THE USE OF GIS IN ASSESSING EXPOSURE TO AIRBORNE POLLUTANTS

Emilie Stroh
To my Parents
Contents

Contents ........................................................................................................... V
Papers .............................................................................................................. VII
Populärvetenskaplig sammanfattning ........................................................... IX
Aim............................................................................................................... XIII
  Specific aims ................................................................................................ XIII
  Structure of this thesis ................................................................................ XIII
Part I – GIS in the Assessment of Exposure to Air Pollution.................... 1
  Emission Sources....................................................................................... 3
    Air pollutants considered in these studies ........................................... 3
    Anthropogenic sources ........................................................................ 5
  Emission sources considered in these studies ......................................... 6
  Guidelines .................................................................................................. 8
  Dispersion Modelling .............................................................................. 11
    Dispersion models for air pollution used in these studies ............... 12
    Data structure ...................................................................................... 12
    Dispersion model ............................................................................... 13
    Meteorological data ............................................................................ 14
    Background levels ................................................................................ 14
    Validation ............................................................................................ 14
  Population Data......................................................................................... 17
    Data sources ......................................................................................... 18
    Spatially aggregated population data ................................................. 18
    Selection and classification of population data .................................. 19
    Population data used in these studies ................................................ 20
  Exposure Modelling ............................................................................... 23
Proximity-based methods ................................................................. 24
Interpolation techniques .................................................................... 24
Land use regression (LUR) models .................................................... 25
Dispersion models ............................................................................. 26
Integrated emission-meteorological models ..................................... 26
Hybrid models .................................................................................. 26
Time-space models .......................................................................... 27
Methods of exposure estimation used in these studies ..................... 28

Part II – The Present Studies ........................................................... 31

  The Study Area .................................................................................. 31

  Levels of air pollutants ...................................................................... 33

  Study I – The effect of study area size and socio-economic factors ...... 35
  Study II – Investigation of spatial and temporal resolution ............... 35
  Study III – GIS modelling of NO\textsubscript{2} exposure ....................... 36
  Study IV – Spatial pattern of levels of blood lead in children ............ 37

  General Discussion ............................................................................ 39
  Conclusions and implications .......................................................... 44
  Future research ................................................................................ 45

Acknowledgements ........................................................................... 46

References ....................................................................................... 49

Part III - Papers ................................................................................. 57
This thesis is based on the following papers, which are referred to in the text by their Roman numerals:


Populärvetenskaplig sammanfattning

Kartor har använts som verktyg för att lokalisera exponeringskällor och rumsliga exponeringsmönster i knappt tvåhundra år. Idag har användandet av kartor för att kunna visualisera sjukdomsmönster och ohälsa utvecklats med en rasande fart, främst tack vare utvecklingen av GIS.

Även om förkortningen GIS börjar bli alltmer känd, i och med GPS:er och interaktiva kartjänters frammarsch, så brukar detta ord ändå kräva sin förklaring. GIS står för ”Geografiska Informationssystem” och är en teknik som har möjliggjorts tack vare vårt inträde i dataåldern. Mycket kortfattat kan GIS beskrivas som en datoriserad variant av kartanvändning, där man med hjälp av datorer och digitaliserat kartmaterial kan analysera konkreta objekts och företeelseras rumliga fördelning och relationer beroende på deras rumliga egenskaper och kännetecken. Möjligheterna att synliggöra rumliga fördelningar och relationer gäller även abstrakta eller svårobservera fenomen, till exempel socioekonomiska mönster eller halter av luftföroreningar. Inom medicinsk forskning används GIS numera inom en mängd olika inriktningar, allt ifrån studier av smittspridning till hur tillgång till hälso- och sjukvård påverkar sjukdomsmönstret i befolkningen. I våra studier användes GIS för att beräkna befolkningens exponering för luftföroreningar så som kväveoxider och bly. Särskild vikt lades vid att se på faktorer som påverkar tillförlitligheten och precisionen i exponeringsbedömningen.

kan man sedan analysera hur individer bor och arbetar i förhållande till, exempelvis, olika former av exponeringskällor, smittohärdar, demografiska grupperingar eller andra företeelser som kan variera i tid och rum.


Att med hjälp av GIS få fram både kvalitativ och kvantitativ data över hur individer blir exponerade i relation till hur de bor eller arbetar i förhållande till olika emissionskällor kan verka tämligen enkelt. Dessvärre öppnar GIS-användningen inte bara dörrarna för nya möjligheter utan även för ytterligare komplikationer. Ett exempel är att varken luftföroreningar eller individer är stationära, vare sig i tid eller rum eller i förhållande till varandra. Emissionerna från biltrafik är som högst under dagtid då de flesta transporter sker. Att då modellera individuell exponering baserat på var individer bor kan bli missvisande, eftersom de allra flesta inte befinner sig på sin hemadress dagtid utan är på sin arbetsplats eller skola. Ytterligare en begränsande faktor är att utomhushalter av luftföroreningar inte alltid reflekterar individuell exponering eller inomhusexponering (som påverkas av exempelvis rökning och användandet av gasspis). Då de flesta av oss tillbringar omkring 70 % eller mer av sin tid inomhus riskerar utomhushalter att ge en felaktig exponeringsbild. Dessa problemställningar behandlas i Arbete 3; även om vi med hjälp av GIS kunde modellera utomhushalter av kvävedioxid (NO\textsubscript{2}) väl, så överensstämde utomhushalterna dåligt med den personliga exponeringen (mått med individurna provtagare). Överensstämmelsen blev inte heller bättre då vi lade till utomhushalter vid arbetsplatsen under arbetstid. Detta tyder på det krävs förfina modelleringssolider, särskilt av tid i miljöer där exponeringen är som högst, dvs. tid i trafik och trafikerade miljöer.

Bortsett från svårigheten att kompensera för dynamiken och interaktionen mellan luftföroreningar och individer förekommer även rent modelleringstekniska problem, såsom hur tids- och rumsupplösning ska väljas vid modellering av luftföroreningar. Detta studerades i Arbete 2; den spridningsmodell som användes kunde tillfredsställande modellera dygnsmedelvärden av kväveoxid (NO\textsubscript{x}). Överensstämmelsen mellan modellera halter och uppklätrade halter ökade ju längre (medelvärdesbildade) tidsperioder som studerades. Även den rumsliga
upplösningen påverkade överensstämmelsen - ju högre rumslig upplösning man modellerade med desto bättre överensstämmelse. Kraven på hög rumslig upplösning minskade däremot med minskad tidsupplösning (längre medelvärdesbildade perioder). I stadsmiljö med höga variationer av NO\textsubscript{x} behövdes högre rumslig upplösning för att korrekt fånga dessa. I landsbygdsmiljö, där både halter och halternas variation är mindre, kan man modellera med lägre rumslig upplösning.

I Arbete 2 visas att det finns samband mellan befolkningens socioekonomiska position (mätt som utbildningsnivå och födelseland) och exponering för NO\textsubscript{2}. Sambandet förändrades dock om man analyserade materialet för hela Skåne jämfört med analyser för enskilda städer. Resultaten från både Arbete 1 och 2 visar tydligt att val av storlek på studieområde spelar stor roll, inte bara när det gäller modellering av luftföroringar utan även beroende på vilka befolkningsfaktorer man är intresserad av att studera.

Sammanfattningsvis visas i dessa fyra studier att GIS är ett användbart verktyg vid exponeringsstudier för luftföroringar. Valet av upplösning i tid och rum samt studieområde och befolkning har däremot en stor betydelse för resultaten, och bör därför noga undersökas innan man påbörjar exponeringsstudier.
Aim

The general aim of the research described in this thesis was to investigate how GIS could be used in exposure assessment studies. The focus was on how factors, such as spatial and temporal resolution, study area and population characteristics, affects the validity and precision in exposure assessment of airborne pollution.

Specific aims

The aims of the first study was to use GIS to investigate associations between the mean annual concentration of NO$_2$ and two socio-economic indices in the region of Scania, and to investigate the influence of differences in size of the study area on any associations observed.

The aim of the second study was to identify the optimal spatial resolution for a GIS-based dispersion model, with respect to different temporal resolutions, using a NO$_x$ emission database for the region of Scania, Sweden.

The aim of the third study was to determine how accurately a GIS-based dispersion model and an emission database could model residential outdoor levels and personal exposure of NO$_2$ using static and dynamic GIS-modelling.

The aim of the final study was to evaluate geographical patterns of lead exposure among children, with special reference to changes in industrial emissions and traffic over time, using GIS.

Structure of this thesis

This thesis is divided into three parts. The first part introduces the reader to the assessment of air pollution exposure through the use of GIS. This part is structured according to the different subjects that form the basis for these kinds of studies, namely: the pollutants and their emission sources, dispersion modelling, population data and exposure modelling, as illustrated in Figure 1. Each of these subjects is extensive, and an overview of each field is provided, focusing on the
pollutants and applications used in the present studies. Alternative methods that may be used for exposure assessment are described briefly.

The second part presents the study area and summarizes the various studies and findings. In a general discussion, the papers are put in perspective. The third part consists of the four papers describing the work making up this thesis.

Figure 1: Schematic illustration of the procedure adopted in the present work: modelling pollutant values for an area of emission sources and estimating the exposure of the population.
Part I – GIS in the Assessment of Exposure to Air Pollution

“Human exposure assessment is a key step in estimating the environmental and public health burden […]” [1]. Exposure assessment is the process of measuring and/or modelling the magnitude, frequency, and duration of contact between a potentially harmful agent and a target population, including the size and characteristics of that population [1]. In the field of epidemiology, exposure assessment is crucial in the study of health/illness patterns within a population.

GIS is the abbreviation for Geographical Information System, although it has recently been used to denote Geographical Information Science. GIS provides means of storing, processing and analysing spatial data digitally. The use of GIS in exposure assessment studies has increased rapidly during recent decades. Improved accessibility to geocoded data, together with faster computers with large storage capacity, has made it feasible to conduct studies over large areas including a vast number of people. Improvements in computer capacity have also made it possible to conduct exposure modelling with increased spatial and temporal resolution. However, it is not only developments in computer science that have led to the application of GIS within this area, but also an awareness of the potential of GIS, especially in the field of epidemiology, where the need to assess the exposure of large populations or areas makes GIS ideal.

Numerous GIS-based methods can be applied within this field, either as the basic concept, or in combination with other GIS-related methods. GIS can also be applied in all the different steps of exposure assessment, such as creating spatial databases of emission sources, locating population groups or areas, modelling levels of air pollutants, estimating exposure levels, locating emission sources and identifying exposure patterns.

In the studies presented in this thesis, GIS has been applied for modelling concentrations of NO\textsubscript{x}, evaluating the spatial and temporal resolution of an emission database, estimating exposure to NO\textsubscript{x} and NO\textsubscript{2}, and estimating the exposure pattern from emissions of lead. These studies illustrate a variety of ways in which GIS can be used within the field of exposure assessment to provide insight into the potential of this tool.
Air pollution is the mixture of solid particles and gases in the air. Depending on their dose and exposure pathway these may cause harm or discomfort to humans or other living organisms. Most people associate “air pollution” with human activities but, although this is justified in many cases, vast quantities of air pollutants are also emitted by natural occurrences, such as lightning, volcanic eruptions, wildfires, and dust storms, as well as swamps and bacteria. These natural sources are impossible to regulate and, therefore, anthropogenic sources, such as fuel combustion and industrial emissions, are those we should try to reduce.

Pollutants can be defined as “primary” pollutants (those directly emitted), or “secondary” pollutants (those formed when primary pollutants react with other pollutants or substances) [2], but may also be a combination of these forms. They can also be classified into “indoor” or “outdoor” air pollution, and although the substances may be the same, the sources may differ, as well as the dose and health consequences.

**Air pollutants considered in these studies**

In the studies described here, two primary outdoor air pollutants of major concern for human health, namely, nitrogen oxides (NO\textsubscript{x}, including NO\textsubscript{2}) and lead (Pb), have been investigated. The following section will therefore focus on these pollutants and their anthropogenic emission sources.

**Nitrogen oxides**

“Atmospheric nitrogen oxides” is a generic term for the nitrogen oxides (NO\textsubscript{x}) that are formed through combustion. NO\textsubscript{x} are formed by the combustion of all types of fuel at high temperatures. The main product is nitrogen oxide (NO), which becomes nitrogen dioxide (NO\textsubscript{2}) when NO is oxidized by air.
In most urban locations, the nitrogen oxides that yield NO$_2$ are emitted primarily by motor vehicles, making it a strong indicator of vehicle emissions in general. NO$_2$ and other nitrogen oxides are also precursors for a number of secondary pollutants. Through photochemical reactions, initiated by solar-radiation-induced activation of NO$_2$, the generated pollutants formed are an important source of nitrate, sulphate and organic aerosols that contribute significantly to particulate mass [3]. NO$_3$ and the total particle number (i.e. reflecting the ultrafine particles) are also well correlated at urban and near-city level and show a distinct diurnal variation, representative to the common traffic source [4].

Under the influence of ultraviolet light, NO$_2$ can form tropospheric ozone (O$_3$) through photolysis. However, this ozone is usually converted back into NO$_2$ and O$_2$ by the surplus NO from this reaction [5]. In highly polluted areas other air pollutants, such as volatile organic compounds, can also react with NO, preventing the reduction of O$_3$, and high concentrations of tropospheric ozone can therefore be formed on sunny days in highly polluted areas. Tropospheric ozone is one of the major substances in photochemical smog.

Health risks from nitrogen oxides may potentially result from NO$_2$ itself or secondary products (such as O$_3$). Epidemiological studies on NO$_2$ and NO$_x$ exposure from outdoor air cannot separate these effects – nitrogen oxide levels should therefore rather be seen as a reasonable marker for exposure to the general cocktail of traffic related emissions [6]. Evidence of the health effects of NO$_2$ comes largely from toxicological studies.

Short-term chamber studies with humans as well as animals supports the view that NO$_2$ levels, generally encountered in ambient outdoor air in Europe, have minor (if any) direct effects on the respiratory system [6]. In contrast, epidemiological studies have shown consistent associations between long-term exposure to NO$_2$ and decreased lung function in children as well as respiratory symptoms and diseases in adults. These associations was found even at levels below the current air quality guidelines, and without evidence of a threshold [6]. However, as mentioned earlier, these effects cannot be attributed to the nitrogen oxide exposure per se, but the nitrogen oxide as a general indicator of traffic-related air pollution.

Both NO$_2$ and NO can cause harm through their reactivity with other atmospheric substances. When NO$_2$ reacts with water molecules nitric and nitrous acid can be formed, causing acidification of the ground and watercourses through rainfall. This illustrates the complexity of the influences that nitrogen oxides can have on the environment, concerning not only air quality but other environmental aspects.
Lead

Lead (Pb) is a highly toxic, metallic element that has a tendency to accumulate in the human body. The developing central nervous system is especially sensitive, and impairment of mental development has been associated with lead exposure, without evidence of any threshold level [7, 8]. Not only foetal and infant exposure but also exposure later during childhood may affect cognitive development. Recent data suggest that subtle neurodevelopment and cognitive effects can be observed at exposure levels as low as 50-60 µg Pb/l of blood [7, 8].

Anthropogenic sources

Sources of Nitrogen oxides

Most NO\(_x\) and NO\(_2\) emitted by anthropogenic sources originate from the combustion of fuels by vehicles and heating and power plants [6]. Emissions in Sweden have decreased by almost 50% during the past 20 years, from 301,000 tons in 1990 to 155,000 tons in 2008, mainly due to regulations and the introduction of catalytic converters in vehicles [9]. However, regulations and improvements are being counteracted by the increase in traffic. Apart from road traffic, international transport by rail, air and sea contribute considerably to nitrogen oxide emissions [9]. However, these emissions are difficult to measure and control and, therefore, many nations, including Sweden, do not include these in their emission levels.

Major indoor sources of NO\(_x\) are gas and oil stoves, as well as smoking [6, 10]. A chamber study in 2006 by Lee et al. [11] also showed that nitrogen oxides were the most abundant gas related to the burning of candles (out of NO\(_x\), NO\(_2\), NO, carbon oxide (CO), methane (CH\(_4\)), and non-methane hydrocarbons (NMHC). The levels emitted from one candle, burning for approximately 1½-2 hours, was too small to allow the quantification of an emission rate or factor [11]. However, the time an individual is exposed to burning candles has been shown to have a significant impact on personal exposure to NO\(_2\) [5]. These results suggest that extensive candle burning may be a non-negligible source of NO\(_x\) and NO\(_2\) exposure.

During the past 50 years, changes in house design and lifestyle habits, such as the introduction of effective cooker hood ventilators and less indoor smoking, have decreased the indoor levels of NO\(_x\) considerably [12]. In a national health survey conducted in 2010 by the Swedish National Institute of Public Health (Statens folkhälsoinstitut) the proportion of daily smokers was only 13% and 83 % of the adult population reported that they were not exposed to environmental tobacco smoke (passive smoking) [13].
Sources of lead

Lead is mainly emitted by the combustion of fossil fuels, but also by industrial activities such as non-ferrous metal, iron and steel production, and recycling activities. The major source of lead exposure in many parts of the world was, until recently, the organic lead added to petrol to increase the octane rating [14], however, many nations have forbidden or restricted the use of organic lead in petrol (Sweden started to phase out lead in petrol in 1988, and it was forbidden in 1995). Although the deposition of lead has decreased considerably during recent decades due to the regulation of lead emissions, this decrease is relatively small compared with the huge increase in airborne lead over the past centuries [15].

In many parts of the world, the major indoor source of lead has, until recently, been leaded paint (which has also been an outdoor source). These paints are now forbidden in most parts of the world, resulting in dramatically decreased blood-lead levels, especially in children [12]. In Sweden, leaded paint is not considered a contemporary source of lead. Household exposure through glazed pottery is another well-known source of lead, which in Sweden is observed nowadays only when imported pottery is used. Another source may be the inhalation of cigarette smoke [14].

Inhalation of airborne lead constitutes a minor exposure pathway for the general population. However, atmospheric lead tends to be bound to ultrafine particles (aerodynamic diameter 0.2-1.0 µm), which are removed from the atmosphere by wet or dry deposition and accumulated in soil and plants [14]. The ingestion of lead through soil, due to hand-mouth activities, is considered one of the most important pathways in children. Even in countries where considerable efforts have been made to control lead, vast reservoirs may still exist in soil, dust and house paint, and these will continue to affect the population for many years [16].

Emission sources considered in these studies

The sources of NO\textsubscript{x} and lead in these studies differ depending on the different study designs and the availability of data. Data on the sources of NO\textsubscript{x} were taken from a detailed emission database covering the whole region of Scania, while the data for lead emission apply only to the two municipalities of Landskrona and Trelleborg, as described below.

Nitrogen oxides

The database used to calculate NO\textsubscript{x} emissions has been described in detail previously [17], and will therefore only be described in general terms in this section. The main source of NO\textsubscript{x} is combustion engines, and the emission database
therefore consists of line sources following roads, railways and shipping routes. In total, there are approximately 24,000 line sources corresponding to these routes. Each road segment was classified into fourteen different road types, corresponding to different traffic situations and flows. Depending on the classification, each line segment was then assigned a daily mean value (based on data for an average year) for the traffic flow and proportion of private cars, buses, lorries without trailers, lorries with trailers and natural gas buses. An emission of NO\textsubscript{x}, originating from the Swedish Transport Administration (Trafikverket) was then assigned to each of these vehicle types, fuel consumption (petrol or diesel) and road classification. For railways, NO\textsubscript{x} emissions were assigned for routes that are not electrified and where diesel locomotives are in use. Shipping and ship routes were divided into five classes depending on the type of vessel, namely pleasure boats, working ships, transport ships, ferries and passing ships (those that do not anchor at Scania’s ports). Emissions of NO\textsubscript{x} have been assigned to each route classification and type of vessel.

Emissions from the fourth major type of transport, planes, were assigned to area segments representing LTO zones (Landing and Take-Off zones), which correspond to the area covered by planes flying at altitudes of about 3000 ft. (914 m). Thus, these emissions represent plane emissions during take-off and landing only, but include emissions from machines and other vehicles in these zones. The emissions are not based on the type of plane, but provide general data on emissions originating from these zones. The LTO zone for the Danish airport at Copenhagen (Kastrup) outside Scania was also included as an emission source in the emission database.

Industrial emissions of NO\textsubscript{x} were collected from a national register of industries and heating plants having a major impact on the environment: EMIR (Emissions registret). The data were treated as point sources corresponding to the location of the plant or the plant’s chimney, together with data on chimney height, industrial activities, emission factors and quantity. The emission database consists of approximately 550 such point sources for industrial plants and heat and power plants.

The contribution from domestic heating appliances in the area was calculated using information from the chimney register (Sotarregistret) managed by the National Rescue Services Agency (Räddningsverket). This register is based on information reported by chimney sweeps about the type of boilers used and how often the chimneys are swept. Based on estimates of the total use of fuel, fuel type and boiler type, the total emissions were calculated for each municipality.

Emissions from machinery and tools are based on emission models from the Swedish Environmental Research Institute (IVL) for machinery in different sectors (e.g. farming, forestry, shipping, etc.). Grids, with a resolution of 1km, were
constructed approximating the emissions from different numbers and types of machinery, except for shipping machinery, where a local grid with a resolution of 100 m was used over the port area. Each grid was then located over the region of Scania according to the infrastructure or land use, i.e. the grid for emissions from farming machinery was placed over agricultural land, etc.

Except local emission sources, the major contribution to air pollution in Scania originates from the densely populated area of Zealand (Själland) in Denmark. The emissions from Zealand have been roughly estimated from an inventory made by the County Administrative Board in Scania (Länsstyrelsen i Skåne) and allocated as a grid over the area. This grid contains a coarse estimate of the Zealandic emissions from the same sources as mentioned earlier, apart from the Danish airport at Copenhagen, which is stored as an area source, and four large power plants that are registered as point sources.

**Lead**

The study in this thesis concerning lead emissions and exposure was focused on two municipalities in Scania: Landskrona and Trelleborg. The municipality of Landskrona was selected due to the location of a lead smelter (battery recycling factory) in the town of Landskrona, and the municipality of Trelleborg was chosen as a reference area of similar size and population structure, but without any major industrial lead source. The sources of lead emission in this study were thus the lead smelter in Landskrona and major roads and urban areas, as possible sources of increased lead exposure through traffic, in both municipalities.

**Guidelines**

The physician Paracelsus (1493-1541) stated: “What is there that is not poison? All things are poison and nothing is without poison. Solely the dose determines that a thing is not a poison”. However, since many air pollutants occur together, and since their effects on health may be difficult to separate from effects caused by other substances, or lifestyle factors, it is often difficult to identify a toxic level for a specific substance.

**International guidelines**

In 1987, WHO introduced their first guidelines for the reduction of health problems related to air pollution. These guidelines were updated in 1997, and their most recent guidelines for air pollution reduction were published in 2005 [6]. These guidelines are based on the most recent findings at that time regarding the
health effects resulting from exposure to particulate matter, ozone, nitrogen dioxide and sulphur dioxide worldwide. Since NO$_2$, by tradition, has been routinely used as a marker for combustion-related air pollution, the current guidelines apply to NO$_2$, and not the primary combustion-related nitrogen oxide products NO$_x$ or NO [6].

According to the WHO’s “Air Quality Guidelines 2005” the hourly mean value for NO$_2$ should not exceed 200µg/m$^3$. This level is based on experimental animal and human studies, which have shown that short-term concentrations exceeding 200µg/m$^3$ increases bronchial responsiveness in asthmatics, but far higher concentrations are needed to elicit symptoms or respiratory functional changes in healthy subjects. The guideline for long-term exposure is an annual mean of 40µg/m$^3$. In population studies, nitrogen dioxide has been associated with adverse health effects even when the annual average nitrogen dioxide concentration was lower. However, it is difficult to determine whether the effects observed for nitrogen dioxide is independent of other pollutants – rather, NO$_2$ should be seen as an indicator of the complex gas–particle mixture that originates from vehicular traffic.

The first international WHO guidelines for lead concentrations in air date back to 1987, and are recommendations for mean annual levels ranging from 0.5 to 1 µg/m$^3$ [18]. These are based on the concentrations of lead in blood, assuming that 1 µg/m$^3$ lead in air correspond to approximately 19 µg Pb/l blood in children and 16 µg Pb/l in adults (the relationship between these is curvilinear, and is mainly applicable for low blood lead levels) [19]. However, exposure to lead in the air also contributes to the uptake of lead through other environmental pathways and, therefore, 1 µg Pb/m$^3$ air is generally assumed to lead to 50 µg Pb/l blood [19]. The WHO guidelines for blood lead levels are at present 100 µg/l for children and women of reproductive age, since these groups are most vulnerable to high lead exposure [14]. However, recent research clearly indicates that neurodevelopmental effects may occur at much lower blood lead levels [7, 8, 20].

**European regulations**

The European Union has based its limitations on air pollution concentrations on the WHO guidelines. The limits for NO$_2$ came into force on 1$^{st}$ January, 2010, declaring each member state was not to exceed an annual mean level of 40 µg/m$^3$ of NO$_2$ and an hourly mean of 200 µg/m$^3$ (the hourly concentration may be exceeded up to 18 times/year).

The limit for lead is an annual average of 0.5 µg/m$^3$ air, which came into force on 1$^{st}$ January, 2005. The annual limit for lead in the immediate vicinity of industries
was 1 µg/m³ until 31st December, 2009, since then, the lower limit also applies at these sites [21].

These limits are legally binding for each member state, and they are required to monitor these values and verify that they are not exceeded by dividing their territory into a number of zones and agglomerates, in which air pollution levels are assessed by measurements or modelling [21].

**Swedish regulations**

Sweden has environmental quality standards (*Miljökvalitetsnormer*) that are legally binding policy instruments. These were first introduced in 1999 as a means of controlling sources of diffuse emission (such as those from traffic or agriculture). These environmental quality standards apply in four different areas; outdoor air pollution being one. These standards are usually set by the government, and should be in line with current directives from the European Commission. Authorities and municipalities are required to make sure these regulations are followed [22].

Sweden has hourly, daily and annual mean limits for NO₂ to protect human health, and an annual limit for NOₓ for environmental protection. The hourly mean value for NO₂ is 90 µg/m³, which may be exceeded 175 times/year as long as it does not exceed the EU limit. The daily mean value is 60 µg/m³ (which may be exceeded up to 7 times/year), and the annual mean value for NO₂ is 40 µg/m³ (which may not be exceeded). For NOₓ, the annual mean value is 30 µg/m³, however, this value only applies in areas at least 20 km from the nearest urban area, or 5 km away from another settlement, industrial site or motorway [23].

The Swedish environmental quality standard for lead is an annual mean of 0.5 µg/m³. This value may not be exceeded in any way [24].

Apart from the environmental quality standards, there are also a number of other policy instruments, such as environmental quality goals (*Miljökvalitetsmål*) and environmental standards (*Riktvärden*), however, these are not legally binding, but should be used as guidance by the authorities responsible for air quality.

The environmental quality goals for NO₂ are an hourly mean of 60 µg/m³ (which may be exceeded 175 times/year), and a yearly mean of 20 µg/m³. These levels should not be exceeded after the end of 2010, but Scania will not be able to fulfil these goals, and according to current forecasts, there is doubt whether these levels will be reached before 2020 [25].
Dispersion Modelling

Particles and chemical compounds dispersed in the air can travel with air masses and winds, and within days be in another geographical region or even on the other side of the globe. During transportation air pollutants may interact with other chemical compounds to form new compounds. Consequently, it is not only the air pollutants emitted in our immediate surroundings that we inhale, but also those from distant sources, and with other chemical compositions.

There are several ways of using GIS to estimate the dispersion of air pollutants over a given area (as described in the chapter entitled “Exposure Modelling”).

The objective is to model the dispersion and concentrations of predefined chemical compounds resulting from a number of emission sources. Today, several dispersion models are in use [26]. The number of emission sources used in the different models can vary from one to several thousands in complex cases [17]. The complexity of the models also varies. Some models are statistical, using measured and established relations between sources and the air pollutants being emitted. Others are based on emission databases with information regarding the quantity, composition and time interval of emissions, as well as meteorological parameters, street-canyon effects, the height of emission sources or the elevation of the landscape. There are also chemical dispersion models that calculate the interactions between chemical compounds or the way in which they change over time or with sunlight intensity, etc. And, of course, there are also many hybrid models involving several of the above-mentioned models.

The choice of model depends on factors such as: the complexity of the terrain, the size of the study area, the air pollutants of interest, the temporal resolution, the complexity of the meteorological conditions in the study area, and the data available on emission sources and meteorological conditions.
Dispersion models for air pollution used in these studies

The software used for emission data management and dispersion modelling in these studies is ENVIMAN (Environment Manager), developed by the company OPSIS AB in Sweden. The program was developed for air quality surveillance, and consists of integrated tools for emission database development and dispersion modelling. The program has been described in detail in a licentiate thesis by Gustafsson at Lund University [17], and will therefore only be described in general terms here.

Data structure

The emission sources included in the emission database are described by four different types of spatial objects: points, lines, areas and grids. The emissions from these different objects are defined so that point sources emit from their exact coordinates (these usually represent chimneys or the location of a factory), while line sources describe the emissions along a line (e.g. road) or a section of a line (i.e. between two crossroads). Area sources represent a zone in which the emissions are assumed to be the same over the whole area, while grid sources divide an area into cells, where the emission can vary between different cells. For all the spatial objects except the grid sources the positioning error may be up to 100 m.

Apart from the spatial structure and aspects associated with the emission sources, it is also necessary to determine the temporal pattern of the emissions and, in some cases, also physical factors that influence the emissions (such as chimney height). This is done by applying the formula [17]:

\[
EDB_X = ESV_X \times ETV_X \times EPV_X \times E_X
\]

where:
- \( EDB_X \) is the emission database/pattern for object X
- \( ESV_X \) is the space variable for object X (spatial profile)
- \( ETV_X \) is the time variable for object X (time profile)
- \( EPV_X \) is the physical variable for object X and
- \( E_X \) is the emission from object X.

There are three ways of calculating emissions, depending on the structure of the data in the emission database.
1. Bottom-up – the emission database consists of different emission sources described in detail, and all the emissions are added to provide an estimate of their total effect.

2. Top-down – the total emission over a large area is divided and down-scaled to fit smaller areas (this is often done to divide emissions calculated or measured on national level into smaller administrative areas).

3. Top-down-Bottom-up – this method applies emissions obtained for larger areas to smaller areas (Top-down), but divides the emissions more thoroughly by using other factors such as land use, population density, etc.

If possible, the bottom-up method should be applied in the first place, followed by the top-down-bottom-up approach and lastly, if neither of these is possible, the top-down method is used. In the emission database used in these studies, the bottom-up principle was applied to all spatial objects except grid sources, where the top-down-bottom-up principle was applied.

Since the emissions in the database are stored as NO\textsubscript{x}, the output from the model are levels of NO\textsubscript{x}. For transformation of NO\textsubscript{x} into NO\textsubscript{2} concentration an empirically developed equation from the Environmental Department of Malmö City (Miljöförvaltningen i Malmö) was applied. This equation has been modified during the years, and the two equations applied in these studies are described in detail in Papers I and III.

**Dispersion model**

The dispersion program ENVIMAN is a combination and modification of the dispersion model AERMOD, developed by the US Environmental Protection Agency [27] and a street canyon model: OSPM (Operational Street Pollution Model), developed by the Danish National Environmental Research Institute.

AERMOD is a Gaussian plume air dispersion model, which assumes the concentration distribution to be Gaussian in both the vertical and horizontal directions. The model is a flat 2-dimensional model, i.e. topography and buildings are not taken into consideration but the height of the emission source is incorporated in the model. The lack of topographic information may be a problem, especially if the Gaussian model is applied over a large area (such as the region of Scania). However, due to the relative topographic flatness of the region in question, the model was considered to be applicable [17].

OSPM is a street canyon model, i.e. it accounts for the recirculation of air and pollutants that may occur in streets with high buildings due to the lack of ventilation. It was not possible to apply this model in the present dispersion calculations, due to a shortage of data regarding building heights and street width.
Modelling in ENVIMAN was performed on a grid with adjustable resolution. All the emissions, those obtained from the emission database and the calculated concentrations were summed and assigned to the centroid of each cell in the grid, but the model can also calculate concentrations at receptor points, along lines and over areas.

**Meteorological data**

The meteorological data used for dispersion modelling in ENVIMAN can be either climatological data or hourly time series. The climatological data provide a meteorological average over a number of years, while hourly time series consists of measured meteorological values for each hour. The required meteorological parameters that are needed are: temperature, wind direction, wind speed and global radiation. These data were collected from a meteorology station in Malmö. The use of a single meteorological station may constitute a major source of inaccuracy in the dispersion modelling since the conditions in Malmö (situated on the southwest coast of Scania) may differ considerably from those in other parts of Scania. Unfortunately, this meteorological dataset was the only one available with complete and high-quality data, and the dispersion program is also limited to the use of only one dataset of meteorological parameters.

**Background levels**

The emission database used includes data on emissions from the land areas of Scania and Zealand, as well as The Sound (Öresund) between Sweden and Denmark, Skagerrak and the Baltic Sea (Figure 2 in Part II, chapter: “Study Area”). To estimate background levels (i.e. long-range contributions of air pollution from outside the area covered by the emission database), data were used from between one and seven meteorological background stations [17]. These stations are situated in rural area, as far as possible from major emission sources, where local and regional contributions should be small. To estimate background levels, local and regional emissions were modelled at the location of the meteorological background station, and the difference between the concentration measured at the meteorological station and the modelled concentration is assumed to be a measure of the level of long-range transportation of the pollutant in question.

**Validation**

The dispersion model was validated against 14 urban background stations during the period 1999-2002. These stations were placed in 12 different towns or cities.
(Burlöv, Helsingborg, Hässleholm, Höganäs, Hörby, Kävlinge, Landskrona, Lund, Malmö, Osby, Trelleborg and Örkelljunga), all with large differences in population size and emission patterns. Unfortunately, these background stations were unevenly distributed geographically, and thus the south-eastern part of Scania was not covered, which may cause a geographical bias. A comparison between measured and modelled concentrations at these sites showed that the difference was approximately 10% (mean: 1.6µg/m³, median: 1.1µg/m³ range: -0.9-7.9µg/m³), and the correlation coefficient was 0.96 [17]. The two towns that differed most (the modelled values overestimated the measured levels) were both harbour towns (Helsingborg and Trelleborg) with extensive shipping and ferry traffic. This suggests that the emissions from shipping traffic may be overestimated in the emission database.

To compare the temporal accuracy between measured and modelled values, hourly levels of NO₂ were modelled for the period, 2003-2004, and compared with levels measured at the meteorological station in Malmö (from which meteorological data were collected for the emission database). The hourly levels showed a correlation coefficient (r) of 0.53, but this was considerably improved by temporal aggregation: r_day=0.71, r_week=0.93 and r_year=0.96. These results suggest that the dispersion program may not be able to reflect the finer temporal variations correctly (hourly variations), but reflects the general temporal trend well when aggregated in time.
Population Data

To estimate the exposure of a population correctly, it is important to know the exact location of the individuals in question. The level of generalisation of both the population and the methodology has a significant influence on the results. Therefore, a variety of parameters is of interest, for example, household and workplace density or, for more detailed studies, specific addresses of homes and workplaces. If the individual’s mode of commuting to work, and the location of leisure activities, etc. are known, these factors can also be included in an exposure study, enabling the creation of time-activity models [28]. Apart from data that enable positioning of individuals in time and space, there is also a need to obtain data describing various personal and area characteristics.

Population data are usually administrated by national and local authorities, but they can also be obtained from private companies. Data can either be collected from official sources such as population, tax or health registers, or can be gathered from directed or undirected surveys. The availability of register-based population data and its quality differ between countries and even smaller administrative regions. Apart from offering the possibility to conduct studies on large populations, the advantage of register data is that the risk of sampling bias (some individuals are less/more likely to take part in a survey) and response bias (some individuals are less/more likely to respond) is small. The disadvantages, on the other hand, are that registers seldom contain all the information required for a study. This can be compensated for, to a certain degree, by linking different registers to each other. There may also be errors in the registers which have not been properly documented.

Directed and undirected surveys have the advantage that researchers are able to customise the study according to the hypothesis being tested. However, these studies are expensive and time consuming, limiting the possibility to study large populations. It may also be difficult to enrol a sufficient number of subjects, especially if the study is demanding, or involves biological sampling, which may intrude on the individual’s privacy. Hence, there is a risk of both sampling and response bias.
Data sources

Statistics Sweden (SCB) has been assigned the responsibility of collecting, analysing and presenting statistics regarding the Swedish population. Due to the system of personal identification numbers (ID number) in Sweden, the possibility of combining data for specific individuals from different sources is substantially better than in many other countries. Each ID number is unique, and is assigned to each Swedish resident at birth (or upon immigration). By combining ID numbers with the Swedish National Land Survey’s (Lantmäteriet) property register, the coordinates of, and information on, an individual’s residential address can be obtained. This kind of data is often desirable when conducting exposure studies with high spatial resolution. However, the disadvantage is that this kind of study poses a threat to personal privacy due to the possibility of geographical identification in addition to sensitive information, such as health and socio-economic factors. Therefore, access to these data is often restricted.

If the individuals’ addresses are known (through registers or surveys) these can be positioned in a GIS by linking the addresses to a geo-coordinated address database. The disadvantages of this approach are that these databases must be up-to-date and are often expensive. The method also requires that the addresses are written in the same way, without abbreviations or differences in spelling, in both registers, otherwise it will not be possible to link them.

Directed and undirected surveys are usually set up by the researchers themselves, which gives them full control of and access to the data and information gathered. These methods have the advantage that it is possible, although expensive, to position the study subjects by equipping them with a GPS device. The research team often owns the information obtained, but due to privacy regulations the use of the data is often restricted.

Spatially aggregated population data

To avoid the problems associated with personal privacy, or when a lack of geographical coordinates makes it impossible to conduct studies with high spatial resolution, population data are often aggregated into administrative areas or grid cells of varying sizes. It is important to be aware that these two methods of aggregating data are quite different. Administrative areas are often created based on infrastructure or the population pattern within an area, for example, postcodes or population density. This method of aggregating population data is advantageous in studies where there is a need to describe the population structure within an area. However, a serious disadvantage of administrative areas is that they might introduce the modifiable areal unit problem (MAUP) [29], i.e. point-based
measure of, for example, population density is aggregated into areas causing the summary values to be dependent on both aggregation level and the chosen administrative level. Also they often change due to changes in the population structure which makes it hard to conduct longitudinal studies. Another problem regarding statistical areas in the same dataset is that the geographical size can vary considerably, i.e. statistical areas in city centres tend to be relatively small, due to the high density of population, in comparison with statistical areas in rural areas, where the density of residents is lower [30].

The problem of varying sizes does not affect grids since the cell size is well defined. In certain studies, this may be a disadvantage as blocks and neighbourhoods are divided into grid cells without taking population density or physical/population characteristics of the area into account. This makes it difficult to geographically identify an area or present statistical results for different areas within a region. Varying numbers of individuals in different areas or grids may also cause problems due to the fact that in areas with few individuals, a difference in one or two cases can make a substantial difference, compared with more densely populated areas (i.e. the small number problem) [31].

**Selection and classification of population data**

In exposure studies, the estimates can either be collected on individual level (i.e. each individual in the study population is assigned a “unique” level of exposure, either by monitoring or by modelling), or by exposure grouping (i.e. the study population is divided into subgroups based on exposure or exposure determinants) [32]. These subgroups can also be divided into smaller subgroups based on biological and physiological determinants, socio-economic characteristics or other health outcome determinants relevant for the epidemiological study in question. When performing exposure studies using exposure grouping, it is assumed that each member of a particular subgroup is exposed to similar levels due to similarities in the exposure characteristics.

The choice of determinants for subgroups is important and should be performed with care. For example, the classification of a study population into age groups could have a huge impact on the results due to widely differing exposure profiles, as well as differences in susceptibility between children, adolescents, adults and the elderly [33]. A recent review of the influence of gender in air pollution epidemiological studies showed that there are gender- (and age-) related differences in susceptibility to air pollution [34]. However, whether these are due to biological differences, exposure differences or an interaction between the two is still unknown.
The choice of determinants and subgroups is even more critical when conducting GIS-based exposure studies. For example, determinants such as urban/rural population, smoking habits, age, income and level of education, etc. may be spatially correlated to each other and the levels of air pollutants [35]. The well-known differences in daily smoking prevalence by age and gender is such an example, analysed with recent data from southern Sweden by Ali et al. [36]. Such a correlation between smoking habits, age and gender could influence the effects in an epidemiological study since these factors (studied separately or in combination) might be spatially auto correlated to certain areas and thereby to differing levels of air pollution. The effect of such spatial correlation could also differ depending on the size and heterogeneity of the study area (as reported in Paper I). This sort of spatial correlation between determinants and exposure variables is complex and could bias the results in epidemiological studies through confounding and effect modification.

A factor acts as a confounder when it is associated with the outcome and the exposure (although it may not be an effect of the exposure) [37]. For example, smokers are more disposed to develop lung cancer, and the habit of smoking may be more prevalent among individuals in the lower socio-economic classes which, in turn, usually live in poorer and more polluted areas. In a study on the development of lung cancer due to exposure to outdoor air pollution, smoking may therefore act as a confounder, biasing the results by giving a false impression of the importance of outdoor air pollution in the development of lung cancer.

Effect modification occurs when a factor modifies the effect of the exposure of interest [33, 38]. For example, if there are gender- or age-related differences in susceptibility to air pollution, these might act as effect modifiers in highly polluted areas where there are a high proportion of these susceptible subgroups.

Population data used in these studies

The population data used in the studies described in Papers I, III and IV originate from three different sources. In Study I, two sets of population data provided by the County Administrative Board of Scania (collected and analysed by Statistics Sweden) were used. The first dataset contained attributes concerning the individuals’ age and sex. The individuals’ places of residence were represented as points at the centre of their property/dwelling (listed in the National Registry). It should be noted that this kind of localisation may cause a certain misplacement of individuals living in high-rise apartments, since the real estate may consist of several buildings. This may also be the case for large estates in rural areas, where the individuals will be positioned in the middle of the real estate instead of at the location of their house. The second dataset contained more detailed personal
attributes such as sex, year of birth, country of origin, marital status, income, etc. In order to protect personal privacy, the individuals were located at the centroid of the 1x1 km grid cell in which their property/dwelling was located.

The population data used in Study III originated from three different measurement campaigns carried out in 2003, 2005-2006 and 2008. The subjects’ residential and workplace addresses were recorded, and later geocoded, together with various personal data such as smoking habits, use of a gas stove, occupation, household characteristics, etc. (different data were collected during the different measurement campaigns).

The population in Study IV consisted of children aged between 8 and 10 years during the years 1978-2007 in the municipalities of Landskrona and Trelleborg. In addition, children aged 10-17 participated in 1978, and in 1986 preschool children from 3 years of age were also included. The location of the children participating between 1983 and 2007 was obtained by linking their personal identification number to the centre coordinate of their home, as listed in the Regional Population Register. Due to a lack of housing information in the Regional Population Registry before 1983, addresses were obtained from school catalogues and geocoded manually. The locations of the schools were also geocoded manually.
Exposure Modelling

Accurate measurements, or modelling, of the concentrations of pollutants over a defined area and the positioning of individuals constitute the first part of an exposure assessment study. The next part consists of assigning this exposure to a population or group of individuals. The exposure of individuals to air pollution is highly dependent on where they live and work, and their time-activity pattern. The GIS technique used to assign exposure is highly dependent on the steps mentioned earlier and the nature of the data, e.g. the quality and aggregation of the population data used, or the kind of pollutant values and measurements that have been made. Thus, modelling in exposure assessment studies is complex, and no single process is best.

Human exposure to a substance can be measured using two types of techniques: direct and indirect [5], also called personal exposure monitoring and environmental monitoring [32]. Direct methods involve personal monitoring (e.g. the subject carries a portable monitor during a study campaign), or the collection of biomarkers (e.g. levels of lead in blood). Direct methods are usually expensive and resource-demanding and, therefore, mainly used in small cohorts and during short campaigns [5]. For specific airborne pollutants, exposure assessment studies using biomarkers may also need to consider multiple routes of human exposure. Since, in addition to inhalation, dermal absorption and oral ingestion may be important pathways of exposure to these pollutants [39]. Direct methods are, however, crucial for the evaluation of the reliability of exposure models.

Indirect methods are used to estimate human exposure by combining information on pollutant concentrations in specific areas or environments with the time spent in that area/environment [5]. The majority of early exposure studies relied solely on data from central monitoring stations, thus ignoring small-scale variations in pollutant concentrations [2, 40-42], and reducing variations in exposure by falsely assigning the same concentration to a large number of people [43]. During recent years, a large number of studies have been conducted to increase the accuracy in describing spatial variations in air pollution in relation to population location, in order to better estimate levels of exposure. General ways of tackling these problems were classified by Jerrett et al.[44], who divided existing methods into
six classes of exposure methods, depending on their level of complexity. The following sections will focus on these six classes, as well as a seventh: time-space models. All these methods except time-space models are usually combined with some sort of “overlay analysis” in GIS, i.e. the resulting dataset containing the level of exposure is usually combined, by spatial overlay, with a dataset containing population data to obtain levels of exposure for individuals or subgroups.

**Proximity-based methods**

These are the most basic exposure assessment methods. The assessment is based on the distance between a pollution source and the individual, assuming that proximity to emission sources is proxy for the exposure of human populations. This technique has been adopted in a number of exposure response studies to examine health outcome in relation to a nearby road or industrial plant. The advantage of these methods is that they are fairly straightforward and easy to use. The disadvantages, however, are many since they are crude proxies for the processes that determine the distribution of pollutants in the environment [26].

Despite these shortcomings, the proximity to roads with specified traffic intensity has been used successfully for exposure assessment in air pollution epidemiology studies. Examples are studies showing an increased risk of mortality due to stroke near main roads in England and Wales [45] and that local exposure to traffic on a freeway had adverse effects on children’s lung development, which were independent of regional air quality [46]. In Scania, the prevalence of adult asthma and chronic bronchitis was found to be increased in subjects living near roads with intense traffic [47, 48].

**Interpolation techniques**

These methods interpolate the levels of air pollutants recorded for an area (or a population) based on measurements obtained at one or several monitoring sites. The simplest method is the so-called “point in polygon” method, where measurements at a single station in a city are used to assign exposure values to the whole population of that city [26]. This method was applied in the well-known “Six Cities Study” by Dockery et al. [49], and although it is a somewhat naïve method, it can be applied successfully, especially in areas with small variations in pollution levels.

To take the spatial variation of pollution into account, more sophisticated deterministic and stochastic geostatistical interpolation techniques must be used. Deterministic methods are based on the theory of spatial autocorrelation, i.e. it is assumed that monitoring stations close to an area/individual provide a better
measure of exposure than monitoring stations further away [26]. The most common GIS models applied for this kind of interpolation are different forms of inverse-distance-weighted models, such as triangulated irregular networks, but moving windows can also be applied. Although these methods are relatively simple to use, they have the disadvantage of being highly dependent on the set-up and choices made by the user. They also require that the monitoring stations are evenly distributed in space and (preferably) located in a rather dense network, which is seldom the case.

To avoid errors resulting from the choices made by the user, stochastic geostatistical interpolation techniques may be applied. A wide variety of geostatistical methods are used within GIS, but they all have in common that the result depends on the data, rather than the choices made by the user. One of the most well-known methods is kriging, which assumes that the spatial variation of an entity (in this case an air pollutant) is dependent on systematic variation of the subject, random but spatially correlated variation of the subject and random but spatially uncorrelated variation of the subject [26]. Apart from predicting levels of air pollutants, this method has the advantage that it also generates their standard error, providing the possibility to identify where the interpolation tends to be less reliable [35].

Compared with proximity-based methods, all these interpolation methods have the advantages that they use real measurements of air pollution in the computations, and that the difference in the level of exposure between individuals can be quantified. A major disadvantage, however, is the fact that interpolation techniques have a tendency to smooth data [41], and are highly dependent on the number of monitoring sites and the quality of the data available for the interpolation [43, 50]. This is especially problematic regarding pollutants that vary significantly over small areas, such as NO$_2$, as these variations are poorly captured by monitoring sites [44].

**Land use regression (LUR) models**

This method has become increasingly popular during recent years, and is considered to be better than, or at least as good as, interpolation methods [51]. These methods employ regression models to calculate levels of air pollution based on factors such as emission sources, land use, population density, altitude, meteorology etc. [44, 51]. Monitored levels of air pollution at a number of sites are combined with predictor variables, usually collected through GIS, to estimate the pollutant levels over the study area through a regression model [51]. The advantages of this method are that it is associated with relatively low costs and that it can be used without the need for additional monitoring or data collection.
However, studies have shown that the method requires monitoring data from a relatively large number of sites (40-80) [51], and that it may produce poor correlations when used to study areas with diverse land use or topography [44]. A recent development of the LUR model is the use of modelled data to increase the number of sites [41].

**Dispersion models**

A variety of dispersion models are in use (see the previous chapter on the subject). These are based on assumptions about deterministic processes, and data on, for example, emission sources and meteorological and topographic conditions, to estimate levels of air pollution. The strength of dispersion models is that spatial and temporal variations in air pollutants can be combined without using a network of monitoring stations. Among the disadvantages, however, are the fact that data and hardware are costly, and the implementation of these models requires a high level of programming and GIS expertise. Due to these costly and demanding requirements, and also lack of awareness and distrust, these models are rarely used within epidemiological studies [52].

Dispersion modelling using the emission database and dispersion model presented in this thesis has been used in epidemiological studies investigating respiratory diseases [47, 48, 53] and stroke [54, 55].

**Integrated emission-meteorological models**

These models combine meteorological and chemical modules that are related to each other to enable the simulation of dynamic atmospheric pollutants. The meteorological data in these models are used as input for the chemistry modules during simulation [44]. These models have a considerable potential to model complex dynamic atmospheric processes. The possibility of incorporating chemical transport and pathways allows the development of secondary pollutants to be simulated and the more precise estimation of the likely pollutant mixture. However, the disadvantage is that these models are costly due to the need for complex computational facilities, software and highly trained personnel. Hence, there has been little use of these complex models in epidemiological studies to date.

**Hybrid models**

Since all the methods described above are highly dependent on the availability and quality of the input data, hybrid methods have been developed in an attempt to
make use of all available data (source-based and monitored) [26]. These can be divided into two types [44]:

1) hybrid models combining personal or home exposure monitoring (i.e. personal monitoring equipment attached to the study subject’s clothes or a fixed monitoring station placed near or in the person’s home) with one of the six methods described above, or

2) hybrid models combining one or more of the preceding six methods with regional monitoring.

The benefit of using hybrid models is that they provide measurement validation. The weaknesses depend on the combination of models used.

**Time-space models**

Time-space models (also denoted “Activity-based models or “Dynamic models”) differ from the previously mentioned methods by incorporating the location of the individuals. The methods described above either apply the location of the population in question as a statistical measure (such as the population density for an area), or the individuals are simply assigned an exposure based on the levels at their place of residence. Neither of these approaches reflects the true exposure well, since most individuals spend most of their waking time somewhere else or, when they are at home, they tend to spend most of their time indoors where pollutant levels do not correspond to the outdoor levels most often applied in exposure studies. Many researchers carrying out exposure studies have recognised the complexity of this problem, and have thus begun to apply time-space models in which the exposure at different locations or in so-called microenvironments (such as the kitchen, bedroom, inside the car, etc.) and the time and duration spent in these environments are taken into account [26]. Measured or modelled levels of pollutants can be used in these models to estimate the exposure, and the output is given in the form of either time-weighted exposure (aggregating the pollutant levels measured or modelled for different microenvironments), or as detailed exposure profiles for each individual or population [26]. An example of such a model is STEMS (Space-Time Exposure Modelling System). This model was designed to simulate the exposure of people as they move through a changing air pollution field. The model integrates data on source activity, pollutant dispersion and individual travel behaviour to derive individual or group-level exposures to air pollution during journeys [56].

To obtain data on time-activity patterns requires “time-activity diaries” to be filled in by the study subjects [57], or the use of theoretical time-activity patterns on group level, or individual GPS devices to record the activity pattern of each individual [58]. Each of these methods has their own drawbacks. The reliability of
time-activity diaries may be difficult to evaluate, due to individual differences in the exactness of recording activities, for example. A study conducted by Elgethun et al.[59] in which time-activity diaries for children filled in by their parents were compared with locations recorded by GPS units carried by the children, showed a disagreement of 48%; the time spent at some locations being highly underestimated in the time-activity diaries (travelling time and time spent outdoors at home). Theoretical time-activity patterns are often developed from large surveys, and will reflect the population’s activity pattern in general, thus providing a general estimate, not an individual one. Equipping each study subject with a GPS device provides the best available log of the individual’s time-activity pattern; however, this can only be done in small studies due to the cost of equipment and the willingness of subjects to participate.

However, the development and use of time-space models has recently been expanded by the use of mobile phones with GPS capabilities. This allows the tracking of a large number of people, and will enable the development and use of dynamic modelling in the future [60]. Although this technique have only recently begun to gain ground within exposure studies the use mobile phones for positioning individuals was integrated in AirGIS [61], a Danish GIS-model for exposure studies, already in 2004 [44] and will hopefully be adopted in future exposure studies in a higher extent.

**Methods of exposure estimation used in these studies**

After using GIS for the creation of the emission database, GIS was applied in study I-III using dispersion modelling.

In the first study (Paper I) GIS was used to model annual air pollution levels of NO$_2$ with a grid resolution of 250 metres, and to assign these levels to the population through bilinear interpolation (i.e. the closest four grid cells were assumed to influence the exposure). Individuals in the population dataset used (PD1) were positioned in the centroid of the 1x1 km grid cell in which their residence was situated. To enable the use of a higher resolution in the dispersion modelling, their position was refined using a second population dataset (PD2) in which the individuals were positioned according to the centroid of their home. Since the individuals within each 1x1 km grid cell were not evenly distributed in space, the population centroid of the grid cell was relocated by combining PD1 with PD2 to find the geographical centre that better reflected the centre coordinate of the population density within each specific square kilometre.

In Study III (Paper III), a combination of direct and indirect methods was used. Measured levels of NO$_2$, collected through personal monitoring and stationary monitors outside the study subject’s homes during a week, were compared with
modelled levels of NO$_2$ for the same time period. Comparisons were made using measurements at the subject’s homes as well as a time-space approach combining modelled levels at the home address (Monday to Friday 6 p.m. – 7 a.m., and all day Saturday and Sunday) with modelled levels at their place of work (Monday to Friday, 8 a.m. – 5 p.m.).

A direct method of measuring exposure was used in Study IV (Paper IV), namely biomonitoring of blood lead in children. Information on blood lead levels and each child’s geographical location (home and school) was used to study the relationship/geographical pattern between blood lead level and the proximity to a major lead smelter and major roads in the study area. A proximity-based method was used to calculate the Euclidian distance of each child from the lead smelter and the closest major road. This was then combined with a time-space approach by constructing a time-weighted average for each child based on their home and the school location. For visualisation purposes, the children were divided into five different groups based on their blood lead level. The mean geographical position (based on the location of the home of each child) was then calculated for each of these groups and visualised in relation to the lead smelter.
Part II – The Present Studies

The studies presented in this thesis are examples of how GIS can be successfully implemented in the field of exposure assessment. The investigations focused on different ways of improving and developing the assessment of exposure to air pollution. The studies cover a wide range of aspects within this field, such as the importance of the size of the study area and socio-economic factors, as well as spatial and temporal resolution, for the development of a pollution database. The agreement between modelled levels of NO$_2$ and measured levels was investigated, as well as whether personal exposure corresponded to outdoor levels in general, but also with respect to temporal and spatial differences between home and working location. GIS was also used to analyse the spatial interaction between lead sources and levels of blood lead in children.

The Study Area

All the studies presented in this thesis were carried out in the county of Scania, in southern Sweden (Figure 2). Although the region is relatively small, with a land area of 11,350 km$^2$, corresponding to 2% of the total Swedish land area, the number of inhabitants is $\approx$1,200,000, i.e. 13% of the total Swedish population. By Swedish measures, it is therefore a relatively densely populated area (112 inhabitants/km$^2$) [62]. Around 60% of the population lives along the west coast of the region, where Sweden’s third largest city, Malmö, is situated, with a total population of $\approx$ 286,000 inhabitants. A more detailed description of the largest towns and cities in the region (Malmö, Helsingborg, Lund, Kristianstad and Trelleborg) is presented in Paper I.

Due to its location, there is a great deal of traffic travelling through the region, due to the transport of goods between Scandinavia and Continental Europe. There are also several major ports in the area, with cargo and ferry traffic to and from the European continent (e.g. Denmark, Germany, Poland and the Baltic states). As a result of this, the road network in Scania is relatively dense, as well as the traffic, especially along the major highways, causing certain areas of the region to be rather heavily affected by air pollutants originating from vehicles.
Figure 2: The study area Scania and its vicinity
Levels of air pollutants

Compared with international values, the levels of air pollution in Scania, and in Sweden in general, are rather low. A study of the NO$_2$ concentrations in Sweden, carried out in 1999 by the Swedish Environmental Research Institute revealed that most parts of Sweden have low NO$_2$ urban background concentrations [63]. In the sparsely populated northern part of Sweden the annual mean level reaches 3 µg/m$^3$, while small and medium-sized towns typically have NO$_2$ concentration of 10-15 µg/m$^3$. The most densely populated areas, such as the major cities of Stockholm and Gothenburg and Scania’s west coast, have levels around 20 µg/m$^3$. This far from exceeds the environmental quality standards (annual mean of 40 µg/m$^3$). A report by The Swedish Environmental Research Institute in 2008 also shows that the urban level of NO$_2$ is decreasing annually by approximately 0.6µg/m$^3$ [64]. However, these values are urban background values, and the environmental air quality limits are sometimes exceeded in hotspots [63], as well as street canyons in major cities [9].

The combustion-related levels of air pollutants in Scania are, as already mentioned, higher than in most other parts of Sweden (apart from Stockholm and Gothenburg). These high levels are a result of the extensive road network, as well as shipping activities along the coastline and the transportation of goods from Denmark and the European continent. There is a clear geographical gradient in NO$_x$ concentration in the east-west direction, the most polluted part of the region being the densely populated western coastline (Figure 3).
Figure 3: The annual NO\textsubscript{x} level in the study area (Scania)
Study I – The effect of study area size and socio-economic factors

The first study (Paper I) was designed to examine whether or not there is an association between levels of NO\textsubscript{2} and the socio-economic indices “country of birth” and “level of education” in Scania. The possible influence of differences in the size of study area on any associations observed was also investigated.

The study included the whole population of Scania in 2001 (≈1.2 million), and mean annual outdoor levels of NO\textsubscript{2} were modelled. The levels of the pollutant were modelled as a grid with a resolution of 250x250 m and thereby linked with the population data to provide a measure of individual exposure to NO\textsubscript{2}. The data were then divided into socio-economic groups and analysed using weighted correlation analysis.

The study revealed that there was an association between the level of NO\textsubscript{2} and the socio-economic indices “country of birth” and “level of education” in the region of Scania and some of the towns and cities studied. However, these associations varied depending on study area, and studying the towns/cities together and separately yielded contradictory results. Moreover, the performance of the two indices for socio-economy sometimes differed between cities.

The region of Scania thus seems to be heterogeneous regarding the associations between air pollution and socio-economy. This indicates that it is inadvisable to analyse associations between socio-economy and exposure to air pollutants on a regional level. The choice of socio-economic indices is crucial when adjusting for socio-economic status in a study of health effects of air pollution.

Study II – Investigation of spatial and temporal resolution

The objective of the second study (Paper II) was to determine the optimal spatial resolution, with respect to different temporal resolutions, for a NO\textsubscript{x} database for the region of Scania, Sweden.

Firstly, modelled NO\textsubscript{x} values were compared with measured NO\textsubscript{x} values to evaluate the general accuracy of the model. Thereafter, pollutant grids with different spatial resolutions (100, 200, 400, 800, and 1600 m) and a time resolution of one hour were generated for a period of four weeks. Mean values of the original data were used to create new pollutant grids with temporal resolutions of one day
and one week. The pollutant grid with the highest spatial resolution (100 m) was used as a reference grid to compare the loss of information resulting from using a coarser grid in relation to the different time resolutions. The results were then evaluated to determine the optimal spatial resolution for the pollutant database.

The results of this study showed that the mean difference between measured and modelled hourly NO\textsubscript{x} values was low, although the standard deviation was rather high. The agreement was much improved when the data were aggregated in time (days and weeks). The results also showed that the agreement between the reference grid and coarser grids improved with increased spatial resolution. However, when pollutant values were aggregated in time, the agreement between the reference grid and the grids with coarser resolutions was considerably improved. The agreement between different resolutions and the reference values was also dependent on the characteristics of the study area (urban/rural).

Based on these findings, it was concluded that the maximal error due to the choice of resolution should not exceed 1 µg/m\textsuperscript{3}, when using a time resolution of one day. Temporal aggregation allowed lower spatial resolution without significant loss of detail. A pollutant database should also allow different resolutions depending on the characteristics of the area (e.g. urban or rural).

**Study III – GIS modelling of NO\textsubscript{2} exposure**

The purpose of this study (Paper III) was to investigate how accurately a GIS-based dispersion model and emission database could model residential outdoor levels of NO\textsubscript{2}, and to investigate how well residential outdoor levels of NO\textsubscript{2} were correlated to measured personal exposure to NO\textsubscript{2} during a period of seven days. It was also determined whether modelled exposure combining residential and workplace outdoor levels of NO\textsubscript{2} better reflected personal exposure.

The study was conducted using 165 subjects who wore a personal diffusion sampler around their neck for seven days. In addition, NO\textsubscript{2} levels were measured on the façades of 86 of the subjects’ homes to provide outdoor levels. Modelled NO\textsubscript{2} levels were determined for each subject’s place of residence and work (when provided) with a resolution of 100x100 m. The NO\textsubscript{2} values for the home and work locations were then combined for each subject by replacing the NO\textsubscript{2} values for the home location from Monday to Friday, from 8 am to 5 pm (i.e. working hours) with the NO\textsubscript{2} values for their working location.

The results showed that modelling levels of NO\textsubscript{2} is a useful method of estimating outdoor concentrations. However, the agreement between measured outdoor levels and personal exposure was poor. The agreement between modelled outdoor levels
and personal exposure did not improve when compensating for the location of the workplace.

It was concluded that emission databases and GIS-based models can be successfully applied to estimate levels of outdoor NO₂. However, the use of measured or modelled outdoor concentrations of air pollution as a proxy for personal exposure during a short time period should be undertaken with caution even when compensating for outdoor exposure at the work location.

**Study IV – Spatial pattern of levels of blood lead in children**

The aim of the final study (Paper IV) was to evaluate the geographical pattern of lead exposure among children with special reference to industrial emissions and traffic during 1978-2007.

Blood lead levels were used for exposure assessment. The home address and school location of the participating children (N=3879) were geocoded, and their proximity to a lead smelter and major roads was calculated using GIS. Sex, school year, lead-exposing hobby, country of birth, parents’ smoking habits and temporal differences due to the regulation of lead in petrol were taken into consideration in the statistical analyses.

The children’s proximity to a lead smelter affected their levels of blood lead throughout the whole study period, even several years after a reduction in the emissions. The children’s proximity to major roads also noticeably affected their concentrations of blood lead during the period 1978-1987. Thereafter, the banning of lead in petrol led to reduced levels of lead in blood, especially among urban children.

Based on these findings it was concluded that there are spatial gradients in children’s blood lead concentrations based on the proximity to a lead smelter. The use of GIS revealed that this spatial pattern could be detected even for emission levels that are, in an international context, relatively low.
General Discussion

Pollutants in the environment can only affect a person’s health if that person is sufficiently exposed to harmful substances [31]. Exposure to a pollutant, and its possible health effects, depend not only on the concentration, but also on the temporal and spatial variation of the pollutant in question, and its interaction with the target population or individual. Furthermore, the difficulty in correctly estimating the concentration and dispersion of air pollution over a specific area is increased by the difficulty in combining these estimates with estimates of the location of individuals or a population, which may also vary in time and space. Thus, the geographical variation is crucial in environmental exposure assessment studies.

GIS analyses can facilitate these studies by modelling the distribution of both pollutants and populations and combining these models to quantify exposure in both space and time. However, although GIS considerably facilitates the performance of exposure assessment studies, it does not reduce the complexity of the problem. The improved possibility of conducting more precise exposure studies over larger areas and populations is also associated with new difficulties, such as the choice of temporal resolution and the spatial aggregation of both air pollution and population (as shown in Papers I and II).

The dispersion of pollutants can be modelled using GIS in a number of ways. In the studies presented in this thesis, a high-resolution emission database and a dispersion modelling program were used. The choice of both the temporal and spatial resolution is crucial for the accuracy of dispersion modelling. However, choosing the appropriate temporal scale in an exposure assessment study is not an easy task, since there may be a conflict between the effective exposure time (i.e. the minimum exposure time required to produce a health effect [31]) and the temporal accuracy of the emission database used. The assessment of exposure might be intended to be used for a study of acute asthma effects resulting from exposure to vehicle emissions, in which case the effective exposure time would be a couple of hours. However, the temporal accuracy of the emission data may be coarser, and more suitable for modelling daily or monthly values. Using a higher resolution (in either time or space) than the input data allow gives a false sense of exactitude, since the quality of the underlying data does not match the apparent quality of the result. It is difficult to model values of air pollution with high accuracy over short time periods [65]. However, previous studies have shown that aggregating modelled dispersion data in time improves the accuracy, when compared with measured values [17, 65, 66]. This was also shown in Study II, where the accuracy of the dispersion model was improved when data were
aggregated to give daily or weekly values, instead of hourly values. These results confirm the results obtained by Gustafsson [17], who used the same dispersion model and emission database. The program ENVIMAN, used for dispersion modelling in these studies, is a modification of the AERMOD dispersion model. Zou et al.[65] evaluated the performance of AERMOD on different time scales and also found that the model performed better over longer periods. Especially monthly values proved to be better at reflecting high concentrations. Although these studies differ regarding emission sources and study area, their results clearly indicate how critical the choice of temporal resolution may be in ensuring reliable results. As pointed out by Zou et al. [65], these results are of interest for researchers and authorities using dispersion models for the assessment of population exposure in environmental epidemiology studies on different time scales [65]. However, surprisingly few pre-studies have been conducted in environmental epidemiological exposure assessment studies to identify the most suitable temporal resolution.

Apart from the temporal resolution, the choice of spatial resolution will also affect the result. The possibility of detecting an association depends not only on the data used, but also on the scale chosen. Thus, choosing the “wrong” spatial resolution in a study may result in the rejection of a hypothesis and the elimination of factors from further analysis [67]. This was illustrated in the study described in Paper I, where the effects of air pollution exposure of population subgroups differed depending on whether the data were analysed on county or city level. These results are in line with the associations found in a study on environmental injustice and exposure to air pollution by Briggs et al. [68]. They concluded that their results were sensitive to both the scale of analysis and the geographical context (urban/rural areas) [68]. Similarly, diverging patterns have been seen in urban and rural settings in an epidemiological study in Scania using modelled NO\textsubscript{x} exposure [47].

The expected variation in the levels of pollution over a specific area dictates the appropriate resolution in exposure modelling. Large variations within a small area will require high-resolution modelling, and vice versa. In a review from 2005, Wilson et al. [42] concluded that modelling with low spatial resolution in an urban area is inappropriate since it may cause exposure misclassification. As seen in Paper I, the study of too large areas may obscure associations, or give a false image of homogeneity due to the concealing of underlying spatial differences. These problems may be of special concern in areas covering both rural and urban settlements, since there are often considerable differences in both lifestyle factors and levels of pollution (as shown in Papers I and II). On the other hand, using too small an area (i.e. population base) for such studies may lead to insufficient data, leading to unreliable associations.
As for the choice of temporal resolution, the spatial resolution is not only dictated by the dispersion modelling but also by the resolution of the available population data. While modelled air pollution data are often available (though not always suitable) at high temporal and spatial resolution, the population data are usually based on administrative areas [69]. Thus, modelling the dispersion of pollutants with high resolution would be a complete waste of time, and may even give erroneous results if the population data were more general, and had a coarser resolution. Zou et al. [70] concluded that the grid size of the population density raster could affect the values of the population exposure. They also established that the distribution of the population within the study area (i.e. the population density) could affect the exposure outcome in relation to the scale chosen for modelling air pollution [70]. However they did not further examine the influences of these factors as the objective was not to study scale effects [70]. João commented already in 2002 that the issue of scale and how it could affect the outcome of environmental impact assessment was rarely addressed in the assessment of environmental impact [71]. Unfortunately, this seems still to be the case, since many environmental health studies, especially those concerned with air pollution exposure, ignore spatial and temporal variations in population density [72] or have (pre-)evaluated how scale could affect their results. Many studies even fail to provide information on the scale used [71].

Traditionally, outdoor concentrations of air pollution at the place of residence have been used to determine the mean human exposure [50, 69]. However, an individual is exposed at other locations than their home location, and the outdoor air quality it not always a suitable proxy for the individual exposure to air pollution, as illustrated by Paper III. To better explain personal exposure, so-called “time-space” models or activity models have been developed, taking the variation in location (home, work, etc.) and/or microenvironments (kitchen, bedroom, car) of an individual into consideration. The results given in Paper III indicated that it was possible to model outdoor levels of NO\textsubscript{2} with satisfactory agreement, but that compensating for time spent at the home and workplace did not improve the agreement between personal exposure and outdoor levels of NO\textsubscript{2}. Therefore, there must be other factors causing this divergence. An example of such a factor could be the difference between indoor and outdoor levels, since individuals tend to spend approximately two-thirds of their time indoors [10, 73]. According to Valero et al. [74] a multitude of studies on the relationship of indoor, outdoor and personal exposure to NO\textsubscript{2} have shown that indoor concentrations are better correlated with personal exposure to NO\textsubscript{2}. However, outdoor levels of NO\textsubscript{2} and gas stoves are determinants for indoor concentrations [10, 74].

Previous studies in this field have mainly focused on either improving the accuracy of estimating population exposure through dynamic activity models, or identifying microenvironments of importance regarding total population exposure
and, thus, the possibility of differentiating the proportion and importance of population exposure resulting from various emission sources has been somewhat neglected [70]. Since outdoor NO\textsubscript{2} concentration at the home has been shown to be associated with traffic-related variables [10, 74], and since exposure from the indoor environment is difficult to incorporate into a GIS to model, the focus should be placed on using GIS models to estimate outdoor exposure. Due to the large contribution from traffic, the time-activity pattern of individuals in traffic environments should be further evaluated.

In most epidemiological studies, adjusting for socio-economic indices is of great importance so as not to bias the results. One reason for this is that air pollution is considered to contribute to social inequalities in health, either through differential exposure (disadvantaged groups are more exposed to air pollution) or differential susceptibility (these groups are more susceptible to resultant health effects), or as a combination of the two [75]. A study conducted in Malmö provides an example of social inequality, showing that children from neighbourhoods with low socio-economic status were more exposed both at their place of residence and at school [76].

As can be seen from Paper I (as well as the spatial issues addressed in the section “Spatially aggregated population data”), the choice of socio-economic indices and the size or administrative level of the study area should not be ignored, as both can have considerable influence on the results [68, 77]. This is a complex problem since both the choice of socio-economic indices and the choice of area might be spatially correlated with levels of air pollution (or the health outcome of interest), which could cause confounding effects or effect modification. Some associations may even be the result of contextual effects, where the neighbourhood itself might affect the outcome [78]. A recent activity-based exposure study conducted in The Netherlands [72] revealed that individuals from low socio-economic groups (income < average) had a higher residential exposure to traffic-related NO\textsubscript{2} than the highest socio-economic group (income twice the average). However, when adopting an activity-based approach, this difference was counterbalanced during morning rush hour, since people in the higher income group were more likely to be exposed to high levels of traffic-generated NO\textsubscript{2} through commuting to work by car [72]. An Australian study compared the exposure to various air pollutants (among them NO\textsubscript{2}) by equipping individuals commuting by car, train, bus, bicycle or walking with monitors. Their results revealed that bus commuters were exposed to the highest levels of NO\textsubscript{2}, followed by cars, walking, bicycles and trains [79]. However, other studies have shown that cyclists inhale a much higher dose than other commuters due to their higher ventilation rate [80, 81], and they are thus exposed to higher concentrations of pollution. Once again, this shows how complex and interlinked the different factors affecting precision and validity are,
and the necessity to further evaluate the effects of exposure in traffic environments.

GIS can be used not only to estimate personal exposure, but also to reveal geographical pollution dispersion patterns, and thus possible sources of pollution. This approach was applied in the study presented in Paper IV by studying point patterns using a proximity method as well as a variation of clustering (i.e. the detection of geographical means for different subgroups). This approach assumes that point clustering (i.e. the density of the measure of interest) should differ significantly from random patterns due to the proximity to a pollution source or other risk factors [35]. Examining children’s blood lead concentrations in relation to the distance from sources of lead exposure (in this case a lead smelter and major roads) made it possible to determine whether these sources had any effect on the children’s blood lead levels or not. The disadvantage of this kind of study is that the measure (in this case blood lead level) is not compared or correlated to measures of actual emissions or the spatial distribution, but simply implies a spatial autocorrelation, and a clear a priori hypothesis is therefore necessary [26]. To avoid potential effects caused by lead exposure through other sources, the study was adjusted for additional potential lead sources such as lead-exposing hobbies and parental smoking. Previous studies have evaluated the effects of parental work exposure and found them to be insignificant in this dataset [82, 83]. Although earlier exposure to lead and potential effects of mitigation were not adjusted for, most other sources of lead should have been controlled for in this study. The results of this study are in line with those of previous studies using geographical approaches, which identified an association between proximity to lead sources and elevated blood lead levels [84-86]. However, the present study differs from these others in the long period of measurement, which made it possible to analyse the influence of the lead smelter after the emissions had been reduced, as well as the introduction of unleaded petrol. The present study showed that although emissions had been reduced, it was still possible to detect a spatial pattern of elevated blood lead levels in relation to proximity to the lead smelter.

This thesis shows the usefulness of GIS when assessing exposure to airborne pollutants and emphasizes the necessity of carefully considering and evaluating the interaction of spatial and temporal resolution before conducting this sort of studies. The choice is complicated by the fact that these factors interact with each other in many different ways, depending on the data and methodology, the characteristics of the population and the size and type of study area, thus making it almost impossible to create any universal rule of thumb for dealing with these different factors.
Conclusions and implications

The main conclusions and implications of this work are summarised below.

✔ When modelling air pollution, temporal and spatial resolution as well as area characteristics should be carefully chosen so as not to affect the accuracy of the modelled values.

✔ GIS methodology can be successfully applied in exposure assessment studies yet the choice of spatial and temporal resolution as well as area and population characteristics are of vast importance and should carefully be pre-examined.

✔ The region of Scania seems to be heterogeneous regarding the association between air pollution and socio-economic status. It is therefore inadvisable to determine and analyse associations between socio-economic factors and exposure to air pollution on a regional level.

✔ Modelling levels of NO\textsubscript{2} is a useful way of estimating outdoor concentrations. However, the agreement between outdoor concentrations of NO\textsubscript{2} and personal exposure is poor. Compensating for outdoor levels at the workplace during working hours did not improve the agreement between outdoor levels and personal exposure to NO\textsubscript{2}. This suggests that exposure in heavy traffic environment or commuting might be a major contributor to individual NO\textsubscript{2} exposure.

✔ In low dose areas, proximity to a lead smelter can still pose exposure threat years after emission reduction.
Future research

The results of this work have revealed the need for continued studies in this field, examples of which are given below.

The effects of temporal and spatial resolution on the modelling of pollution in areas with different characteristics must be systematically evaluated to improve the accuracy of modelled values.

The choice of socio-economic indices and the size and homogeneity of the study area should be carefully considered in future studies on health effects resulting from exposure to airborne pollution.

The effect of outdoor levels of air pollution on personal exposure should focus on modelling individual time-activity patterns, especially time spent in traffic. This kind of studies can be facilitated by incorporating new technology, such as mobile phones, for positioning and measuring.

Previous sources of lead may pose a threat of lead exposure years after emission reductions. Future studies should therefore focus on estimating the duration and effects of long-term exposure to such sources, in order to take remedial measures.
Acknowledgements

This work would not have been completed if I had not been able to stand on the shoulders, or at least lean on them from time to time, of the following people.

I wish to express my gratitude to my supervisors: Dr Kristina Jakobsson, for her experienced supervision which illustrated the diversity and depth of her knowledge; Dr Håkan Tinnerberg, for his expertise and guidance, as well as for knowing when encouragement was needed; Dr Jonas Ardö, whose wise counsel I wish I had followed from the beginning, but which I now on will nail to the wall (with a 4-inch nail); and Dr Jonas Björk for his ability to share his vast knowledge with someone less experienced.

“My dispersion modellers”: Ralf Rittner, whose patience and skill I have tested again and again (and whom I will probably continue to torment); and Susanna Gustafsson and her colleagues at The Environmental Department of Malmö City for their guidance in dispersion modelling.

The data providers for these studies: Daniel Nilsson and his colleagues at The Department of “Samhällsanalys in Region Skåne for their timely, friendly and professional support.

My colleagues: Anna Oudin, whose vast knowledge and friendship are two of the greatest treasures I take with me from my time as a PhD student; my fellow PhD students, Ebba Malmqvist and Anna Lindgren, who not only endured sharing a room with me, but also enhanced this academic retreat solely by their presence; all my wonderful supportive colleagues at The Department of Occupational and Environment Medicine, especially Lars Rylander, Anna Rignell-Hydgbom, Anna Axmon, Ulf Strömberg, Maria Albin, Zoli Mikoczy and Frida Eek, for all their help and for making work such a pleasant place to be. I also wish to thank the administrative heroes: Mona Frick, Yulia Lindholm, Lisbeth Prahl and Monica Hansson for all their help, which made it possible to concentrate on work and less on administrative matters; as well as, my former colleagues at the Department of Physical Geography and Ecosystem Analysis and GIS Centre for all their help and support throughout the years.

I would also like to mention Helen Sheppard, who patiently battled with my inability to remember fourth grade English grammar.
Last, but not least, my wonderful Family, for always being there, and for finally realising that I am not studying geology; and my beloved husband Niclas, who has played a major part in the realization this thesis by making sure that neither I nor our children starved or suffer any discomfort while I was engaged in this work.
References


