R., Fischer, P., Hoek, G., Briggs, D., 2010. Comparison of land-use regression models between Great Britain and the Netherlands. Atmos. Environ. 44, 688–696. doi:10.1016/j.atmosenv.2009.11.016

Vinceti, M., Malagoli, C., Fabbi, S., Teggi, S., Rodolfi, R., Garavelli, L., Astolfi, G., Rivieri, F., 2009. Risk of congenital anomalies around a municipal solid waste incinerator: a GIS-based case-control study. Int. J. Health Geogr. 8, 8. doi:10.1186/1476-072X-8-8

Vinceti, M., Malagoli, C., Teggi, S., Fabbi, S., Goldoni, C., Girolamo, G. De, Ferrari, P., Astolfi, G., Rivieri, F., Bergomi, M., 2008. Adverse pregnancy outcomes in a population exposed to the emissions of a municipal waste incinerator. Sci. Total Environ. 407, 116–121. doi:10.1016/j.scitotenv.2008.08.027

Walpole, S.C., Prieto-merino, D., Edwards, P., Cleland, J., Stevens, G., Roberts, I., 2012. The weight of nations: an estimation of adult human biomass. BMC Public Health 12, 439. doi:10.1186/1471-2458-12-439

Walraven, C. Van, Hart, G., 2008. Leave'em Alone - Why Continuous Variables Should Be Analyzed as Such. Neuroepidemiology 9, 138–139. doi:10.1159/000126908

Wang, M., Beelen, R., Eeftens, M., Meliefste, K., Hoek, G., Brunekreef, B., 2012. Systematic Evaluation of Land Use Regression Models for NO2. Environ. Sci. Technol. 46, 4481–4489.

Weiland, S.K., Mundt, K.A., Rückmann, A., Keil, U., 1994. Self-reported wheezing and allergic rhinitis in children and traffic density on street of residence. Ann. Epidemiol. 4, 243–247. doi:10.1016/1047-2797(94)90103-1

Weinberg, C.R., Umbach, D.M., 1994. When Will Nondifferential Misclassification of an Exposure Preserve the Direction of a Trend? Am. J. Epidemiol. 140, 565–571.

Whitsel, E.A., Quibrera, P.M., Smith, R.L., Catellier, D.J., Liao, D., Henley, A.C., Heiss, G., 2006. Accuracy of commercial geocoding: assessment and implications. Epidemiol. Perspect. Innov. 3, 8. doi:10.1186/1742-5573-3-8

WHO, 2001. Report on Integrated Risk Assessment. WHO/IPCS/IRA/10/12. Geneva, Switzerland.

WHO, 2013a. Review of Evidence on Health Aspects of Air Pollution–REVIHAAP Project. Copenhagen.

WHO, 2013b. Global Environment Monitoring System - Food Contamination Monitoring and Assessment Programme (GEMS/Food).

WHO, 2014. Ambient (outdoor) air pollution in cities database 2014 [WWW Document]. URL http://www.who.int/phe/health_topics/outdoorair/databases/cities/en/ (accessed 6.4.14).

WHO Europe, 2000. Air Quality Guidelines for Europe. Second edition. European series N.91. WHO Regional Office for Europe, Copenhagen.

WHO Europe, 2003. Health Aspects of Air Pollution with Particulate Matter , Ozone and Nitrogen Dioxide.

WHO Europe, 2006. Air Quality Guidelines for europe. Global update 2005. WHO Regional Office for Europe, Copenhagen.

WHO Europe, 2007. Population health and waste management: scientific data and policy options. Copenhagen.

Williams, F.L.R., Lawson, A.B., 1992. Low Sex Ratios of Births in Areas at Risk from Air Pollution from Incinerators , as Shown by Geographical Analysis and 3-Dimensional Mapping. Int. J. Epidemiol. 21, 311–319.

Willis, A.W., Evanoff, B.A., Lian, M., Galarza, A., Wegrzyn, A., Schootman, M., Racette, B.A., 2010. Metal Emissions and Urban Incident Parkinson Disease: A Community Health Study of Medicare Beneficiaries by Using Geographic Information Systems. Am. J. Epidemiol. 172, 1357–1363. doi:10.1093/aje/kwq303

Wilson, J.G., Kingham, S., Pearce, J., Sturman, A.P., 2005. A review of intraurban variations in particulate air pollution: Implications for epidemiological research. Atmos. Environ. 39, 6444–6462. doi:10.1016/j.atmosenv.2005.07.030

Wong, D.W., Yuan, L., Perlin, S. a, 2004. Comparison of spatial interpolation methods for the estimation of air quality data. J. Expo. Anal. Environ. Epidemiol. 14, 404–15. doi:10.1038/sj.jea.7500338

Wu, X., Bennett, D.H., Lee, K., Cassady, D.L., Ritz, B., Hertz-picciotto, I., 2011. Longitudinal variability of time-location/activity patterns of population at different ages□: a longitudinal study in California. Environ. Heal. 10, 80. doi:10.1186/1476-069X-10-80

Wynder, E.L., 1994. Investigator bias and interviewer bias: The problem of reporting systematic error in epidemiology. J. Clin. Epidemiol. 47, 825–827.

Yegnan, A., Williamson, D.G., Graettinger, A.J., 2002. Uncertainty analysis in air dispersion modeling. Environ. Model. Softw. 17, 639–649.

Yu, H.L., Chen, J.C., Christakos, G., Jerrett, M., 2009. BME estimation of residential exposure to ambient PM10 and ozone at multiple time scales. Environ. Health Perspect. 117, 537–544. doi:10.1289/ehp.0800089

Zambon, P., Ricci, P., Bovo, E., Casula, A., Gattolin, M., Fiore, A.R., Chiosi, F., Guzzinati, S., 2007. Sarcoma risk and dioxin emissions from incinerators and industrial plants: a population-based case-control study (Italy). Environ. Heal. 6, 19. doi:10.1186/1476-069X-6-19

Zandbergen, P.A., 2007. Influence of geocoding quality on environmental exposure assessment of children living near high traffic roads. BMC Public Health 7, 37. doi:10.1186/1471-2458-7-37

Zandbergen, P.A., 2009. Geocoding Quality and Implications for Spatial Analysis. Geogr. Compass 3, 647–680.

Zartarian, V., Ott, W., Duan, N., 1997. A quantitative definition of exposure and related concepts. J. Expo. Anal. Environ. Epidemiol. 7.

Zhang, Y.X., Tao, S., 2009. Global atmospheric emission inventory of polycyclic aromatic hydrocarbons (PAHs) for 2004. Atmos. Environ. 43, 812–819. doi:10.1016/j.atmosenv.2008.10.050

Zmirou, D., Parent, B., Potelon, J.L., 1984. Epidemiologic study of the health effects of atmospheric waste from an industrial and household refuse incineration plant. Rev. depidemiologie sante publique 32, 391–397.

Zou, B., Wilson, J.G., Zhan, F.B., Zeng, Y., Benjamin, F., 2009. Air pollution exposure assessment methods utilized in epidemiological studies. J. Environ. Monit. 11, 475–490. doi:10.1039/b813889c

APPENDIX A – Literature review on incinerators

Table S1 - Reviewed studies, classified by exposure method (see Table 1 in the paper). See paper bibliography for references.

Study	Outcome	Country	Population	Industrial sources	Exposure measure in the model	Other exposure information	Covariates	Class
(Zmirou et al., 1984)	Consumption of medicines for respiratory symptoms	France		1 industrial incinerator	Distance of the area of residence from the plant (200,1000,2000 m)			**2.1.1**
(Lloyd et al., 1988)	Twinning frequency	United Kingdom	258 mothers 470 twins	1 MSWI 1 CWI	Residence in postcode sectors at low VS medium/high risk (qualitatively based on wind direction and distance)	Polychlorinated hydrocarbons in soil	Maternal age	**1.1.1**
(Hsiue et al., 1991)	Pulmonary function in children	Taiwan	382 children	Open wire reclamation incineration	Children from schools in 3 impacted areas (qualitatively defined) VS 1 school in control area.	Measures of NO_2 and SO_2 , higher in impacted areas.	Questionnaire (passive smoking)	**1.1.1**
(Williams et al. 1992)	Sex ratio	United Kingdom	3,576 births	2 incinerators	Birth in a postcode district classified as at risk VS low risk (qualitatively, based on wind direction, topography, people perception and soil PAH concentrations)			**1.1.1**
(Elliott et al. 1992)	Larynx and lung cancer	United Kingdom		10 incinerators of solvents and olis	Distance of postcode zone of residence from the sources, observed VS expected in circles of 0-3 and 3-10 km		Socioeconomic status	**2.2.1**

(Gray et al. 1994)	Childhood asthma or allergy	Australia	713 exposed, 626 control children	High temperature sludge burning	Two regions close to the incinerators VS one region without incinerators	air pollution measures of NO_2,, SO_2, HF e PM (no differences between areas)		**1.1.1**
(Barbone et al. 1995)	Lung cancer mortality	Italy	755 cases, 755 controls	1 incinerator, 1 iron foundry, shipyard, city centre	Distance of the last residence from the industrial sources + wind direction	PM deposition measurement in 28 stations + PM concentrations from 4 stations	Questionnaire (smoke, occupation)	**2.3.1**
(Shy et al. 1995)	Chronic and acute respiratory effect	North Carolina, US	6,963 participants	1 MSWI, 1 BWI 1 HWI	Residence in a community within 2x5 km ellipse (oriented along principal wind direction) from the incinerator VS upwind outside 5 km	12h pollution measurement in each community, CMB estimation of incinerator contribution	Questionnaire	**2.1.1**
(Biggeri et al. 1996)	Lung cancer	Italy	755 cases, 755 controls	1 incinerator, 1 iron foundry, shipyard, city centre	Distance and direction between the last residence and the sources	PM concentrations	Smoke, occupation, PM concentrations	**2.3.1**
(Elliott et al. 1996)	Cancer incidence	United Kingdom	14 million people	72 MSWI	Distance from postcode zone of residence, observed VS expected in 0-3km, 0-7.5 km,0.5-1-2-3-4.6-5.7-6.7-7.5 km radius buffers		Deprivation score	**2.2.1**
(Michelozzi et al. 1998)	Cancer mortality	Italy	341,389 residents	1 landfill, 1 incinerator, 1 refinery	Distance between centroid of census block and centroid of the industrial area, SMR for 3-8-10km buffers + 1km buffers		Deprivation score (census block)	**2.2.1**

(Rydhstroem et al. 1998)	Twin deliveries	Sweden	17,067 twin deliveries	14 incinerators	Expected VS observed twinning in municipalities before/after the activation of incinerators			**1.1.1**
(Lee et al. 1999)	Pulmonary function	North Carolina, US	480 residents	1 MSWI 1 BWI 1 HWI	Residence in a community within 2miles from the incinerator VS outside 2 miles + Longitudinal analysis with PM_{10}	24h PM_{10} concentration in each community (no differences)	Questionnaire	**2.1.1**
(Tusscher et al. 2000)	Orofacial cleft in children	Nederland	8,803 children	1 chemical open burning site	Pregnancies at a clinic near the burning site VS pregnancies at a clinic 12km far	Mother's residences dispersion around the burning site (qualitative). Tons of waste burned. Some measures of environmental contamination		**1.1.1**
(Mohan et al. 2000)	Respiratory symptoms	North and South Carolina, US	4,025 subjects	1 MSWI 1 BWI 1 HWI 1 IWI	4 communities within 2x5 (or 8x5) km ellipse VS 4 communities 5 to 25 km upwind from exposed communities	PM_{10} average concentration per community	Questionnaire (demographic, house characteristic)	**2.1.1**
(Knox, 2000)	Childhood cancer	United Kingdom	22,458 children	70 MSWI 307 BWI 460 hazard waste landfills	Distance (at birth VS at death) of the postcode of residence from sources			**2.2.1**
(Viel et al. 2000)	Soft tissue sarcoma Non Hodgkin Lymphoma	France	110 STS 803 NHL, 176 HL	1 incinerator	Cluster analysis based on residence in a "canton"	Information on dioxin emissions and concentrations in milk		**2.1.1**

(Hu et al. 2001)	Pulmonary function	North Carolina, US	10,187 subjects	1 MSWI 1 BWI 1 HWI	Residence in a comunity with incinerator VS with no incinerator (>5km) + distance from the address of residence to the incinerator + exposure index (function of distance, direction, n. days downwind, n. hours outside home)		Questionnaire (smoking, house characteristics)	**2.3.1**
(Comba et al. 2003)	Soft tissue sarcoma	Italy	37 STS cases 171 controls	1 IWI 1 oil refinery, 1 paper and 1 chemical industry	Distance of the main residence from the incinerator, buffer 1 km (<2, 3, 4, 5, >5km)			**2.3.1**
(Dummer et al., 2003)	Adverse pregnancy outcomes	United Kingdom	244,758 births	3 crematorium, 4 incinerators	Distance of mother residence to the plants		Social class, year, birth order, multiple births	**2.2.1**
(Floret et al. 2003)	Non Hodgkin Lymphoma	France	225 cases (match 10:1)	1 MSWI	PCDD/F average atmospheric concentrations from dispersion modelling (APC3 Gaussian), 4 classes of exposure (very low to high)	Dioxin concentrations in soil [32]	Education, occupation, social class and household-based indicators	**3.3.1**
(Fukuda et al. 2003)	Mortality from major diseases	Japan	803 municipalities		Residence in a municipality with an incinerator VS without. Indices of incineration activity at municipality level derived from emission inventory.		7 socio-economic indicators	**1.1.1**

(Cordier et al. 2004)	Congenital anomalies	France	2,872 communities with less than 50'000 inhabitants	70 incinerators	PCDD/F concentrations in most populate areas from dispersion model (POLAIR, Gaussian), multiplied by the num. of years of functioning of the plant. 3 classes of exposure		Pop. density and income at community level, local traffic, presence of industries, mother age	**3.1.1**
(Parodi et al. 2004)	Lung cancer mortality	Italy	118749 residents	1 waste incinerator, 1 power plant, 1 lead oxide industry	Residence in industrial area VS residence in urban/semi-urban areas	Measures of heavy metals in lichens higher in industrial areas	Age, deprivation index	**1.1.1**
(Tango et al. 2004)	Adverse reproductiv e outcomes	Japan	225,215 live births, 3,387 foetal deaths, 835 infant deaths	63 incinerators	Distance of maternal residence from incinerators, 1 km buffers	Measurement of dioxin in soil for two plants	Stratification by vital statistics (mother age, weight, father employment)	**2.3.1**
(Miyake et al., 2005)	Allergic disorders and general symptoms	Japan	450,807 children in 996 schools	37 MSWI	Distance from the school to the nearest incinerator		Socioeconomic status, healthcare access in the municipality	**2.1.1**
(Biggeri et al. 2005)	Soft tissue sarcoma, Non Hodgkin Lymphoma	Italy		1 MSWI	Residence in the municipality with the incinerator VS residence in other municipalities 80 km around		Age	**1.1.1**
(Bianchi et al. 2006)	Non Hodgkin Linfoma	Italy	1,830 NHL	25 incinerators	Residence in a municipality with an incinerator VS residence in municipalities 30-50 km around		Age	**1.1.1**

(Biggeri and Catelan 2006)	Non Hodgkin Lymphoma	Italy	1,119 NHL	17 incinerators	Residence in a municipality with an incinerator VS residence in municipalities 50-80 km around		**1.1.1**
(Tessari et al. 2006)	Neoplasias	Italy	188 sarcomas, 134 HL, 774 NHL, 24,184 cancers	Various industrial sources of PCDD/F (3 incinerators)	Concentration of PCDD/F at the residence address from a model for all sources in the area. 4 exposure categories.	Sex, age	**3.3.1**
(Zambon et al. 2007)	Sarcoma	Italy	171 cases 405 controls	26 incinerators 7 industrial plants	PCDD/F dispersion model (ISC3, Gaussian,), average concentration weighted by length of each residence. Reconstruction of 40 years of emission data.		**3.3.2**
(Viel et al. 2008)	Breast cancer	France	434 cases 2,170 controls	1 MSWI	PCDD/F dispersion model (APC3, Gaussian) concentration at census block. 4 classes of exposure	Socioeconomic at census level	**3.2.1**
(Viel et al. 2008b)	Non Hodgkin Lymphoma	France	3974 NHL cases	13 incinerators	PCDD/F dispersion model (ADMS3 Gaussian), cumulative deposition to soil, median of all receptors in each census block	Pop. density, urbanisation, socio-economic level, traffic and industrial pollution	**3.2.1**

Vinceti et al. (2008)	Abortion and congenital anomalies	Italy	23 cases	1 MSWI	PCDD/F dispersion model (ISC3, Gaussian + Spray, Lagrangian), concentration at mother address and place of work, 2 classes of increasing exposure			**3.3.1**
(Goria et al. 2009)	Cancer	France	135,000 cases	16 incinerators	PCDD/F dispersion model (ADMS3, Gaussian) cumulative depositions to soil, median of receptors at census block level.		Pop. density, rural/urban indicator, socio-economic score, road traffic pollution and other polluting industries	**3.2.1**
(Vinceti et al., 2009)	Congenital anomalies	Italy	228 cases	1 MSWI	PCDD/F and heavy metal dispersion modelling (Windimula, Gaussian), concentration at mother address. 3 classes of exposure.		Mother education and age	**3.3.1**
(Cordier et al. 2010)	Urinary tract birth defect	France	324 cases 226 controls	21 MSWI	PCDD/F and heavy metal dispersion modelling (ADMS3), concentration and deposition at home address, 3 classes of exposure.	Consumption of local foods enhance risk in the most exposed areas while lower risk in less exposed areas	Deprivation index, pop.density, questionnaire, other pollution sources at municipality level, old incinerators	**3.3.1**
(Federico et al. 2010)	Cancer incidence	Italy	16,443 cancer cases	1 MSWI	Distance between residence census block centroid and incinerator's census block centroid (2-3.5-5 km buffer). Cluster analysis		Deprivation index at census block	**2.2.1**

(Ranzi et al. 2011)	Mortality and morbidity	Italy	31'374 individuals	1 MSW incinerator	Heavy metal dispersion model (ADMS Urban, Gaussian), concentrations at the geocoded address of residence, 4 classes of exposure plus analysis population inside 3.5 km buffer VS Regional population	Measurement of various pollutants in air and soil determined heavy metals as preferred indicator of MSWI impact	Dispersion model for pollution from other sources (industry, traffic, heating), deprivation index at census block level	**3.3.1**
(Garcia-Pèrez et al. 2013)	Cancer mortality	Spain	8098 municipalities	129 installations for waste treatment 14 incinerators	Distance from the town centre to the facility. Near (<5km) VS far analysis.		Socioeconomic status at municipality level	**2.1.1**
(Reeve et al. 2013)	Cancer and all cause mortality	United Kingdom		5 incinerators (> 150'000)	Distance between the incinerator and the centroid of census unit (LSOA) of residence. Comparison of exposed (<10km) and control (>10 km) areas.		Age, sex, deprivation index at census level	**2.2.1**
(Candela et al. 2013)	Adverse pregnancy outcome	Italy	21,517 births between 2003 and 2010	1 MSW incinerator	Dispersion model for PM10 (ADMS Urban, Gaussian), weighted average of monthly concentrations at the geocoded address of residence over the 9 month of pregnancy, 4 classes of exposure		Dispersion model for pollution from other sources (industry, traffic, heating), maternal characteristics (age, country, education, socioeconomic status, hospitalization)	**3.3.2**

ABBREVIATIONS

MSWI = municipal solid waste incinerator; CWI = chemical waste incinerator; PAH = polycyclic aromatic hydrocarbons; NO_2, SO_2 = nitrous and sulphur dioxide; HF = fluoridic acid; PM = particulate matter; BWI = biomedical waste incinerator; HWI = hazardous waste incinerator; CMB = chemical mass balance model; IWI = industrial waste incinerator; PCDD/F = dioxins and furans; LSOA = Lower Layer Super Output Area

APPENDIX B - Risk assessment

Health risk assessment for criteria pollutants

To estimate the number E of cases per unit of time (case year^{-1}) due to a projected increase in air pollution we use the following equation (Forastiere et al., 2012; Kunzli et al., 2000; Martuzzi et al., 2006):

$$E = I_0 \cdot \Delta_{\%} I \cdot \Delta C^{+} \cdot P \qquad \text{(D.1)}$$

where:

- I_0 is the baseline incidence of the pathology of interest in the unexposed population (case per person-year),
- $\Delta_{\%} I$ is the % change in the baseline incidence rate due to an unitary change in atmospheric concentration ([μg m^{-3}]$^{-1}$),
- ΔC^{+} is the projected increase in atmospheric concentration (μg m^{-3}),
- P is the exposed population (number of people).

We obtained values of $\Delta_{\%} I$ from the epidemiological literature (Martuzzi et al., 2006, Torfs et al., 2007), calculated as relative risks (RR = incidence rate between exposed and unexposed) for a concentration increase of 10 μg m^{-3} of pollutant.

When DRFs are reported in terms of relative risk (RR = incidence rate between exposed and unexposed) for a concentration increase of 10 μg m^{-3} of pollutant, then $\Delta_{\%} I = RR\text{-}1$ and E is computed as follows:

$$E = I_0 \cdot [(RR\text{-}1) \cdot (\Delta C^{+} / 10)] \cdot P \qquad \text{(D.2)}$$

Our analysis retained two assumptions typically made in the epidemiological literature: first, dose-response relationships are linear, at least over limited ranges of exposure concentrations (Kunzli et al., 2000; Ostro, 2004; Pope III et al., 2011). Second, there are no positive or negative interactions between pollutants in terms of their health consequences. Consequently, in case of a reduction in atmospheric concentrations, following Martuzzi et al. (2006), first the baseline incidence rate in the unexposed I_0 is computed as follows:

$$I_0 = I_E / [1 + (RR\text{-}1) \cdot (\Delta C^{-} / 10)] \qquad \text{(D.3)}$$

where:

- I_E is the incidence at the exposed level
- ΔC^{-} is the reduction in air concentration

and then the number of attributable cases is computed with Eq.(2) by setting $\Delta C^{+} = \Delta C^{-}$.

Health risk assessment for micropollutants

For this group of pollutants, we referred to the HHRAP model proposed by the United States Environmental Protection Agency (EPA, 2005a).

Table S2 reports a more detailed reference for the specific equations of the model.

TABLE S2 – References for model equation

Equations	**Reference in EPA (2005a)**
COPC concentration in soil	Section 5.2, Eq. 5-1D/5-1E
COPC concentration in produce	Section 5.3, Eq. 5-14/5-18/5-20A/5-20B
COPC concentration in beef and dairy products	Section 5.4, Eq. 5-22/5-24
COPC concentration in pork	Section 5.5, Eq. 5-25
COPC concentration in chicken and eggs	Section 5.6, Eq. 5-26
Inhalation exposure	Section 6.1, Table B-5-1
Soil ingestion exposure	Section 6.2, Table C-1-1
Food ingestion exposure	**See discussion below**
Cancer risk calculation	Section 7.1, Table C-1-7/C-1-9/C-1-12
Noncancer hazard calculation	Section 7.2, Table C-1-8/C-1-11

As described in section 2.2.1 of the article, we modified the equations for computing the total daily intake I_{tot} (mg kg_{weight}^{-1} day^{-1}) of a specific contaminant due to food ingestion, both for farmed animals and for humans. We computed I_{tot} as follows:

$$I_{tot}=\sum_i (F_{loc\text{-}i}\cdot P_{loc\text{-}i}+F_{avg\text{-}i}\cdot P_{avg\text{-}i})\cdot CR_i \; I_{tot}=\sum_i (Floc_i\cdot Ploc_i+ Favg_i\cdot Pavg_i)\cdot CR_i \quad \text{(D.4)}$$

where:

- $Itot$ is the total daily intake due to food ingestion (mg kg_{weight}^{-1} day^{-1})
- $Floc_i$ is the fraction of i-th food that is home-grown, i.e. produced in the cell of residence (%)
- $Ploc_i$ is the concentration of a pollutant in the home-grown food i (mg kg_{food}^{-1})
- $Favg_i$ is the fraction of i-th food that is not home-grown but come from the study area (%)
- $Pavg_i$ is the average concentration of pollutant in the food i in the agricultural areas (mg kg_{food}^{-1})
- CR_i is the consumption rate for food i (kg_{food} kg_{weight}^{-1} day^{-1}).

The categories i of food accounted for in the EPA model (2005a) for human ingestion are: aboveground exposed, aboveground protected vegetables (i.e., fruits or vegetables with non-edible protecting covering, such as peas, corn or melons), belowground vegetables, beef, milk and derivatives, pork, poultry and eggs. Here, we did not consider the consumption of contaminated fish and water, due to the specific characteristic of the area.

For each food category i, $Ploc$ was computed following equations reported in EPA (2005a, section 5), while $Pavg$ was computed as a weighted average on the study area:

$$Pavg \;\; =\sum_j (Ploc_j\cdot A_j)/\sum_j A_j \; P_{avg}=\sum_j (P_{loc\text{-}j}\cdot A_j)/\sum_j A_j \quad \text{(D.5)}$$

where:

- $Ploc_j$ is the concentration of pollutant (mg kg_{food}^{-1}) in home-grown food produced in the cell j
- A_j is the weighting factor that account for the j-cell's importance in determining the contamination of food and depends on the food category i.

For vegetables, A was here assumed to be the number of hectares cultivated in a cell, while for animal products the tons of cattle bred in a cell.

TABLE S3 – Exposure-Response functions (ERF) for particulate matter (referred to an increase of 10 µg m^{-3}). Some ERF are expressed as relative risk (RR) while other as increase in number of events.

Sanitary Outcome	Dose-Response function (CI 95 %)	Exp. measure	Age	Primary reference [a]
All causes mortality	RR [b] = 1.06 (CI95% : 1.02 – 1.11)	$PM_{2.5}$	≥ 30 years	Pope *et al*. 2002
Lung cancer	RR = 1.08 (CI95% : 1.01 – 1.16)	$PM_{2.5}$	≥ 30 years	Pope *et al*. 2002
Infraction	RR = 1.18 (CI95% : 1.14 – 1.23)	$PM_{2.5}$	≥ 30 years	Pope *et al*. 2002
Stroke	RR = 1.02 (CI95% : 0.95 – 1.10)	$PM_{2.5}$	≥ 30 years	Pope *et al*. 2002
Acute bronchitis	26.5 (CI 95%: -1.9 / 54.1) cases per year per 100'000 people	PM_{10}	≥ 27 years	Abbey *et al*. 1993 Abbey *et al*. 1994
Asthma in children	180 (CI 95% : -690 / 1060) extra days of bronchodilator usage per year per 1000 children	PM_{10}	6-7 and 13-14 years	Anderson *et al*. 2004
Asthma in adults	912 (CI 95%: -912 / 2774) extra days of bronchodilator usage per year per 1000 adults	PM_{10}	≥ 15 years	Anderson *et al*. 2004
Restricted Activity Days (RADs)	902 (CI 95%: 792 / 1013) RADs per year per 1000 adults	$PM_{2.5}$	15-64 years	Ostro 1987
Work Lost Days (WLDs)	222 (CI 95%: 188 / 256) WLDs per year per 1000 adults	$PM_{2.5}$	15-64 years	Ostro 1987
Minor Restricted Activity Days (MRADs)	577 (95% CI: 468 / 686) MRADs per year per 1000 adults	$PM_{2.5}$	18-64 years	Ostro e Rothschild 1989
Lower Respiratory Symptoms (LRS) in children	1.86 (CI 95%: 0.92 / 2.77) extra days with symptoms per child per year	PM_{10}	5 – 14 years	Ward e Ayres 2004
Lower Respiratory Symptoms (LRS) in adults	1.30 (CI 95%: 0.15 / 2.43) extra days with symptoms per child per year	PM_{10}	≥ 15 years	Hurley *et al*. 2005

[a] Reference details available in Martuzzi et al. (2006).

[b] RR = Relative Risk

TABLE S4 – Dose-response functions (DRF) for micropollutants as reported in the Integrated Risk Information System (IRIS) (EPA, 2008) and the Risk Assessment Information System (RAIS) (Oak Ridge National Laboratory, 2008).

	Carcinogenetic effect		Toxic effect	
COC	URF m^3 μg^{-1}	CSF kg day mg^{-1}	RfC mg m^{-3}	RfD mg kg^{-1} day^{-1}
Cd	1.8E-03	-	-	1.0E-03
PCDD	3.3E+01	1.5E+05	-	-
PAH	1.1E-03	7.3E+00	-	-
Hg^{2+}	-	-	1.1E-03	3.0E-04
Hg^0	-	-	3.0E-04	-
MeHg	-	-	-	1.0E-04

URF = unitary risk factor for inhalation, CSF = cancer slope factor for ingestion, RfC = reference concentration for inhalation, RfD = reference dose for ingestion

TABLE S5 – Assumptions about chemical partitioning of pollutants at incinerator's stack exit.

COCP	Gas phase (% total emission)	Particle phase (% total emission)	Particle properties	Gas properties
PCDD/F	66%	34%		Df^c = 0.104 cm^2 sec^{-1}
PAH	29%	70%		Df= 0.043 cm^2 sec^{-1}
Cd	1%	99%	Dp^a =2.5 g cm^{-3}	Df= 0.077 cm^2 sec^{-1}
Hg			Gd^b = 0.48 µm	Df= 0.011(Hg^0) cm^2 sec^{-1}
	20%(Hg^0) 60%(Hg^{2+})	20%(Hg^{2+})		Df= 0.045 (Hg^{2+}) cm^2 sec^{-1}

[a] Particle density, [b] Geometric mass mean diameter, [c] Gas diffusivity

APPENDIX C – Sensitivity analysis for the simulation study of Chapter 4

To test if the results of the simulation presented in Chapter 4 remain valid when using different criteria for categorization of exposure variables, I developed a second simulation using some *a priori* defined classes for the exposure variables ADCO, ADDE and ADDI.

The design of the simulation remained the same described in Chapter 4. The only difference was in the cut-offs values chosen for categorization. Instead of quintiles of the distributions, the exposure variables were categorized as follows:

- 0.3-0.8-1.6-3.2-6.4-25.2 ng m^{-3} for address concentration (ADCO.arb)
- 21-40-60-150-300-2114 ng m^2 h^{-1} for address deposition (ADDE.arb)
- 0.8-1.6-2.4-3.2-4.0 and 1.5-2.5-3.0-3.5-4 km for address distance (respectively ADDI.rego, ADDI.arb)

In the *real risk* scenario the reference exposure method became ADCO.arb, while in the *apparent risk* scenario the reference method is ADDI.arb.

The results of the simulations are reported in Figure 45 for the *real risk* scenario and Figure 46 for the *apparent risk* scenario. The simulations confirmed a trend similar to that highlighted in Chapter 4, with the distance measure being unable to measure the *real risk* and atmospheric concentrations not confirming the *apparent risk* associated with distance. The use of arbitrary chosen classes, with very different population sizes in each class, makes the results more unstable. This is especially true for the regular classification of distance (ADDI.reg), which have few people in the reference (0) and most exposed (4) classes.

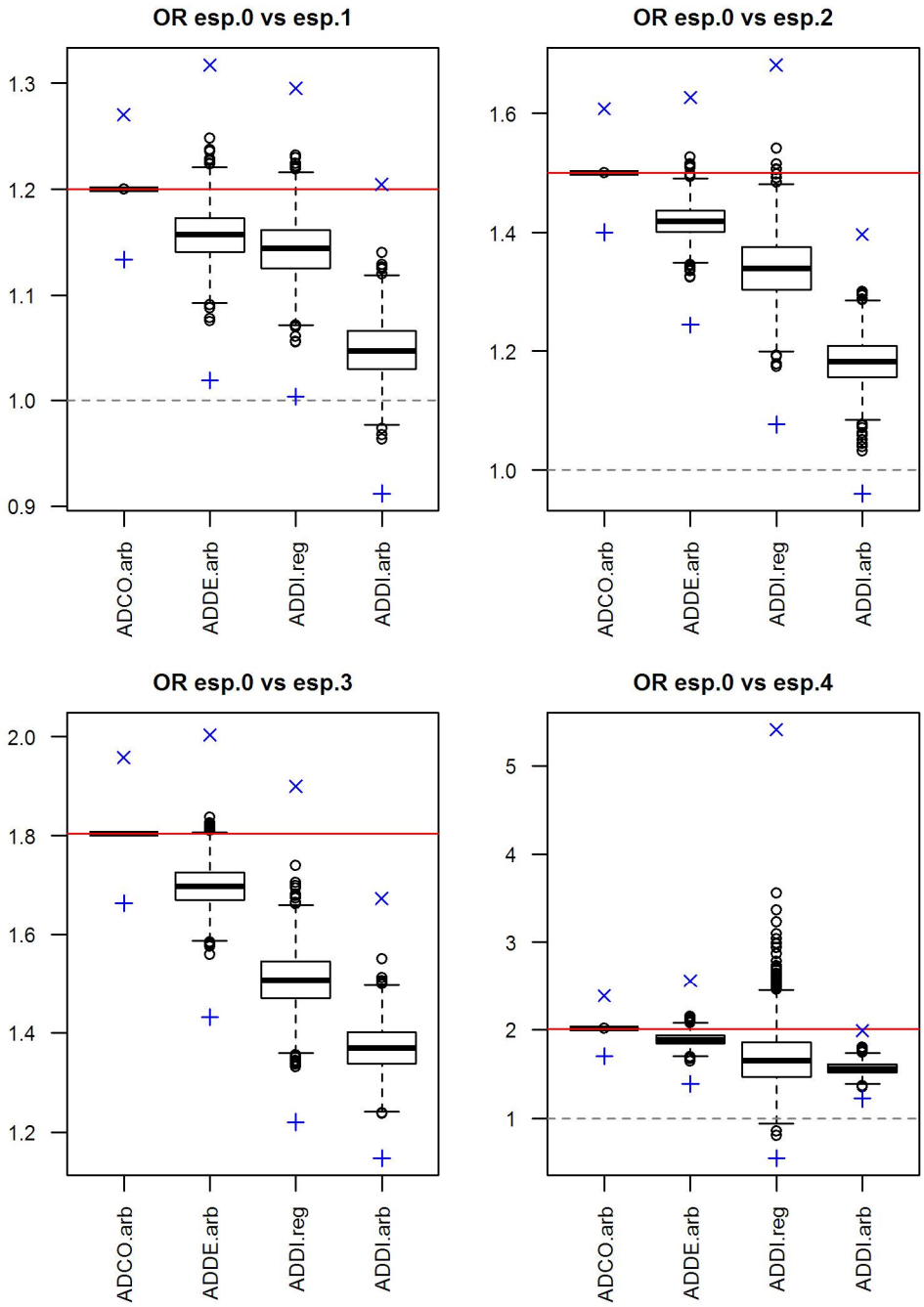

Figure 45 - Distributions of simulated Odds Ratios (OR) using different exposure assessment methods in the *real risk* scenario. The box represent the inter-quartile range (IQR) of the distribution, the horizontal line inside the box is the median value, the whiskers extend to 1.5 times the IQR from the box .The red line represent the predefined OR based on ADCO.reg exposure. The dashed line represent the "no effect" value. The blue markers represent the minimum value of the lower bound confidence interval (-) and the maximum value of the upper bound confidence interval (+) for the calculated ORs.

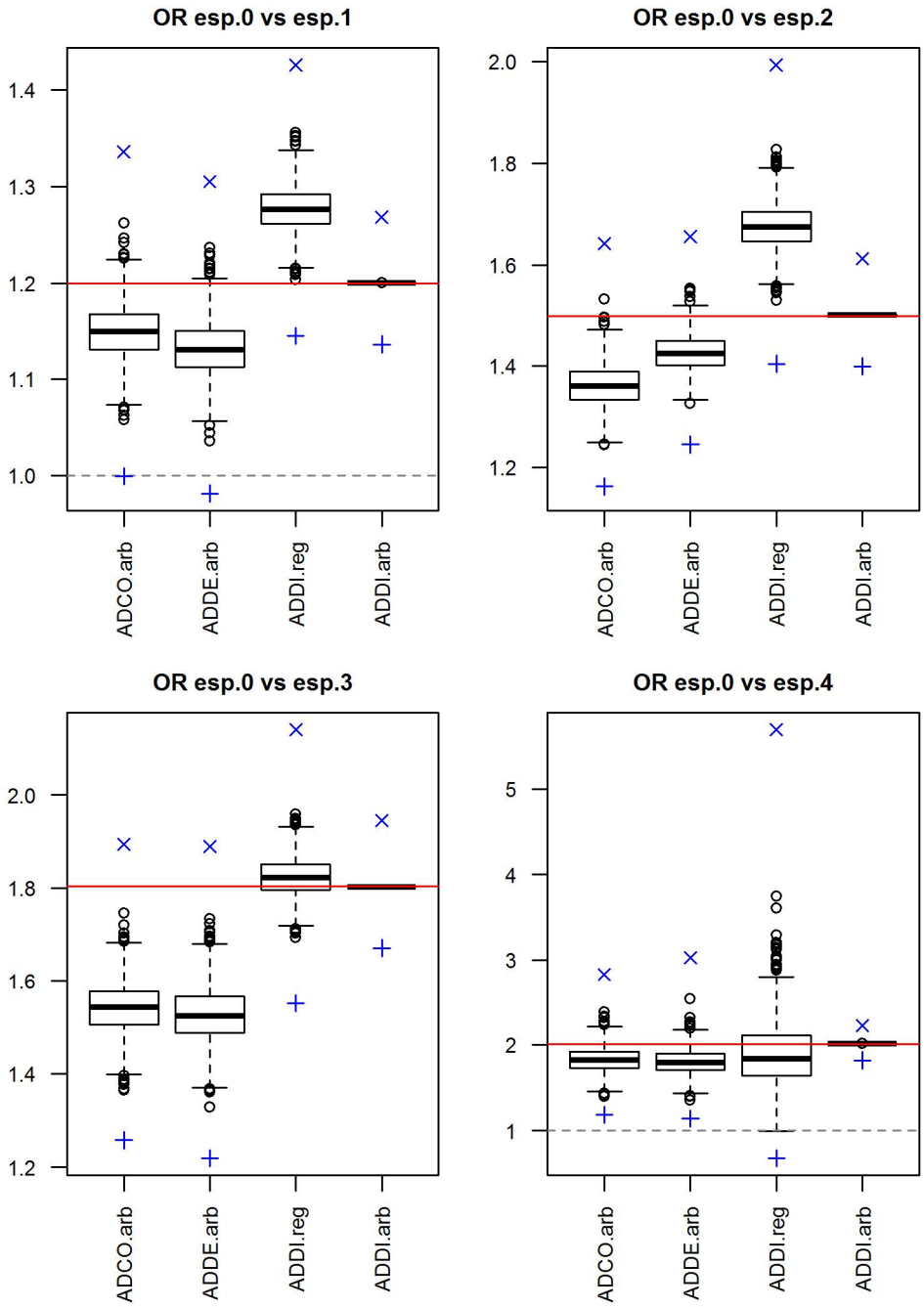

Figure 46 - Distributions of simulated Odds Ratios (OR) using different exposure assessment methods in the *apparent risk* scenario. The box represent the inter-quartile range (IQR) of the distribution, the horizontal line inside the box is the median value, the whiskers extend to 1.5 times the IQR from the box . The red line represent the predefined OR based on ADDI.arb exposure. The dashed line represent the "no effect" value. The blue markers represent the minimum value of the lower bound confidence interval (-) and the maximum value of the upper bound confidence interval (+) for the calculated ORs.