Université catholique de Louvain
Faculté des Sciences, École de Géographie
Center for Operations Research and Econometrics

Geographies of green and health in Belgium: measurements, opportunities, and challenges

Doctoral Dissertation presented by

Sonia Trabelsi

in fulfillment of the requirements
for the degree of Doctor in Sciences

Thesis committee:

Prof.	Isabelle Thomas (Supervisor)	Université catholique de Louvain
Prof.	Philippe Chevalier (Chair)	Université catholique de Louvain
Prof.	Sophie Vanwambeke (Secretary)	Université catholique de Louvain
Prof.	Catherine Bouland	Université Libre de Bruxelles
Dr.	Lidia Casas	Katholieke Universiteit Leuven
Dr.	Hans Keune	Universiteit Antwerpen
Dr.	Camille Perchoux	Luxembourg Institute of Socio-Economic Research

Louvain-la-Neuve, Belgium
February 2020

Acknowledgements

This dissertation is the result of several months of questioning, meetings, making mistakes and making decisions. This would have not been possible without the material support of BELSPO which funded the research projects GRESP-HEALTH (grant BR/143/A3) and NAMED (grant BR/175/A3) in which I took part.

I consider myself lucky to have been able to carry out my research under the supervision of *Prof. Isabelle Thomas*. She has been an encouraging and inspiring mentor in these incredible years. I am glad to have had the chance to work such a passionate and tireless mentor. I am thankful to her for trusting me with my research choices and for advising me when needed. The road has not always been easy, but that did not stop us.

During the years spent working within the GRESP-HEALTH and NAMED projects, I have been surrounded by two great (multidisciplinary) teams of researchers, from whom I learnt a lot. In particular I would like to thank here *Dr. Lidia Casas* for following my research developments from the very beginning and for patiently sharing knowledge. This work has benefited a lot from our meetings and discussions. I am also thankful to *Prof. Catherine Bouland* for encouraging my research throughout the years and pushing me further, to dare more. I also had the chance to meet and work with *Dr. Hans Keune* whose questions and critical approaches have offered great food for thought and inspired part of this work.

I would like to take the opportunity to thank *Prof. Sophie Vanwambeke* and *Dr. Camille Perchoux* for accepting to be part of my Jury and offering extremely valuable comments and recommendations. I would like to thank *Prof. Philippe Chevalier*, who has not hesitated a second to join the Jury, and whom I had the chance to get to know during several board meetings concerning the present and future of LIDAM and CORE. Last but not least, I would like to thank once more *Prof. Geoffrey Caruso* who has been a source of motivation and a great advisor throughout my Master and my PhD.

Conducting research at CORE would have not been the same without a "dream team" (cit.), for which I have been grateful every day, for the past many days (give or take)... I found at the office a friendly, lively, stimulating and supportive working environment. The Geo-en-CORE team has turned out to be a group of amazing colleagues, lunch companions, and travel mates. I could have not asked for more. It is also thanks to *Arnaud*, *Gaëtan*, *Justin*, *Madeleine*, *Olivier*, *Adeline* and *Jonathan* that I am now here writing these words. Thanks for the journey (from Siberia to New Orleans, via LLN).

Life at CORE would have just been less loud and entertaining without coffee, pizza and board-meeting mates and friends. *Huan*, *Dagmara*, *Erika*, *Sabina*, *Fabrizio*, *Lorenzo*, *Leonardo*, *Angelo*, *Paolo*, *Sefane*, *Hassan*, *Rim*, thank you.

I would like to thank the GRESP-HEALTH and NAMED teams for the fruitful meetings and discussions over the past few years. Multidisciplinary work is not easy challenge but is extremely rewarding and satisfying.

I would like to thank the SMCS team of UCLouvain, for their extremely valuable help. Their courses and their support have proven to be fundamental for my research and for the successful completion of this work.

E da ultimo, ma non per importanza, vorrei dedicare qualche riga per ringraziare le mie famiglie (geograficamente) lontane. Grazie a *Andrea*, *Elisa*, *Giulia* e *Tommaso* per gli ultimi dieci anni di discussioni e di spritz. Lavorare e vivere con voi mi ha fatto capire che, con la volontà, anche il più aspro dissenso può trasformarsi in un tesoro.

Grazie *Sara* per questi vent'anni di amicizia, ricchi di momenti kafkiani, quotidianità e avventure "giovani". Non mi dilungo qui (scriveremo un libro, un giorno) ma vorrei ringraziarti per esserci stata, avermi supportata e sopportata ovunque fossimo nel mondo.

Il più grande, ma non scontato, ringraziamento lo vorrei fare ai miei genitori *Maria Teresa* e *Fathi*. Grazie per avermi sempre spronata a fare quello che desideravo, dove lo desideravo. Grazie per avermi fatta viaggiare e scoprire il mondo fin dai primi giorni. Grazie per essere stati sempre vicini, nonostante la distanza e per avermi ascoltata e incoraggiata sempre.

Dedico questo *kteb* a mia nonna *Wasila*.

Contents

Pars destruens,
Pars construens.

Francis Bacon, *Novum Organum*, 1620

Abbreviations

ASMP: Age Standardised Medication reimbursement Prevalence
BC: Black Carbon
BCR: Brussels Capital Region
BMI: Body Mass Index
CLC: Corine Land Cover
CNS: Central Nervous System
COPD: Chronic Obstructive Pulmonary Disease
CVS: Cardiovascular System
DA: Depression and Anxiety disorders
ENN: Euclidean Nearest Neighbor
GDPR: General Data Protection Regulation
GIS: Geographical Information System
GS: Green Space
HIS: Health Interview Survey
IS: Immune System
LIP: Lipids
MAUP: Modifiable Area Unit Problem
NDVI: Normalised Difference Vegetation Index
$\mathbf{NO_2}$: Nitrogen Dioxide
P95: 95^{th} Percentile
PCA: Principal Component Analysis
PLADJ: Patch Like-Adjacencies
$\mathbf{PM_{2.5}}$: Particulate Matter 2.5
$\mathbf{PM_{10}}$: Particulate Matter 10
RS: Respiratory System
SD: Self-developed
UA: Urban Atlas
UHI: Urban Heat Island

I

Introduction

1

Background

This chapter introduces the context and the driving motivation of this thesis. The first section of this chapter introduces the general context in which this work stands, by presenting current challenges of increasing urban population, the need to create more livable and green environments, and to reduce health expenditures. In section 1.2 the characteristics of green spaces are then presented to provide an overview of the intricacy of the issue. The following section (section 1.3) tackles the question of socio-economic inequalities associated to green space provision. The complexity of measuring and assessing health is then discussed in section 1.4, including an overview on the use of medication data and on the limitations of spatial analysis in health studies. Section 1.5 provides an overview of existing data on landuse, and the data limitations for Belgium. The last section (section 1.6) contextualises the works carried out in this thesis first within the Belgian healthcare system, and second within two research projects. The final section outlines the research question and the structure of this thesis.

1.1 General context

The burden of health expenses worldwide is significant, and keeps increasing. In 2016 costs for health amounted to almost 10% of the global Gross Domestic Product (GDP), circa 7.5 trillion $ (Xu *et al.*, 2018), making the health sector one of the leading sectors of the global economy. Figures for Belgium, a high-income European country, are consistent: in 2016 the healthcare expenditure accounted for 10% of the GDP (OECD/European Observatory on Health System and Policies, 2019). In Belgium healthcare expenses have witnessed almost a 12% increase between 2011 and 2016 (Eurostat, 2019). Governments are therefore seeking for ways to decrease expenditures and alleviate the healthcare load. Investments and attentions are directed to improving living environments as a way to ameliorate the quality of life of the citizens. It is an attempt to tackle the problem "upstream", in a preventive way, in order to decrease the costs of treatment (WHO - Regional Office for Europe, 2016).

Many studies have shown that a favorable living environment has the potential to mitigate some health problems. It is the case for some non-communicable diseases (NCD), such as cardiovascular diseases, diabetes type 2, or chronic respiratory diseases. Green spaces, in particular, have been at the center of the interest as possible mitigators. As a matter of facts, links between green spaces and health have been recognized through history, but it was only in the 21st century that the methods and outcomes have become more rigorous (WHO - Regional Office for Europe, 2016). In spite of the number of studies carried out on the subject, the associations between nature and health have proved to be characterised by numerous challenges in assessment (Hartig *et al.*, 2014). Urban green spaces, especially, are raising most of the attention. According to the United Nations, today over half of the world population resides in cities, and the number is prospected to increase to almost 70% by 2050 (UN/DESA/Population Division, 2019). Population health, well-being and sustainability of the living environment are clearly a key concern. Green spaces, in this context, are considered to play a fundamental role to improve both urban sustainability perspective as well as human health. The presence of green spaces in cities can improve general environmental conditions by, for example, mitigating the Urban Heat Island (UHI) effect (Bowler *et al.*, 2010), filtering noise and air pollution, and improving the water inflow (Nowak *et al.*, 2006; Paoletti *et al.*, 2011), which in its turn, improves citizens' mental and physical health.

1.2 Green: the good, the bad and the ugly

A more livable environment affects human well-being and can improve people's life. In particular, studies have shown that green spaces (GS) play a relevant role for human well-being. In 1984 Ulrich published a study demonstrating the beneficial effects of view of green spaces for hospitalised patients recovering from surgery. Since then, the link between green spaces and numerous aspects of well-being have been widely investigated. Several research studies have explored further relations between the surrounding (natural) environment and both physical and mental health. Green spaces have been shown to have direct and indirect salutogenic effects.

Presence of urban GS provides increased opportunities to do **physical activity** (Foster *et al.*, 2009; Astell-Burt *et al.*, 2014a), which in turn reduces risks of obesity (Ellaway *et al.*, 2005; Dadvand *et al.*, 2015), and offers a profitable environment for attention **restoration** (Kaplan, 1995; Berman *et al.*, 2008). Gamble and Howard (2016) show that the positive effects of green spaces can be noticed even by just viewing pictures of nature. Green spaces are also associated with increased opportunity for social cohesion (Kemperman and Timmermans, 2014; Jennings and Bamkole, 2019), thus improving the mood and mitigating chronic stress, depression or anxiety (Maas *et al.*, 2009; de Vries *et al.*, 2016; Nutsford *et al.*, 2016). Takano *et al.* (2002) have reported improved longevity of senior citizens associated to walkable green spaces in mega-cities. While there is a general consensus on the importance of green spaces for improving health conditions, results are not always significant (Di Nardo *et al.*, 2010; Huynh *et al.*, 2013; Saw *et al.*, 2015). This suggests that other factors may intervene and play a role on health status.

A growing urbanisation entails a higher vulnerability to environmental as well as social risk factors in relation to human health: an increased exposure to traffic and therefore air and noise pollution, can negatively affect human well-being and physical health (Gruebner *et al.*, 2017; Thurston *et al.*, 2017). Air pollution exposure has several potential negative outcomes on health, among which increasing risk factor for cardiovascular morbidity and mortality (Peters *et al.*, 2000), lung cancer (Raaschou-Nielsen *et al.*, 2013), type 2 diabetes (Yang *et al.*, 2018), respiratory conditions (Dominici *et al.*, 2006), as well as premature birth (Mendola *et al.*, 2019). Green spaces, in this case, can have positive indirect effects by **mitigating** environmental externalities, through the provision of ecosystem services (Coutts and Hahn, 2015). The green space cooling effect has proved to reduce the UHI effect, thus providing comfort to the inhabitants (Bowler *et al.*, 2010; Buyadi *et al.*, 2015). Also, green spaces help removing air pollution, providing for a cleaner air and indirectly ensuring a better living and healthier environment for the citizens (Nowak *et al.*, 2006). Some studies have also shown that species richness in green spaces can improve human well-being (Aerts *et al.*, 2018). Natural environments can further have a positive impact on health, by acting on some trivial stimuli, and through cultural pathways (Clark *et al.*, 2014). McKinney (2008) has shown that green spaces in urban areas can include a variety of biodiversity and species diversity that can be compared to natural areas.

The past decades have witnessed increased investments and development of green roofs and green walls in urban areas. Such greening approach not only maximise the space available in dense urban areas, but have also direct and indirect economic, environmental and health benefits. Green roofs help rainwater drainage, and can provide space for amenities, as well as reducing the UHI effect, improving roofs' durability and retaining air pollution (Feitosa and Wilkinson, 2018; Köhler *et al.*, 2002). Green walls as well have a cooling effect, by insulating the building, reducing energy costs for indoor temperatures, act in carbon sequestration, as well as providing visual benefits on the exterior (Feitosa and Wilkinson, 2018; Li and Yeung, 2014). Currently Germany is the country who has most invested in green roofs, where more than 10% of residential buildings have green roofs (Köhler *et al.*, 2002), but several European countries are planning in that direction.

While some research has focused on beneficial effects of natural environments on health, some studies have instead investigated the potential adverse effects of such spaces. Green spaces have also been looked at in terms of sense of safety (Troy *et al.*, 2012; Gatersleben and Andrews, 2013), and it has been shown that those areas may increase the sense of insecurity or the crime rate. Poorly maintained green areas (uncut grass, presence of litter, or graffitis) may deter their use (Gobster, 2002). A higher vegetation density may increase the concentration of allergens in the air and therefore the prevalence of allergic symptoms (Cariñanos and Casares-Porcel, 2011). While green spaces can support **biodiversity** (Lepczyk *et al.*, 2017), they can also be a favorable habitat for the proliferation and existence of some disease vectors such as ticks and rats, and thus spreading infectious diseases (Lõhmus and Balbus, 2015). Fear and feeling of insecurity are also associated to park settings. Big or open parks lacking light and surveillance can be easy locations for crime perpetuation (Michael *et al.*, 2001; Bogar and Beyer, 2015). In rural areas, where green is wider and more frequent, people are more exposed to pesticides and herbicides used for agriculture and farming (WHO - Regional Office for Europe, 2016; Mamane *et al.*, 2015) or to endotoxins, that could trigger respiratory disorders in adults (de Rooij *et al.*, 2017). Some studies have shown that exposure to a diverse microbial environment from an early age can act as protective factor against asthma, for example (Müller-Rompa *et al.*, 2018)

Since green spaces are shown to have effects, positive or negative, on well-being, it is hence important to have a good understanding of how green is measured and defined, in order to have a better understanding of the geographical effect. The use of data on green spaces in the literature includes both **subjective** assessment of the areas, through self-reported quality or safety of the environment, or by providing information on the use for example, and **objective** measures such as quantity and typology of vegetation, or size of the green areas, for example.

The availability, access and usage of landuse data has (tremendously) increased in the past decade(s). Landuse data is getting more and more easily and freely available online. More satellite images, more software to visualise and interpret the data are available and accessible to a large share of the population. This also means that a wider variety of ways to capture and define vegetation is available, potentially harming comparability of different studies, if the study methods are not properly defined (see section 1.5, for further discussion).

1.3 Just environment

Socio-economic characteristics have long been associated to health status. Several studies have shown links between socio-economic deprivation and poorer health. Rocha *et al.* (2017) have found lower physical health-related quality of life for people living in more deprived neighborhoods in one municipality in Portugal. Results are consistent with several other studies carried out in different geographical areas and for targeted population groups. A study conducted on elderly population in Britain demonstrates how poor personal and neighborhood socio-economic characteristics are associated with general lower quality of life (poor self-care, social interactions and home management) (Breeze *et al.*, 2005). In Belgium, a study conducted by Bossuyt *et al.* (2004) showed consistent results with existing international literature on the links between socio-economic deprivation and poorer health. In a meta-analysis of the literature Lorant *et al.* (2003) found evidence of socio-economic inequalities for depressive disorders.

As presented in the previous section, mental and physical health can both benefit from green spaces. Green spaces can play a positive role on health for different groups of the population (children, adults and elderly) through different pathways. Nonetheless, literature has also shown that green space provision and effects are often associated to individual socio-economic factors (Dadvand *et al.*, 2012b; Maas *et al.*, 2009; Hoffimann *et al.*, 2017). Several measures and actions are currently being taken by citizens and governments in order to create a more livable and sustainable urban environment for everybody. In Europe, for example, the European Commission has deployed numerous funding opportunities for green investments (European

Commission, 2020). While, as previously mentioned, green urban areas can mitigate the impacts of climate change, foster physical activity and mental and physical restoration, urban greening can also have negative consequences on citizens if not properly managed. Green spaces can have in fact indirect repercussions on different aspects of life. Although such repercussions are not directly imputable to green, these can have tremendous consequences on citizens' life and well-being, and on further urban development. Investments in green space development is often followed by increased property prices and hence foster a gentrification process (Morancho, 2003; Jim and Chen, 2010). Negative effects of green spaces are rather to be found in public policies not properly tackling gentrification or overlooking the potential of existing green spaces. Rupprecht and Byrne (2018), for example, study the effects of the creation of "informal green spaces" on environmental gentrification and housing price in Australia and Japan. They find that such green areas can improve access to green spaces, community cohesion and participation while reducing costs for policies and avoiding gentrification processes. Nonetheless, results on the effects of increased green availability does not always have the expected outcomes. Droomers *et al.* (2015), for example, have analysed whether an increase in green space availability in deprived neighborhoods had a positive health effect, showing no short-term health benefits in the studied area. It is also important to keep in mind the residential self-selection bias, that is the tendency of people to choose the residential location based on personal needs and preferences (Boone-Heinonen *et al.*, 2011). Allen and Balfour (2012) concludes that it is less likely for people from deprived neighborhoods to live in green areas, while it is the opposite for better off people living in wealthier and greener neighborhood, who in turn experience better health. van Lenthe *et al.* (2005) demonstrate that environmental characteristics of the neighborhood not only affect the level of physical activity but also contribute to social inequalities. Several studies have shown a disproportionate distribution of such spaces in urban environment, benefiting the more well-off part of the population. Therefore, the presence and distribution of green spaces is nowadays seen as an environmental justice issue (Wolch *et al.*, 2014). Maantay (2002) looks at several studies measuring exposure to environmental hazards and linking race and socio-economic deprivation to poorer environmental quality.

1.4 Health in space: a multifaceted problem

The World Health Organisation defined health as *"a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity"* (World Health Organization, 1948).

Health is a relevant, complex and delicate subject to tackle. While governments make efforts to improve the well-being of the population, there are many mechanisms, more or less easy to identify, that come into play and that can alter it. Three main factors determine health and well-being: socio-economic environment, physical environment and personal characteristics and behaviour (World Health Organization, 2020). Most of health determinants are not or cannot be under direct control of the individual, that is why dealing with "health" can be considered a complex matter. Moreover, health is a very sensitive topic, and data is subject to strict rules and its use is limited by the legislation in order to protect the privacy of the citizens (European Parliament and Council, 2016). Therefore, oftentimes information on health is made available only after a cleaning, aggregation or anonymising process. Data gathering in health can be done through several research methods, depending on the research question. Research methods in health studies include qualitative, quantitative, analytical and descriptive approaches (Bowling, 2002). For simplicity, the terms "*subjective*" and "*objective*" will be used here to identify two common types of data to be found in the literature. The first are obtained through surveys and interviews, prior agreement of the respondent. In spite of the authorisation on the collection of the data, the use of such information can also be limited. Health data collected using these methods commonly cover a selected part of the population (a sample), and sometimes require a long time or several people for gathering and compiling the information. Objective measures of health, instead, rely on information gathered from healthcare specialist, hospitals, or insurance companies, for examples. With the evolution of technology, collection and storage of such data has improved, and it allows to have a wide span of information for the whole population. Nonetheless, such information is usually more difficult to obtain and its use and publication is also conditioned to some regulations. Obtaining data at a fine geographical scale and for a large study area enables researchers to investigate spatial trends and variations.

In the literature exploring the relationships between the environment and health, mostly subjective data has been used. Carter and Horwitz (2014) and de Vries *et al.* (2013) for example analyse the associations between health and greenness by using data on self-reported general health. Self-reported walking and physical activities have also been used to assess the effects of green spaces on increased physical activity and mortality (Coutts *et al.*, 2013; Lachowycz and Jones, 2014). de Vries *et al.* (2016), in an analysis on prevalence of mental health disorders and local green space availability, rely on self-reported information collected through interviews. Song *et al.* (2019) use Korean National Survey data to show significant protective effect of different indicators of green spaces on depressive disorders. Self-reported mood has been accounted for in order to estimate the restorative effects of green areas (van den Berg *et al.*, 2014). Tischer *et al.* (2017) use parent-reported wheezing to study the association of respiratory disorders and surrounding greenness and grayness. Questionnaires and interviews data have the advantage to enable researchers to conduct the investigations at the individual level, at a fine scale, and therefore go into detail concerning the potential factors affecting health. Using such data allows to capture some characteristics of the individual and of its surrounding environment, useful to best investigate the complex relationship between environment and health. Nonetheless, this type of data has the limit of being representative only of a sample of the population involved in the interviewing process and the results of the analyses cannot be generalised. Moreover it is possible to fall into self-reporting biases, and comparability with other studies on the same topic can be undermined due to differences in variables' cut-offs (Manea *et al.*, 2012).

The use of objectively-measured data permit to conduct analysis on a larger scale, as well as enables comparability. Generally, as previously mentioned, this type of data is protected and therefore is less commonly used. Objective data such as hospital admissions are principally used to investigate the associations between cardiovascular diseases or mortality and green areas (Pereira *et al.*, 2012; Wilker *et al.*, 2014). Crouse *et al.* (2017) rely on Canadian census data to analyse effects of greenness on cause-specific mortalities. Mortality data has also been recently used by Orioli *et al.* (2019) to investigate the effect of residential greenness on cause-specific mortality in Rome. Focusing on longitudinal studies, Rojas-Rueda *et al.* (2019) have found consistent inverse associations between greenness and all-cause mortality. Hospitalisation records have also been recently used by Alcock *et al.* (2017) to investigate associations between hospital admissions and green spaces and pollution in urban areas in England. Grazuleviciene *et al.* (2015) and Laurent *et al.* (2013) refer to medical records to look into associations between greenness and birth outcomes. Objective data such as medication prescriptions have also been adopted, although the use of this type of data is quite recent. The use of prescription data is further discussed in the following section (section 1.4.1).

Analysis on the associations between health and green environments is generally limited to a restricted area, and studies looking at the spatial distribution of a specific morbidity or mortality cause are very rare in the literature. This type of analysis can provide new perspectives and can contribute to inform public health and policy. Section 1.4.2 will have a closer look at spatial issues linked to health analyses.

All the studies presented until here have used measures of the surrounding environment to analyse possible associations with health. Environmental exposure has often been assessed around the home address or within the census tract of residence. Chaix *et al.* (2012) emphasises how non-residential exposure is key to the individual and is important to be accounted for to better assess environmental exposure. Daily commute to work, work and leisure environments, are part of individuals' daily life and can play a role in determining a person's well-being. Recent literature has investigated different and multiple exposures in the analysis of associations between health and the environment (Almanza *et al.*, 2012; Honold *et al.*, 2014; Burgoine *et al.*, 2015) in order to overcome limitations of single-exposure or residential exposure assessments. The use of GPS tracking and accelerometers has enabled researchers to collect data for multiple locations and therefore account for more factors in environment-health analysis (Chaix *et al.*, 2012; Perchoux, 2016; Tamura *et al.*, 2019). Collection of this type of data is only possible with specific material (GPS trackers, tablets, etc). As it is the case for subjective data, the acquisition of such data is restricted by privacy regulations.

Figure 1.1 summarises what has been presented in the previous sections, about how green and health are linked and how they are individually addressed in the literature. The scheme sums up the two approaches used in the literature in order to assess green on the one side, and health on the other side. Both green and health can be appraised through objective and/or subjective methods. Some examples of the methods used for the assessment are added to the scheme, but more methods exist (represented by the "..." box). In the center, the pathways linking green and health, as presented in the previous sections, are reported (opportunity for physical activity, physical and mental restoration, weather and pollution mitigation, and biodiversity).

1.4.1 Medication data

As previously mentioned, the study of association between environment and health using objective health data is less frequent than the use of subjective data. This choice can be due to difficulty to obtain the information, to lengthy process of data cleaning, or to anonymisation procedure that reduces the final available information, for example. This is especially the case for the use of data on prescribed medication. A few studies using such variable exists but the analyses of its association with environmental factors is still limited. Prevalence of medication prescription has been recently used by Aerts *et al.* (2020) to analyse associations between cardiovascular disease and residential green spaces in Belgium at a fine administrative scale. Results point to stronger prediction of medication sales due to socio-economic status rather than to greenness. Also Gidlow *et al.* (2016) have recently used medication prescription data. In their study in England, they look at the associations between cardiovascular and depression medication prescriptions and green spaces. In their study as well, they did not find an association between higher

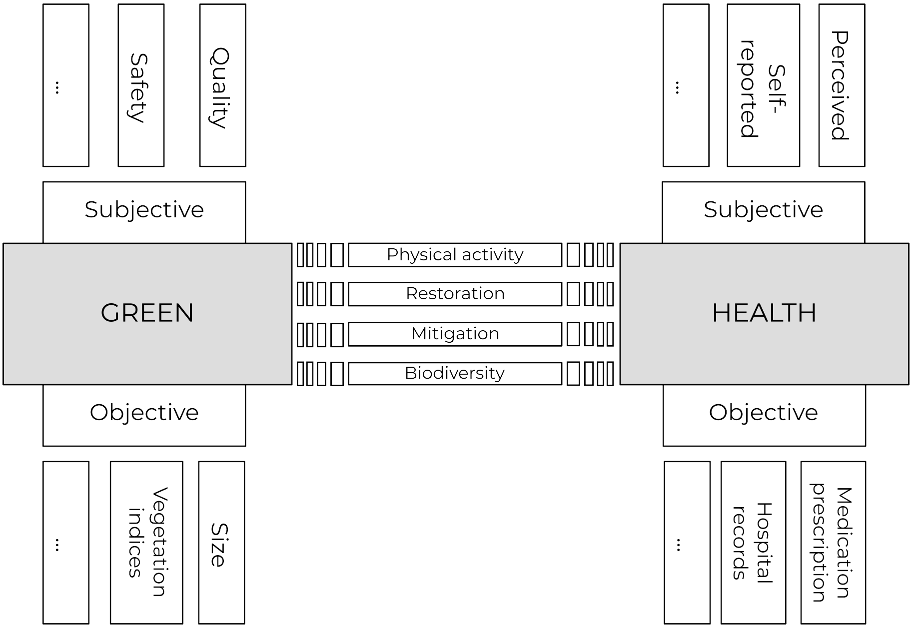


Figure 1.1. Approaches to green and health in the literature and pathways linking the two elements.

density of green areas and lower prescription or costs. Using the same type of prescription data, Taylor *et al.* (2015) reported associations between street tree-density and anti-depressants' prescription in London. Also Helbich *et al.* (2018b), found some associations with green space availability and decrease in medication prescription rates for antidepressants in one municipality in the Netherlands. They also define a threshold at which green spaces have the most beneficial effects on mental health. As for associations of green spaces with other types of health data, prescription data also presents inconsistent findings. It should be kept in mind that prescription data is only partially representative of the population health. In fact, as stated by Helbich *et al.* (2018b), such data only represent the part of the population who seek help. Moreover, prescription does not imply purchase nor consumption. Self-reported information on medication consumption has been recently used in the literature. Both Triguero-Mas *et al.* (2015) and Gascon *et al.* (2018) analyse associations between surrounding greenness and mental health using data on self-reported intake of tranquilizers or sedatives, antidepressants, or sleeping medication, along with data on perceived depression and/or anxiety. They report benefits or protective effects of some measures of greenness (surrounding green spaces) and self-reported consumption of medication, although results are not always statistically significant. Some caution should also be used when handling prescription data, as the same group of medication can be used to treat symptoms linked to different disorders. For example the R03 group (for respiratory disorders) is used to treat asthma if prescribed in the young group of population and Chronic Obstructive Pulmonary Disease (COPD) if prescribed to population aged 40 or older. Also the problem of co-morbidity exists: the concomitant appearance of two illnesses of disorders in the same person as, for example, is the shared risk factors of COPD and some cardiovascular diseases.

1.4.2 Spatial issues associated with health analyses

Spatial analysis not only aims at revealing spatial structures and organisations of some phenomena, but also at understanding the underlying processes that lead to those structures.

The most famous, pioneering and ground-breaking study using spatial analysis with health data is the study from John Snow ("the Broad Street pump" and the cholera epidemics). The English physician managed to detect the location of the cholera outbreak that hit hard the City of London in 1854. By drawing a grid to keep track of cholera victims, he was able to identify the source of the epidemics: the Broad Street pump, and demonstrate the geographical clustering of the cases around the pump itself. Since then epidemiology, namely the study of how often disease occur in a population and why (Porta, 2014), has developed further and numerous have been the studies looking at the relation between space and health.

While spatial analysis has extensively been used to explore and explain spatial variations in many disciplines (ecology, economy, and planning for example), it has not been the case for health studies, where this type of approach is still constrained. Analysis of spatial variations of health is only possible and significant when data is provided at a fine scale for a large study area. By using spatially aggregated data we may incur into the Modifiable Area Unit Problem and into ecological fallacy. The Modifiable Area Unit Problem (MAUP) is a well-known geographical problem, first formally defined by Openshaw (2009) in 1979. MAUP refers to the spatial boundaries, that are artificially created to delimit the space. We can think of administrative units, census tracts, national boundaries, that is to say all geographical partitions that can be subject to change but that have a practical use (Fotheringham and Wong, 1991). Considering four observations fixed in space, the change in the delineation of the boundaries has an effect on the spatial unit to which it is assigned. Statistical data is often collected and analysed at the level of these artificially defined units. The change in the shape of such spatial units, and the choice of the partition can therefore have consequences on the final results.

Another common problem when working with aggregated data is the so-called "ecological fallacy" (Piantadosi *et al.*, 1988). The ecological fallacy is a methodological issue that appears when performing ecological analysis, meaning when using data describing groups of individuals. The risk here is to infer individuals' behaviours or causality from information describing the group that they are part of. Ecological data are often used when data at the individual level is not available due to gathering data problem, privacy limitations or large scale studies. It is therefore not possible to infer that associations found at the group-level persist at the individual-level.

1.5 Green spaces and where to find them

The way green space is defined in some studies is partly due to the way green space itself is initially constructed and represented in landuse data. Landuse is used for multiple purposes: vegetation assessment, landuse and landscape planning, or to monitor landuse changes over the years, for example. Depending on the intention, different data can be used. Landuse data can vary depending according to the way they have been collected, their spatial extent and the resolution. Landuse information can be retrieved from the satellite images, the most known being SPOT and LANDSAT (Satellite Imaging Corporation, 2020), or can be developed by institutional agencies, as is the case for the Urban Atlas (European Union, 2011), or else it can be created by the community, as for Open Street Map (Open Street Map Community, 2020). With new technologies and scientific possibilities, accessing and using landuse data has become quite simple. Online it is possible to find numerous landuse data sources, and several online and offline mapping tools exist to conveniently analyse them. The proliferation of such data and tools, has made working and anlysing landuse data potentially accessible to anyone (Maantay, 2002). Landuse data is available at different spatial resolutions and at different scales (local, national or continent). Also, depending on the purpose of the database, the specification and number of classes, and therefore of discrimination of landuse types, can vary.

In this thesis four different landuse data sources have been used and compared: Corine Land Cover (European-level land cover data), Urban Atlas (municipality level, for over 300 European Urban Areas), Normalised Difference Vegetation Index (NDVI, landuse data derived from satellite images by the author), and a self-developed data source, for one Belgian municipality, partially based on OpenStreetMap data, and existing cadastral data.

Because of the political and administrative structure of Belgium, obtaining consistent and homogeneous landuse data has revealed some limitations. The three Regions (Brussels Capital Region, Flanders and Wallonia) have independent authorities for specific matters, among which the environmental data collection and management. Since each Region has its own rules for landuse data collection and codification, combining the three datasets proves to be an arduous project. While the Regions benefit each of data with a good level of resolution and detail of information, this could not be extended to the level of the whole country. For this reason, in the analyses at the national level, the European Corine Land Cover data has been used, while more detailed data sources have been chosen for studies at the local level.

1.6 Context of this thesis and health data

Healthcare in Belgium

As previously mentioned, the cost for healthcare in Belgium accounts for 10% of the Gross Domestic Product. Healthcare expenses have witnessed a 12% increase in 5 years (between 2011 and 2016), according to the European statistical office of the European Union (Eurostat, 2019). Public healthcare in Belgium is subsidised by both social security and a healthcare insurance paid by the citizens. Seven health insurances, called *mutuelle* in French or *ziekenfonds* in Dutch, exist in Belgium and citizens can choose freely to which insurance they want to subscribe. The part of unsubsidised care is covered by the citizens (Federal Public Service, 2018). In general the funds cover up to 75% of hospital, medical visits costs; around 20% of prescribed medication costs; and maternity and basic dental expenses. All people registered to a health insurance in Belgium (around 98% of the population) is entitled to the same subsidies, whether they are working, unemployed, retired or students. In 2018 Belgium ranked 5th for in the Euro Health Consumer Index (EHCI), a relative measure of healthcare performance based on over forty indicators (Björnberg and Phang, 2018).

Research projects

This thesis was developed in the context of two research projects funded by the Belgian Science Policy Office (BELSPO). The first project (GRESP-HEALTH, BR/143/A3) had the objective of assessing the impact of the natural environment (green and blue spaces) on specific morbidity and mortality causes in Belgium. While the initial aim of the project was to analyse both green and blue spaces, it has become clear that the two environmental features entailed a degree of complexity to properly define and measure them. Furthermore, the data available on blue spaces for Belgium is limited and not sufficient to completely grasp the effects of such landuse on health. Therefore for the purpose of this thesis the focus has been on exploring and understanding green spaces only. The Belgian Intermutualistic

Agency (IMA-AIM) provided aggregated data on medication reimbursed to citizens residing in Belgium enrolled in the National Health System (circa 98% of the population). The second project (NAMED, BR/175/A3) has the aim of understanding the associations between urban nature and mental health distribution. Individual data on health was obtained from the National Health Interview Survey (HIS) (Lauwers *et al.*, 2020).

Two types of health data were available and have been used in this thesis. First, data collected through a national Health Interview Survey (HIS) by Sciensano, a public health institution in Belgium. The HIS is administered to individuals and households, with the aim of *"assessing the health status of the population and identifying the main health problems as well as the determinants and behaviours that could influence them"*. The survey includes a sample of 10,000 people (3500 in the Flemish and Walloon Regions, and 3,000 in the BCR), from selected municipalities (Sciensano, 2019). Population is selected through a stratified sampling process. Since information contained in the survey is sensitive, its use is bound to a strict anonymising process, as illustrated in figure 1.2. As the scheme shows, the first step of extracting the relevant participants' information from the HIS and the creation of an alternative ID number for each participant is performed by Sciensano, who then sends these data to Statistics Belgium (StatBel) who is in charge of converting the data to be sent to the collaborating universities (UCLouvain and UHasselt) for further treatment (step 2). The conversion consists of the appointment of a new ID number, and the removal of all the identification details (national register number, home address, and the answers to the survey). Home addresses are converted to x and y coordinates, that are then provided to UCLouvain and UHasselt together with a new identification number. The Universities can therefore use the x and y coordinates to develop environmental indicators about the residential surroundings (step 3). ID, x and y coordinates are then sent back to StatBel for a second round of anonymising: ID numbers are changed back to match the ones provided by Sciensano and the x and y coordinates are removed. Sciensano, in step 5, gets the final set of data including the environmental indicators and the ID numbers to be matched with the HIS answers. They can therefore use this final database to run the analysis on the associations of some mental disorders and the surrounding environment.

The second set of health data used in this thesis is the one for medication reimbursed to Belgian citizens. For over a decade, the Belgian health security agency (IMA-AIM) has been systematically collecting data on medication reimbursement. The objective here is to analyse the medication reimbursement from a geographical perspective. The systematic collection of data from IMA-AIM enables us to have a (partial) view on Belgians' public health system. While over 98% of the Belgian residents are registered in the

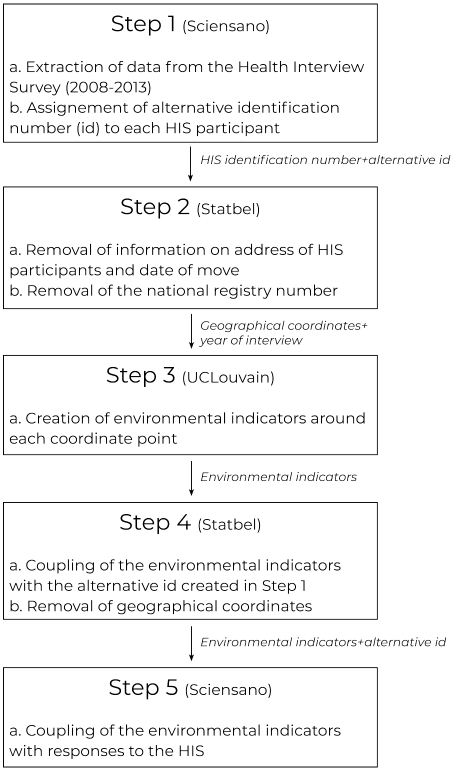


Figure 1.2. Anonymisation process for the Health Interview Survey (HIS) responses

health system, the data represents the "sick" population only partially. Of the population accusing any ill-symptom, only a part goes to a medical doctor to get a diagnosis. Of those who go to the medical doctor and get a prescription for drugs, only a part buys the medication prescribed, and of that part, not everybody asks for reimbursement for the medication. Therefore what IMA-AIM data is representative of the part of the population who asks for reimbursement. No information is available, nor can be provided, on the actual consumption of the prescribed drug. While information is lost in the path from patient to consumption, it is possible to notice how this data is as close as it is possible to get to how the individual benefits of the health system. For time reasons, within the GRESP-HEALTH project, and for this thesis, only aggregated data have been made available. Individual data can be requested to IMA-AIM following the authorisation of the privacy committee.

1.7 Outline of the thesis

The associations between green spaces and health have broadly been explored in the literature in the past decades. This thesis aims at taking one step back and investigate "green" and "health" more in depth, understanding how they are defined, assessed and used. Special attention is given to the complexity of the definition and interpretation of the two elements under examination. Green spaces and health are both the sum of a number of factors that come into play and that are hard to untangle.

It is clear that the relationship between human health and its surrounding (built or natural) environment is a key question of research, policy makers and planners. It mainly raises questions about the definition of the concept of green spaces and health, and of the processes that link the two together, also controlling for exogenous factors. The second chapter of this thesis explores, through a systematic review of the literature, the definitions of green spaces and how they are objectively measured and used across disciplines. Drawing from the conclusions of the first chapter, chapter 3 compares four landuse data sources on their objective similarities and differences in the representation of landuse classes such as green and built areas. The following chapter analyses, using two sources and types of health data (medication reimbursement and perceived health), whether the health outcomes are affected by the choice of the landuse data source, responding to the concerns of the sensitivity of the results when using different sources of data. The question of spatial variations of health are then tackled at the level of Belgium by means of medication reimbursement data. Chapter 5 explores the spatial patters of medication reimbursements for four groups of medications in Belgium. The following chapter then focuses on the prescription for asthma medications to pre-schoolers at the scale of three administrative levels. The last chapter (Chapter 7) summarises the findings and the contribution of this thesis. Limitations and implications for policy and further studies conclude this thesis.

Figure 1.3 provides a schematic overview of this thesis.

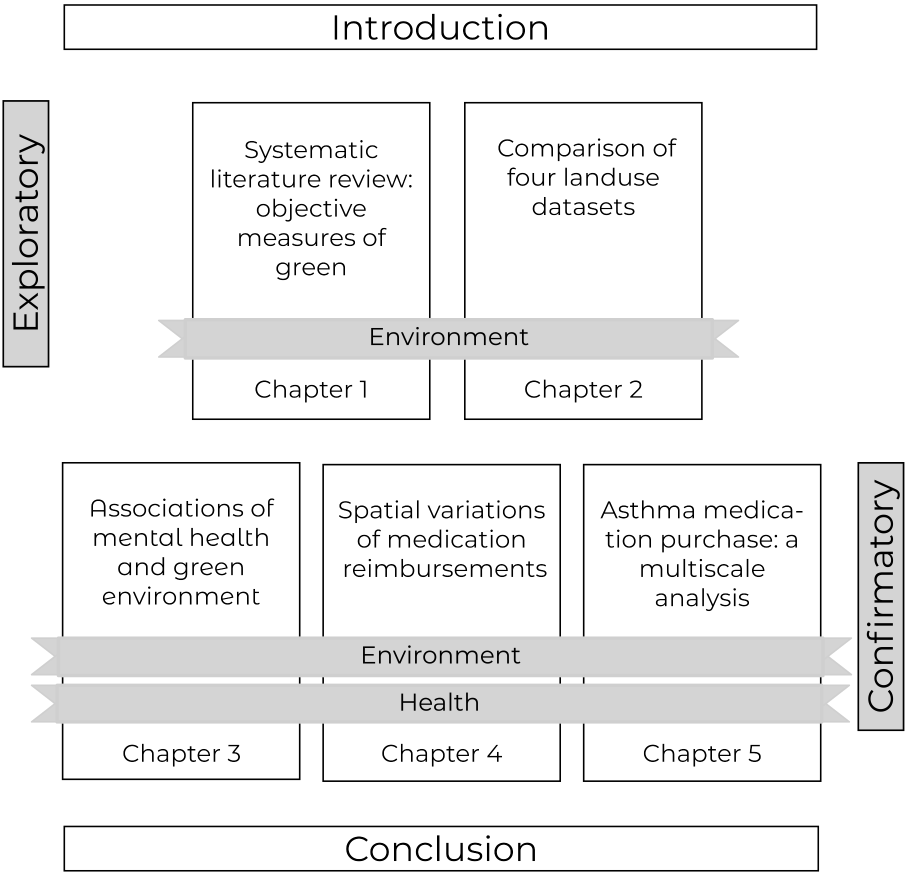


Figure 1.3. Outline of the thesis

II

Exploratory analyses

2

On the measures of "green": an interdisciplinary systematic review of objective methods

Over the past forty years, there has been a growing interest on the role of green spaces in urban and non-urban areas. This chapter aims at exploring and understanding trends and specificity in measuring and assessing green across disciplines by means of a systematic review. Following the PRISMA method, eighty-six relevant papers are retained for the analysis. Among the selected papers are included studies carried out in over thirty countries and five main disciplines. Overall, ten recurrent methods of objectively measuring green spaces are observed: proximity, accessibility, size, fragmentation, use, spatial pattern, shape, quality, view, and fractality. The selected papers also reveal a clear preference and interest for urban areas. A clustering analysis shows a clear tendency in the use of some measures for specific issues and contexts. We here reveal that there is no agreement across disciplines and authors on the definition, delineation, characterization of green. This discordance leads to questioning many causation effects discussed in the literature.

Chapter based on a research conducted in collaboration with Dr. L. Casas and Prof. I. Thomas.

2.1 Introduction

The past decades have witnessed a growing interest for environmental issues. Ecologists, economists, geographers, epidemiologists, and environmental scientists have shown (renewed) curiosity for the environment. The interest and debates have mainly been directed towards the role of green spaces. They attempted to assess its value, both economic and ecological, to identify its benefits for human and ecosystem health, or to criticize its distribution, in and outside urban environments. Notwithstanding the larger interest and concern on the role and importance of green environment, it is still difficult to find a consistent and common definition of what a green space is. Existing literature on "green spaces", uses the term for both natural and artificial green, both urban and rural, both accessible and not.

Over the years and across continents the studies tackling the topic have been numerous. Descriptions vary across studies, including or excluding certain features of green, even showing sometimes gaps among the different definitions. Gascon *et al.* (2015) in their systematic review of the literature on the benefits on mental health of green and blue spaces conclude that in the articles analysed there was a lack of homogeneity in the assessment methods. Also Taylor and Hochuli (2017) have reported a lack of consensus in the use of the terminology "green space", or "greenspace", in the literature and recommend to systematically provide a definition of the term to simplify the interpretation and progress in the research.

2.1.1 On the definitions of green

While some differences in the definition of "green" depend on the type of vegetation that is being considered, in other cases the definition revolves around some physical characteristics of the space itself. For the first case, for example, some studies include in the definition of "green" riversides and wasteplaces (Van Herzele and Wiedemann, 2003), overgrown gardens (Venn and Niemelä, 2004; Gupta *et al.*, 2012), or else forests, agricultural land and wetlands (Xu *et al.*, 2011), among others. Some authors also include those areas where vegetation has taken over dismissed buildings or areas awaiting redevelopment or construction (Nicol and Blake, 2000). Nonetheless, these areas are only temporarily vegetated (Rößler, 2008) and therefore cannot be considered as "green", or providing benefits, in the long term.

Concerning the physical characteristics that are used to define "green spaces", various definitions exist. For some, as for Maruani and Amit-Cohen (2007), "green" is included in the larger category of "open space", meaning all these areas that cannot be considered "built environment", and green areas are differentiated depending on the level of human intervention. Together with urbanisation, other features of green spaces are used to characterise them. Among others, it is possible to find distinctions based on quantity of green (Gupta *et al.*, 2012), on its size (Talen, 2010), on the location and services available. Some examples of differentiation of green spaces can be based on whether it is municipal, neighborhood, regional or theme park (Frumkin, 2001; Wendel *et al.*, 2011; Lee and Hong, 2013; Ibes, 2015).

Some studies instead consider that urban green spaces are all publicly owned and publicly accessible open spaces having a high degree of vegetation cover, introducing the question of "accessibility" to green (Schipperijn *et al.*, 2010). While Schipperijn *et al.* (2010) consider in the same group both the "planned" and the more "natural" green, Baycan-Levent and Nijkamp (2009) further differentiate between "natural" and "urban" green, the former being described as forests and agricultural land, and the latter being urban green as parks, gardens and water places. From an ecological perspective, the elements defining green spaces instead focus on the structural complexity of the vegetation, the connectedness of the natural areas, the number of trees or the depth of the leaf litter (McKinney, 2008; Philpott *et al.*, 2014). The different characterisations of green spaces go beyond the physical and measurable features of the landuse to consider, including the perception of safety, the usability, the quality, the target population. A variety of measures are therefore used to take into account the possible characteristics defining green spaces.

2.1.2 On measuring green

To a variety of definitions of green spaces, corresponds a (logical) variety in the methods used to measure it. Methods to identify and work on green spaces vary and include objective, subjective, and mixed methods. The latter combine objective and subjective methods (Carter and Horwitz, 2014), to capture different aspects of "green" (Leslie *et al.*, 2010). Objective methods measure the characteristics of the green space independently from how green is experienced (e.g. quantification of size, proximity, criminality rates…). This method makes use of green space data sources such as NDVI or of Geographic Information System. Subjective methods provide information on the perceived characteristics of green (e.g. feeling of safety, cleanliness or usability) and are based on questionnaires or in-depth interviews administered to (potential) users. This review is based on objective measures of green only. So far, a number of objective methods has been used to assess green spaces. A simple equation of size/person has been broadly used in the past to identify a comfortable green area per person (Haq, 2011). Nowadays, debates on whether this type of measure still fits modern planning and needs are numerous. To its advantage there is the applicability and comparability in very different environments.

Other methods used to characterize green spaces are not of such straightforward applicability. Proximity and accessibility, for example, are two very different concepts that are often confused and used as synonyms. Both consider distance to be the main factor influencing the use of green (Van Herzele and Wiedemann, 2003; Schipperijn *et al.*, 2010). While the first is often defined by the Euclidean distance, the second rather looks at network/road distance between two points. This assumes that green areas can be accessed always and from every point, simplifying the issue (Van Herzele and Wiedemann, 2003). At the same time proximity, for example, has different meanings throughout disciplines. In ecological studies, it is a measure of distance between green spaces. In disciplines where humans are the focus, proximity is the distance between, for example, residence and green spaces. Also on the distance to the closest green area there is no agreement or rule. For some (Boone *et al.*, 2009; Van Herzele and de Vries, 2012) a distance of 400 meters, equal to a five-minute-walk, is considered to be the point after which the frequency of use starts to decline. Instead, other studies revealed that people are willing to walk up to 1.15 kilometers to access, see, or experience a good quality green area (Natural England, 2010). The use of green areas is not only determined by measures such as proximity and accessibility but also by their number or their connectedness.

In addition to proximity and accessibility, disciplines such as ecology make use of measures (landscape metrics) that take into account aspects such as fragmentation or spatial pattern of the green areas, considering ecological change for the flora and the fauna. Both measures are typically used in landscape ecology and look into the relation between patches of green: while the first measures on the number and surface of the patches, the second focuses on how they are distributed in the landscape. Aware of the variety and complexity of existing definitions and measures of green spaces, we here look into creating a useful document of synthesis of the literature drawing from multiple discipline and therefore, possibly, different measures of green. We consider papers coming from different fields and having different study purposes of green measurements in order to obtain an interdisciplinary review of the existing measures of green spaces and their trends of use to make informed decisions. We further perform a network analysis based on the journals and the authors of the papers used in this study, with the objective to explore whether there is collaboration between authors on the topic.

This chapter is structured as follows: section 2.2 presents the methodology used for the systematic review, with the description of the selection process and the statistical analysis; section 2.3 presents the results of the work, followed by a discussion on the findings and a conclusion.

2.2 Methodology

2.2.1 Data collection

The bibliographic search was carried through two databases: *Pubmed* (medicine-focused) and *Scopus* (interdisciplinary scientific database). The selected keywords were: *"((measure OR assess* OR indic* OR metric) AND green AND (space OR area) AND (surface OR shape OR fractal* OR fragment* OR proximity OR access* OR ndvi) AND environment)"*. The choice of the keywords was derived from the literature. The aim was to keep the search as interdisciplinary as possible and account for diversity in terminology across disciplines. The inclusion of selected quantitative measure of green spaces aimed at steering the results, and not limiting them. The selected words were based on existing and known literature quantitatively assessing green spaces. The authors aimed at finding additional methods used in combination with the ones used as keywords.

The search was conducted on title, abstract and keywords. Results of the research were limited to scientific articles published in international journals, in English, using at least one of the following objective measures of green: proximity, shape, size, fragmentation, view, spatial pattern, access, use, quality or fractality. Articles using only subjective measures of green (i.e. questionnaires or in-depth interviews) were excluded. Results included papers published until October 2015.

Figure 2.1 summarises the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) results, a four-step process guiding to the selection of relevant articles (Moher *et al.*, 2009). The first step of the selection process was the "identification", which included the 1159 papers found through the keyword search. Results of the search were screened according to title and abstract pertinence to the research (step 2). The selection of papers on title and abstract was conducted independently by two researchers following the inclusion and exclusion criteria previously described. In case of disagreements, the results of the selection were compared and discussed. The aforementioned databases were combined excluding duplicates. Eligibility (step 3) was determined by a full-text assessment of the articles. Articles were deemed eligible if they presented at least one of the objective measures of green presented previously in this section (i.e. proximity, accessibility, etc.). A total of eighty-six papers (full list in Appendix A.1) responded to the whole selection criteria.

2.2.2 Characterisation of the selected articles

To answer the research question, all of the 86 selected articles have been further screened to extract some relevant information. These included

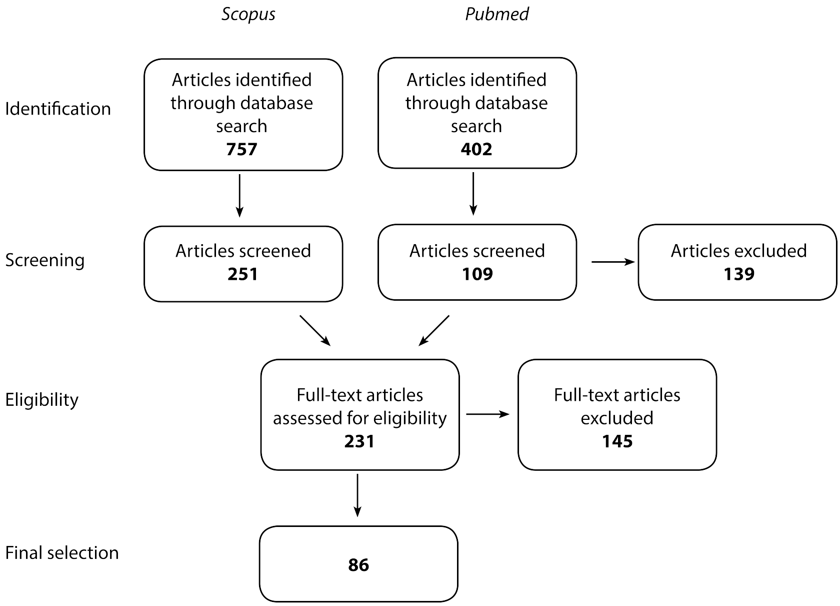


Figure 2.1. Selection of the 86 studied papers: PRISMA flow diagram

identification criteria (bibliographic database, authors, year and journal of publication, title), study focus variables (discipline, study population, purpose of the study), study area variables (country, region, location, urbanisation level), green identification and data source (green database, resolution of the data, measurement period, type of green, unit of measurement of green database) and measurement variables (including the family measures and specific method used to measure green). We end up with a characterisation of the eighty-six selected papers according to forty-two variables. The full list of the variables extracted and their description is available in Appendix A.2.

2.2.3 Statistical analysis

To assess the similarities between the selected scientific articles we used the variables described in the previous section (2.2.2). All the identification variables as well as the variable concerning the discipline of the were excluded for this part of the analysis. Such choice enables to obtain a clustering solely based on the characteristics of the study itself. In order to compute the distance between papers, all the categorical variables were transformed into dummy variables. Hamming distance (i.e. the number of changes needed for papers to be equal) was computed for each pair of articles on the set of variables chosen. Then, papers were clustered through a hierarchical clustering using Ward's criterion. This method allows for a grouping of the data minimising the losses related to the clustering. Ward's criterion groups the data in an iterative process starting from one-observation-clusters, and then combining each one with the closest observation. The number of clusters is automatically generated and not pre-set, allowing for a natural clustering. The advantage of such hierarchical methods lies in the output: a dendrogram represents the (nested) partitioning of the data, showing the distance and number of clusters of the data. To determine the ideal number of clusters for the data, we have compared the results of three methods: the Cubic Clustering Criterion (CCC), the pseudo F test and the elbow method. The first method suggested two clusters, the second five, while the third method suggests the clustering of the data into fifteen groups. To further investigate the collaboration of authors and permeability of the different measures of green, a network analysis has been performed. The network has been created using Cortext Manager (https://www.cortext.net/) through the "data analysis" and "network mapping" tools. The analysis has been based on the list of journals names of the 86 papers used in this study, as well as the full list of authors for each article.

2.3 Results

2.3.1 Methods for measuring green

In the process of the characterisation of the selected articles (see section 2.2.2), ten families of measures of green spaces have been identified: proximity, shape, size, fragmentation, view, spatial pattern, access, use, quality and fractality. Most of these families include one or more specific methods used to further characterise green, except for size, spatial pattern, fragmentation and view. **Proximity**, for example, can be assessed using five methods: buffers (e.g. Becker *et al.* (2012); Dadvand *et al.* (2014a); McMorris *et al.* (2015)), Euclidean distance (e.g. Coombes *et al.* (2010); Michael *et al.* (2014); Grazuleviciene *et al.* (2015)), grid (Grigsby-Toussaint *et al.*, 2011), nearest neighbour (Liu and Shen, 2014), block or self-rated distance (e.g. Wilker *et al.* (2014); Pietilä *et al.* (2015)). Methods to measure **accessibility** include road network (minimum distance) (Macintyre *et al.*, 2008; Potestio *et al.*, 2009; Dai, 2011), walking distance (La Rosa *et al.*, 2014; Houston, 2014; Saw *et al.*, 2015), street connectivity and slope (Chaix *et al.*, 2014), presence of barriers (Rupprecht and Byrne, 2014), time-cost distance (Kong *et al.*, 2007), inverse distance weighting (Burgoine *et al.*, 2015). **Shape** is assessed using area-weighted mean shape (e.g. Rafiee *et al.* (2009); Becker *et al.* (2012); Weber *et al.* (2014)) or landscape shape index (e.g. Coskun Hepcan

(2013); Olad Ghaffari and Monavari (2013). Assessing the **use** often recurs to frequency of use of the green area (Nielsen and Hansen, 2007), the time spent (Reklaitiene *et al.*, 2014), the location of the activity (Broberg *et al.*, 2013), or the potential use (Koppen *et al.*, 2014). **Quality** of green spaces is assessed through inspections (Hillsdon *et al.*, 2006), remote sensing images (Taylor *et al.*, 2011), residential preferences (Hill *et al.*, 2009) or surface of the area (Kruize *et al.*, 2007). **Fractality**, a measure of self-similarity pattern of a series of elements, is measured using Bovill's box-counting method (Liang *et al.*, 2012), area-mean-patch fractal dimension (Liang *et al.*, 2012), or perimeter-area fractal dimension (Coskun Hepcan, 2013).

The mosaic plot in figure 2.2 summarises the frequency of use of each of the ten families of measures (width of the column), as well as their relative use in combination with other measures (height of the box). We see a prevalence of use of measures such as proximity and accessibility, and size. These correspond to the most widely recognised measures related to green, followed by fragmentation and spatial pattern, commonly used in ecological studies. Looking at the first column, for example, we can see that proximity, accessibility, size and use are often used together. At the same time, if we look at the last column, we can see that fractality, although rarely present, is more commonly used in combination with other ecological measures such as (in order of frequency) size, shape, spatial pattern and fragmentation. An objective assessment of the quality of the green environment, as well as the view and the fractality of the area, are measures not often called upon.

The ten measures described in this section will be used throughout the chapter for further analyses.

2.3.2 Descriptive analyses

While no constraint on the year of publication was set in the search phase, the papers included in this review were published from the year 2000 up to 2015, with a peak of publications in 2014, with 29 articles. A total of 31 articles were present in both search databases (*Scopus* and *Pubmed*), 45 only in *Scopus*, and 10 only in *Pubmed*. The majority of the 86 papers were epidemiological studies (N=43), followed by urban and environmental geography studies (N=21), and ecology (N=11). The subjects/objects studied in the selected articles were humans (children, adolescents, adults, elderly, N=97), landuse types (forests, parks, general vegetation, N=62), and animals (bees, arthropods, N=6).

Table 2.1 provides an overview of the measures of green spaces used according to the study subject. A clear preference can be observed for the use of proximity and accessibility in studies involving humans, while the studies focusing on the landuse tend to use measures such as fragmentation, spatial pattern and size of the green areas. Nevertheless, size is frequently used with both human populations and landuse as study objects.

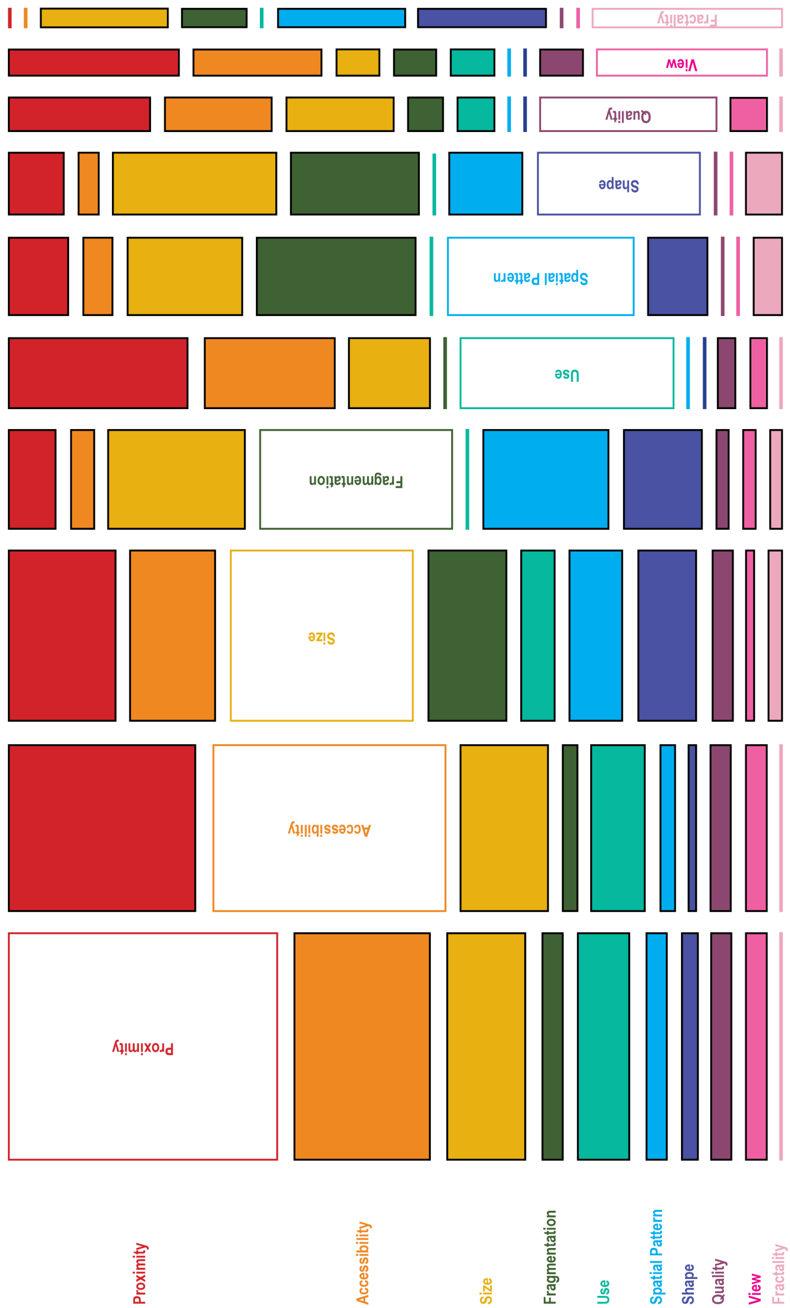


Figure 2.2. Mosaic plot of the assessment methods related to green used in the selected papers
Each measure is assigned a colour. Horizontally, the width of column represents the frequency of use of each measure (proportion); vertically, the height of each box represents the frequency of combination of the measure with the one of the respective column (proportion).

Table 2.1. Number of papers per family of measures of green by studied population (highest values in bold)

	Human	Landuse	Animals
Proximity	**43**	7	1
Accessibility	**28**	4	0
Size	10	**14**	1
Fragmentation	0	**14**	**2**
Use	8	1	0
Spatial pattern	1	**11**	1
Shape	1	6	1
Quality	3	2	0
View	3	0	0
Fractality	0	3	0

2.3.3 Study areas

Based on the level of urbanisation, we can categorise the papers in three main groups: highly urban (urban and suburban), mixed areas (combining urban, suburban and rural), and highly rural areas (focusing on suburban and rural only). Across all the selected studies, urban and suburban areas appear to have obtained the most interest, with forty studies on both urban and suburban areas and twenty-six on urban areas only. The trend to study urban areas is consistent across all continents (see table 2.2). Rural areas have only been considered in three of the selected papers (Bajwa and Vories, 2007; Becker *et al.*, 2012; Duguy *et al.*, 2012). Also considering the methods used to assess green spaces, as it is possibe to see from table 2.3, all measures are primarily used within highly urbanised areas. Measures belonging to the landscape metrics (fragmentation, spatial pattern, shape) are often associated to urban environments due to the concern on the ecological changes brought by an increasing urbanisation, and thus the destruction of natural habitats.

Table 2.2. Number of articles per level of urbanisation across continents (highest values in bold)

	Highly urbanised	Mixed areas	Highly rural
Americas	**19***	2	2
Europe	**23**	11	1
Africa	0	1	0
Asia	**14**	6	0
Oceania	7	0	0

*Includes one study conducted in both USA and China

Table 2.3. Distribution of the 86 articles by measures of green and level of urbanisation of the studied area (highest values in bold)

	Highly urbanised	Mixed areas	Highly rural
Proximity	**41**	8	2
Accessibility	**27**	4	1
Size	**20**	4	1
Fragmentation	10	**5**	1
Use	5	3	1
Spatial pattern	9	3	1
Shape	6	1	1
Quality	4	1	0
View	3	1	0
Fractality	2	1	0

Moreover, a trend between the purpose of the study and the location of the study area is noticeable. Studies conducted in Asia have a higher interest in landuse pattern and changes (e.g. Rafiee *et al.* (2009); Byomkesh *et al.* (2012); Rahman *et al.* (2011); Maimaitiyiming *et al.* (2014); Chi *et al.* (2015) ...), and in ecological issues in general, while American and European analyses focus more on epidemiological issues.

2.3.4 Data sources used for measuring green

A classification of the data sources used in the articles is performed. Three categories of data sources are identified: undefined, partially defined, defined/delimited. Within the **"undefined"** group there are data sources such as aerial photographs, orthophotos, or Google Earth images. Here the identification of "green" is the result of the interpretation of the images by the author. In the second group, the **"partially defined"** sources, are included satellite data such as NDVI (Normalised Difference Vegetation Index), Landsat, Quickbird, Ikonos, Aster and Spot images. In this case green areas are defined starting from information of percentage of greenery measured by satellites. This data is not always accurate due to mis-measurements of the satellites due to the presence of clouds or the light of the image (elements might have the same green percentage values despite not being vegetation), and of the pixels size leading to not very accurate information due to low image resolution. The last group that has been identified, the **"defined/delimited"** group, includes those data sources where green is pre-defined, by local authorities or institutions for example. Here, green spaces are then usually represented as patches or parcels. Typical examples are local, regional and national landuse maps, which might be based on remote sensing images. These data present landuse as defined by an external authority, following precise criteria and definitions.

Among the included studies, NDVI data and local landuse maps are the most often used reference data sources (24% and 21%, respectively). Various Landsat images, aerial photographs and orthophotos have been used for measuring green spaces, followed by Google Earth, Corine Land Cover or Urban Atlas even if in smaller share. Each data source only captures some aspects of green areas: either the density of vegetation or the physical limits of green spaces or their classification according to their characteristics and facilities. To provide for the partiality of information yielded by each database, some of the selected articles use up to five different green data sources, in order to have as much information and coverage of the area of interest as possible (Attwell, 2000; Byomkesh *et al.*, 2012; Coskun Hepcan, 2013; Chi *et al.*, 2015).

2.3.5 Buffers

In figure 2.2 we see that proximity is the most frequently found family of measures of green spaces. Within this family, buffers are the most recurrent measure used. The number of buffers and their sizes vary according to the purpose of the study: researches measuring proximity to green spaces for children or elderly adopt more frequent and smaller buffers than studies with an ecological focus. Buffers around home or school are often drawn to evaluate a more or less satisfactory proximity to green spaces for doing physical activities (Coombes *et al.*, 2010; Hanibuchi *et al.*, 2011; Chaix *et al.*, 2014; Janssen and Rosu, 2015).

Buffers are also used when looking at the association of green spaces with allergies or asthma (Fuertes *et al.*, 2014; Dadvand *et al.*, 2014a), or else to measure its impact on mental health (Annerstedt *et al.*, 2012; Amoly *et al.*, 2015; Huynh *et al.*, 2013; Markevych *et al.*, 2014a; Saw *et al.*, 2015; Triguero-Mas *et al.*, 2015). The definition of acceptable distance to green varies according to the studied population: children and elderly do not have the same mobility and perception of distances as adults. Table 2.4 shows, for each buffer radius, how often it has been used. When more than one buffer was used in the same study, we have noted whether it was used as 1st, 2nd, 3rd, or 4th buffer. It is possible to see that small buffers, associated to accepted walking distances, are preferred to large ones (>1000 meters), which are generally used in ecology studies.

2.3.6 Clustering analysis

From the results presented previously on the use of data sources and measures, a clustering technique (see section "Statistical analysis") was applied to check

Table 2.4. Number of studies for buffer radia according to buffer relative use in the study (1^{st}, 2^{nd}, 3^{rd}, 4^{th}). Highest values in bold

	1^{st} buffer	**2^{nd} buffer**	**3^{rd} buffer**	**4^{th} buffer**
<50m	7	0	0	0
100-250m	**10**	5	2	0
300-500m	**24**	**10**	3	1
600-1000m	5	6	7	2
1200-2000m	2	1	5	3
3000-5000m	2	2	0	1
>10000m	0	1	0	0

whether there are resemblances within the 86 selected papers. Results of the clustering are reported in form of a dendrogram in figure 2.3. Each article is identified by a number (from 1 to 86). The list of the articles with the corresponding number can be found in appendix A.1. Variables 1-7 were excluded from this analysis. From the dendrogram it is possible to identify two main clusters of articles. Cluster 1, rectangle on the left side, includes 19 papers mainly from ecology and geography. Cluster 2 instead, includes papers mainly dealing with the effects of green on health (epidemiology). **Cluster 1** is not divided into sub-clusters and primarily consists of papers available on *Scopus*, especially based on Asian urban and suburban areas with a focus on landuse patterns and changes. The preferred data sources of the articles part of this cluster are Landsat and aerial photographs (56%) and Google Earth (43%). The families of measures found in this cluster are mainly used to analyse the fragmentation, spatial pattern, size and fractality of green areas. **Cluster 2** is generally characterised by epidemiological studies, with a focus on the effects of the environment on human health using proximity (through buffers), accessibility, view, use and quality. These studies are mainly conducted in Europe.

On a lower height of the dendrogram it can be observed that Cluster 2 is further subdivided into four sub-clusters, called hereafter (2a), (2b), (2c), and (2d). Here, the study area and the methods of measurement of green are determinant in discriminating the selected papers and clustering them.

2a groups epidemiological papers available on *Scopus*, typically using local landuse maps and focusing on urban and suburban areas. Research topics include BMI and physical activity investigation, as well as informing policies. Frequently measuring proximity, but not using buffers.

2b also groups mainly epidemiological papers but often available both on *Pubmed* and *Scopus*. The study subjects are adults and elderly, in European urban and suburban areas and the interest is mainly on physical activity, and mental health. In this case as well there is a recurrent use of local landuse

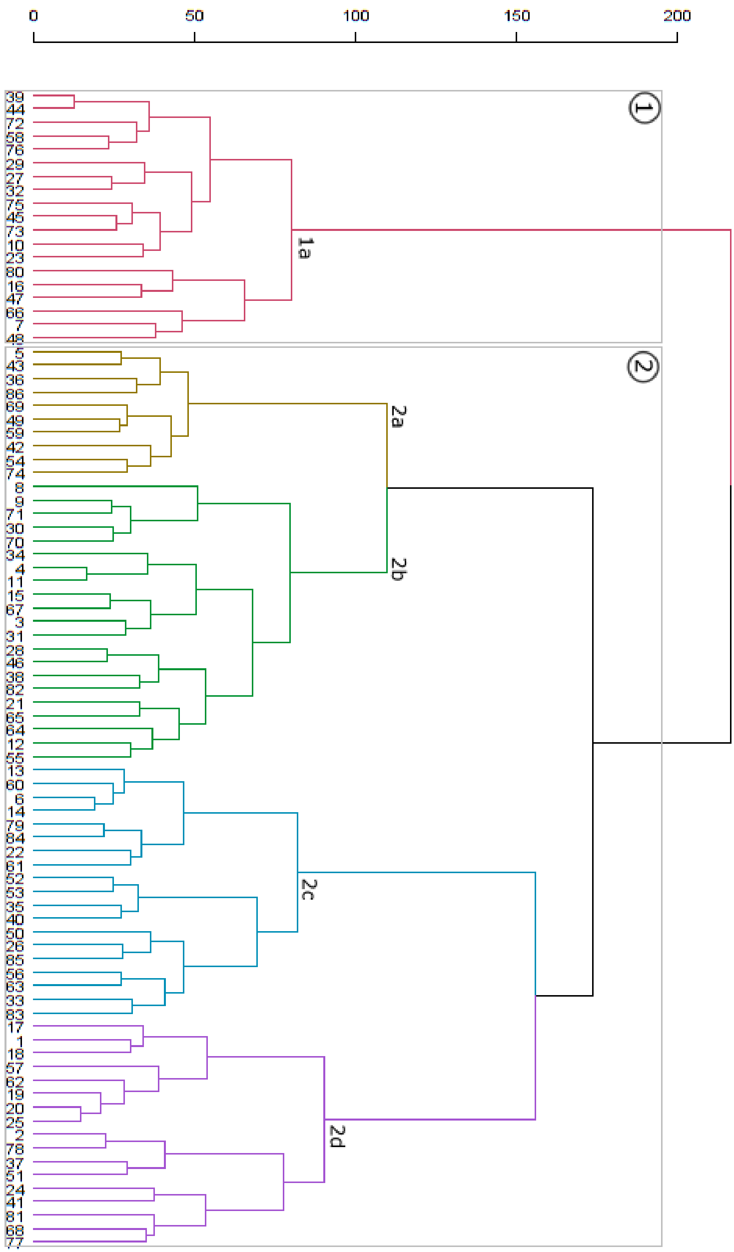


Figure 2.3. Clustering of the 86 selected papers. Numbers represent paper ID (full list in appendix A.1)

maps to measure accessibility, view, use, quality and proximity of green areas. Proximity is measured through buffers in this case.

2c clusters epidemiological studies with particular interest in physical activity and cardiovascular diseases on children and adults. These studies make use of NDVI. The resolution of the data is often very precise (<50m resolution). Proximity through buffers is again the preferred measure.

2d gathers epidemiological studies conducted mostly in European urban and suburban areas. The focus of these studies is on birth outcomes, mental health and access to green areas. Proximity again is one of the preferred family of measures. These articles also look into accessibility, size, use and quality of green areas.

The clustering analysis shows that, despite the multitude and diversity of assessment methods, data sources, and study locations, there is a pattern in their use. Each of them is actually embedded in a clear context and often combined in the same way to other elements, leaving little space to variations. The clusters we identify support the assumption that some families of measures are typically and strongly associated to specific disciplines.

2.3.7 Collaboration network

In order to further investigate to what extent the topic is cross-disciplinary and whether there is collaboration between authors tackling the issue, we perform a network analysis. Collaboration of authors is here proxied by co-authorship. We use here information on the journals of publication and the full list of authors who have dealt with the topic of measuring green spaces. The 86 articles used in this systematic review have been published in 40 different journals, from different disciplines. By using the full list of authors we look where they publish and how the journals are close in terms of publications. CorTexT Manager has been used to create and visualise the co-citation network. The variables used for the clustering are "Full list of authors" and "Journal". We can identify nine large clusters of authors that are connected (figure 2.4). In the figure, journals are identified by a triangle and each author by a circle. Author who have published together, based on the 86 studies used for the systematic review, are connected by a straight line. The first and largest cluster that can be identified is composed of epidemiology journals as: *"BMC Public Health", "Health and Place"*, *"Journal of Epidemiology and Community Health"*, and *"American Journal of Preventive Medicine"*. This cluster is the one that is the most connected to the other clusters. First, it is linked to the group composed by *"International Journal of Behavioral Nutrition and Physical Activity"*, *"Journal of Urban Health: Bulletin of the New York Academy of Medicine"* and *"Science of the Total Environment"* by one author. This cluster does not share any author with other clusters. Several authors from the first cluster have published on the same topic in *"Environment International"*. This last journal is part of an environment and health oriented cluster, together with *"Environmental Health Perspectives"*, *"International Journal of Hygiene and Environmental Health"*, and *"Scandinavian Journal of Public Health"*. These two clusters are in turn linked to publications in the multidisciplinary journal *"Plos One"* and *"Environmental Research"*. Two authors from this last cluster have also published in more geography-oriented journals as *"Agricultural and Forest Meteorology"*. This journal is in the same cluster as *"GIScience and Remote Sensing"*. These two journals are connected by one link to the cluster of landscape and planning journals: *"Urban Forestry and Urban Greening"*, *"Journal of the Housing and the Built Environment"*, and *"Landscape and Urban Planning"*. One author from this cluster is the link to the last connected cluster composed of *"Landscape and Ecological Engineering"* and *"Regional Environmental Change"*.

2.4 Discussion

Theories and studies on the relation between green spaces and human health or ecosystem services have been present in the literature for almost forty years, and the role of the environment is still a topical issue. But it is only recently that new technologies and availability of landuse data have made the topic more accessible. Scientific literature as well as political interests have brought to the attention the necessary and relevant ecological and health role of green areas, especially in an increasingly urban environment. Our study explores the scientific literature on green spaces, with a focus on the data, measures and contexts used to describe such landuse. The result is a review of objective methods of "green" measurements applied in different fields of the scientific literature. We do not filter according to specific purpose of measurement but rather take all relevant studies irrespective of the discipline.

The selection criteria that we use to reach our objective results in a collection of literature published between 2000 and 2015 only, as studies conducted before that did not respond to the selection requirements. Therefore our analysis, although broad, does not cover the entirety of the literature available on the topic. The presence of thirty-one papers available in both *Scopus* and *PubMed* is representative of how interdisciplinary the topic is. Within the 86 selected papers there is a large variety of measures of green spaces that we group a priori into ten families: proximity, accessibility, size, fragmentation, use, spatial pattern, shape, quality, view and fractality. This categorisation enables us to see a trend in the use of some families of measures within some disciplines (e.g. proximity and accessibility for epidemiology; shape, fragmentation and spatial pattern for ecology studies).

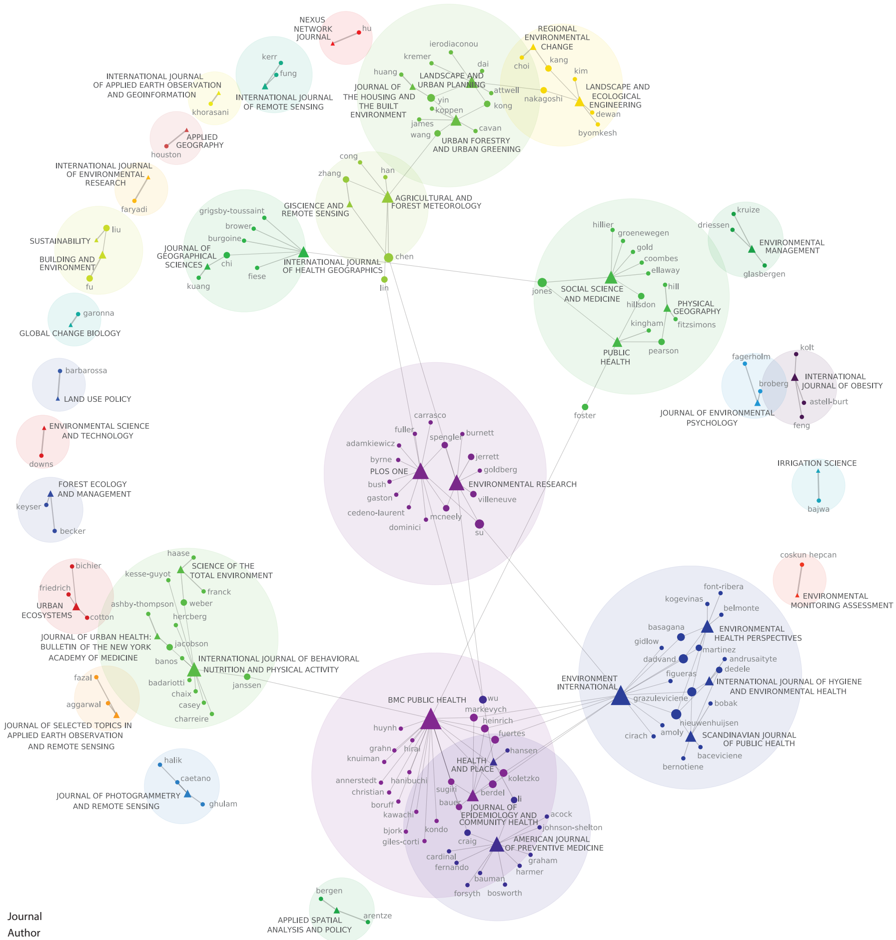


Figure 2.4. Network of co-authorship of the authors of the 86 selected articles

Nonetheless, the variety of measures included in most methods are many and the comparability of studies, although with the same purpose, is complex. It is therefore difficult to find homogeneity across studies. Moreover, definitions of the measures are often not provided in the studies leading sometimes to a mismatch between the purpose of the study and measure used. For example, Triguero-Mas *et al.* (2015) uses a measure generally meant for assessing proximity to measure the accessibility to green areas. Examples as such show that there is still no agreement on the definitions and measures used and maybe a lack of attention on the specific implications of each measure.

We notice how some measures are more frequently used than others. Proximity or accessibility are the most used methods of measurement across studies and disciplines, and in particular epidemiology and geography (Arentze *et al.*, 2009; Houston, 2014; Wendel *et al.*, 2011; Nutsford *et al.*, 2013). Other measures such as view, quality or use, are rarely used, probably because of the limited number of existing examples in the literature or the difficulty of an objective assessment of the measure (Hill *et al.*, 2009; Kruize *et al.*, 2007). We can also assume that the focus on "proximity measures" might be due to an interest from institutions and policy makers into the topic aiming at the further creation of "nearby" green areas and infrastructures in cities. This might have also shifted the interest from the possible benefits of green at a distance, its view, to the benefits of the experience of green for city development and spatial planning purposes (Attwell, 2000; Hill *et al.*, 2009).

We see how the choice of the measures is also determined by the study population. Studies on human beings make great use of proximity and accessibility (Coombes *et al.*, 2010; Dadvand *et al.*, 2014a), and studies on the landuse look into the surface and the fragmentation of the study areas (Coskun Hepcan, 2013); moreover a general preference for the use of buffers can be explained by the facility of drawing and computing them with Geographical Information Systems (GIS) and therefore are accessible to a wide public. On a larger scale, we can easily recognise a relation between the discipline and the family of measures used. Heterogeneity of results is also determined by the use of different and multiple landuse data sources making comparability among studies even more complex. To conduct measurements of green, we notice the use of a variety of landuse data sources, which differ in the information provided (e.g. different classes of landuse), the resolution of the data, or the purpose of the database. While some of them allow for potential comparability across countries as they are "standardised" (e.g. Vegetation indices, Landsat, ...) (Kang and Choi, 2014; Michael *et al.*, 2014), comparability of others is very limited as data are created by local institutions following local criteria or needs (Potestio *et al.*, 2009; Lovasi *et al.*, 2011; Amoly *et al.*, 2015). Also, the adoption of multiple data sources within the same study across the years, supports the idea that no dataset is a complete or self-standing source of information (Huynh *et al.*, 2013; Nutsford *et al.*, 2013). The difference in the content of the landuse data sources leads to the use of different measures or to different results using the same measure.

Generally, despite the differences in the source data, we notice a particular interest for measuring green in urbanised areas rather than in rural ones. In fact, out of the 86 selected articles only three have tackled the issue of measuring green in rural areas (Bajwa and Vories, 2007; Becker *et al.*, 2012; Duguy *et al.*, 2012). An interest that could be explained by an increasing shift of world population to urban and suburban areas, but also meaning a knowledge gap on the characteristics of green in rural areas. We also remark a different study interest according to the location of the case study. Papers based in Asia consider mainly ecology issues and the impact of landuse changes and patterns (Byomkesh *et al.*, 2012; Rahman *et al.*, 2011; Liu and Shen, 2014; Maimaitiyiming *et al.*, 2014). Studies in the Americas and Europe instead show a greater interest on the relation between green and human health, although the specific interest still differs in the two continents. While in the Americas the main focus is on the Body Mass Index (BMI) and physical activity (Li *et al.*, 2008; Grigsby-Toussaint *et al.*, 2011; Burgoine *et al.*, 2015), in European studies there is also a focus on mental health and birth outcomes (Annerstedt *et al.*, 2012; Grazuleviciene *et al.*, 2015; Markevych *et al.*, 2014a). The geographical difference of interests may be determined by specific local issues, as for obesity and physical activities in North America, or the change in landuses with increasing sprawl and pollution assessment in Asian cities.

This review shows the strong dependency of measure used to the context and discipline in which it is inscribed, presenting a variety of cases, proving to be a fundamental element for discriminating between disciplines and study purpose. Overall, we notice that very little attention is given to measuring green when dealing with its landuse changes and pattern in relation to animal survival and habitat. Our clustering analysis has revealed that there is a "natural" grouping of articles measuring green. Study areas, landuse data source, family of measures and methods used, are determinant elements to identify the discipline of pertinence for the majority of the studies.

The choice of using specific keywords related to objective measures of green spaces can be seen as a limitation to this review. While possibly steering the results of the bibliographic search towards articles containing the selected terms, the choice has enabled the authors to identify a number of methods to use and interpret the different measures.

This chapter also reveals that objective measures of green spaces have been used in several disciplines. In this study ten macro-areas of research have been identified, ranging from epidemiology, psychology, ecology, geography or planning, for example. When looking at the journals of the published articles and the full list of authors, it is possible to notice that while some authors have published research with objective measures of green only in one journal, there is a part that appears to be linked to more than one journal. These authors are the linking nodes to different clusters of journals (and in some cases also different disciplines). While not having included the "discipline" within the variables used for the clustering, this can be detected quite clearly in the network, meaning that authors collaborate with researchers that are close to them in terms of discipline. Although the number of authors who are linked to more than one journal is still small, it is still a sign that a dialogue between disciplines is possible and that measuring green is an interdisciplinary issue.

Systematic reviews on green spaces till now have mainly been devoted to the relation of the environment to human health (Van den Berg *et al.*, 2015; Gascon *et al.*, 2015; Lambert *et al.*, 2017). This work instead included studies coming from other disciplines with a focus on the measures used, regardless of the purpose of the study and provides a good overview of the state of the art concerning the measurement of green spaces within different fields. Our selection of studies includes only articles published in the last fifteen years, showing what are the current issues and perspectives on green measurements suggesting paths for further research, as for more extensive research in rural areas. Also, a recent and increased interest in perception of the surrounding environment and on its quality, might call for a study on the subjective measures used to assess green.

Eventually, a general definition of green space cannot be derived from the studies as it is very context-dependent, as well as suggestions on best measures or data sources to be used for specific studies cannot be provided, as these are determined by specific characteristics of the data availability, location, purpose of the study and population characteristics. The heterogeneity of the measures makes the comparability of studies complex if not impossible.

2.5 Conclusion

Measures of green appear to be a matter of difficult labeling within a structured process as the choice is determined by a series of factors/variables such as the data availability or type of study, as well as preferences of the author(s). Here we have offered a comprehensive overview of the existing measures and the context in which they are used. We show that choices on the measures to use are often linked to specific factors, such as study area or discipline of the author(s), with not much transfer of existing methods. We expect to see advances in the topic, with future studies further testing the use of new ways of quantitatively measuring green, and borrowing from other disciplines. The big variety of definitions and interpretation of green spaces as well of the measures used for the same study purpose lead to questioning the reliability of results linked to green measurements in a context in which the use of a "right" measure is arbitrary.

3

Landuse data sources choice: are they interchangeable? The case of Namur (BE)

This chapter builds on the conclusions derived from the systematic review presented in the previous chapter. Technological advances have lead to an increase of available data in many fields. This is also the case for spatial and satellite data. In the scientific literature it is possible to find a constellation of landuse data sources used. Here we test the equivalence of four landuse data sources: Corine Land Cover, Urban Atlas, NDVI-based and a self-developed dataset. Similarities and differences between four landuse sources on the same study area are highlighted using both descriptive statistics and more in depth analysis of the overlap and spatial structure of the data. A selected number of measures, including surface and fragmentation, are computed over the four databases to measure the difference and the variation in the results. Results show a significant variation, not only on total surface for some landuses, but also in fragmentation and spatial pattern of the landuse patches, mainly due to the level of aggregation of the data. This chapter looks at the characteristics of the four data sources, highlighting similarities and substantial differences. Outcomes provide a reflection on limits and advantages of the data sources allowing for a more informed choice and more accurate results.

3.1 Introduction

Interest for environmental issues in urban areas has increased in the past decades. It has become a hot topic in several disciplines, starting from ecology and geography, but also getting the attention of economists, epidemiologists, and sociologists, among others. Numerous are the studies and debates on the role of urban green spaces in particular: whether about their value, economic or ecological, or about its benefits, for human or ecosystem health, or about its presence/absence, within or outside urban areas. From the analysis of the existing literature quantitatively measuring green spaces performed in the previous chapter (chapter 2), it appears that the majority of the studies are conducted within urban areas.

Green urban areas are increasingly seen as an asset in cities: they are considered fundamental to preserve ecosystem health and biodiversity (Fortel *et al.*, 2014), as well as being generally considered beneficial for human health (Nielsen and Hansen, 2007; Nutsford *et al.*, 2013; Van den Berg *et al.*, 2015). Environmental data is frequently used to question the relation between urban environment and health in different populations. Many elements have been considered to impact health either positively or negatively: view, proximity, accessibility, safety, or availability, for example. Nonetheless, it should not be forgotten that green spaces in cities can also have negative externalities, since presence, and proximity, to green spaces impacts housing prices (Morancho, 2003; Kong *et al.*, 2007), sometimes enhancing socio-economic disparities (Dai, 2011; Dadvand *et al.*, 2014a). Nonetheless a common definition of green space among and within disciplines, as seen in the previous chapter, is still nonexistent (Taylor and Hochuli, 2017).

The lack of a common lexical definition translates in a great variety of accepted and used representations of green spaces on maps and landuse data sources. Nowadays, databases representing landuses are increasing in number and improving in quality. They are developed at local as well as at national scales by a number of different institutions, or directly by the individuals. Raw data is in fact more and more easily and freely accessible now, making it possible for the end user, with a basic knowledge of the process, to extract useful data for its research purposes. Markevych *et al.* (2017) highlight the variability of spatial patterns and frequency of vegetation pixels across three vegetation indexes (NDVI, SAVI or GRVI), and raise the question of the sensitivity of the associations between health and different vegetation indices.

As shown in the previous chapter (chapter 2) most of the literature objectively measuring green spaces belongs to epidemiological research. Within these studies, numerous are the criteria and data sources used to look at green spaces, even with researches having the same study-purpose. For example, out of 8 studies investigating BMI and proximity to green, by means of walking-distance buffer (1 in Australia, 3 in the USA, 3 in Europe, and 1 in Canada) (Li *et al.*, 2008; Potestio *et al.*, 2009; Wall *et al.*, 2012; Charreire *et al.*, 2012; Astell-Burt *et al.*, 2014b; Dadvand *et al.*, 2014a; Halonen *et al.*, 2014; Michael *et al.*, 2014), none of them uses the same source of data for measuring green spaces (among others: National Land Survey of Finland, Landsat TM images, IAU-idf, US National Land Cover Database, etc.). Although, in general, studies have proven the beneficial effects of the environment on general and perceived health (Van den Berg *et al.*, 2015), the comparison of studies and results is difficult as for the use of different data sources of landuse, different scales of study, or different definitions of green spaces. In order to compare findings and make results and policy-making effective and significant, a comparable and consistent approach is needed (McDonnell and Hahs, 2009). Le Texier *et al.* (2018), Rupprecht and Byrne (2014), and Taylor *et al.* (2011) have explored the effects of using different landuse data sources on green spaces availability, characteristics and quality.

The previous chapter (chapter 2 has highlighted the number of objective measures of green spaces used in the scientific literature. Of the ten measure presented earlier, four were used in this study to compare the similarities/divergences of the four landuse data sources. The choice of the measures to reproduce here has been based on the characteristics and the quality of the data. Measuring "accessibility", based on road network required information on roads, which was not available for all the data sources. To assess the "quality" of green, information on the services provided in each green space, for each landuse source would have deemed necessary. This also applies for measuring "view": information on height or type of vegetation is needed to estimate visible greenness. No measure of "fractality" has not been used in this study, as it provides information on the overall urban structure, rather than on the specificity of the green spaces.

Does the use of different landuse data sources really have an impact on measurements and results? Our objective is to explore and measure the difference in landuse representation through four landuse data sources to understand whether the choice of different data can lead to obtaining similar and/or comparable results. We here consider three typologies of green spaces (agriculture and pasture lands, forest, and urban green spaces), as well as other landuse types existing in the municipality in order to account for their relation. In the following sections we motivate and describe the chosen data sources, the methodology used, and the results obtained from the analysis. Flaws and advantages of the different data sources are then presented and discussed.

3.2 Data

3.2.1 The four data sources

Based on the literature, we identify the most commonly used landuse data sources for measuring green spaces across disciplines. Among them, we identify three that are freely available in Europe, and more specifically in Belgium: Corine Land Cover, Urban Atlas, and NDVI. To those three we add a self-developed dataset aiming at providing a more accurate representation of the landuses in the selected study area. The main features of the four data sources are summarised in table 3.1.

Table 3.1. Comparison of the main features of the four data sources

	CLC	UA	NDVI	Self-dev
Free online data	✓	✓	✓	x
Satellite image	Landsat & SPOT	Earth observation	Landsat	Landsat & Copernicus
Treatment needed	x	x	✓	✓
Predefined class.	✓	✓	x	x
Raster format	✓	✓	✓	x
Vector format	✓	✓	x	✓
Regular Updates	✓	✓	✓	x
Resolution	25ha/100m	0.25ha	30x30m	ca. 10m

The **Corine Land Cover** (CLC) is a European landuse database developed by the European Environment Agency (EEA) (European Environmental Agency, 2007) in 1985. It was created with the purpose of working on different environmental issues across countries using a homogeneous database divided in 42 classes (see appendix A.3 for the full list). The majority of the classes are associated to natural features (as opposed to artificial ones). The data is regularly updated. The data resolution is not very detailed, as the Minimum Mapping Unit (MMU) is of 25 hectares for aerial phenomena and 100 meters for linear phenomena. Therefore, only areas and segments having the minimum area and length mentioned above are detected, smaller elements are assigned to the closest "prevailing" landuse. This leads to a low precision of the shape and size of landuses, or even a total absence of some landuse patches, if smaller than the target MMU. The data we used for the comparison is the Corine Land Cover of 2012.

The **Urban Atlas** (UA) (European Union, 2011) is also produced by the EEA. It was created for urban areas to enable comparability in Europe: it includes 305 Large Urban Areas of more than 100,000 inhabitants. The landuse information is classified in 20 classes (full list in appendix A.3): the classes are for the majority linked to urban and artificial features. The level of resolution of the data is of 0.25 hectares for the urban fabric (Class 1), and of 1 hectare for the other classes. In spite of the fact that the focus is on the built environment, the vegetated areas are detected and properly classified. Unfortunately, this type of data is only available for urban areas, excluding peri-urban and rural areas that could also benefit of this definition of landuses.

The **Normalised Difference Vegetation Index**, better known as NDVI, is an index of vegetation density. This information is derived from remote sensing images, such as the Landsat images, and can be computed for the whole world. NDVI is also used to detect not only the presence of vegetation but also its state of health. The index can be computed by any user applying the formula $[(NIR - Red)/(NIR + Red)]$ using the red (Red) and near infrared (NIR) bands of the preferred satellite images. The result is a continuous value, ranging between -1 and 1, that defines the vegetation level for each pixel of the image. Positive values are associated with vegetated areas, while negative values are associated with water bodies, values around 0 are usually associated with built areas but also rocky, and sandy areas (Weier and Herring, 2000). In general, the cut-offs can be identified at 0-0.2 for soil, 0.21-0.03 for shrubs and grassland, 0.31-0.6 moderate vegetation, and the values $\geq$ 0.6 represent dense vegetation canopy. These thresholds are indicative, and can and should be adapted according to the image quality (snow, clouds, season). Because of the possible adaptation of the thresholds to the study area conditions at the time the images were taken, the thresholds can vary between different studies. Although Landsat images have a good resolution, 30x30 meters, the use of pixels instead of polygons decreases the level of detail of the final dataset. In our case, the data for the NDVI has been derived using two Landsat 8 images (February 2017 and July 2017), and the images have been corrected for sun reflectance. While NDVI is usually treated in the literature as a continuous variable, here it has been classified to enable comparability with the other landuse data sources selected for this study. For this study, cut-offs for defining the six landuse classes, were initially based on the thresholds presented earlier, and then adjusted manually using visual assessment of the area to improve the accuracy. The NDVI for the city of Namur was comprised between -0.6 and 0.49. The thresholds defined for the six classes were then defined as follows: (-0.6; -0.3) for water bodies, (-0.29; -0.015) for built-up areas, (-0.014; 0.07) for roads and infrastructures, (0.07; 0.12) for green urban areas, (0.13; 0.25) for agricultural areas, and (0.26-0.49) for forests.

The **self-developed database** (hereafter *"self-dev"*) was created by the author. The aim is to provide an accurate representation of landuses and features in the study area, to overcome at least some of the limitations of the other data sources. It was developed to answer to some of the critics raised over the previously presented data sources. To do so, Landsat images and Google Street View have been used in order to draw polygons of the city features (houses, parks, industrial areas, parking etc.). Information on the built up elements has been retrieved from the Belgian Cadastral Administration (Administration du Cadastre de l'Enregistrement et des Domaines, 2005), road network data was derived from Open Street Map (OSM) (Open Street Map Community, 2020). A specificity of the *self-dev* landuse dataset, over being developed at a very fine resolution, is the detection of private green spaces, for example. Cadastral data and Open street Maps data on roads have been used to identify built-up areas (residential, industrial and commercial) as well as to define the roads within the municipality. This data has been taken as such, as it is already available at the finest level required for the analysis. Cadastral information has then been split into three typologies of buildings: residential, industrial and commercial, based on surface and Google Street View images. All other features have been classified following a visual identification, based on the scheme in figure 3.1: first identifying whether the element was artificial or natural, and then whether it was a linear feature or an areal one. These two steps helped narrowing down the classification options. The full list of the initial classes created is available in the appendix in the section A.3. The final classification of the polygons into six classes has been an iterative process, aiming at obtaining a comparable group of classes for the four databases studied here. Polygons have been assigned to the following six classes: built up, road and rail network, green urban areas (including private green areas), forest, agriculture and pasture, and water bodies.

To make the comparability possible, the four data sources have been homogenized in a raster format. The transformation of Corine Land Cover, Urban Atlas, and the self-developed map from vector to raster, with a 30x30 meters grid, enables for a cell-to-cell comparison of the maps.

3.2.2 The study area

The municipality of Namur, of approximately 175 km^2 and with a population of circa 110,000 inhabitants, is located in the Walloon Region, and is situated 63 kilometers south of the capital city Brussels (see figure A.1 in appendix). We choose this city as study area for its characteristics: it includes urban, suburban and rural features, as well as a diversity of landcovers. The rivers Sambre and Meuse cross and meet within the city. Moreover, several landuse data sources are available for the study area, which enables us to perform our comparison study. The relatively small size of the municipality enabled for a manual digitisation of the area, as well as providing a variety of landuse types interesting for the purpose of this study. Figure 3.2 shows the study area of Namur as it is represented in the four selected data sources, after having been reclassified into six classes.

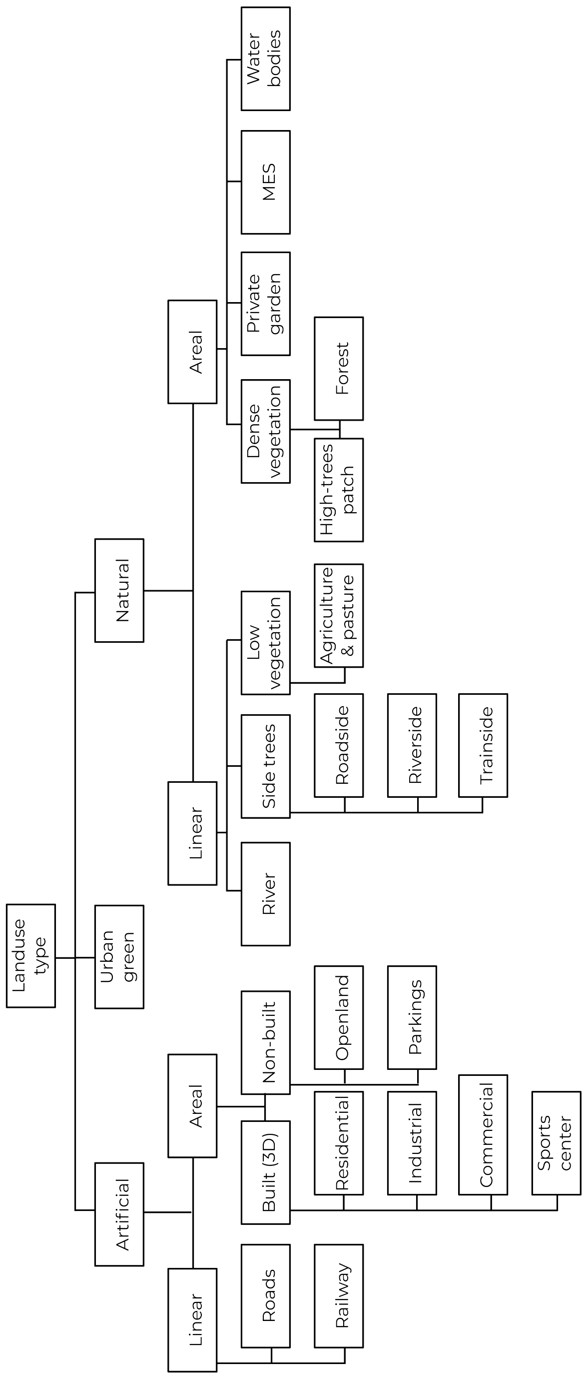
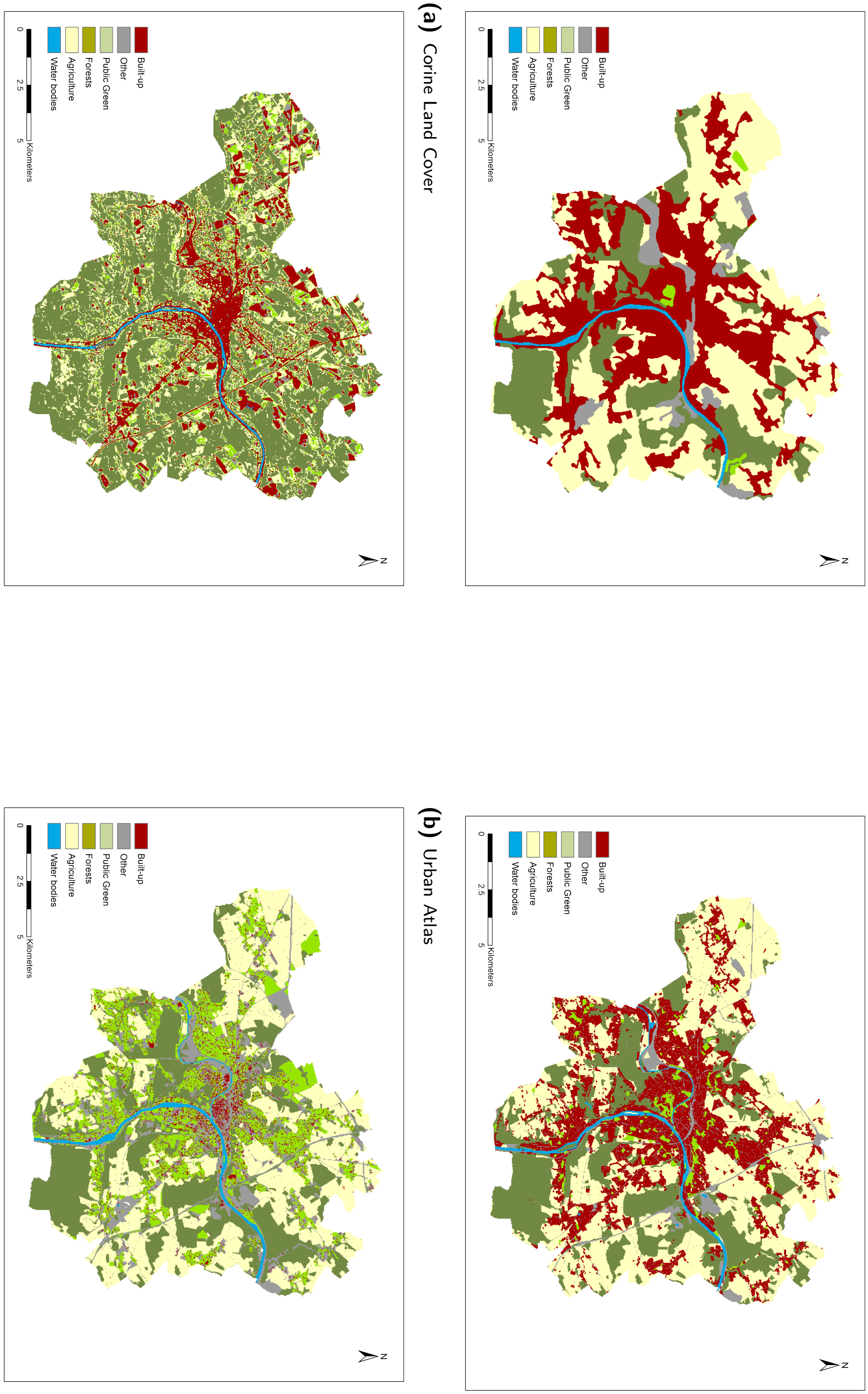


Figure 3.1. Classification process scheme used for the creation of the self-developed map



(a) Corine Land Cover

(b) Urban Atlas

(c) Normalised Difference Vegetation Index

(d) Self-developed map

Figure 3.2. Representation of Namur by means of the four selected data sources using the same classification (6 classes: Built-up, Other, Public Green, Forest, Agriculture, and Water bodies)

3.3 Methods

The first step in the comparison is a visual assessment of the differences of the four data sources. This approach is a quick method to detect the main similarities and differences between maps. Although it does not provide a quantifiable and reproducible output, compared to other methods, it enables the user to account for different elements of the image at the same time (Visser and De Nijs, 2006). This approach is complemented by descriptive statistics and a cell-to-cell comparison. The first informs on the part of landuse allocated to each category, on the average distance to the nearest patch, on the spatial distribution of the patches; the second enables for a cell-to-cell comparison of the maps.

3.4 Results

Differences in landuse share and map representation can already be detected with a simple visual assessment (figure 3.2): the frequency and the extent of the different landuses in the four maps already depicts a different structure and image of the same municipality. A statistical analysis further shows the differences and similarities among the maps.

Table 3.2 summarizes the proportion of land occupied by each landuse and the corresponding number of patches. A patch here is defined as a homogeneous area that differs from its surrounding. Analyses on the four maps indicate that over half of the municipality of Namur is covered in vegetation (CLC 62%, UA 67%, NDVI 79%, Self-developed 79%). Built up and residential areas are subject to bigger variations: while Corine Land Cover presents a high percentage of built environment (32.2%), the Self-developed map gives a very different representation of the municipality, with only 3.4% of the total surface occupied by buildings. NDVI reveals that half of the surface of the municipality is covered by forests (50.4%), while the other data sources show a presence of forests close to one fourth of the total surface of the municipality. In contrast, water bodies appear to be consistently detected and defined across the four data sources (min 1%, max 1.4%). The total area dedicated to infrastructures (other landuses), and to urban green (both public and private green areas), is larger in the self-developed map.

The number of patches for each class of landuse show a consistent trend, emphasizing the level of aggregation of the data. In fact, the lower the number of patches and higher the total class surface, the higher is the level of aggregation of the landuse. As expected from a first visual assessment of figure 3.2, Corine Land Cover has the least number of patches overall (139). Urban Atlas and the Self-developed map have 4399 and 7571 patches respectively, while the NDVI map represents a more scattered landuse distribution (with over 16,000 patches). The higher is the number of patches, the higher is the level of fragmentation of the different landuses.

Table 3.2. Percentage of each landuse and number of patches for each of the four data sources

	CLC	UA	NDVI	Self-dev
Built up	32.2 (26)	25.1 (417)	14.3 (1370)	3.4 (2916)
Roads and infrastructure	4.5 (14)	6.3 (3497)	5.2 (4058)	15.9 (3324)
Urban green	0.7 (6)	2.4 (141)	8.7 (5255)	17.9 (713)
Forest	20.9 (38)	24.5 (141)	50.4 (1089)	27.05 (464)
Agriculture	40.3 (54)	40.2 (189)	20.35 (4706)	34.5 (147)
Water	1.4 (1)	1.4 (14)	1 (50)	1.2 (7)

Each value represents the part of the total landuse occupied by each class (in %). The value within brackets indicates the total number of patches for each class (absolute number).

The level of aggregation of the patches and their compactness has been further assessed through a measure of like-adjacencies. The Patch Like-Adjacencies (PLADJ) is based on the equality of neighboring cells, meaning that it summarises the adjacencies of cell of the same class to each other. If a cell of class "1" is surrounded by cells of class "1", then its PLADJ value is high. If instead the cell of class "1" is surrounded by cells of different classes, then its like-adjacency value will tend to 0. This indicator is affected by the shape of the patch: a patch with a circular or squared shape will have a higher number of like-adjacencies than an elongated patch. This indicator, in some way, is also representative of the compactness of the patches. The PLADJ values range between 0 and 100, as they represent the ratio of the number of cells of the same class over the total number of neighboring cells, multiplied by 100. Table 3.3 summarises the results of like-adjacencies for the six landuses for each landuse data source. The results of this analysis reinforce the expectations derived from the visual appraisal and also the outcomes of the previous analysis. The high values of like-adjacencies for the CLC dataset confirm the high level of compactness and aggregation of all the classes. Conversely, NDVI has lower values of like-adjacencies, confirming a more scattered representation of the landscape. Across the four landuse data sources, forests and water bodies are the two landuses that always have high values of PLADJ (>84.9% and >73.9%, respectively), appearing as a compact landuse area. The other landuses are less similar among each other, with very large differences in percentage. Results of the built-up are in line with the results of the total percentage of landuse, suggesting an aggregation and approximation effect of landuse data sources such as CLC and UA.

The Euclidean Nearest Neighbor (ENN) is a measure of the concentration, or dispersion, of the patches. Measured in meters, the ENN provides information on the average minimum distance between patches of the same

Table 3.3. Percentage of like-adjacencies of pixels of the same class for the four landuse data sources

	CLC	UA	NDVI	Self-dev
Built-up	94.3	81	71.6	23.1
Roads and infrastructure	93.1	44.6	28.6	58.9
Urban green	88.7	72.9	33	70.3
Forest	94.9	90.9	84.9	88.8
Agriculture	94.7	90.2	49.4	90.6
Water	86.3	77.6	73.9	80.7

class. The interpretation is intuitive: the smaller the distance, the closer and more compact the patches are; the higher the average distance between patches, the more dispersed the patches of the same class are. As expected, the few patches within Corine Land Cover are generally further away from each other compared to the patches of the other data sources.

Table 3.4. Average distance (in meters) to the closest patch of the same type for each landuse class, and standard deviation

	CLC	UA	NDVI	Self-dev
Built up	409.1 (341.3)	121.4 (121.8)	103 (69.3)	85.6 (54.1)
Roads and infrastructure	507 (575.8)	87.4 (37.3)	85.4 (41.2)	81.2 (36.6)
Urban green	2703 (2514.6)	279.1 (313.2)	77 (31.3)	83.5 (81.8)
Forest	298.7 (266.2)	143.5 (145.6)	78.4 (29.6)	111.1 (93.6)
Agriculture	196.6 (175.3)	115.9 (111.7)	69.7 (22.7)	135.5 (134.9)
Water	NA	399.1 (726.7)	417 (867.9)	968.8 (1224.1)

We further perform a cell-to-cell comparison in order to further investigate similarities and differences between the maps. The self-developed dataset is used as reference for pairwise comparison. The choice of the reference map is based on data accuracy. In a pairwise comparison of maps, the use of kappa statistics enables us to quantify the similarity between two maps. The value of Kappa (hereafter κ) accounts for pixel location (κLoc) and histogram distribution ($\kappa Histo$) similarities (κ = κLoc * κHisto). κLoc is sensitive to differences in location of pixels for each landuse class, while κHisto measures the frequency of pixels for each class. Both κLoc and κHisto are based on the agreement between two maps, corrected for the fraction of agreement that would be expected by a random distribution of the pixels (Visser and De Nijs, 2006). In general, κ results can be interpreted as follows: values between 0 and 0.2 can be seen as none to slight agreement, 0.21 to 0.4 can be considered to be a "fair" agreement, if the values are between 0.41 and 0.6 the agreement is moderate, between 0.61 and 0.8 the agreement is substantial, and values of κ ranging between 0.81 and 1 are a sign of an almost perfect agreement (McHugh, 2012). Clearly the interpretation of the agreement has to be based on the purpose of the study. A medical trial with a non-agreement percentage of circa 40% ($\kappa = 0.63$, for example), might not be cannot be considered as having a "substantial" agreement. For the purpose of this study, the interpretation of the thresholds described earlier can conveniently be used. Table 3.5 shows the variations of the agreements between the maps. NDVI presents the lowest values of κ, meaning that NDVI is the data source having the lowest agreement value with the Self-developed map (0.15 = none to slight agreement), while the Urban Atlas is the one with the highest agreement (0.55 = moderate agreement). In particular, the similarity of the two maps is noticeable when looking at the congruence of pixels location. The κLoc for the Urban Atlas has a very high value (0.85), meaning an almost perfect agreement.

Table 3.5. Results of the κ statistics for a pairwise comparison of the maps

	CLC	**UA**	**NDVI**
κ	0.35	0.55	0.15
κ**Loc**	0.63	0.85	0.28
κ**Histo**	0.56	0.64	0.55

As previously mentioned, the literature often uses measures such as proximity or accessibility to assess the association between human (proxied by their residential location) and the green environment. Here we use a commonly adopted measure of proximity: we create a 300 meter buffer around the built-up centroid, as a proxy for the place of residence. We compute the percentage of residential patches having a green area within the buffer. Our results show little proximity to green using the Corine Land Cover data source: only 1 built-up area out of 20 is considered to be at "walking distance" to a green area (within a 300 meters buffer). The Urban Atlas, made of smaller and more fragmented patches, has 58% (1036/1785) of its residential areas in proximity to a green area. NDVI and the self-developed map have over 99% (2040/2040 and 47,559/47,642, respectively) of their patches in close proximity to a green space.

Lastly, figure 3.3 represents the pixels that are assigned to the same landuse across the four maps. A pixel in a darker shade of grey is a pixel that always belongs to the same landuse class in all the maps, meaning that the landuse converge in all the maps. On the contrary a light shade of grey is given to those pixels which are associated to different landuses. Agricultural and forested areas are the landuses most consistently located in the four data sources, represented by the most compact and dark set of points. The built-up area, in the centre of the municipality, north of the point where the Sambre and the Meuse river meet, is also characterised by a high convergence of identification throughout the four data sources.

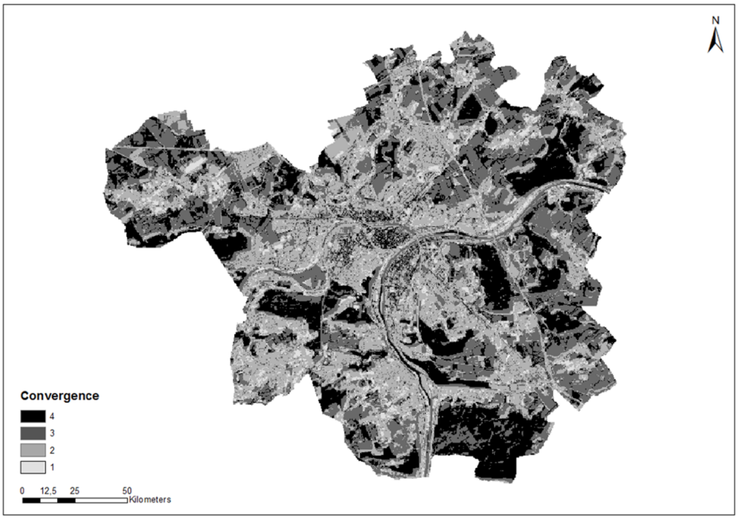


Figure 3.3. Convergence of pixels from the four data sources. Pixels of the same class, located in the same position have higher values (and are represented in darker shades of grey)

3.5 Discussion

This study explores similarities and divergences of four landuse data sources, in order to understand whether datasets can be interchangeable for measuring green. This study has been motivated by the variety of the number of sources of landuse data used in the literature for spatial analysis in different fields of research, and in particular when studying green spaces, as seen in chapter 2. No formal rule or set of recommendations exists for the choice of the most suitable data according to the study purpose and study area. The lack of a set of recommendations hinders the comparability and generalisation as well as have repercussions on the accuracy of the results. For example, analyzing residential proximity to some green spaces using data sources developed with requirements (scale, aggregation, landuse definitions) means that the elements considered differ, and this may undermine the solidity of the results (Markevych *et al.*, 2017).

In this study, we select four data sources developed by different providers to compare the landuse representation and definition over the same study area.

Conversely to what has been done in the other studies comparing different data sources, we here include man-made landuses such as buildings and urban infrastructures (Rupprecht and Byrne, 2014; Taylor *et al.*, 2011; Le Texier *et al.*, 2018). We choose three commonly used data sources, and we developed the fourth in order to use data at finer scale (including private gardens). Namur, the selected study area, has been chosen for the different levels of urbanisation it includes (urban, peri-urban and rural), as well as different typologies of green spaces (urban, forest and agricultural areas). We use several measures to compare the data.

Descriptive statistics of the share of each landuse in the data sources and a visual appraisal show us that differences are important. Landuses that are characterized by large patches, such as agricultural areas, are overall consistently represented in the four data sources. Landuses with smaller patches are instead subject to higher variations in representation, as they are generally joined to the closest and larger landuse patches. The urban development of the municipality of Namur is characterised by a dense city center, with blocks of houses with a private garden or an internal open space, and by a more sprawled peri-urban area with semi-detached and detached houses, with private gardens or agricultural areas. Small green spaces and private gardens are generally missing in landuse datasets due to resolution size. In fact, small green spaces and private gardens, due to their size they are generally associated to the closest and more prevalent landuse type. In urban areas, this phenomenon is common when looking at built-up: private or small public green spaces are usually surrounded by buildings (larger in size and more dense).

Corine Land Cover is a large-scale landuse data source that conveniently represents land cover in Europe. At a fine scale, it presents highly aggregated information. The lack of precision and detail leads to an approximate representation of urban environments (level of accuracy is higher in rural areas where patches are generally larger). The low level of resolution, or the high level of aggregation, in the representation can be observed when comparing the share and the number of green patches with the other landuse data sources. Due to its high level of aggregation, it provides an image of a very compact and non-fragmented municipality. When looking at a measure of proximity to green areas from the residential location, it appears clearly that Corine Land Cover is not a fit data source for such type of analysis, as it overestimates built up areas and underestimates green urban areas. Overall, when looking at the different measures considered in this study, CLC appears to be the most dissimilar data source compared to the other three analysed here. Such difference in the quality of the information provided by the data source is justified by the purpose for which it has been developed, namely the representation of the European land cover. The inclusion criteria of the landuse elements is mainly based on the size: small urban and non-urban features are not detected and are therefore aggregated to the larger neighboring landuse.

The Urban Atlas is a landuse data source specifically developed to capture the diverse features existing in urban areas. The accuracy of identification and representation of the elements is higher than in Corine Land Cover, and this is particularly true for built-up areas and infrastructures. The main and bigger public green areas are present in the data source, while smaller green patches, such as the small public green areas, road-side green patches or private green are not represented. The Urban Atlas offers a convenient source of data to account for the role of big green spaces (urban and rural) in relation to the built-up features. In terms of landuse detection, the Urban Atlas is the most accurate in correctly allocating the pixels in the right location (kLoc) compared to the self-developed dataset. Results are less similar to the ones obtained with the self-developed map when looking at residential proximity to green spaces. This dataset has been developed with the purpose of enabling comparability of studies carried out in urban areas, and therefore the level of detail for urban features is higher than for the CLC, for example. This data source bases its inclusion of the different urban features on the definitions and delineations given by the authorities (e.g. green spaces are the ones listed by the public authorities).

Normalized Difference Vegetation Index information are derived from satellite images. Season and weather conditions affect the image characteristics and quality, and can therefore result in variations in the overall vegetation index. The lack of a fixed threshold to define the different landuses leads to an "arbitrary" choice of the author, based on the characteristics of the specific NDVI image. Fuzzy thresholds for landuse classification can undermine comparability and replicability of studies. The low confidence in the correct landuse classification, as well as the use of images from different days, makes the NDVI data source difficult to compare, even on the same study area. This data source is characterised by a low shape compactness of the patches, and generally a low distance between patches of the same class. This is due to the range of values possible to determine the landuses, based on the vegetation index. While the previously discussed data sources are based on institutionalised definition of some urban features, NDVI instead is only based on the color of the satellite image. This means that it is possible to capture more elements, especially if we think about vegetation: small, scattered patches of green spaces can be identified using NDVI. It should be noted, nonetheless, that this data source has proved to be the closest to the self-developed map (that was used as reference dataset) when measuring residential proximity to green spaces. This can be explained by the purpose of the data source itself: NDVI is aimed at detecting vegetation and has therefore a higher quality and resolution for vegetated areas.

The self-developed dataset combines data at a fine scale from existing datasets, such as road maps, cadastral data, and completes it with other unique characteristics of the municipality. In particular, private green areas and small vegetated patches have been detected and added to the dataset. This dataset is characterised by a low level of aggregation of the different landuses and patches as well as a higher fragmentation, represented by low values of distance between the patches. While the comparability can be assured by the creation of a codebook, the manual digitalization of the municipality features results in a time consuming method of data acquisition. The presence of both public and private green enables a more accurate estimate of green spaces. It is a convenient data source to take into consideration when looking at green spaces and health in epidemiological, or psychology research at fine scales where there is a focus on green availability in the individual's surroundings.

Our results are in line with other studies showing differences in the detection and definition of green spaces in different landuse data sources (Le Texier *et al.*, 2018; Rupprecht and Byrne, 2014).

3.6 Conclusion

The use of different data sources, for the same study purpose, makes the comparability of studies and results complicated and can undermine the research and possibly a correct decision making process. The choice of conducting our study on the municipality of Namur enabled us to work on a mixed environment: urban, peri-urban, and rural. Thus, different urban typologies and morphologies, as well as a diversity of green spaces have been considered. We have demonstrated here, through different measures, that differences among data sources are significant. The purpose and development method of the different landuse data sources leads to a different representation of the same study area and highlights different elements of the environment, shifting the focus from more urban or vegetation characteristics.

If the tendency of using different types of data sources for the same purposes is maintained, it will contribute to a scattered and hard-to-compare literature. A clear justification of the chosen data and of the methodology (i.e. NDVI threshold values) should be presented. This will simplify, encourage and hopefully foster comparison and takeaways among studies and across disciplines. Attention to the type of data used should be paid not only for the generalisation of empirical studies, but also to be able to provide an accurate image of the environment and better inform policy makers and planners. To have relevant and comparable results, as well as to better inform policy makers, a clear justification and description of the data is necessary.

III

Confirmatory analyses

4

On the impact of the environment on mental health: outcomes using four landuse data sources in Namur

This chapter builds on the conclusion and questions presented in the previous chapters. The variety of definitions of green spaces and the number of potential landuse data sources available makes comparability of studies a difficult task. As shown in chapter 3, the most commonly used landuse data sources provide quite different representations of the same area. Results may be influenced by different definitions and landuse data sources. Here, the associations between two mental health indicators and green spaces are tested over the four landuse data sources presented in the previous chapter. The aim is to investigate whether the results differ through the datasets, and how large is the gap. This chapter explores to what extent environmental factors are associated with mental health. Despite the sign of the association being consistent through the four landuse data sources, the significance is not always the same. This study reveals a consistent and significant association between mental health and socio-economic factors rather than environmental ones.

The study presented in this chapter has been carried out in collaboration with Ingrid Pelgrims (Sciensano)

4.1 Introduction

The raising interest on the potential effects of the environment on human physical and mental health has led to an increase in the scientific literature investigating the subject. In the last decades, the relationship between green space and human health has been widely investigated. Several studies have shown that closeness to green spaces plays a positive effect on physical and mental health (Maas *et al.*, 2009; de Vries *et al.*, 2003; Nielsen and Hansen, 2007; Grahn and Stigsdotter, 2003).

Depression is a common and serious mental health problem. It is characterised by a loss of interest, a persistent sadness, and/or a feeling of guilt, limiting the ability to carry out daily activities. The population suffering from depressive disorders has been increasing worldwide in the past decades (over 18% increase between 2005 and 2015) (World Health Organization, 2017b). In Belgium, the burden of depression and anxiety disorders was estimated to around 500,000 cases each (that is around 5% of the population) in 2015 (World Health Organization, 2017a). Depression and anxiety are considered to be stress-related mental disorders, making the contact with nature in these cases, relevant for health.

Several studies have looked into the salutogenic effects of natural spaces on mental distress around the world (Gascon *et al.*, 2018; Melis *et al.*, 2015; de Vries *et al.*, 2016; Helbich *et al.*, 2018a; Sarkar *et al.*, 2018). Green and Benzeval (2013) and Muntaner *et al.* (2004) point to higher odds of suffering from depression or anxiety according to the socio-economic position. As seen already in chapter 2, the majority of the studies dealing with the relation between health and the environment, are conducted in urban areas (Pereira *et al.*, 2012; McMorris *et al.*, 2015; Wall *et al.*, 2012; Potestio *et al.*, 2009; Villeneuve *et al.*, 2012). This also concerns studies related to mental health, as seen earlier in this thesis (Maas *et al.*, 2009; Amoly *et al.*, 2015; Nutsford *et al.*, 2013; Saw *et al.*, 2015). Furthermore, indicators of green and health vary in the literature, leading to difficulties in comparing methods and outcomes (Gascon *et al.*, 2018). As for data on the vegetation, authors mainly use the data and the tools that are at their disposal without any critical appraisal. For landuse data, for example, it is common to find studies using the Normalized Difference Vegetation Index (NDVI), as seen in chapter 2. Since the thresholds of the vegetation classes are manually defined by the authors, the comparability and reproducibility is still limited (Cheng *et al.*, 2008). As discussed earlier in this thesis, measures of green, and the landuse data choice, vary across different studies. The previous chapter highlights similarities and differences among four selected landuse data sources. The same ones are employed here to test the variance in the results when observing the relationship between different aspects of the environment and specific indicators of mental health.

The objective is here simply to understand to what extent the choice of the landuse data has an effect on the associations measured between the environmental indicators and the outcomes of a health research study. Some relevant demographic and socio-economic indicators are used as control variables. We want to test whether the strength and significance level of the association between the environmental indicator and the mental health outcomes varies depending on the landuse data set used. Based on the results of the comparison of data sources carried out in the previous chapter, the hypothesis is that significance and strength of the associations will be different for the four data sources.

4.2 Methodology and data

4.2.1 Study area and environmental indicators

The four landuse data sources employed in this study are Corine Land Cover, Urban Atlas, NDVI and a Self-developped map. The description of the main features and characteristics of the four landuse maps as well as the discussion on their similarities and differences can be found in chapter 3. The description of the study area, the municipality of Namur, can be found in section 3.2.2 in the previous chapter. This Belgian medium-sized municipality has been chosen as it embeds urban, suburban and rural features, and thus includes different types of green spaces. Figure 3.2, in chapter 3, illustrates the study area as represented by the four different landuse data sources. The environmental characteristics have been developed by computing the percentage of each landuse in each statistical section. Concerning the analysis at the individual level, x and y coordinates of the participants' home address have been used to construct indicators of the surrounding environment. The share of each landuse around each home address is computed for the four data sources within two circular buffers: 500 meters and 1000 meters. The choice of the buffer sizes is derived from the existing literature addressing the relationship between environment and mental health at different levels of urbanicity (Maas *et al.*, 2009; van den Berg *et al.*, 2010; Markevych *et al.*, 2014b; Dadvand *et al.*, 2014a). A 500 meters buffer is usually associated to a 10 minutes walking area for children and elderly people, or for people with limited mobility (Markevych *et al.*, 2014b), while 1000 meters buffer considers greenness in proximity to residential place but not immediately around the individual, and is generally associated to a desirable walking distance to services in urban environments (Chaix *et al.*, 2014). For this study only the share of public and private green (hereafter noted "green"), forests, and agricultural and pasture land (hereafter noted "agri") are considered, as defined on the previous chapter.

4.2.2 Health indicators

Two mental health indicators, anxiety and depression, are assessed both at the statistical sector level and at the individual level. In a first stage, data on medication reimbursement for depression and anxiety is used at the level of the statistical sections (see table A.1 in the Appendix for the list of Belgian administrative units). Then, the same associations are tested at the individual level (within a buffer distance from the place of residence).

Data on medication reimbursement has been provided by the Belgian Intermutualistic Agency (IMA-AIM). IMA-AIM is a Belgian agency that centralises information on reimbursed medication to Belgian residents (98% of the population) who are enrolled in the Health Security System. Here, the prevalence of medication reimbursed for depression and anxiety, proxied by Psycholeptics and Psychoanaleptics, is used. These correspond to the Anatomical Therapeutic Chemicals (ATC) code N05 and N06 (WHO Collaborating Centre for Drug Statistics Methodology, 2018). The prevalence (in %) is calculated by dividing the number of cases[1] of reimbursed medication in 2014, by the population registered in the Health System, and then multiplied by 100. The population included here encompasses all age groups: age-standardised yearly prevalence of medication reimbursement has been used here. For privacy reasons, the "health" data adopted here, is aggregated at the administrative level. If the number of cases of medication reimbursed in one year, in one administrative unit, is less than 5, then the information for the spatial unit is considered as missing. For this study, only 3 out of 65 administrative units had missing information in the municipality of Namur.

The second set of data, is obtained from the Health Interview Survey (HIS) carried out in Belgium in 2008 and 2013. The HIS is a nationwide representative epidemiological survey aiming at informing on the mental and physical status of the general population in Belgium. Respondents are selected based on a multistage stratified cluster sample. Individual and household characteristics are assessed using questionnaires and face-to-face interviews. Questions pertain to the topics of general health status, reduced mobility, use of healthcare services, mental health and use of drugs or alcohol. For this study, answers pertaining to the issue of depression and anxiety are considered. The indicator was developed using the answer to questions SL02 ("Have you ever seriously thought of ending your life?") and SL03 ("Did you have such thoughts in the past 12 months?") (Sciensano, 2019), based on the answers to a validated questionnaire SCL 90 R. To safeguard the privacy of the respondents, the data has been handled in 5 steps and with the collaboration of three institutes (Sciensano, Statbel, and UCLouvain), as described in chapter 1, section 1.6. A

[1] "case" is defined as the number of citizens who have benefited at least once of medication reimbursement for depression and anxiety in 2014

total of 150 interviews were retained for the municipality of Namur.

Table 4.1 reports the main characteristics of the two datasets used in this chapter.

Table 4.1. Characteristics of the two health datasets used for this study

	Prevalence CNS	**Perceived DA**
Provider	IMA-AIM	Sciensano
Type of data	%	Yes/No
Observations	62 adm. units	150 respondents
Aggregation level	Statistical section	Buffer around individual at 500m & 1000m (residence)

4.2.3 Demographic and socio-economic indicators

As seen in chapter 1 some socio-economic factors can play a role in the association between green spaces and health (Allen and Balfour, 2012; van Lenthe *et al.*, 2005). In particular, a study conducted in Belgium showed worse perceived health and poorer health for people with a low level of education (Bossuyt *et al.*, 2004). Here information on some relevant demographic and socio-economic characteristics have been included. Data was derived from the latest Belgian Census for the analyses at the level of the statistical sections, and from the HIS survey for the analyses at the individual level. For the first part of the analysis, data on population density, highest level of education (percentage of population with "Primary education", "Secondary education"), and employment status (percentage of "Inactive population") have been used. Socio-economic data is only available at the level of the statistical sectors (finest administrative unit), and therefore has been aggregated to the statistical section level for the purpose of the study. For the second part of the study, two relevant socio-economic indicators have been acquired from the Health Interview Survey: gender and education level has been collected for each individual. Education level has been classified in three categories: "No diploma or lower education" including people who do not have a diploma or a primary education only, "Higher secondary" including people with a high school diploma, and "University degree" including people who obtained a university degree.

4.2.4 Statistical analysis

Linear regressions are used to compare at the association of prevalence of medication reimbursement, in the municipality of Namur, at the level of the statistical sections for the four data sources. Adjusted logistic regression models were performed to assess the association between the environmental and socio-economic variables and likelihood to present depressive and anxiety disorders at the individual level. The use of ordinary logistic regression with a low number of observations and/or high number of "0" or missing values can entail extremely high Odd-Ratios and Confidence Intervals. In such cases a Penalised Logistic Regression (PLR) can limit the overestimation of the OR and the CI (Devika *et al.*, 2016). Socio-economic variables such as employment, age and income, have been omitted due to collinearity. The use of logistic regressions and PLR have both been tested and the results obtained were comparable. Ordinary logistic regression has then been chosen for this study.

We also compute Receiver Operating Characteristics (ROC) curves. The ROC curve is a popular tool for the evaluation of the diagnostic tests, but it can also be used to demonstrate associations between a continuous variable for a binary outcome. Based on HIS mental health data, the ROC curves enable to determine the prediction capability of the model. ROC curves plot the True Positive Rate (TPR) on the y axis ("Sensitivity TPR"), against the False Positive Rate (FPR) of the model on the x axis ("1-Specificity FPR"). If True Positives and False Positives are correctly identified by the model, that is the ideal case, the value of the Area Under the Curve is the maximum (1). The lower the AUC value, the lower is the discrimination capacity of the model between positives and negatives. A model with an AUC of 0.5 is generally considered to have no discrimination capacity.

4.3 Results

First, we present the results of the analysis on the medication reimbursements, then the results of the analysis on the likelihood to present depressive and anxiety disorders. Descriptive statistics of the data are available in tables 4.2, and 4.3.

Table 4.2. Descriptive statistics of medication reimbursement for the Central Nervous System (CNS) and socio-economic variables employed for the analyses at the level of the statistical sections

	Units	Mean	Std dev	Min	Max	Median	P95
CNS	%	24.30	7.3	0	51.83	24.07	34.36
Pop density	Inh/km^2	1272.1	1575.9	68.4	6953.6	623.5	4943.03
Inactive pop	%	55.4	5.3	42.11	78	54.7	63.02
Primary edu	%	0.7	3.1	0.13	16.2	0.6	12.41
Secondary edu	%	19.8	2.5	12.1	26.15	20.1	23.66

Table 4.3. Characteristics of the HIS samples for depressive and for anxiety disorders (number of cases)

	Depressive disorder		Observations
	Yes	**No**	
	18	132	150
Gender			
Male	8	59	67 (44.6%)
Female	10	73	83 (55.3%)
Diploma			
No diploma or lower education	3	9	37 (24.6%)
Higher secondary education	6	20	26 (17.3%)
University degree	6	81	87 (58%)
	Anxiety disorder		**Observations**
	Yes	**No**	
	14	135	149
Gender			
Male	4	60	64 (42.9%)
Female	10	75	85 (57.1%)
Diploma			
No diploma or lower education	2	10	36 (24.1%)
Higher secondary education	6	20	26 (17.5%)
University degree	6	81	87 (58.4%)

4.3.1 Medication reimbursement at the statistical section level

Pearson correlation coefficients have been computed between depression and anxiety (CNS) medication reimbursements and landuse classes referring to green spaces, for each of the four landuse data sources (see chapter 3). Results, presented in table 4.4, show that correlations between the medication reimbursements and the three selected types of green spaces are almost never strongly significant, with the exception of urban green in CLC that appears to be negatively correlated to prescription. The sign of the correlation is not always consistent. The sign remains the same (negative) only for forest areas; for urban green and agricultural areas the sign changes. Pearson correlations between environmental, demographic and socio-economic factor have also been computed (table 4.5). Population density is, predictably, always negatively correlated with forest and agricultural areas, but strongly positively correlated with public green areas from the Urban Atlas. Forest and agricultural areas are generally negatively correlated to the share of inactive population and of people with only a primary education. The share of people with a higher education appear not be significantly correlated to any green space. Signs of the correlations are mostly consistent, except for the public green spaces defined by the Urban Atlas. It appears that urban green spaces are positively correlated to population density, and with the share of inactive people. With the exception of the Urban Atlas, public green spaces are generally not significantly correlated with the selected demographic and socio-economic variables.

Table A.2, in Appendix, presents the results of the Pearson correlations computed between the medication reimbursements for disorders linked to the Central Nervous System (CNS), and the selected demographic and socio-economic variables. Results show strong positive correlations with the part of inactive population (students, retired people, unemployed).

Linear regressions were performed for each landuse data source independently (Corine Land Cover, Urban Atlas, NDVI, and the self-developed map), using the age-standardised reimbursed medication (CNS) for depression and anxiety as dependent variable. Results are presented in table 4.6. The results of the four regressions highlight some consistencies. The percentage of inactive population appears to be always positive and significant in the four models, and the population with a secondary education degree, although the latter is slightly less significant. A higher share of unemployed population, corresponds to an increase in reimbursements for medications targeted at the Central Nervous System. In none of the four models the environment appears to have a significant association with the reimbursement of medications for the Central Nervous System. All models explain over the 31% of the variability. This means that two thirds of the spatial variation depends on factors not included here. Information on noise and air pollution, not available for this study, might help further improve the model. Such similarity in R^2 for the four data sources, using the same explanatory variables, suggests that for the municipality of Namur, medication reimbursement is only partially explained by the green environment. Consistency of results clearly point to the need of including correlated environmental exposures and socio-economic variables.

Table 4.4. Pearson correlation coefficients between medication reimbursement for CNS, and green landuse classes for the four datasets in the municipality of Namur

	CLC	UA	NDVI	SD
Green	-0.38**	0.19	0.16	0.09
Forest	-0.23•	-0.31*	-0.22•	-0.27*
Agri	0.02	0.05	-0.02	0.04

P-value : 0 <'***' <0.001 <'**' <0.01 <'*' <0.05 <'•' <0.1

Table 4.5. Pearson correlation coefficients between the selected socio-economic variables and the three typologies of green for each landuse database, at the level of the statistical sections

	CLC			UA		
	Green	**Forest**	**Agri**	**Green**	**Forest**	**Agri**
Pop density	0.08	-0.46***	-0.55***	0.64***	-0.5***	-0.61***
Inactive pop	-0.05	-0.1	-0.37**	0.39**	-0.11	-0.42***
Primary edu	-0.15	-0.18	-0.37**	0.5***	-0.3*	-0.37**
Secondary edu	-0.04	0.04	0.52	-0.44*	0.14	0.48*
	NDVI			**GE**		
	Green	**Forest**	**Agri**	**Green**	**Forest**	**Agri**
Pop density	-0.04	-0.79***	-0.5***	0.16	-0.46***	-0.57***
Inactive pop	-0.10	-0.29*	-0.3*	0.1	-0.04	-0.41**
Primary edu	-0.04	-0.51***	-0.46***	-0.077	-0.19	-0.37*
Secondary edu	0.18•	0.52	0.58	0.19	0.05	0.48•

P-value : 0 <'***' <0.001 <'**' <0.01 <'*' <0.05 <'•' <0.1

Table 4.6. Linear regression coefficients by landuse data source, between medication reimbursement for CNS and environmental and socio-economic variables at the level of the statistical sections in Namur

	CLC	**UA**	**NDVI**	**SD**
Intercept	-25.20	-42.36•	-32.41	-41.30•
Green	-0.87*	0.15	0.19	0.11
Forests	-0.044	-0.03	-0.007	0.003
Agri	0.05	0.13•	0.19	0.15*
Pop density (log)	-0.31	0.25	-0.55	0.47
Inactive pop	0.71**	0.85**	0.69*	0.72**
Primary edu	0.47	0.36	0.72•	0.78•
Secondary edu	0.42	0.59	0.56	0.53
Adj. R^2	0.39	0.38	0.31	0.40

P-value : 0 <'***' <0.001 <'**' <0.01 <'*' <0.05 <'•' <0.1

4.3.2 Depressive and anxiety symptoms at the individual level

Logistic regressions have also been performed using individual data on the likelihood of presenting depressive and anxiety disorders, as dependent variable. Data has been derived from the HIS database, and included 150 complete observations within the municipality of Namur. A series of logistic regression models have been performed for each mental health outcome and Corine Land Cover, Urban Atlas, NDVI and the self-developed datasets. The analysis has been carried out for the two buffer sizes defined earlier (500 meters and 1000 meters). Results are reported in tables 4.7, 4.8, 4.9 and 4.10. The Akaike Information Criterion (AIC) is reported for each association.

Table 4.7. Associations between depressive disorder and four landuse sources (Corine Land Cover, Urban Atlas, NDVI and Self-developed data). Associations computed at 500m buffer

	CLC Adj. OR (95% CI)	UA Adj. OR (95% CI)	NDVI Adj. OR (95% CI)	Self-dev Adj. OR (95% CI)
Green	0.74 (0.24, 2.34)	0.99 (0.9, 1.08)	1.04 (0.95, 1.14)	1.01 (0.98, 1.05)
Forest	0.93 (0.84, 1.02)	0.95 (0.89, 1.02)	0.99 (0.96, 1.03)	0.99 (0.94, 1.03)
Agri	**1.03 (1.01, 1.06)**	1.03 (1, 1.06)	1.06 (0.97, 1.16)	1.03 (1, 1.06)
Diploma (vs. University degree)				
No diploma or lower edu	3.03 (0.86, 10.69)	3.4 (0.92, 12.56)	3.07 (0.87, 10.8)	3.51 (0.99, 12.52)
Higher secondary	**4.73 (1.25, 17.81)**	**4.9 (1.31, 18.4)**	**4.34 (1.14, 16.54)**	**5.2 (1.42, 19.08)**
Gender (F vs. M)	1.27 (0.43, 3.74)	1.2 (0.42, 3.48)	1.1 (0.39, 3.1)	1.22 (0.43, 3.48)
Log-likelihood	-46.65	-48.79	-49.40	-50.13
AIC	107.31	111.59	116.80	114.26

Table 4.8. Associations between depressive disorder and four landuse sources (Corine Land Cover, Urban Atlas, NDVI and Self-developed data). Associations computed at 1000m buffer

	CLC Adj. OR (95% CI)	UA Adj. OR (95% CI)	NDVI Adj. OR (95% CI)	Self-dev Adj. OR (95% CI)
Green	0.94 (0.72, 1.24)	0.87 (0.68, 1.12)	1.4 (0.97, 2.03)	1.02 (0.97, 1.08)
Forest	0.92 (0.83, 1.01)	0.95 (0.89, 1.02)	1.02 (0.96, 1.07)	0.98 (0.93, 1.04)
Agri	**1.04 (1.01, 1.07)**	1.01 (0.96, 1.05)	0.96 (0.77, 1.2)	1.03 (0.99, 1.06)
Diploma (vs. University degree)				
No diploma or lower edu	3.22 (0.9, 11.5)	**3.99 (1.08, 14.74)**	2.5 (0.7, 8.89)	3.27 (0.93, 11.52)
Higher secondary	**5.89 (1.56, 22.24)**	**5.92 (1.56, 22.42)**	3.54 (0.97, 12.94)	**5.27 (1.43, 19.51)**
Gender (F vs. M)	1.24 (0.43, 3.59)	1.22 (0.43, 3.47)	1.17 (0.41, 3.35)	1.15 (0.41, 3.22)
Log-likelihood	-47.27	-49.67	-48.33	-50.78
AIC	108.54	113.34	110.67	115.55

Table 4.9. Associations between anxiety disorder and four landuse sources (Corine Land Cover, Urban Atlas, NDVI and Self-developed data). Associations computed at 500m buffer

	CLC Adj. OR (95% CI)	UA Adj. OR (95% CI)	NDVI Adj. OR (95% CI)	Self-dev Adj. OR (95% CI)
Green	1.02 (0.9, 1.17)	0.96 (0.87, 1.06)	0.9 (0.77, 1.06)	0.99 (0.95, 1.03)
Forest	0.98 (0.92, 1.04)	0.99 (0.94, 1.04)	0.98 (0.94, 1.02)	1.01 (0.96, 1.05)
Agri	1.01 (0.97, 1.03)	0.99 (0.96, 1.02)	1.08 (0.97, 1.2)	0.99 (0.96, 1.03)
Diploma (vs. University degree)				
No diploma or lower edu	2.03 (0.51, 8.09)	2.4 (0.57, 10.03)	2.42 (0.58, 10.15)	1.98 (0.49, 8.11)
Higher secondary	**4.21 (1.07, 16.56)**	**4.61 (1.15, 18.38)**	**5.46 (1.34, 22.28)**	3.76 (0.96, 14.8)
Gender (F vs. M)	1.43 (0.44, 4.6)	1.41 (0.44, 4.53)	1.58 (0.48, 5.17)	1.44 (0.45, 4.64)
Log-likelihood	-43.96	-43.78	-43.01	-44.13
AIC	101.93	101.57	100.02	102.26

Table 4.10. Associations between anxiety disorder and four landuse sources (Corine Land Cover, Urban Atlas, NDVI and Self-developed data). Associations computed at 1000m buffer

	CLC Adj. OR (95% CI)	UA Adj. OR (95% CI)	NDVI Adj. OR (95% CI)	Self-dev Adj. OR (95% CI)
Green	1.05 (0.86, 1.27)	0.91 (0.71, 1.17)	1.49 (0.93, 2.39)	0.98 (0.93, 1.04)
Forest	0.99 (0.93, 1.05)	1.01 (0.94, 1.06)	1.04 (0.98, 1.09)	1.01 (0.95, 1.07)
Agri	1.005 (0.97, 1.04)	0.97 (0.92, 1.03)	0.8 (0.59, 1.07)	0.99 (0.95, 1.03)
Diploma (vs. University degree)				
No diploma or lower edu	1.97 (0.48, 7.98)	2.37 (0.53, 9.6)	1.75 (0.43, 7.1)	1.9 (0.47, 7.71)
Higher secondary	**4.11 (1.07, 15.85)**	**6.93 (1.07, 17.19)**	2.97 (0.7,5 11.83)	3.64 (0.92, 14.34)
Gender (F vs. M)	1.45 (0.45, 4.64)	1.44 (0.45, 4.64)	1.37 (0.42, 4.48)	1.41 (0.44, 4.51)
Log-likelihood	-44.01	-43.6	-42.61	-43.91
AIC	102.02	101.19	99.21	101.83

Results of the logistic regression models show almost no significant associations between green areas and depressive and anxiety disorders. Agricultural areas, and therefore more rural and green settings, are the only green space associated to the two mental health disorders analysed here. Results of this association are significant, in particular when looking at depressive disorders, compared to anxiety disorders. Outcomes are stronger for smaller buffer distances (within 500 meters from the place of residence) in the case of depressive disorders: agricultural land is associated with mental health for Corine Land Cover only. Odds-ratios show that the presence of agricultural areas around the place of residence increases the likelihood of presenting depressive disorders increases by a factor of 1.03 (OR= 1.03; 95%CI: 1.01, 1.06). Other types of green spaces instead do not seem to be associated to the likelihood of presenting depressive or anxiety disorders. Looking at larger distances from the place of residence (1000 meters buffer), outcomes are consistent: the presence of agricultural areas increases the risk of depressive disorders for CLC only (OR= 1.04; 95%CI: 1.01, 1.07). The possession of a higher secondary degree increases the likelihood of suffering from depressive disorders at the two buffer distances (at the exception of NDVI). Not being in possession of a diploma or having a lower education, compared to having a university degree, is associated to an increased likelihood of presenting depressive disorders when using the Urban Atlas data source (OR= 3.99; 95%CI: 1.08, 14.74). Concerning the likelihood of presenting anxiety disorders, no association with the environmental factors appears to be significant. At a closer distance from the residence (500 meters), as for the depressive disorder previously described, a higher education level ("Diploma, University degree"), is associated with an increased risk of presenting anxiety disorders for all data sources except for the Self-developed one. The association still holds at a 1000 meters buffer for just Corine Land Cover (OR= 4.11; 95%CI: 1.07, 15.85) and Urban Atlas (OR= 6.93; 95%CI: 1.07, 17.19). The risk for people with a higher secondary education is four-to-six times higher than for people with a university degree.

Figure 4.1 shows the results of the Receiver Operating Characteristics analysis and the curve paths. Results of the ROC curves for both health outcomes (medication reimbursement and self-perceived depression and anxiety) are displayed for two buffer sizes (500 meters, 1000 meters). The horizontal axes, "1-Specificity FPR", represent the False Positivity Rate Values of the Area Under the Curve (AUC), are computed for each ROC curve are included in each figure. ROC curves and AUC valued for each landuse dataset are compared. As it is possible to see from figure 4.1, the curves for each dataset often overlap, and the AUC results are close. The highest variability can be found in the curves for anxiety disorder at 1000m buffer distance. The trajectory of the curves appears to be consistent at the two different buffer distances.

All areas under the ROC curves are significantly greater than 0.5 (min 0.66, max 0.79), meaning that each model gives results that are better than guesswork. Although Corine Land Cover appears to always have the highest AUC values, the values are not significantly different from the other landuse data sources. There is no consistency on which set of data predicts best depressive or anxiety disorders, as the ranking order changes in the four ROC models. Nonetheless the trajectory of the four curves, as well as the AUC values for both buffer sizes are close, meaning that the difference in prediction is small. These results are in line with the odd-ratios outcomes, where the associations between depressive and anxiety disorders and environmental and socio-economic variables seemed to be consistent in sign and significance across the four landuse data sources.

4.4 Discussion

We test here sensitivity of associations between environment and health using four landuse data sources. We use linear and logistic regressions to look at similarities in the associations between green spaces and two indicators of mental health (depression and anxiety). The study is conducted within the municipality of Namur, Belgium, at two different spatial levels: statistical sections for medication reimbursement for the entire population, and buffer rings around the individual place of residence for a sample of surveyed people. Results of the analyses show divergent outcomes and therefore lead to unstable conclusions. Demographic and socio-economic factors have been used as control variables for the two analyses.

First, we compare the association of some environmental, demographic and socio-economic variables with medication reimbursement prevalence for the Central Nervous System at the level of the statistical sections in the municipality of Namur.

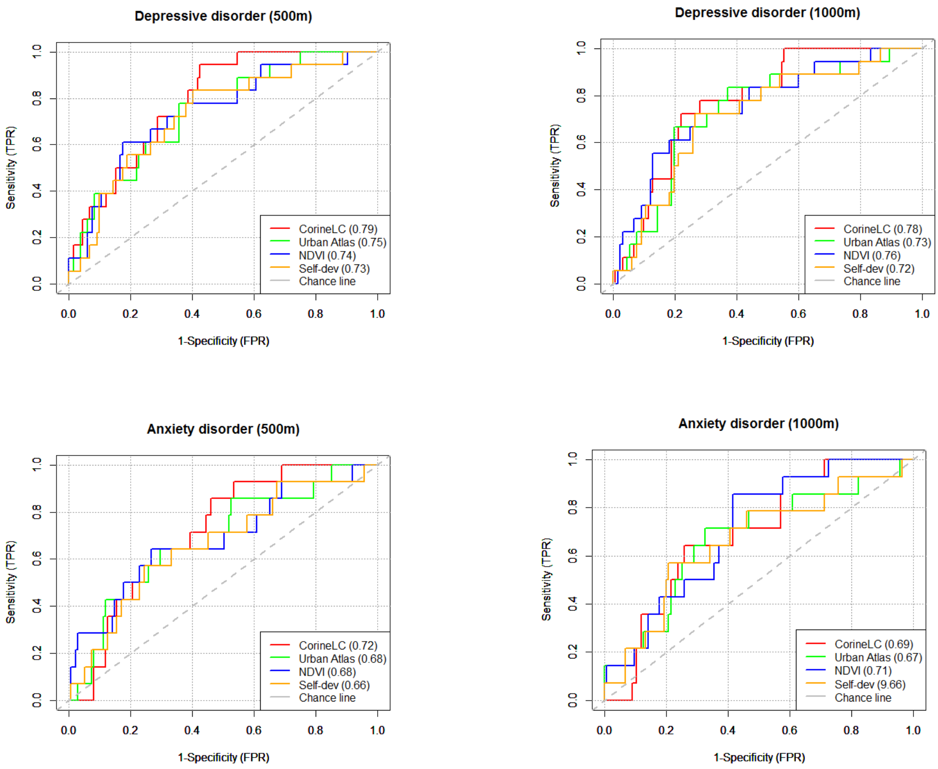


Figure 4.1. Receiver Operating Characteristics (ROC) curves for depression and anxiety disorders, for four landuse data sources. The Area Under the Curve (AUC) value of each data source is included between brackets.

When looking at the prevalence of medication reimbursement for depressive and anxiety disorders (CNS), we see a weak association with the environmental variables used. Inverse association between percentage of urban green and medication reimbursement can only be found when using the Corine Land Cover data source, suggesting to a higher prevalence of reimbursements with a decrease of green areas. The same typology of green from the other data sources is not associated with medication reimbursement and present an inverse sign compared to CLC. From the regression analysis, agricultural areas appear to be positively associated, although weakly, to medication reimbursement for the Urban Atlas and the self-developed database. The association with medication reimbursement is instead positive and significant for the share of people who are not in employment across all landuse data sources. A lower socio-economic condition is positively associated to a higher prevalence of medication reimbursement. This is in line with what Green and Benzeval (2013), Muntaner *et al.* (2004) show: a positive association between depressive disorders and socio-economic position. These findings are also in line with the results of the meta-analysis carried out by Lorant *et al.* (2003) on socio-economic inequalities in depression: lower socio-economic strata are generally disfavored.
Secondly, we use data on perceived mental health obtained from the Health Interview Survey. Data collected from the survey is at the individual level. The results obtained, again show differences in associations among the four data sources. Agricultural areas, as identified through the Corine Land Cover data source, appears to increase the likelihood of presenting depressive disorders, at both 500 meters and 1000 meters from the place of residence. Associations with environmental factors are instead not significant for the other data sources. Our results show an opposite effect of green spaces on depression and anxiety compared to the ones presented in the existing literature. Sarkar *et al.* (2018) and Gascon *et al.* (2018), for example, show a potential protective effect of green spaces for depressive and anxiety disorders. Both studies use buffer rings around the place of residence. Here we show that at the scale of the municipality of Namur, the association between green spaces and the investigated health disorder, is not always significant. Socio-economic factors, such as education and employment, are consistently associated with depressive and anxiety disorders.

We here use two sets of data on depressive disorders to investigate whether the choice of one landuse data source over the other leads to different results when looking at the associations between environment and health. Our results show that sign and strength of the associations across the different landuse data sources differ. Two green space types (green urban areas and agricultural areas), as defined by Corine Land Cover, have appeared in both ecological studies to be associated with depressive and anxiety disorders. The presence of agricultural areas in the surrounding neighborhood appears to be weakly and positively associated to medication reimbursement when using two out of four data sources (Urban Atlas and the Self-developed dataset). NDVI, the most commonly used data source in studies investigating associations between green environment and health, surprisingly does not hold any significant associations with the health measures used here. Results hint toward individual and social factors affecting depressive and anxiety disorders, while green environment does not play a direct role on the association. ROC curves' trajectories and AUC values appear to be generally consistent for the different landuses. No landuse data source appears to better predict cases of depression and/or anxiety.

The contrasting results obtained using two health indicators point clearly to a diverging way of assessing and looking at depression and anxiety. The use of data on medication prescription and perceived mental health require handling with care, as the perception of health seem not to be consistent with medication purchase for the same symptoms when looking at the associations with the environment. The author is aware that due to the different ecological design the two studies cannot be compared. The aim here is to warn on the difference of message that the two studies convey, especially in the perspective of providing recommendations to decision makers. A limitation of this study lies in the type of health data available: the different spatial extent of the datasets does not allow for a comparison of the outcomes. Here we have employed both population aggregated data at the administrative level (statistical sections), and a sample of data at the individual level. One risk when using aggregated data is to fall into ecological fallacy and MAUP issues. On the other hand, non-localised individual data do not allow for more spatial-oriented analysis, such as spatial autocorrelation, for example, and not enabling for a more in-depth investigation of the spatial factors that may come into play.

Based on the results of the data source comparison presented in the previous chapter (chapter 3), strong inconsistencies in the outcomes of associations were expected. This study shows that strength and effect of the associations between environmental factors and health vary across the data sources, although the results were expected to be less consistent across data sources. Care should be used when interpreting the results, as associations between green spaces and health are not clear and consistent across data sources.

4.5 Conclusion

This chapter compares the association between environmental factors and mental health across four different landuse data sources, and two different epidemiological designs. The results obtained confirm the hypothesis that the choice of the landuse data source leads to different strengths in the associations between green environment and health, also answering the concerns raised by Markevych *et al.* (2017). We show that differences are detected in the results of the analysis when using different landuse data sources for the same health indicator. Nonetheless differences are stronger and striking when looking at two indicators of the same mental health disorder. The differences in the construction of the indicators, as well as on their spatial extents, enables us to explore the discrepancies in the measures of green spaces through different landuse data sources. Our findings point towards a more important association between socio-economic factors and health, rather than environmental characteristics.

We show that the choice of the landuse data source, at the scale of the municipality, are not interchangeable. In particular, the differences in outcomes are stronger when using highly spatially aggregated landuse data at a small scale of analysis. We also demonstrate that the choice of the mental health indicators, and thus the ecological design, lead to partially different associations. Further research should analyse more in-depth gaps in findings using different landuse data sources with more accurate and detailed health data and including further control variables. When studying the role and impact of green spaces on health it is fundamental to have a complete understanding of the data that is being handled. It is necessary to be careful when interpreting the results, which may be telling/unveiling just a part of the whole story.

5

Spatial variations of medication reimbursements in Belgium: environmental, socio-economic or healthcare determinants?

This chapter uses medication reimbursement data to analyze associations between health status and environmental and socio-economic factors in Belgium. Five groups of medications, corresponding to health disorders potentially related to environmental conditions, are analyzed. The prevalences of medication reimbursements are positively correlated to each other. Reimbursements are instead negatively correlated to air pollution and distance to green spaces, as opposed to what has been found in previous studies. Principal component analysis and hierarchical clustering reinforce these results. The spatial variations of medication reimbursements have been analysed here at the scale of the municipalities, for the whole country. Regional differences appear clearly, although the determinants are hard to quantify (education of medical doctors, pharmaceutical commercial activities, for example). This chapter raises the question on what medication reimbursement data really represents. It warns on using care when working and interpreting such datasets.

This chapter is based on research carried out with Dr. Lidia Casas, Prof. Benoît Nemery and Prof. Isabelle Thomas, submitted to the Journal "Cybergeo" with the title "On the geography of medication reimbursements in Belgium: environmental, social or healthcare determinants?"

5.1 Introduction

Living environment and health issues have long and often been associated, positively or negatively. Among others, salutogenic effects of proximity and contact with green spaces have been reported on physical and mental health (Gascon *et al.*, 2015; Nielsen and Hansen, 2007; Sacker and Cable, 2005; Ulrich, 1984). In addition, negative externalities of some environmental determinants have been demonstrated: the adverse effects of ambient air pollution on respiratory and cardiovascular problems have been studied for decades (Dominici *et al.*, 2006). It has been shown that high levels of air pollutants such as PM_{10} or NO_2 are associated with increased respiratory and cardiovascular problems and medication intake (Casas *et al.*, 2016; Sofianopoulou *et al.*, 2014). Recent studies show additional adverse effects of air pollution on the brain, birth outcomes (e.g. low birth weight) or metabolic disorders like diabetes mellitus (Landrigan *et al.*, 2017; Thurston *et al.*, 2017). Positive and negative externalities of the environment are observed in both urban (traffic-related air pollution) and rural areas (pesticides).

The relationship between environment and human health has most often been investigated at the local level (Cheng *et al.*, 2011) and most often within urban areas: in particular when tackling issues of increasing urban population, for targeted policies and investments on green infrastructures (Adli *et al.*, 2017; Conway *et al.*, 2010; Coutts and Hahn, 2015). City dwellers are generally considered as living in less healthy environments, but enjoying the proximity of a large variety of facilities. Non-urban areas, although experiencing low built-up densities are generally considered as healthier environments, but they also may suffer from some types of pollution due to agricultural practices (Paulot and Jacob, 2014), including intensive animal breeding or the use of pesticides, as well as poorer access to healthcare services. Studies analyzing spatial variations at country level (Henau *et al.*, 2015), are generally lacking. Presence and proximity to green areas has been often associated to better/improved health status. In the last decade numerous studies have looked at the relationship between surrounding green spaces and specific morbidities or well-being (Maas *et al.*, 2009; Mukherjee *et al.*, 2017; Pietilä *et al.*, 2015; Van Herzele and de Vries, 2012).

In the literature, health status is generally proxied by different types of data such as individual (survey) perceived health (Maas *et al.*, 2006; Triguero-Mas *et al.*, 2015), hospital records (Dadvand *et al.*, 2012a; Pereira *et al.*, 2012), more recently by medication prescriptions (Halasa *et al.*, 2002; Naughton *et al.*, 2006; Bianchi *et al.*, 2009), or derived from time-series and cohort studies, in the case of mortality (Samet and Krewski, 2007). Although medication prescriptions have been used in the literature for a variety of specific medications, mapping their spatial variations is novel (Cheng *et al.*, 2011). We here refer to Wangia and Shireman (2013) for a meta-analysis on the application of GIS tools for analyzing spatial variations of medication prescriptions. The vast majority of studies has investigated the spatial variations of health by looking at individual morbidity or mortality causes, rather than looking at their covariation in space (Trabelsi *et al.*, 2019; Waldhoer *et al.*, 2008). Looking at medication prescription data some studies have analysed differences in medication prescription and practice behaviour due to training and education of the physicians (Ekedahl *et al.*, 1995; Figueiras *et al.*, 2000).

This study aims at mapping the spatial patterns of medication reimbursement prevalence for five different medication groups at the level of the municipalities in Belgium. The association between medication reimbursement prevalence and environmental or socio-economic indicators are further analyzed. A geographical interpretation of the link between medication reimbursements and environmental and socio-economic factors is proposed here.

5.2 Data and methods

5.2.1 Medication purchases

Health insurance is compulsory in Belgium and thus nearly all residents in the country are enrolled in the Health Security System (98%). In this System, big part of healthcare expenditure, including most medication prescribed by physicians, is reimbursed by one of the seven Belgian health insurance funds. The Intermutualistic Agency (IMA-AIM) is an umbrella organization that centralizes all information on healthcare reimbursements collected by all Belgian health insurance funds. Information on reimbursed medication is coded using the ATC (Anatomical Therapeutic Chemical) code (World Health Organization and Norwegian Institute of Public Health, 2013). After medication prescription by a medical doctor and the purchase of that medication at a Belgian pharmacy (including hospital pharmacies), the ATC codes are registered in the health insurance system. It should be noted that a small part (2%) of the population living in Belgium is not covered by the social security system and, therefore, not included in this study. These correspond to undocumented immigrants, people with complex social problems (like homeless), or employees of international bodies (EU, NATO, etc.). The database does not include non-reimbursable medications. It is worth pointing out that the purchase of a medication does not necessarily imply its consumption.

In this study, medication reimbursement information is provided at the municipality level for all 589 municipalities in Belgium for each individual who was reimbursed after purchasing a specific medication at least once in a calendar year. Information on its actual use, frequency of use, or about comorbidities, meaning, in this case, the presence of the same individual in more than one medication groups, cannot be deducted from the data. Also, this data does not provide information on the frequency of purchase of the medication in one year. An average prevalence over 3 years (2012-2013-2014) is here considered to minimize the effect of small numbers (statistical and privacy issues). Prevalence is here defined as the number of cases divided by the total population registered in the Belgian health security system within the municipality. Age-standardised yearly prevalence is expressed by 100 inhabitants (Age-Standardised Medication reimbursement Prevalence, noted here as ASMP), and is computed separately for men and women. Details on the methodology used to create the ASMP can be found in Costa *et al.* (2019).

We here limit our investigation to five medication groups corresponding to health disorders potentially related to environmental conditions, according to the scientific literature. Table 5.1 presents the target system, the ATC codes, the name of the codes and the abbreviations used in this paper for each medication group. The first group (RS) includes medications for respiratory obstructive diseases such as asthma or Chronic Obstructive Pulmonary Disease (COPD), the second (CVS) includes medications used to reduce blood pressure and to treat other cardiovascular problems like ischemic diseases; the third (LIP) includes medications prescribed to reduce hyperlipidemia, the fourth (IS) includes local and systemic medication for treating allergy symptoms (topic and systemic), and the fifth (CNS) includes drugs acting on the central nervous system mainly indicated to treat symptoms of anxiety and depression.

5.2.2 Study area

Belgium is a small but densely populated Western European country (374 inh/km^2). It is divided into three administrative regions: Flanders in the North (tight urban network, high level of urbanization), Wallonia in the South (greener, less densely populated, looser city network) and the Brussels Capital Region (BCR) in the center. The BCR corresponds to the city center of Brussels, while the urban agglomeration sprawls into both Flanders, in the Province of Vlaams-Brabant, and Wallonia, in the Province of Brabant Wallon. Vlaams-Brabant and Brabant Wallon were once part of the same province (former Brabant), but now they are separated and belong to Flanders and Wallonia respectively, and are administratively, politically and linguistically different. Belgium is further divided into provinces (10 provinces + BCR) illustrated in figure A.1 (in the Appendix), and 589 municipalities. The municipality of Herstappe was removed from the analyses because of its small size (88 inhabitants in 2018) (StatBel, 2018).

Table 5.1. Groups of medications used in this study

	Target system	ATC code	Name
RS	Respiratory system	R03	Drugs for obstructive airway diseases
CVS	Cardiovascular system	B01A	Antithrombotic agents
		C01	Cardiac therapy
		C02	Antihypertensives
		C03	Diuretics
		C07	Betablocking agents
		C08	Calcium channel blockers
		C09	Agents acting on renin-angiotensin system
LIP	Lipids	C10	Lipid modifying agents
IS	Immune system (allergy)	D04A	Antipruritics [a]
		R06	Antihistamines for systemic use [b]
CNS	Central nervous system	N05	Psycholeptics
		N06	Psychoanaleptics

[a]incl. antihistamines, anesthetics, etc. (Dermatologicals) [b](Respiratory)

The three Regions have separate political administrations, different language (Dutch in Flanders, French in Wallonia, German in the German-speaking community in Wallonia, and both Dutch and French in the BCR) and different universities delivering medical degrees. Healthcare policies and reimbursement criteria are defined at the national level.

5.2.3 Environmental and socio-economic variables

We gathered information on several relevant environmental (ENV) and socio-economic (SE) characteristics at the level of the municipalities. Both data on air pollution and green spaces were available environmental indicators.

Data on air pollution was provided by the Belgian Interregional Environment Agency (Irceline, 2018), including data on the average values of Particulate Matter 2.5 ($PM_{2.5}$), Particulate Matter 10 (PM_{10}), Nitrogen Dioxide (NO_2), and Black Carbon (BC) in 2015 are used, as measured by over 100 fixed monitoring stations in Belgium. The data is then further modelled using 4x4km cells. Such monitoring stations are more dense in urban areas than in rural areas. The use of such large cells smooths the effects of the extremes values, and has therefore to be interpreted with caution. Variables on green spaces were obtained from the Corine Land Cover dataset (European landuse dataset) developed by the European Environment

Agency (European Environmental Agency, 2007), more details about this data source can be found in chapter 3, in section 3.2.1. Here we used the percentage of green spaces (hereafter *"Total GS"*) in each municipality, and the average distance to the nearest green space (hereafter "Ave Near GS"). For the purposes of this study, four typologies of green spaces were summed: urban green, forests, pastures and agricultural land, and semi-natural areas.

Socio-economic variables were limited to a set often used in epidemiology and health geography. We include here variables associated to socio-economic deprivation. We used the share of unemployed persons in the 18-64 year-old population (noted *"Unemployed"*), the share of individuals aged 25-34 with primary education only (noted *"Primary edu 25-34y"*), the share of individuals aged 25-34 with a higher education degree (noted *"Higher edu 25-34y"*), the share of manual workers (noted *"Manual workers"*) and the share of foreign-borns coming from low income countries (noted *"Foreign-born"*). The Belgian Statistical Office (StatBel) provided data on median income and population density on 2014, while socio-economic variables were derived from the Belgian Census of 2001: we expect that trends kept consistent. The descriptive statistics about these variables are reported in table A.3 in appendix.

5.2.4 Methodology

The age-standardised medication reimbursement prevalence data for each of the five groups of medications were first mapped at the municipality level. For all these maps, Jenks discretization was applied in order to minimize the intra-class differences and maximize inter-class differences. The visual output given by the map is accompanied by the computation of a Moran I index for quantifying the spatial autocorrelation, and hence how far contiguous municipalities look alike.

Correlation between all five ASMP variables was then assessed using Pearson correlation (table 5.3). A hierarchical clustering of the five types of medications was then computed using Ward's method to identify clusters of municipalities with similar medication reimbursement patterns (figure 5.3).

To extract a reduced and relevant set of correlated variables, a Principal Component Analysis (PCA) was also performed on all the variables presented in table A.3. The PCA allows reducing the number of variables into a small set of correlated variables. Due to the independence of the ASMP variables in the first PCA (all grouped into one factor), these were excluded from a second PCA computation. Thus, a PCA including only the environmental and socio-economic variables was performed and three factors were extracted. These factors were then combined with the five groups of medications to perform a clustering and identify municipalities that were similar in terms of medication reimbursement practice and environmental and socio-economic characteristics.

A summary of the data used and of the analysis carried out in the paper can be found in figure 5.1: the scheme summarizes the sets of data used (ASMP, environmental and socio-economic (ENV-SE), and 3 factors derived from a PCA on ENV-SE (3F ENV-SE)), the type of analysis carried out, and the reference to the outcome of the analysis (figure or table numbers)

By adding environmental and socio-economic variables to the reimbursement data the effect of such exogenous factors could clarify spatial variations in medication reimbursement.

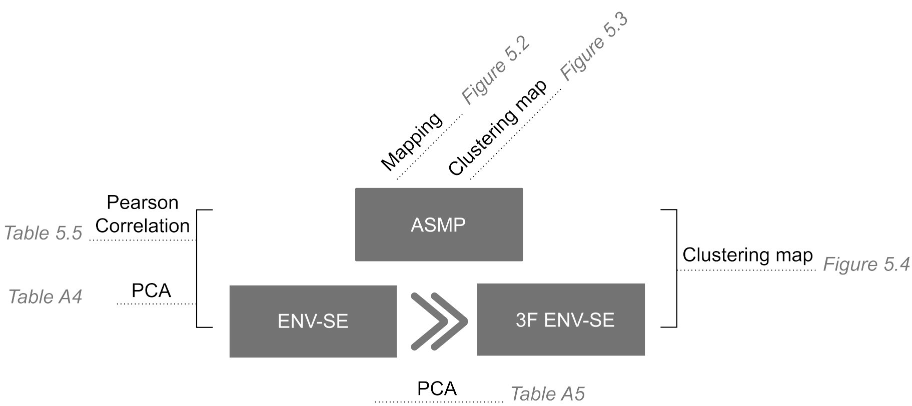


Figure 5.1. Schematic outline of the data included and the analyses performed in the study

5.3 Analyses and results

5.3.1 Mapping the Age Standardised Medication reimbursement Prevalence

Basic descriptive statistics of the five ASMP variables are reported in table 5.2. Results for both men and women are presented separately. Since the gender-difference is not significant (similar descriptive statistics) and comparable medication reimbursement patterns (see correlation coefficients noted "r" in table 5.2), we focus on reporting the results for men only (women data and results are not illustrated but available upon request).

The maps of the prevalence of medication reimbursement are shown in figure 5.2. Medication reimbursements for the immune system (IS) are more prevalent in the north east of the country, in the Province of Antwerp, and partially in the South-East near the border with France (South in the Province of Hainaut, and in the western part of the Province of Namur). The Province of Hainaut and Namur are characterized by a high prevalence of medication reimbursements

Table 5.2. Environmental and socio-economic variables employed in the study by gender and medication group.

Medication group	Gender	Mean	Std dev	Min	Max	Median	*r*
RS	men	14.23	2.33	5.44	20.4	14.09	0.95
	women	14.01	2.52	5.34	21.99	13.84	
CVS	men	27.64	2.87	14.47	34.09	27.6	0.95
	women	27.48	2.64	15.05	33.01	27.57	
LIP	men	15.26	1.85	6.51	20.57	15.43	0.86
	women	15.21	2.09	7.23	21.16	15.26	
IS	men	10.78	2.33	3.09	19.44	10.73	0.97
	women	10.67	2.12	2.84	19.61	10.51	
CNS	men	16.7	2.88	7.33	25.54	16.72	0.96
	women	16.83	2.75	8.01	23.59	16.66	

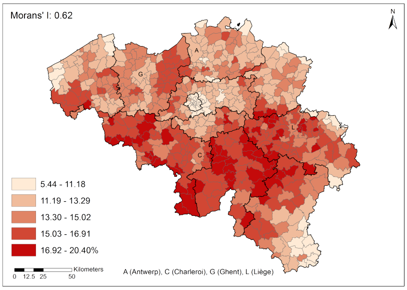
Age-standardised yearly prevalence of reimbursement of 5 categories of medication prescribed in 588 Belgian municipalities between 2012-2014. ("Std dev": standard deviation, "Min": minimum value, "Max": maximum value, "r": Pearson correlation coefficient between ASMP men and ASMP women for each medication group)

for the respiratory system (RS), the cardiovascular system (CVS), and for the central nervous system (CNS), all expanding towards the East. Wallonia, with the exception of Walloon Brabant, is generally characterized by high values of the prevalence of medication reimbursements.

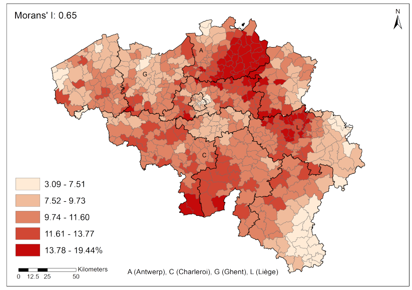
Spatial autocorrelation is measured by the Moran's I index (reported on each map in figure 5.2). Spatial autocorrelation has been assessed considering all municipalities sharing at least one border (direct neighbors). All five ASMP measures are characterized by significant and positive autocorrelation, all varying between +0.5 and +0.67: municipalities that look alike tend to be closely located.

5.3.2 Comparing ASMP

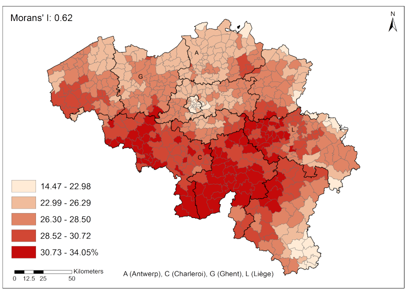
The Pearson correlation coefficients (table 5.3) are positive and statistically significant between the five medication groups. The strongest correlations are observed between the cardiovascular system medication group (CVS) and all other medication groups with the exception of immune system targeted (IS) medications. The cardiovascular medication included here is not only indicated to treat diseases but, like the lipid modifying agents, it is also prescribed to prevent the onset of chronic cardiovascular diseases or acute events. The respiratory medication is mainly used to treat symptoms of asthma (childhood and early adulthood) and COPD (adulthood and elderly), and COPD shares many risk factors with cardiovascular diseases.



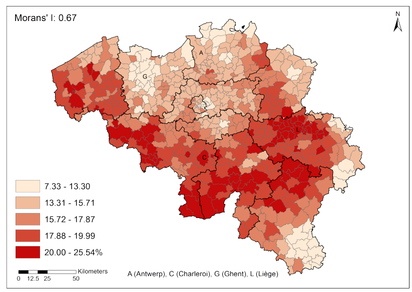
(a) Medication for respiratory obstructive airway diseases such as asthma or COPD (RS)



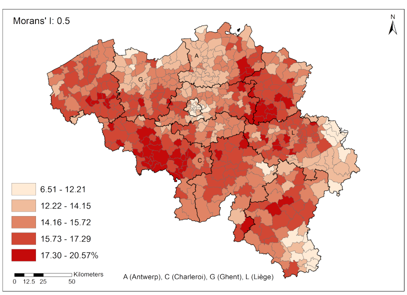
(b) Medication to treat allergy symptoms (topic and systemic) (IS)



(c) Medication used to reduce blood pressure and to treat other cardiovascular problems like ischemic diseases (CVS)



(d) Drugs acting on the central nervous system (CNS)



(e) Medication prescribed to reduce hyperlipidemia (LIP)

Figure 5.2. Spatial variations of medication prescription, and Moran's I index (indicative of spatial auto-correlation between municipalities), for the five medication groups, for men in 588 municipalities in Belgium between 2012-2014

It is therefore expected that COPD medication sales (and reimbursement) are strongly correlated with sales (and reimbursement) of cardiovascular medication. Finally, the strong correlation that we can be observed between CNS medication sales reimbursement mainly used among patients with mood disorders and the cardiovascular medication sales is consistent with the increasing evidence that mental health problems may be risk factors of cardiovascular disease (Cohen *et al.*, 2015).

Table 5.3. Correlation coefficients between the five groups of medication in the 588 municipalities (all significant at 0.001)

	RS	**CVS**	**LIP**	**IS**	**CNS**
CVS	**0.81**	1			
LIP	**0.56**	**0.75**	1		
IS	0.37	0.37	0.27	1	
CNS	**0.65**	**0.78**	**0.58**	0.42	1

5.3.3 Clustering ASMP

Based on the significant results of the correlation analyses, a Ward clustering method is applied on the 5 ASMP variables to understand whether there is a clear spatial pattern in the medication reimbursements of the five groups of medication (figure 5.3). Three clusters are identified. Contrary to our expectations, spatial fragmentation is small: each cluster is rather made of contiguous municipalities. Cluster 1 (yellow) extends primarily in the North of the country (Flanders), in the south-east in the provinces of Liège (including among others the German speaking Community of Belgium) and partly in the Province of Luxembourg. Despite its location within Wallonia, most of the municipalities of the Brabant Wallon are grouped into Cluster 1. Cluster 2 (green) mainly includes: 1) municipalities in and around Brussels, and also 2) municipalities close to the border of Belgium and/or in the vicinity of an important international work place (i.e. Luxembourg). Cluster 3 (red), is strongly present in Wallonia, although it also includes a few municipalities in West-Vlaanderen. Cluster 3, extending in most of the Wallonian region, usually defined by higher percentage of vegetation, is characterised by the highest cluster means for each variable, pointing to a higher prescription of medication. Cluster 2, close to the borders with Luxembourg and Germany, as well as in the southern and eastern municipalities of and around Brussels, instead is characterized by the lowest values of cluster means, suggesting a lower prescription prevalence or possibly purchase from the neighboring countries. Table 5.4 reports the mean values for the five variables in each cluster. While cluster three is characterised by higher average prescription prevalence, cluster two (the smallest cluster comprising only 56 municipalities) is characterised by municipalities with low prescription prevalence on average. The mean values of cluster one, the largest of the three clusters, set in between the values of cluster two and three. Such results are not surprising considering the strong correlation of some of the variables included in the clustering. This clustering enables us to reveal the spatial pattern of medication prescription.

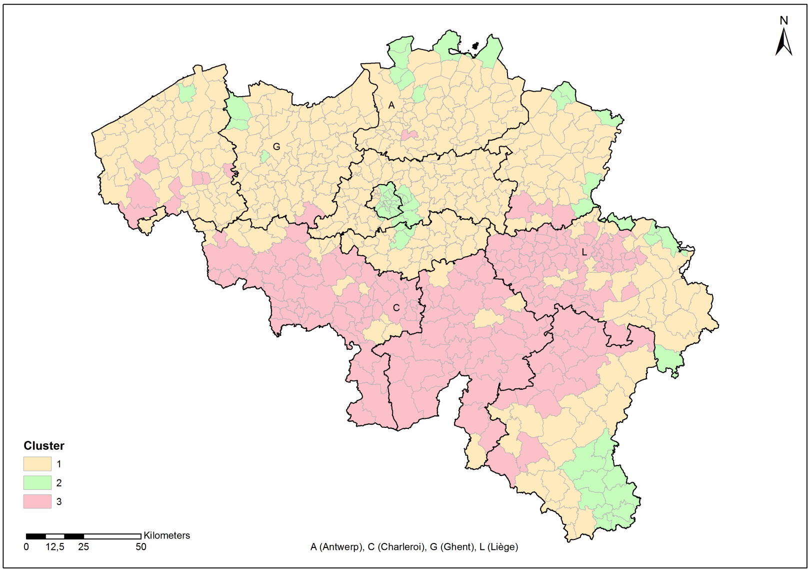


Figure 5.3. Clusters of medication reimbursements for men by municipality in Belgium

Table 5.4. Cluster composition for the five ASMP variables (mean values, and standard deviation in brackets)

	Cluster 1 (N=361)	**Cluster 2 (N=56)**	**Cluster 3 (N=171)**
RS	13.68 (1.44)	10.27 (1.6)	16.68 (1.3)
CVS	27.05(1.55)	22.15 (2.16)	30.67 (1.34)
LIP	15.24 (1.5)	12.21 (1.57)	16.28 (1.45)
IS	10.89 (2.35)	7.46 (1.66)	11.64 (1.32)
CNS	15.87 (1.97)	12.53 (1.84)	19.82 (1.48)

5.3.4 ASMP and environmental and socio-economic variables

The thirteen environmental and socio-economic variables defined in Section 5.2 are now considered. These variables are potentially related to the five ASMP variables as they are a priori assumed to play a (major) role in medication consumption/health issues. Positive correlations between medication reimbursement and environmental variables are limited to the percentage of total green in the municipality and moderately between allergies and $PM_{2.5}$. Relationships between pollutants and medication reimbursements appear not to be dependent of socio-economic indicators. Medication reimbursements are poorly and negatively correlated to environmental variables; this does not mean that pollution and health are weakly associated but at least that the spatial variations of the measures are weakly associated if no control variable is considered. Most ASMP variables are positively correlated (table 5.5) with the proportion of unemployed, while central nervous system medication reimbursement is positively correlated to the share of population (25-34y) with primary education only. Strong negative correlations can be found between the proportion of higher educated people (25-34y), and foreign born and most groups of medications (except for immune system (IS), and central nervous system (CNS), respectively). Also, median income and population density show both a negative sign with asthma and COPD (RS), and cardiovascular problems (CVS), and median income is also negatively correlated with depression and anxiety (CNS).

Table 5.5. Correlation coefficients between the five ASMP variables (RS, CVS, LIP, IS, and CNS) and the environmental and socio-economic variables for the 588 municipalities

	RS	**CVS**	**LIP**	**IS**	**CNS**
$PM_{2.5}$	-0.29***	-0.29***	n.s.	0.15***	-0.3***
PM_{10}	-0.2***	-0.19***	n.s.	0.12**	-0.22***
NO_2	-0.42***	-0.43***	-0.22***	n.s.	-0.29***
BC	-0.33***	-0.32***	-0.18***	n.s.	-0.23***
Manual workers	n.s.	n.s.	0.18***	0.17***	n.s.
Unemployed	0.37***	0.41***	0.16***	n.s.	0.52***
Foreign-born	-0.28***	-0.29***	-0.28***	-0.19***	-0.08*
Primary edu (25-34y)	n.s.	n.s.	n.s.	-0.14***	0.23***
Higher edu (25-34y)	-0.36***	-0.4***	-0.28***	-0.11**	0.4***
Median income	-0.26***	-0.5***	-0.1**	n.s.	-0.5
Population density	-0.3***	-0.31***	-0.26**	-0.13**	-0.13**
Ave near GS	-0.33***	-0.3***	-0.26***	-0.14***	-0.15***
Total GS	0.32***	0.3***	0.16***	n.s.	0.14***

Significance levels: ***: 0.001, **: 0.01, n.s.: non significant)

A PCA is first performed on all environmental (ENV), socio-economic (SE) and ASMP variables in order to see how they co-vary in space (see table A.4 in the Appendix) at the municipal level within Belgium. Environmental, socio-economic and medication reimbursements data are observed to be independent from each other. Environmental variables ($PM_{2.5}$, PM_{10}, NO_2, BC, average nearest green space and percentage of green space) are all part of the first factor (Factor 1) together with population density. Factor 2 is characterized by socio-economic variables, exception for the proportion of manual workers and the share of young people with high education degree; the latter two variables have high loadings in Factor 4. Factor 3 is composed of the five ASMP variables, confirming the independence of the variables. A PCA is also performed separately for the two Belgian Regions (Flanders and Wallonia). Given the geography of the country, Brussels Capital Region is included once in the Flemish Region, and once in the Wallonian Region, to mirror the strong effect it plays in the two Regions. Results of the Anova (figure 5.4), confirm the independence of the three groups of variables (ASMP, environmental and socio-economic) discussed above.

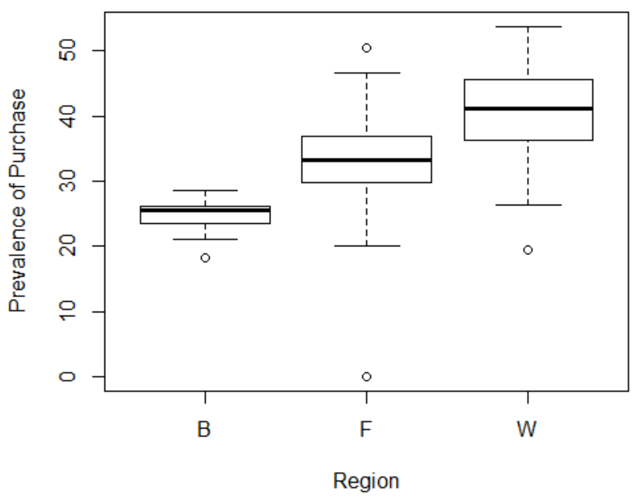


Figure 5.4. Analysis of variance of ASMP, environmental and socio-economic factors among Regions in Belgium (B: BCR; F: Flanders; W: Wallonia)

Considering the independence of the ASMP variables in the PCA results, an additional PCA analysis is conducted, limiting the variables to the thirteen environmental and socio-economic ones. Three factors result from the new PCA (table A.5 in the Appendix). The first factor (Factor 1) is characterized by air pollution ($PM_{2.5}$, PM_{10}, NO_2, and BC), low percentage of green areas and higher distance to green spaces, and by the proportion of individuals born in low-income countries. The second factor (Factor 2) is characterized by lower socio-economic characteristics: higher share of unemployed, lower education (primary education 25-34y), presence of foreigners born in low-income countries, as well as lower median income. The last factor (Factor 3) is composed of better-off socio-economic conditions (i.e. a small share of manual workers and higher education).

When we add the three PCA factors, which summarize the environmental and socio-economic variables, to the five medication groups and we perform a new cluster analysis, we obtain a similar structure, as it can be seen in figure 5.5 (table 5.6 describes the composition of the four clusters). The BCR with its periphery clearly appears in Cluster 3 –in green, alongside some border municipalities, close to Luxembourg and Germany. This cluster is characterized by low prevalence of reimbursement for all groups of medications, as well as a higher education level (factor 3). Cluster 4 (in red) mainly groups Walloon municipalities, following the regional border with one exception in the Province of Brabant Wallon. Wallonia is predominantly characterized by high reimbursement prevalence for medication for the respiratory system (RS), the cardiovascular system (CVS), and the central nervous system (CNS). The Province of Brabant Wallon is part of Cluster 1 (light blue). This cluster covers mainly the north of the country (Flemish Region), with some exceptions in the provinces of Liège and Luxembourg (around the green cluster). It is characterized by low prevalence of reimbursements for medications for the cardiovascular system and the central nervous system, a higher income (low unemployment and primary education). The last cluster (Cluster 2, in orange) spreads primarily to the north east and to the north west of the country, in the Flemish Region. This cluster is characterized by higher medication prescriptions for the immune system and lipids targeted medication, a lower share of higher educated people and more manual workers. This last clustering confirms the strong and clear spatial structure observed in the previous clustering of the ASMP variables without taking the environmental and socio-economic variables into account. With some exceptions, clusters seem to follow, again, the linguistic/regional border between Flanders and Wallonia.

5.4 Discussion

Results show a strong co-variation in space of the five medication groups analysed here, as well as their independence from environmental and

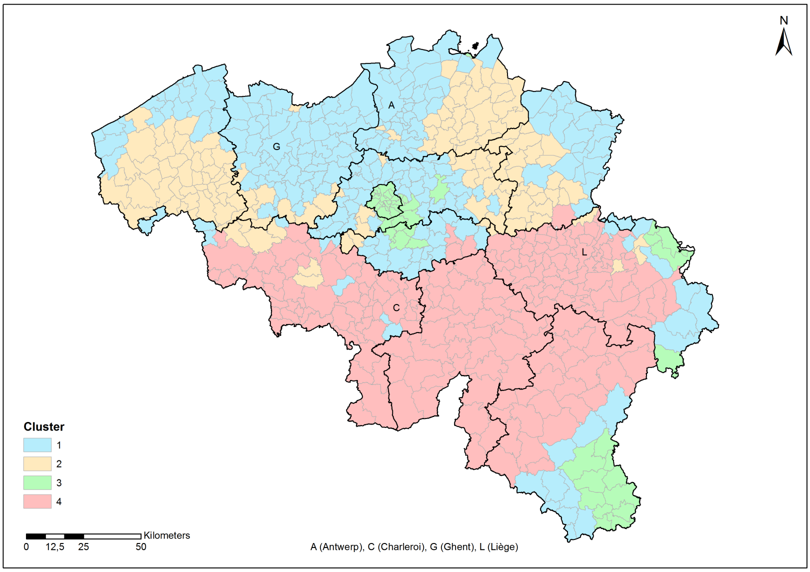


Figure 5.5. Clusters of medication reimbursements for men, and three PCA factors including environmental and socio-economic variables

Table 5.6. Cluster composition for the five ASMP variables and the three PCA factors including environmental and socio-economic variables (standard deviation included in brackets

	Cluster 1 (N=218)	**Cluster 2 (N=134)**	**Cluster 3 (N=49)**	**Cluster 4 N=187**
RS	13.34 (1.52)	14 (1.25)	10.1 (1.53)	16.5 (1.44)
CVS	26.2 (1.4)	28.05 (1.53)	22.24 (2.46)	30.43 (1.52)
LIP	14.58 (1.28)	16.36 (1.45)	12.09 (1.63)	16.09 (1.38)
IS	9.98 (1.75)	12.24 (2.6)	7.53 (1.8)	11.53 (1.5)
CNS	14.76 (1.64)	17.29 (1.86)	13 (2.28)	19.53 (1.66)
Factor 1	0.26 (0.73)	0.23 (0.48)	0.9 (1.97)	-0.7 (0.72)
Factor 2	-0.52 (0.66)	-0.35 (0.46)	0.79 (1.79)	0.65 (0.81)
Factor 3	-0.01 (0.9)	-0.89 (0.59)	1.54 (1.16)	0.24 (0.58)

Factor 1: high pollution and *Ave Near GS*; low *Total GS*. Factor 2: high *unemployment, primary edu, foreign born*; low *median income*. Factor 3: high *higher edu*; low *manual workers*

socio-economic factors. We used age-standardized data at the municipality level: this level of aggregation presents a very low number of missing data, avoids privacy issues and allows comparison with other contextual variables. Mapping medication reimbursements for men, across all medications, shows a very consistent spatial pattern, with a strong North-South difference.

Correlation coefficients between the five medication groups are as expected. Prescription of medication for respiratory disorders (asthma in children and young adults, and COPD in adults and elderly) are highly correlated with medication prescribed to prevent the the onset of chronic cardiovascular diseases or acute events, as they share many risk factors. Also, the correlation between medication for the central nervous system and cardiovascular system can be explained by the increasing evidence of that mental health problems may be risk factors for cardiovascular disease (Cohen *et al.*, 2015).

Medication reimbursement prevalence is generally higher in the southern part of the country, a region characterized by lower pollution levels, and higher share of green spaces. Medication reimbursement for the immune system is the exception, as the highest prevalences are observed in the north-east of the country, in the region of Antwerp. Low prevalence of medication reimbursement can be found in the big cities, presenting an urban-rural difference in purchase and reimbursement. Our results show that all medication reimbursements variables are highly and positively correlated, and unexpectedly co-vary in the Belgian space. All medication reimbursements types are highly autocorrelated (Moran's I ≥ 0.5). Clustering the medication reimbursements for the five medication groups shows a clear spatial pattern. The regional (and linguistic) division appears clearly, as well as the effect of the international borders (with Luxembourg in particular), together with the possible effect of the presence of expats with special status in Brussels. This structure cannot be explained by environmental factors, such as pollution, which are not affected by administrative partitions, but is suggestive of a bias in prescription practices.

To our knowledge, few studies have analyzed the variations in space of several environmentally linked health issues at once. Waldhoer *et al.* (2008) looked at the spatial distribution of infant mortality in Austria for different causes of death. They found a clear non-random spatial distribution of mortality causes, and a non-relevant effect of socio-economic status of the mother. We included in this study a selection of relevant environmental and socio-economic variables in addition to the medication reimbursement data to better understand spatial variations in medication prescription. We have calculated Pearson correlations among ASMP variables and the environmental and socio-economic ones. This yielded unexpected inverse correlations with air pollution ($PM_{2.5}$, PM_{10}, NO_2, and BC) and distance to green spaces, while the correlation was positive with the share of green spaces. These results are contrary to the existing literature supporting the adverse effect of air pollution on health, and the beneficial effects of natural environment (Landrigan *et al.*, 2017; Pereira *et al.*, 2012; Sofianopoulou *et al.*, 2014; Thurston *et al.*, 2017; Triguero-Mas *et al.*, 2015). Low socio-economic position (share of unemployed, manual workers, young people with primary education only) appears to be positively correlated with medication reimbursement, with the exception of individuals coming from low-income countries.

Principal Component Analysis including medication reimbursement data, environmental and socio-economic variables indicates that there is no relationship between the medication reimbursement and all other variables, confirming again the strong correlation and independence of such variables from external factors. The clustering analysis, including the five medication reimbursement groups and three PCA factors of environmental and socio-economic variables, presents a strong spatial structure. The addition of environmental and socio-economic variables does not have a relevant impact on the spatial pattern of medication reimbursements, as ASMP appear to be highly correlated between each other and independent from external factors. Unfortunately, the data available does not allow for a more in-depth analysis at the individual level, entailing a possible ecological fallacy in the results, if we were to look at the determinants of health status. Here, analyzing spatial variations at the national level allows the identification of a regional structure that raises the question of the role of the medical practice on medication prescription, and thus reimbursement in Belgium. Practitioners' behaviour is prescribing medication, as shown in previous studies (Ekedahl *et al.*, 1995; Figueiras *et al.*, 2000), can be determined by differences in education or training.

The data used for this study is based on reimbursed medication by the seven Belgian sickness funds for each individual who is covered by the Belgian security system. The IMA-AIM database has hence limitations that need to be acknowledged. First, 2% of Belgian residents is not covered by the social security system and therefore, not included in our study. Second, non-reimbursable medications is not included in the database, and thirdly, the purchase and reimbursement of a medication does not equate its consumption. These biases can unfortunately not be estimated. Moreover, averaging pollutants such as BC and NO_2 at the municipality level may introduce some estimation biases. The use of aggregated data may induce to ecological fallacy. Last but not least, we here use the latest census data available (2001 and 2014) as source for relevant socio-economic variables: we expect that the trends do not have significantly changed in the past years.

The patterns of medication reimbursement in Belgium highlighted in this paper indicate that the lack of information at the individual level does not really allow us to measure health, and to draw conclusions on the effects of factors such as environmental characteristics and socio-economic position on well-being. Our data rather hints at medical and pharmaceutical practices, and reminds us on keeping in mind the existence of geographical differences in medical practice, when using medication as health proxy. Information on the education and training of medical doctors and their current working location might be useful to further understand medication reimbursements distribution in Belgium.

5.5 Conclusion

Using data on medication prescription reimbursement for five groups medication, this chapter analyses whether spatial variations in medication reimbursements in Belgium are associated to environmental factors and/or socio-economic factors. The use of data at the country level and for several medication groups enables us to get a wider perspective on the factors leading to medication purchase in Belgium for all age groups. Using several medication groups enables us to identify biases of the prescription system (prescription practices rather than specific characteristics of the physical environment and socio-economic position). The study shows that medication prescription and reimbursement is not (only) an environmentally driven problem at least at the level of the municipalities. Variation in medication prescription and reimbursement appear to be linked to curative healthcare matters and language/education of practitioners. Therefore, it is important to be careful when using and interpreting such data as health proxies: this chapter reminds us how important it is to control for the real meaning of the data, as well as the level of data aggregation. Both factors do orient the interpretation of the results but are hard to change due to the type and structure of the data available.

The spatial distribution across the country of the five groups of medication could reflect true spatial variations in morbidity but also geographical variations in prescription practices. In the absence of strong indications for similar independent influences of environmental factors, such as air pollution, or socio-economic factors, as evaluated at the municipality level, we conclude that prescription practices may be a dominant factor.

6

Geographies of asthma medication purchase for pre-schoolers in Belgium

Following the research carried in the previous chapter on the spatial patterns of reimbursements for five groups of medication, this chapter focuses on one drug only and to a restricted group of the population. This study aims to describe the spatial differences in asthma medication sales for children aged 0 to 5 years in Belgium, and to evaluate the correlations with socio-economic and environmental factors. This ecological study includes information on prescribed and sold asthma medication (R03 ATC code) for all the pre-schoolers registered in Belgium, aggregated at the municipality and smaller administrative levels, in 2014. The number of cases, the prevalence of medication purchase and the costs per case (€) are mapped at different administrative levels. Finally, Pearson correlations of medication prevalence and costs per case with socio-economic (population density and area median income), and environmental (PM_{10}, pig density and green spaces) factors are computed. It is shown that in Belgium, the geographical distribution of purchases of prescribed asthma medication for pre-schoolers is only weakly correlated with some environmental factors. The existence of clear differences in prevalence of purchase between Flemish and French-speaking regions suggests that the observed geographical distribution is strongly driven by regional differences in medical practice.

The present chapter is based on the publication from Trabelsi *et al.* (2019) in Respiratory Research Journal

6.1 Introduction

Wheeze is a common symptom among pre-schoolers, with an estimated cumulative prevalence at the age of 6 of up to almost 50% (Martinez *et al.*, 1995; Bisgaard and Szefler, 2007). Wheezing can be a manifestation of many diseases, including respiratory infections and asthma, and it can be triggered by environmental factors such as air pollution, allergens or tobacco smoke (Ducharme *et al.*, 2014; Brand *et al.*, 2008), some of them very much dependent on the geographical location of the residence. The management of pre-school wheezing is challenging and includes measures to avoid environmental triggers and medication (Brand *et al.*, 2008; Global Initiative For Asthma (GINA), 2017). Nevertheless, evidence on which to base recommendations in pre-schoolers is limited (Brand *et al.*, 2008) and this results in high numbers of prescriptions (Bianchi *et al.*, 2009; Ferreira de Magalhães *et al.*, 2017) and substantial costs (e.g. 0.15% of the healthcare budget in the UK (Stevens *et al.*, 2003)). In addition, the choice of a treatment is influenced by parents or caregivers' interpretation of the symptoms as well as medical practice by primary care doctors and paediatricians Levy *et al.* (2004); Michel *et al.* (2006), which may be geographically determined.

The concentration of people and activities in large cities leads to many positive effects (Fujita and Thisse, 2011), but also to negative externalities, including socio-economical segregation or high levels of traffic-related air pollution (World Health Organization United Nations Environment Programme, 1992). Children growing up in rural areas may, in contrast, benefit from protective effects on respiratory health due to lower exposure to traffic-related air pollutants, to a more diverse microbial environment or to specific lifestyles (e.g. consumption of unprocessed milk) (Strachan, 1989; Müller-Rompa *et al.*, 2018). However, living in rural areas may also imply poor access to health care (Bauer *et al.*, 2018) and high exposure to other pollutants (e.g. pesticides) (Mamane *et al.*, 2015) or pro-inflammatory agents like endotoxins (de Rooij *et al.*, 2017). In this regard, previous research suggests that the prevalence of pre-school wheezing, the risk of asthma diagnosis, and that of the use of medication for airway-related symptoms may differ between urban and rural environments. While studies comparing urban and rural prevalence of respiratory symptoms in pre-schoolers show a higher prevalence and risk of asthma diagnosis in urban areas (Marfortt *et al.*, 2018; Kutzora *et al.*, 2018), the few studies investigating differences in the use of asthma medication show inconsistent results (Kutzora *et al.*, 2018; Ferreira-Magalhães *et al.*, 2016). A potential explanation for these inconsistencies may be the differences in health care access or in medical practice across countries or regions.

Here, we describe the spatial differences in asthma medication prescribed to pre-schoolers within Belgium, and evaluate the correlations with some relevant socio-economic and environmental factors, such as population density, income, air pollution, farming activity and greenness.

6.2 Materials and methods

6.2.1 Studied area

Belgium is a small but densely inhabited European country (11.35 million inhabitants over 30,528 km^2), administrated federally into three Regions: the Flemish, the Walloon and the Brussels Capital Region (see figure A.1 in the Appendix). Belgium is further divided administratively into 10 provinces, 43 “arrondissements” and 589 municipalities (see table A.1 in Appendix). Recently, medication reimbursement data have become available at different statistical levels: provinces, municipalities, statistical sections and statistical sectors (Jamagne, 2012), enabling to study spatial variations at different levels of aggregation. The smallest statistical sector is 0.01 km^2 large and the largest one is over 6.3 km^2; 24% of the statistical sectors contain less than 100 inhabitants.

The three Regions differ in language, as well as culture, traditions, living habits, historical, or socio-economic conditions. Although social security and health care are federal matters governed at the scale of the entire country, education (including higher education) is a "community" matter (OECD, 2017) and, thus, governed regionally. Except for Brussels Capital Region, mobility of workers between regions in Belgium is limited (Thomas *et al.*, 2017) and only very few medical doctors practice in another Region than the one where they obtained their degree. Also, very little communication takes place between medical practitioners across the regional (linguistic) divide. It is also important to note that, for historical and geographical reasons the Belgian city network is very tight in the North, leading to high densities of population and activities and, consequently, to a high level of air pollution (Belgian Interregional Environment Agency, 2014), while in the South the city network is - on the average - looser, inter-city distances are much larger and landscapes more hilly and green. In most Belgian cities, better-off people tend to live in the peripheries, and poor people in city centres or in former industrial areas (Mérenne-Schoumaker *et al.*, 2015; Querriau *et al.*, 2011; Brueckner *et al.*, 1999; Dujardin *et al.*, 2008).

We analysed the geographical distribution of prescriptions of asthma medication to pre-schoolers at three levels, with increasing degrees of spatial resolution: first, the whole of Belgium, using its 589 municipalities as spatial units; then, the former province of Brabant, using 375 former municipalities (as they existed before being merged in 1977) as spatial units; and, finally, the Brussels Capital Region (BCR), using its 147 statistical sections as spatial units.

The former province of Brabant is the central area of Belgium that contains, in its centre, the Brussels Capital Region (BCR), a politically defined entity (161 km^2 and 1.2 Mio inhabitants), surrounded by “Vlaams-Brabant” (2,106 km^2, 1.1 Mio inhabitants, Flemish speaking) and “Brabant Wallon” (1,090 km^2, 0.4 Mio inhabitants, French speaking). Whereas the BCR is fully urban, both Flemish and Walloon Brabant contain urban, suburban and some rural areas.

6.2.2 Medication purchases

Almost all Belgian residents (98%) are enrolled in the social security system. The information on individual health care expenditures that are reimbursed by the seven "sickness funds" is centralized by a single agency [Intermutualistisch Agentschap – Agence Inter-Mutualiste (IMA-AIM)]. The agency’s database contains detailed records of all reimbursed drugs by ATC code (Anatomical Therapeutic Chemical (World Health Organization and Norwegian Institute of Public Health, 2013)) over defined calendar periods, for each patient identified by national security number, age, gender and place of residence. The use of these data is governed by the principles of the Declaration of Helsinki. For the present study, we obtained data aggregated by administrative spatial units (provinces, current municipalities, statistical sections and statistical sectors) on medication, reimbursed in 2014, included in the R03 ATC code (i.e. "Drugs for obstructive airway diseases", (WHO Collaborating Centre for Drug Statistics Methodology, 2018)), with the limitation of not having any information on the exact medication type or the frequency of use.

We limited our analyses to children aged 0 to 5, named here pre-schoolers. Our database did not allow us to get finer age groups, but it was possible to compare pre-schoolers to older children: 6-12 years and 13-18 years. The diagnosis of asthma can indeed be established more firmly in children above 6 years, but because wheezing is less frequent among primary schoolchildren and adolescents, the numbers of individuals receiving asthma medication are much lower than among pre-school children, thus leading to statistical representativeness problems, especially in small spatial units. Nevertheless, the geographical distribution of reimbursements for asthma medications for schoolchildren (6-12y) and adolescents (13-18y) exhibited the same spatial structures (in appendix, figures A.2 and A.3), and similar but statistically less representative results were obtained for the Pearson correlations between the prevalence of asthma medication reimbursements and the environmental variables described for pre-schoolers (0-5y) in table 1.

The data used here are: (1) the number of cases in an administrative spatial unit (municipality, former municipality, statistical section), in other words, the number of pre-schoolers aged 0-5 years that benefitted from reimbursement at least once in 2014 for any purchase of a prescribed drug with an R03 ATC code, (2) the total number of pre-schoolers registered in the social security system in 2014 in the administrative spatial unit, and (3) the total cost (reimbursed + out-of-pocket contribution) of the medications reimbursed in 2014 in the administrative spatial unit. It is worth noting that in Belgium the price of the medications is fixed at the federal level (Annemans *et al.*, 1997), and prices are, therefore, identical in all three Regions. For privacy reasons, IMA-AIM does not provide the exact address of the patients but only the code of the administrative spatial unit of his/her current declared residence, and information is not provided if less than 5 cases of registered pre-schoolers are recorded in a spatial unit. The three variables described above ((1) number of cases, (2) number of registered pre-schoolers, and (3) total costs of reimbursed medication) were then further combined into two relative variables: the prevalence of purchase (in %) and the costs per case (in €). The prevalence of purchase was computed as the number of cases (1) divided by the number of registered pre-schoolers (2) and multiplied by 100. To further explore potential geographical differences in the frequency of medication use, we also computed the costs per case for each administrative spatial unit. This is the total costs for reimbursed asthma medication (3) divided by the number of cases (1); it is interpreted as follows: the higher the costs per case, the higher the probability of having chronic users in an administrative unit. Thus, two areas with the same prevalence may be characterised by completely different consumption profiles: sporadic users (low cost per case) and chronic users (high cost per case).

6.2.3 Characterising the environment

Data about green environment was derived from the Corine Land Cover dataset, a European land use dataset, developed by the European Environment Agency (European Environmental Agency, 2007). Four typologies of green spaces were initially considered (urban green, pastures and agricultural land, forests, and semi-natural areas) and further summed up. The percentage of the total surface occupied by these green spaces in each spatial unit was used as a measure of greenness. Data about air pollution was provided by the Belgian Interregional Environment Agency, as daily average Particulate Matter 10 (PM_{10}) values estimated for areas of 4x4 km. These estimates are based on a spatial-temporal (Kriging) interpolation model that combines pollution data collected on a half-hourly basis by a dense network of automatic monitoring stations and satellite-based land cover images (Janssen *et al.*, 2008). These data were used to calculate area-weighted average concentrations of PM_{10} per spatial unit. The value of the 95th percentile (noted $PM_{10}P95$) of the yearly average of PM_{10} per municipality was also considered.

The 2014 data on population density and median income at different spatial levels, as well as the information on the density of pigs, were obtained from the Belgian statistical office (Statistics Belgium).

6.2.4 Methodology

Data were mapped using two different methods: first, a proportional symbol map, where the surface of the circles is proportional to the number of reimbursements. This type of maps gives an idea of the spatial variation of the absolute numbers. Secondly, we used choropleth maps where spatial units are shaded according to prevalence, i.e. the relative numbers of medication reimbursements; the limits of the classes of values are fixed by a natural breaks classification (Slocum, 1999). The latter minimises intra-class variation and maximises inter-class variation. However, for the cross-scale analysis (figure A.4 another discretisation method was used to enable comparisons between maps: the standard deviation method (Slocum, 1999).

To examine the relationships of prevalence of purchase and costs per case with socio-economic and environmental variables at the level of the municipality, we calculated Pearson correlations. Although data were available for all municipalities, considering smaller spatial units, such as former municipalities and statistical sections, automatically lead to missing data problems (privacy and statistical issues) and the non-availability of some variables for very small units. Spatial units with missing data were removed from the analysis.

6.3 Results

Basic descriptive statistics about prevalence of purchase, cost per case and environmental and socio-economic variables at the level of the Belgian municipalities are reported in table A.6 (see Appendix).

6.3.1 Geographic variations at the country level

The spatial distribution of the absolute numbers of pre-school age children that benefitted from at least one prescription of asthma medication in 2014 (ATC code R03) is shown in figure 6.1. Numbers vary from a minimum of 78 to a maximum of 44,144, with an average value of 1,281, and a median value of 744. The map clearly reflects the tight urban network of the northern part of the country and the less populated southern part of the country. Also the big cities (Antwerp, Ghent, Brussels, Charleroi and Liège), clearly appear. In contrast, the map of the prevalence of purchase of asthma medication (figure 6.1b) showed low values in large cities (minimum of 18%), whereas the highest values were observed mainly in Wallonia (with a maximum of 54%), evidencing the country's North-South division with an average prevalence of 33% (SD = 5.5) in the Flemish region, and 41% (SD = 6) in the Walloon region.

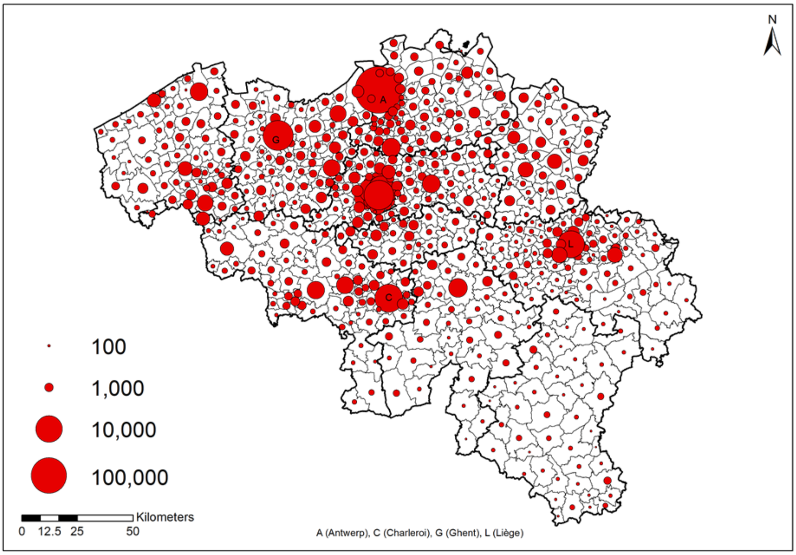
As for the absolute number of cases, total costs (figure 6.1c) were high in urbanized municipalities, with a maximum expenditure of 291,825 € (and a minimum of 735 €); on average the cost per municipality was 17,438 € (median 11,332 €). However, costs were also much higher in the North Western part of the country. This structure persisted and was even stronger when considering the relative expression: cost/case (figure 6.1d). In Belgium, high values of the cost per case were concentrated in areas where the prevalence of purchase (figure 6.1b) was relatively small.

Table 6.1 provides pairwise Pearson correlation coefficients between prevalence of purchase and cost/case on the one hand, and population density, income and the selected environmental factors, on the other hand. Confirming the visual appraisal (figure 6.1b), prevalence of purchase was negatively correlated with population density and, as a consequence, positively related with the percentage of green; prevalence was negatively correlated with income and particulate air pollution. In other words, green areas far away from urban and polluted areas were characterised by a higher prevalence of purchase, confirming the results showed in the maps. In contrast, cost/case was poorly correlated with most variables but the correlation with PM_{10} was positive.

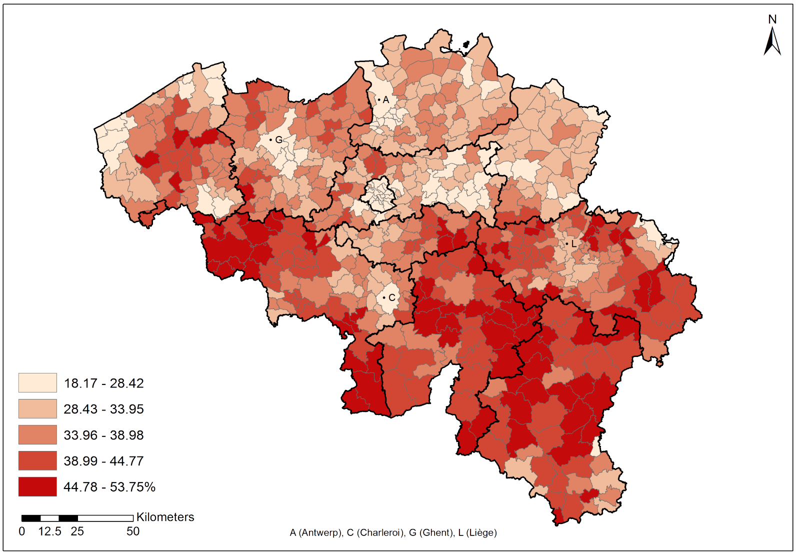
Table 6.1. Correlation coefficients of prevalence of purchase and cost/case and explanatory variables

	Prevalence of purchase	**Cost/case** (log)
Population density (log)	-0.587	-0.057 (n.s.)
Median income/declaration	-0.230	0.034 (n.s.)
$\mathbf{PM_{10}}$ (μ/m^3)	-0.444	0.172
$\mathbf{PM_{10}}$ **P95** (μ/m^3)	-0.470	0.221
Density of pigs (log)	0.379	0.329
% Green	0.544	0.145

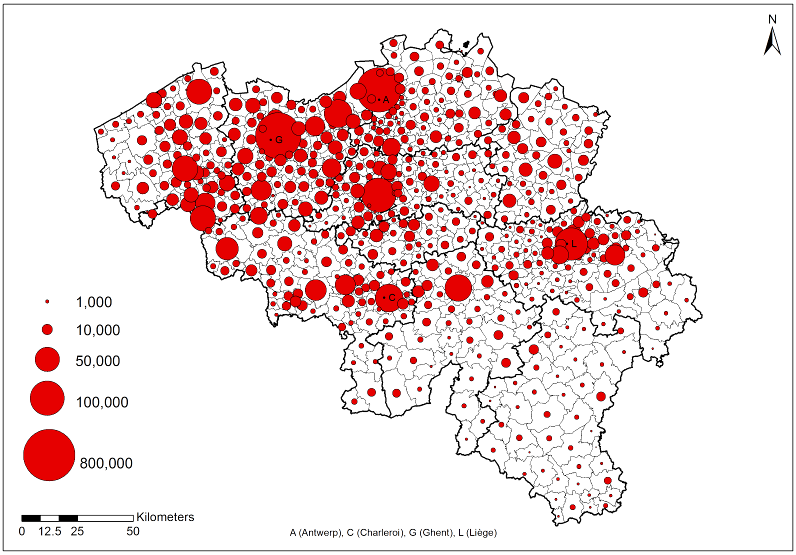
The table reports Pearson correlation coefficients between prevalence of purchase of asthma medication for pre-school children (left column) or cost/case (right column), and selected explanatory variables in municipalities in Belgium (significant α =0.001; *n.s.* Not significant; 588 municipalities)



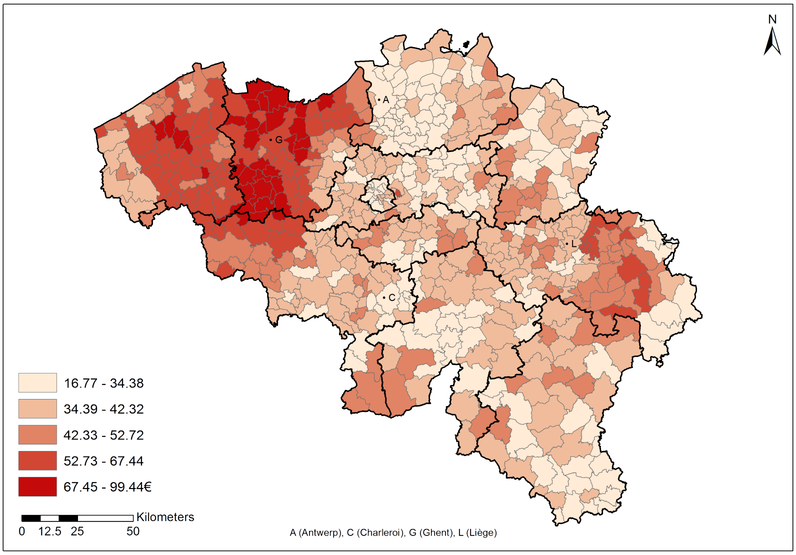
(a) Number of 0-5y children reimbursed at least once in 2014 for asthma medication



(b) Prevalence of purchase of asthma medication among 0-5y children in 2014



(c) Total cost per municipality for asthma medication for 0-5y children



(d) Cost/case per municipality for asthma medication for 0-5y children

Figure 6.1. Asthma medication purchases for pre-schoolers by municipalities (for comparability reasons the 19 municipalities of Brussels are summarised into 1 entity in figures 6.1a and 6.1c)

6.3.2 Zooming into Brabant and the Brussels Capital Region

Zooming into the former province of Brabant, the values of prevalence of purchase ranged between 16% and 64%, with an average of 34% (table 6.2). Spatially, prevalence of purchase was higher outside the BCR (figure 6.2), and mainly concentrated in the Eastern part of Brabant Wallon and in the Western part of Vlaams-Brabant. In contrast, the cost/case distribution (figure 6.3) appeared to occur at random without any spatial autocorrelation of the values.

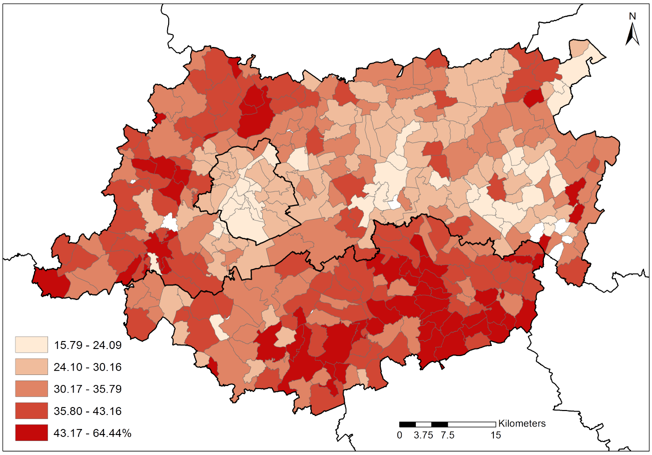


Figure 6.2. Prevalence of purchase of asthma medication in the former Province of Brabant among 0–5 y children (2014)

Figure 6.4 shows the intra-urban variation of prevalence of purchase at the level of the sections within the BCR, with a strong spatial autocorrelation observed in its Western and Eastern parts. Cost/case (not illustrated here) again varied randomly. The municipality of Uccle (the largest municipality in the S-E of the BCR) was conspicuous for its low values. The average value was much lower at the level of Brussels compared to the two Provinces (see table 6.2).

Figure A.4 (in the Appendix) presents the variation in prevalence of purchase for the three studied areas, but with a discretisation method based on standard deviations, allowing better comparisons between the three different scales.

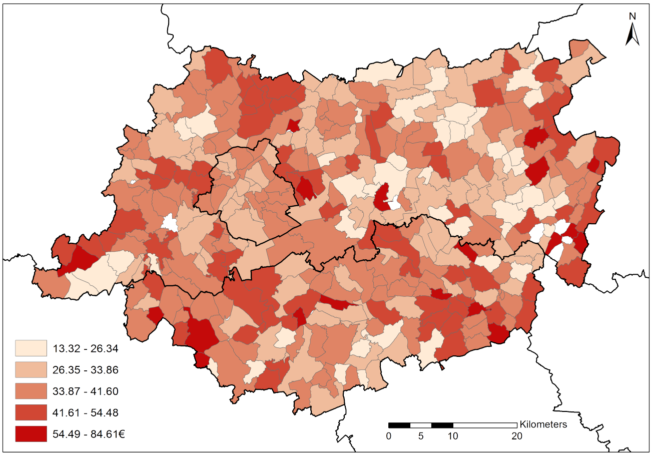


Figure 6.3. Cost/case of asthma medication in the former Province of Brabant among 0–5y children (2014)

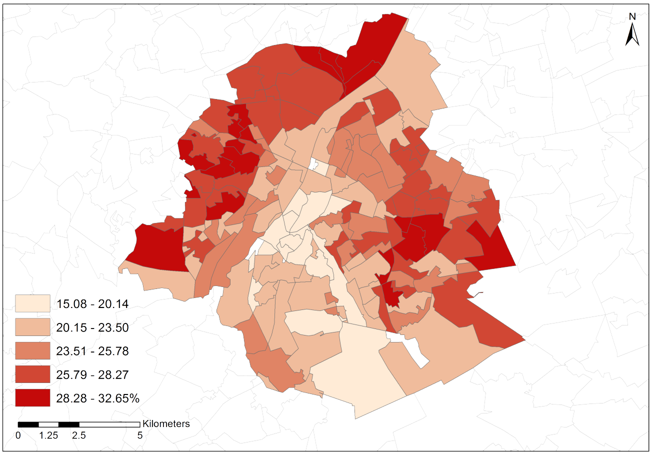


Figure 6.4. Prevalence of purchase of asthma medication among 0–5y children by statistical sections within the Brussels Capital Region (2014)

6.4 Discussion

We described the geographical distribution of asthma medication purchases for pre-schoolers in Belgium. The prevalence of purchase ranged between 18% and 54% in 2014, with higher values observed in the southern part of

Table 6.2. Descriptive statistics of prevalence of purchase and cost/case for the three studied areas

	Administrative level	**Spatial unit** Total N	Retained	**Prevalence of purchase** %	**Cost/case** €
Belgium	Municipalities	589	588	36.40 (7.01)	42.72 (13.05)
Brabant	Former municipalities	375	357	33.62 (8.38)	36.55 (11.15)
BCR	Statistical sections	147	141	24.60 (3.63)	34.32 (6.52)

The table reports descriptive statistics of prevalence of purchase and cost/case for asthma medication in the three nested studied areas (where "Retained" is the number of spatial units for which data were not missing)

Belgium, where population density and air pollution concentrations are low, and the amount of green spaces is the highest. In contrast, the costs per case, considered as a proxy for frequency of use, were higher in the north-western part of the country. To our knowledge, no previous publication has mapped such prescription/sales of asthma medication. Only one study investigated the trends and geographical distribution of hospital admissions due to paediatric asthma or asthma-like symptoms in France (Delmas *et al.*, 2013). The authors observed the highest rates in the North of the country. Although they hypothesized about environmental exposures, primary care attention, or socio-economic conditions as potential explanatory factors, they did not investigate potential correlations.

We calculated correlations with available environmental and socio-economic variables intending to understand the observed geographical distribution of asthma medication purchase. We found inverse correlations between prevalence of purchase and income as well as PM_{10} concentrations, and positive but moderate correlations with green spaces and density of pigs, considered as a proxy for a well-known type of polluting farming activity. Our results are consistent with previous research showing a high prevalence of wheezing in the lowest socio-economic groups (Koopman *et al.*, 2002; Chen *et al.*, 2017). Recent studies on green spaces and asthma-like symptoms showed inconsistent results, suggesting that this relationship is complex, with the concept of green spaces combining a series of protective factors (e.g. low air pollution) and risk factors (e.g. allergens) (de Rooij *et al.*, 2017; Koopman *et al.*, 2002; Chen *et al.*, 2017; Borlee *et al.*, 2017; Sonnenschein-van der Voort *et al.*, 2012). In our study, we did not have detailed enough information to disentangle this. Finally, the positive correlation we observed with farming activities is consistent with previous research conducted in the Netherlands where the number of livestock within a 1000m radius explained 12% of the variation in endotoxin levels in the air, a pro-inflammatory component of the bacterial wall (de Rooij *et al.*, 2017). Although living close to or in farms (Müller-Rompa *et al.*, 2018; Ege *et al.*, 2007) during childhood may protect against the development of asthma, we should keep in mind that only a small proportion of pre-schoolers receiving asthma medication are or will be confidently diagnosed with asthma (Martinez *et al.*, 1995).

So far, only two studies have specifically looked at the geography of asthma medication use and described urban vs rural differences in asthma symptoms and medication use in children. Although both studies showed higher prevalence of symptoms in urban children, their results for medication use are inconsistent (Kutzora *et al.*, 2018; Ferreira-Magalhães *et al.*, 2016). In Portugal, asthmatic children (0-17 years old) from rural areas were less likely to report using any inhaled therapy than children from urban areas (Ferreira-Magalhães *et al.*, 2016). In Bavaria (Germany), asthma medication intake among pre-schoolers was significantly higher in rural compared with urban areas (Kutzora *et al.*, 2018). Our results regarding prevalence of purchase are consistent with those shown in the Bavarian study. A potential explanation for this finding is related to health care access, as distances to health centres or general practitioners are larger in less densely populated areas (Dewulf *et al.*, 2013). However, another potential explanation for the high prevalence of purchases in less densely populated areas in Belgium is the medical practice. Previous research showed that despite the availability of international guidelines for the management of asthma, prescribing practices vary considerably across countries (Jepson *et al.*, 2000). The Belgian primary health care system is mostly based on small single-handed practices. Although international clinical guidelines are available, they are not compulsory and, therefore, they are not well followed. Moreover, guidelines may vary by linguistic region, with Flemish and French-speaking doctors being influenced by recommendations and practice in The Netherlands and France, respectively (Aujoulat *et al.*, 2015). Since the evidence on which to base recommendations in pre-schoolers is limited (Brand *et al.*, 2008), the "school of medicine" may play a role in the management of asthma-like symptoms in this age group. In this regard, a sharp regional (and linguistic) border is clearly apparent on our maps of the prevalence of purchase. Differences in prescription of medication between the two main linguistic communities in Belgium have not been published, however a study showing differences between Flemish-speaking and French-speaking regions in euthanasia practice in Belgium suggests that – admittedly, ill-defined – cultural factors do play a role in medical practice (Cohen *et al.*, 2012).

The distribution of costs per case did not follow the same pattern as that of prevalence. The costs per case were weakly but positively correlated with PM_{10} concentrations and green spaces and moderately with pig density.

Geographically, the area with high costs per case corresponds to the course of a major river (Scheldt) in the north-west, an area characterised by a slightly milder climate, and a high number of days with extreme PM_{10} and $PM_{2.5}$ pollutions (Belgian Interregional Environment Agency, 2014). In addition, the difference in the distribution of prevalence and costs per case may be related to the definition of both variables. While prevalence may be a proxy for both recurrent wheezing and children with only one wheezing episode in 2014, an area with high costs per case can be interpreted as an area with more recurrent wheezing, thus pointing with more confidence to an environmental explanation.

This study relied on information about the medications that are reimbursed by the sickness funds for every person covered by the Belgian social security system. Reimbursement data cover a broader range of disease severity compared with hospital admissions, emergency room visits or mortality (Samet and Krewski, 2007). However, some limitations need to be acknowledged. First, a small part (2%) of the Belgian population is not covered by the social security and, therefore, not included the statistics (i.e. mainly: undocumented immigrants, children of people employed by international bodies - EU, NATO, etc.). Second, the IMA-AIM agency does not provide information if there are less than 5 cases in a spatial unit (privacy issue). For R03 medication in pre-schoolers, this was the case for only one out of 589 municipalities in Belgium, eighteen out of 375 former municipalities in Brabant, and six out of 147 sections in the BCR. Third, we did not have information on asthma medication at individual level, nor for smaller age categories with the group of "pre-schoolers" (1 year for instance). We know that one third of the infants has at least one wheezing episode by the age of 1 (Bisgaard and Szefler, 2007), and a second peak in asthma medication use is observed at the age of 4 (Bianchi *et al.*, 2009). Taking here a 5-years-large group hence smoothens these effects. Analysis for R03 medication reimbursement for older children yielded a structurally similar spatial pattern with regard to the prevalence of purchase, including the striking north-south divide observed at national level. Last but not least, drug purchase does not necessarily equate drug use. Also, we were not able to consider the proportions of specific types of medication included in the R03 code (e.g. short-acting beta agonists, long-acting beta agonists, inhaled corticosteroids,...) which would have helped distinguishing children with recurrent wheezing and potential asthma from those having sporadic wheezing more likely related with a respiratory infection. Finally, although the costs of specific medications do not differ between regions, they do vary between types of medications and between generic vs non-generic types, which could be prescribed differently by region. Unfortunately, we also were not able to document these points with the provided data.

The ecological design of our study entails the possibility of ecological fallacy. We did not have any detail about the season of the treatment, the exact reason of the prescription, or the individual clinical record. In addition, although we did correct for socio-economic status at the administrative level, we did not know the individual's/household's socio-economic characteristics, lifestyle, exposure to pollutants, or distance to the closest healthcare centre. An in-depth individual survey would be necessary to further investigate the determinants of the distribution of asthma medication purchases. Nevertheless, as previously mentioned, the correlations between the studied health outcomes and potential determinants are supported by the results of previous individual studies, thus reducing the risk of presenting an ecological fallacy.

6.5 Conclusion

We showed that, in Belgium, the geographical distribution of purchases of prescribed asthma medication for pre-schoolers was only weakly correlated with some quite commonly used environmental factors. The fact that we observed clear differences in prevalence of purchase between Flemish and French-speaking regions, combined with the fact that evidence for recommendations in the management of pre-school wheezing is limited, suggests that the observed geographical distribution may be driven by differences in medical/pharmaceutical practices, rather than by environmental realities, and that medical purchases can hardly be considered as a good proxy for health.

IV

Conclusion

7

General conclusion

This chapter concludes this doctoral dissertation by giving an overview on the work carried hitherto aiming at exploring spatial variations in the relationship between green spaces and health in Belgium. Section 7.1 summarises the key findings of this contribution in terms of definitions and measures of green spaces, of the sensitivity of associations between green and specific health disorders, and finally of the meaning of medication reimbursement data in spatial analyses. Section 7.2 proposes a critical assessments of the results presented in this work, followed by suggestions of paths for further research to overcome some known limitations (section 7.3). Section 7.4 provides recommendations to policy makers and urban planners to tackle the challenges linked to the assessment of the environment and citizens' well-being. The last section wraps up this multidisciplinary work.

7.1 Summary of findings

Assessing the associations between natural environment (green spaces, mostly) and health status has proven to be a difficult and complex matter. Existing literature on the subject has shown that many factors come into play and affect the relationship in direct or indirect ways. Characteristics of the living environment, demographic and socio-economic factors can influence, in a different measure and manner, the association. This thesis yields results that can be summarised in three main areas: the first, on the definitions and characterisation of green spaces, the second on the sensitivity of associations between different landuse data and health data, and the third on the meaning and representativeness of medication data as a measure of health.

7.1.1 On the definitions

While numerous systematic literature reviews have been carried out to report findings on the relationships between green and health disorders, few have critically discussed the type or the measures of green taken into account for their analyses. The work of this thesis has started by investigating the existing definition of "green space" in order to fully understand what is being measured and used in the current literature. In order to have a coherent overview, definitions have been searched in those studies that had used "objective" measures of green spaces for their investigation. What comes out of our review is that green spaces are catching the interest of researchers from a diversity of disciplines, although it appears clearly that the largest interest is from epidemiological research. Despite the high curiosity that the issue has raised, the cross-disciplinarity and collaboration between the authors (of the selected papers) is limited. A greater collaboration may contribute to the construction of a common ground of research on the topic. Another result provided by the systematic review, is a strong interest towards urban environments and very little for more rural areas, where instead the share of green spaces is bigger. In the researches carried out for this thesis measures of both urban and non-urban (peri-urban and rural) green areas have been included.

Chapter 3 in fact considers an area comprising urban, peri-urban and rural features in order to explore the divergence in representation of landuses (artificial and natural) across four different landuse data sources. Namur is a medium-sized Belgian municipality that includes a variety of landuses and in particular in terms of diversity of green spaces (private, public green, forested areas and agricultural and pasture areas). The selected data sources have been selected based on the ones most often used in the literature, and available for the study area: Corine Land Cover, Urban Atlas, NDVI, and a self-developed dataset. The latter has been developed in order to have information on the environment at a highly detailed scale (e.g. including private gardens, or road-side greenery). We show that the representation of the green spaces, as well as of the other landuses, experience strong variations across the four landuse data sources chosen, raising concerns on the biases in geographical analyses. The results obtained make us question even further the soundness of the conclusions drawn in former epidemiological studies: data sources strongly influence the way green is measures and therefore further associations.

7.1.2 On the sensitivity of the associations

Responding to this last concern, also raised by Markevych *et al.* (2017), we compare outcomes of associations between three typologies of green spaces (urban green, forests, agricultural areas) and some specific health indicators, using the four above-mentioned landuse data sources. In this sense, two different ecological studies are carried out. The first uses data on medication of reimbursement for drugs to treat depressive and anxiety disorders, aggregated at the statistical section level. Results of the associations vary across the four data sources. In particular when looking at the results using Corine Land Cover it is possible to see that green urban spaces and medication reimbursement are negatively associated. This goes against results obtained by (Helbich *et al.*, 2018b), finding lower prescription rates for antidepressants in greener areas, as well as with the association for the other data sources, reported in the study. The second analysis uses survey data on depressive and anxiety disorders, and shows a higher likelihood of suffering from depressive disorders in proximity of agricultural areas, when using the Corine Land Cover data, at 500 meters and 1000 meters from the place of residence. These results are not consistent with findings from Gascon *et al.* (2018) and Reklaitiene *et al.* (2014) who show a protective effect of greenness for depression. No associations is found using Urban Atlas, NDVI and the self-developed map. Our study does not find any significant association between green environment and anxiety. The first major finding of this study lies in the different results obtained using Corine Land Cover when measuring green-health associations. This warns us on the need of a careful choice of landuse data in this type of studies, at this scale of analysis. While associations with green areas are weak and not consistent in both ecological studies, socio-economic factors are instead significant and consistent in sign. This is coherent with the existing literature (Lorant *et al.*, 2003), reporting inequalities in the risk of depression onset. Significant associations between socio-economic variables and health have also been found in the studies presented in chapters 5 and 6.

7.1.3 On the representation of medication data

Significance of socio-economic factors has been observed when looking at spatial patterns of reimbursements for five groups of medications. The ratio of medication reimbursements (age-standardised) for respiratory system, cardiovascular system, lipids, immune system and the central nervous system is generally positively correlated to low socio-economic position (manual work, unemployment, foreigners born in low income countries), and negatively correlated with income. The same relation between drug reimbursement and income holds when looking at reimbursements for asthma medication for pre-schoolers. Both analyses also report a positive correlation between reimbursement and green spaces. In fact, medication reimbursements appear to be more prevalent in the greener municipalities in the south of Belgium, characterised by lower population densities and lower air pollution. The north-east, and more urban, municipalities appear more importantly when looking at expenditures for asthma medication per case: cost/case is higher is the Province of West-Vlaanderen and less pronounced in the southern provinces. When looking at spatial patterns in medication reimbursements the regional structure appears quite clearly. The strong spatial structure, almost exactly following the regional borders, suggests underlying non-environmental determinants of medication reimbursements. Reimbursements seem to be rather linked to political or commercial factors, to a pharmaceutical practices or the education of the practitioners.

7.2 Limitations and critical assessment

When doing research, we are often confronted with some difficult (strategic) decisions. The good ideas have always some flaws, or lack the data needed. For this thesis, it has always been the objective to carry out the analyses at the scale of the whole country. Because of the peculiar political and administrative situation of Belgium, the task has proven to be more difficult than expected. Due to the political setting, the three Belgian Regions (Flanders, Wallonia and Brussels Capital Region) can make autonomous decisions concerning some matters. The collection of environmental data is one of those. In fact, each Region gathers and produces landuse data independently from the others, making the process of combining all the information together demanding and ambitious. Nonetheless, this complication has led to a reflection on the comparability and equivalence of different landuse data sources. Ultimately, data provided by the European Environmental Agency has been used for conducting the analysis at the level of the country. Considering the increasing relevance and attention that environment assessment has gained in the past years, we should aspire to more consistent data at a fine scale, for larger areas than just a city or a country.

In order to comply with the recent General Data Protection Regulation (GDPR) proclaimed by the European Union (EU) to safeguard EU citizens' privacy, the health data obtained had to go through a "cleaning" and anonymising process. Data at the individual level can be obtained following a privacy commission request but the process can be lengthy. Therefore, for time constraints of the GRESP-HEALTH project (BR/143/A3) and this thesis, only aggregated and anonymised data was made available. Medication reimbursement data has been aggregated, and has been made available at different administrative levels. Depending on the number of cases, that is the number of people having benefited of at least one reimbursement within the year, the level of aggregation varied. A low number of cases implied either numerous missing information at the small administrative scales, or an aggregation to big administrative units (such *arrondissements* or provinces, for example). The higher the number of cases instead, enabled for obtaining complete information (no administrative units with missing data) at the finest scale available (see table A.1 for the list of administrative levels in Belgium). For more rare health problems, such as reimbursement for thyroid dysfunctions or Parkinson disorder, the availability of the data was limited to larger administrative units, as municipalities or even *arrondissements*. This constraint has consequences on the type of analyses that can be performed and the interpretation of the results. Moreover, the prescription of a group of medication is not necessarily used for the pathology, but can also be used to treat symptoms linked to a different pathology (co-morbidity), further complicates the reading of the results.

In spite of the interest raised by some causes of morbidity, this thesis has focused on five specific, likewise interesting, groups of medication reimbursements. The poor quality/quantity of data for a few disorders, such as Parkinson, thyroid disfunction, or autism, has raised the question on the possibility of different treatment approaches to the disease. The available data at the individual level has not been aggregated, although information on the location of the person had to be concealed. Characteristics of the surrounding environment of the individuals have therefore been approximated through some indicators, but specific clues on the spatial distribution of the respondents of the survey was not available. This hindered more in-depth spatial analysis and spatial variations of the distribution of the disorders taken into account. Further data collection should account for the possible privacy constraint and aim at gathering meaningful set of information to be able to grasp as much as possible of the number of factors that come into play when studying health.

Furthermore, the use of aggregated data entails the occurrence of some spatial problems. The most well-known problem, the Modifiable Area Unit Problem (MAUP), arises when spatial information are grouped to some pre-defined units. On varying shape or scale of these units, the results can be critically affected. Considering Tobler's first law of geography "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970), we might think that elements spatially close share some properties and therefore loss is minimal in the aggregation. Nonetheless individuals are complex and have unique characteristics that are lost when doing an averaging spatial information. Furthermore the use of aggregated data may lead to ecological fallacy, meaning that individual behaviors are inferred by information retrieved from the group they belong to.

For this thesis we did not have information neither on the exact place of residence, nor on the spatial behaviour of individuals in their daily life. Activities such as work commute, path to chores, dropping/picking up children are part of the daily experience of people, and are important determinants of environmental exposure (Chaix *et al.*, 2012). When using just residential location, we account for only a part of the daily environment, oversimplifying the use of space and possibly ignoring other factors affecting well-being. This limitation on having only a partial representation of individuals' surrounding is valid not only concerning daily mobility, but also for past experiences: information on previous living, working, leisure environments are very often not accounted for. Such places can also play an important role in affecting mental and physical health of individuals, as well as his actions. Assessing the well-being of an individual only based on his present location can lead to falling into a fallacy of assessment. When conducting a study based only on the current or residential location we risk to leave out some relevant information and consider individuals belonging to the same environment as having the same characteristics, and having the same environmental exposure experiences.

7.3 Paths for further research

This thesis has demonstrated that both green spaces and health are difficult to measure, and that there are many shades to consider for both. The existence of numerous endogenous and exogenous variables characterising them and playing a role in the association make the isolation of precise factors challenging. Research should tend beyond a simplistic approach and take into account other confounding elements (socio-economic, social, commercial or political factors) (Lee and Maheswaran, 2011), as well as avoiding the risk of falling into biases in causal inferences due to residential self-selection (Diez Roux and Mair, 2010). Such type of biases can be tackled by considering other elements of life that are not necessarily linked to the residential location, such as working place or leisure activities. Care should be paid not to consider the individuals' "current" situation alone, but also contemplate the role of previous residence or job, for example, as possible factors affecting the general well-being. Numerous are the characteristics of green and non-green spaces that affect daily life, their use and their effect. A more macro perspective should be adopted when discussing the implication of green spaces, and consider such areas in relation to other environments and aspects of life.

Several aspects of the study areas should be further investigated in the future. First, most of the research carried out so far has regarded urban areas (Burgoine *et al.*, 2015; Lovasi *et al.*, 2011; McMorris *et al.*, 2015). While this interest is motivated by an increasing urbanisation in the coming years (UN/DESA/Population Division, 2019), more attention should be given to the unique elements characterising such areas. First, more composite measures of green should be developed and used to be able to grasp more features of the natural environment present in urban areas (Markevych *et al.*, 2017). Second, it appears therefore fundamental to start questioning and exploring the role of man-built landscapes on human well-being. The BELSPO project NAMED (BR/175/A3) set this as one of its objectives (Lauwers *et al.*, 2020), and hopefully the results of the project will set the example and foster more research in that direction. While urban areas are drawing increased curiosity for health research, one should not forget non-urban areas when looking at the effects of proximity to green spaces on health. Research should further investigate rural areas, and further unveil benefits and externalities of such green areas on human well-being. It should not be forgotten that, although in different ways, rural areas are also subject to high levels of air and noise pollution, coming both from anthropogenic as well as natural sources. Examples can be found in the use of pesticides and intensive farming (WHO - Regional Office for Europe, 2016; Mamane *et al.*, 2015), but also from ozone formation (Irceline, 2018) or water quality (Strosnider *et al.*, 2017).

Health benefits are not only due to urban planning and increased green patches, but rather to an interconnected system of physical and non-physical characteristics of the environment, among which we can think of the dispersion of pollutants (urban, industrial or agricultural). A more integrated effort and a larger broadness of view is needed on the issue, in order to capture a clear image of the phenomenon. When dealing with such complex subjects, and especially when individual data is not available, care should be paid to the understanding and description of the group that the individuals' are part of: further exploring the social factors characterising the group, their interactions and their common features (Diez Roux and Mair, 2010).

Several literature reviews have been published in the last years to try and make the point of the state of the art and the findings on the associations between green environment and health. Often they report inconsistent or non-comparable results. Researchers should strive for the development of a framework of research for the study of the relationship(s) between environment and health. Also, cross-country collaborations and studies can foster research on the field: sharing methods and tools to investigate this complex issue enables for obtaining comparable results in different study areas, hopefully shading a light on the mechanisms underlying the possible associations.

Drawing from the results presented in chapters 5 and 6, we recommend that future research considers factors such as education system or healthcare regulation. This is especially relevant for countries where these elements are managed at different levels. Analyses in such cases in particular would benefit of a multi-scale approach taking into account for the nested effects. Such research could give further insights on the factors that drive healthcare choices, and could inform policy makers for tackling health challenges.

7.4 Implications and recommendations

In the perspective of providing more livable and healthy environments, policy makers and urban planners are devoting growing attention on the state of the art of research in terms of salutogenic and negative impact of the environment on health. As shown all along this thesis, the definition and clear understanding of the study object can have significant effects on the results of the investigations. In Belgium the three administrative Regions, the Flemish Region, the Walloon Region and the Brussels Capital Region, have a number of competences. Regions have in fact independent powers on issues relating to their own territory. Among others there are the ones concerning the environment, nature conservation, town and country planning. A dialogue among the parts, and a definition of common methods to describe and measure the environment and/or to collect data can provide a less-scattered image of the country, a coherent and consistent data source to be used to conduct analysis at the country level and better inform policies at the federal level. It is necessary to tend towards a more comprehensive approach to be able to take concrete and effective measures at the federal level.

The results presented in this thesis recall the importance of correctly measuring and interpreting the available data, and of interpreting subtle spatial differences before developing statistical explanatory models and further taking national political decision. The strong structure of the spatial variations of medication reimbursement in Belgium highlights the important effect played by the regional levels in healthcare administration (through the education system, for example). It is important that policy makers and governments consider such elements when analysing health expenditure and public health at the national level.

Urban and country planning should look towards a more integrated approach. Planning healthier cities means putting the people first, considering their needs, preferences and possibilities. It is not only about "blindly" adding patches of green, but rather considering the surrounding environment of the individuals in a more comprehensive way: acting on a mode-choice transition to less polluting (public transport) or more active way of traveling (cycling, walking). It is also about making cities less dense and increasing the quality of contact, and thus improving the social environment, making cities safer. Nonetheless we should keep in mind that, considering the complexity of the task at stake, sometimes the results lie in the process of interdisciplinary dialogue and collaboration itself. Attention should be put on a wider scale, putting efforts in tackling socio-economic disparities and social justice.

Governments should keep investing in research projects studying the links and associations between health status and the environment, encouraging multidisciplinary approaches. Multidisciplinarity can enrich research by providing novel perspectives and methods to define and tackle the issues linked to associations between environment and health. Also, research would benefit from enlarging the scale of analysis, as for example, by investigating the same research question in multiple study areas with different characteristics. The application of the same research methods and using comparable data (census, self-administered questionnaires or survey with same questions and defined thresholds) can lead to comparable and robust results. This can provide relevant information to policy makers. Allowing researchers to use meaningful data at several (nested) scales, can enable to carry analyses of the associations between environment and health, not only as isolated phenomena, but as a part of a complex system.

Last, but not least, technology today enables us to collect a wide amount and type of data. Considering the tools, knowledge and capacities available, efforts should be done to make good use of the data that is gathered. The recommendation here is not to aim at creating even more big-data sets, but rather at perfecting the data collection process in order to have information for which the definition is clear and adapted to the researchers' needs. Especially concerning health data, it is important to understand individual behaviors and the associations to the environment. Attempts should be made to gather meaningful, spatialised, objective individual data.

7.5 Concluding words

We have first gone further in understanding what "green space" is, instead of relying on the definition given by the landuse data source. We have explored the literature and brought to the surface a number of different interpretations of this concept. We highlight a lack of common definition and, with it, a lack of a shared framework for the analysis of green spaces. It is therefore necessary to find a common language and improve communication intra- and inter-disciplines in order to be able to build a strong theoretical background for studies on green spaces.

In the era of big data and advanced technologies, we are submerged by data of all kinds. We have to be increasingly careful about what the available data is really describing, and most of all what we are really measuring. We have brought here the example of data on medication reimbursement in Belgium, where the data rather than informing on the health status of the population tells us about a way of practicing the profession in Belgium. With a nation-wide analysis of patterns of reimbursement, we identify a regional/administrative trend linked to differences in medical or pharmaceutical practices rather than environmental characteristics.

In this thesis we tackle some of the the issues and relationship between green spaces and health from a different perspective than the existing literature, and put an emphasis on the spatial aspects of the two elements. We show that adding a spatial perspective to the investigation is key to unveil latent factors.

V

Appendix

A

Additional tables and figures

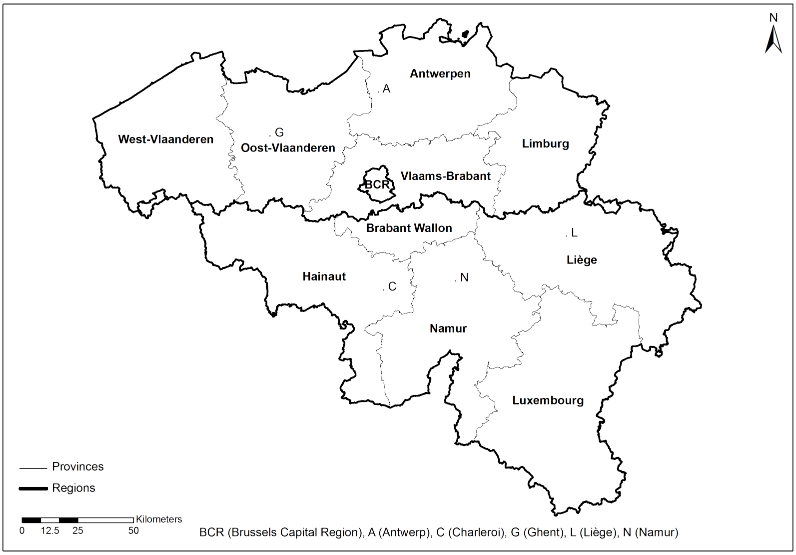


Figure A.1. Administrative partition of Belgium

Table A.1. Belgian statistical levels, in decreasing order of size

Belgian administrative/spatial units	
3	Regions (Flanders, Wallonia, Brussels Capital Region)
10	Provinces
43	Arrondissements
589	Municipalities (until 2018)*
2739	Former municipalities (until 1977)*
6344	Statistical sections*
19782	Statistical sectors

(*) Level used in Chapter 6

A.1 Composition of the clusters of the systematic literature review

Cluster 1a
[39] Olad Ghaffari and Monavari (2013)
[44] Rafiee *et al.* (2009)
[72] Kang and Choi (2014)
[58] Zhang *et al.* (2015)
[76] Kong *et al.* (2014)
[29] Liu and Shen (2014)
[27] Liang *et al.* (2012)
[32] Maimaitiyiming *et al.* (2014)
[75] Kong *et al.* (2007)
[45] Rahman *et al.* (2011)
[73] Kang and Kim (2015)
[10] Byomkesh *et al.* (2012)
[23] Faryadi and Taheri (2009)
[80] La Greca *et al.* (2011)
[16] Coskun Hepcan (2013)
[47] Rupprecht and Byrne (2014)
[66] Hill *et al.* (2009)
[07] Becker *et al.* (2012)
[48] Saw *et al.* (2015)
Cluster 2a
[05] Attwell (2000)
[43] Potestio *et al.* (2009)
[36] Michael *et al.* (2014)
[86] Li *et al.* (2008)
[69] Huang and Yin (2015)
[49] Taylor *et al.* (2011)

[59] Zhao *et al.* (2011)
[42] Pietilä *et al.* (2015)
[54] Weber *et al.* (2014)
[74] Kihal-Talantikite *et al.* (2013)
Cluster 2b
[08] Broberg *et al.* (2013)
[09] Burgoine *et al.* (2015)
[71] Janssen and Rosu (2015)
[30] Lovasi *et al.* (2011)
[70] Huynh *et al.* (2013)
[34] Markevych *et al.* (2014a)
[04] Astell-Burt *et al.* (2014b)
[11] Chaix *et al.* (2014)
[15] Coombes *et al.* (2010)
[67] Hillsdon *et al.* (2006)
[03] Arentze *et al.* (2009)
[31] Macintyre *et al.* (2008)
[28] Lin *et al.* (2014)
[46] Reklaitiene *et al.* (2014)
[38] Nutsford *et al.* (2013)
[82] Lachowycz and Jones (2014)
[21] Dai (2011)
[65] Hanibuchi *et al.* (2011)
[64] Halonen *et al.* (2014)
[12] Charreire *et al.* (2012)
[55] Wendel *et al.* (2011)
Cluster 2c
[13] Chi *et al.* (2015)
[60] Fung and Siu (2000)
[06] Bajwa and Vories (2007)
[14] Cong *et al.* (2012)
[79] Kume *et al.* (2009)
[84] Leprieur *et al.* (2000)
[22] Duguy *et al.* (2012)
[61] Garonna *et al.* (2014)
[52] Villeneuve *et al.* (2012)
[53] Wall *et al.* (2012)
[35] McMorris *et al.* (2015)
[40] Pereira *et al.* (2012)
[50] Triguero-Mas *et al.* (2015)
[26] Li *et al.* (2014)
[85] Leslie *et al.* (2010)

[56] Wilker *et al.* (2014)
[63] Grigsby-Toussaint *et al.* (2011)
[33] Markevych *et al.* (2014a)
[83] Laurent *et al.* (2013)
Cluster 2d
[17] Dadvand *et al.* (2014a)
[01] Amoly *et al.* (2015)
[18] Dadvand *et al.* (2012a)
[57] Wu *et al.* (2014)
[62] Grazuleviciene *et al.* (2015)
[19] Dadvand *et al.* (2012b)
[20] Dadvand *et al.* (2014b)
[25] Fuertes *et al.* (2014)
[2] Annerstedt *et al.* (2012)
[78] Kruize *et al.* (2007)
[37] Nielsen and Hansen (2007)
[51] van den Berg *et al.* (2010)
[24] Fortel *et al.* (2014)
[41] Philpott *et al.* (2014)
[81] La Rosa *et al.* (2014)
[68] Houston (2014)
[77] Koppen *et al.* (2014)

A.2 Variables used for the systematic literature review and collaboration network

Article identification

1. **Bibliographic database**: database where the paper appears (*Scopus, Pubmed* or both)
2. **Author**: name of first author
   Full list of authors: names of first author and all co-authors
3. **Year**: year of publication of the article
4. **Journal**: name of journal where the article is published
5. **Title**: title of the article

Study focus

6. **Discipline**: discipline of the first author - based on the Research Institute of primary affiliation
7. **Population**: population concerned by the study (adults, children, animal species...)
8. **Purpose**: aim of the study (informing policy makers, introducing new tool, delivering health outcomes...)

Study location

9. **Country**: location of case study
10. **Region**: administrative region of case study (if available)
11. **Location**: precise location of case study, as name of the city (if available)
12. **Type of environment**: urbanisation level of the study area (urban, suburban, rural, or combinations)

Green identification

13. **Database**: database(s) used to identify green spaces (NDVI, Google Earth, local landuse maps, SPOT, Corine Land Cover...)
14. **Resolution**: resolution quality of the green database (50 square meters, 1 square kilometer...)
15. **Season of measurement**: season of "capture" of green database (if applicable)

16. **Month of measurement**: reference month of the database (if applicable)
17. **Year of measurement**: reference year of the database
18. **Type of green**: "patch" if borders of green areas are defined a priori, "vegetation" if green areas are assessed from satellite images
19. **Scale**: scale at which the green space is measured and contextualised, i.e. neighborhood, city, census track...
20. **Unit**: units in which the green measure is expressed (percentage, meters...)
21. **Centroid**: starting point of the measurement(s), i.e. school, park, house, etc.

Measures and methods

22. **Number of buffers**: total number of concentric buffers used in the article
23. **Buffer 1st radius**: radius of the smallest buffer
24. **Buffer 2nd radius**: radius of the second buffer, if any
25. **Buffer 3rd radius**: radius of the third buffer, if any
26. **Buffer 4th radius**: radius of the fourth buffer, if any
27. **Proximity**: "1" if author(s) measure(s) proximity to green, otherwise "0"
28. **Criteria of proximity**: method used to measure proximity: average distance, block, buffer, Euclidean distance, Euclidean nearest neighbour, grid, self-rated, service area
29. **Shape**: "1" if author(s) measure(s) green shape, otherwise "0"
30. **Criteria of shape**: specific measure used to assess the shape: area mean weighted shape index, mean shape index, landscape shape index, perimeter/area ratio
31. **Size**: "1" if author(s) measure(s) green surface, otherwise "0"
32. **Fragmentation**: “1” if author(s) measure(s) fragmentation of green patches, otherwise "0"
33. **View**: "1" if author(s) measure(s) view on green, otherwise "0"

34. **Spatial pattern**: "1" if author(s) measure(s) structure and spatial organisation of green areas, otherwise "0"

35. **Access**: "1" if author(s) measure(s) access to green, otherwise "0"

36. **Criteria of access**: specific measure of accessibility of green space: presence of barrier, buffer, street connectivity, slope, inverse distance weighting, activity location with GIS, percentage of green, nearest green space, road distance, road network, service area, time-cost distance, walking distance

37. **Use**: "1" if author(s) measure(s) use of green areas, otherwise "0"

38. **Criteria of use**: method used to assess the use of green spaces: duration, location of activity, number of visits, position, potential use, time spent

39. **Quality**: "1" if author(s) measure(s) quality of green, otherwise "0"

40. **Criteria of quality**: method used to assess the quality of green areas: inspections, remote sensing, residential preference and habitat, size

41. **Fractality**: "1" if author(s) measure(s) fractality of green, otherwise "0"

42. **Criteria of fractality**: specific method used to estimate the fractality of green areas: Bovill's box-counting method, area-weighted mean patch fractal, pafrac

A.3 Landuse classes of the data sources used in chapter 3

Corine Land Cover
1 – Artificial surfaces
1.1. Urban Fabric
1.1.1 Continuous urban fabric
1.1.2 Discontinuous urban fabric
1.2. Industrial, commercial and transport units
1.2.1 Industrial or commercial units
1.2.2 Road and rail networks and associated land
1.2.3 Port areas
1.2.4. Airports
1.3. Mine, dump and construction sites
1.3.1 Mineral extraction sites
1.3.2 Dump sites
1.3.3 Construction sites
1.4. Artificial non-agricultural vegetated areas
1.4.1 Green urban areas
1.4.2 Sport and leisure facilities
2 – Agricultural areas
2.1 Arable land
2.1.1 Non-irrigated arable land
2.1.2 Permanently irrigated land
2.1.3 Rice fields
2.2 Permanent crops
2.2.1 Vineyards
2.2.2 Fruit trees and berry plantations
2.2.3 Olive groves
2.3 Pastures
2.4 Heterogeneous agricultural areas
2.4.1 Annual crops associated with permanent crops
2.4.2 Complex cultivation patterns
2.4.3 Land principally occupied by agriculture, with signs of natural vegetation
2.4.4 Agro-forestry areas
3 – Forests and semi-natural areas
3.1 Forests
3.1.1 Broad-leaved forest
3.1.2 Coniferous forest
3.1.3 Mixed forest
3.2 Shrub and/or herbaceous vegetation association

3.2.1 Natural grassland
3.2.2 Moors and heathland
3.2.3 Sclerophyllous vegetation
3.2.4 Transitional woodland/shrub
3.3 Open spaces with little or no vegetation
3.3.1 Beaches, dunes, sand
3.3.2 Bare rock
3.3.3 Sparsely vegetated areas
3.3.4 Burnt areas
3.3.5 Glaciers and perpetual snow
<u>4 – Wetlands</u>
4.1 Inland wetlands
4.1.1 Inland marshes
4.1.2 Peatbogs
4.2 Coastal wetlands
4.2.1 Salt marshes
4.2.2 Salines
4.2.3 Intertidal flats
<u>5 – Waterbodies</u>
5.1 Inland waters
5.1.1 Water courses
5.1.2 Water bodies
5.2 Marine waters
5.2.1 Coastal lagoons
5.2.2 Estuaries
5.2.3 Sea and Ocean

Urban Atlas
<u>1 – Artificial surfaces</u>
1.1 Urban fabric
1.1.1 Continuous urban fabric
1.1.2 Discontinuous urban fabric
1.1.2.1 Discontinuous dense urban fabric
1.1.2.2 Discontinuous medium density urban fabric
1.1.2.3 Discontinuous low density urban fabric
1.1.2.4 Discontinuous very low density urban fabric
1.1.3 Isolated structures
1.2 Industrial, commercial, public, military, private and transport units
1.2.1 Industrial, commercial, public, military and private units
1.2.2 Road and rail network and associated land
1.2.2.1 Fast transit roads and associated land
1.2.2.2 Other roads and associated land

1.2.2.3 Railways and associated land
1.2.3 Port areas
1.2.4 Airports
1.3 Mine, dump and construction sites
1.3.1 Mineral extraction and dump sites
1.3.3 Construction sites
1.3.4 Land without current use
1.4 Artificial non-agricultural vegetated areas
1.4.1 Green urban areas
1.4.2 Sports and leisure facilities
2 - Agricultural areas, semi-natural areas and wetlands
3 – Forests
5 – Water

Self-developed dataset classes
1 - Agriculture
1.1 Pasture
1.2 Agricultural areas
1.3 Greenhouses
2 - Residential buildings
3 - Non-residential buildings
3.1 Schools
3.2 Churches
3.3 Commercial areas
3.4 Offices
3.5 Hospitals
4 - Cemeteries
5 - Mineral extraction sites
6 - Vegetation
6.1 Roadside green space
6.2 Open green
6.3 Camping
6.4 Abandoned green areas
6.5 Riverside green space
6.6 Trainside green space
7 - Parking and industrial areas
7.1 Parking
7.2 Industrial areas
7.3 Construction sites
8 - Private green
9 - Public green
9.1 Parks

9.2 Playgrounds
9.3 Urban open green spaces – construction areas
10 - Sport facilities
10.1 Basketball, football pitch
10.2 Tennis court
10.3 Sport center
11 - Train and road infrastructures
11.1 Train infrastructure
11.2 Roadways
12 - High vegetation
12.1 High trees
12.2 Forests
13 - Water bodies

Table A.2. Pearson correlation coefficients between medication reimbursement for the Central Nervous System and the selected socio-economic variables at the level of the statistical sections in Namur (chapter 4)

	Pop density	**Inactive pop**	**Primary edu**	**Secondary edu**
CNS	0.16	0.51***	0.50***	-0.10

P-value : 0 <'***' <0.001 <'**' <0.01 <'*' <0.05 <'•' <0.1

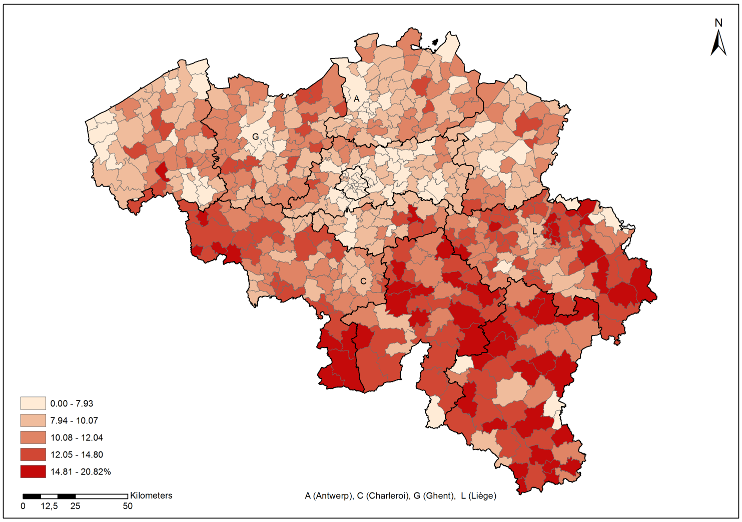


Figure A.2. Prevalence of purchase for asthma medication among 6-12y children in 2014

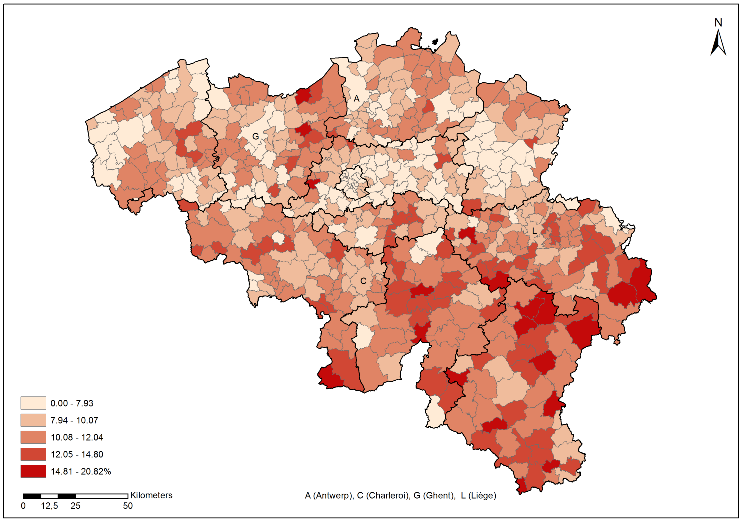


Figure A.3. Prevalence of purchase for asthma medication among 13-18y adolescents in 2014

Table A.3. Descriptive statistics of the environmental and socio-economic variables employed in chapter 5

	Units	Mean	Std dev	Min	Max	Median	P95
$PM_{2.5}$	(μ/m^3)	10.73	1.8	4.45	14.41	11.13	13.22
PM_{10}	(μ/m^3)	16.74	2.96	5.81	22.55	17.23	20.4
NO_2	(μ/m^3)	0.89	0.22	0.42	2.28	0.86	1.23
BC	(μ/m^3)	14.5	5.08	5.25	38.42	13.94	23.71
Manual workers	%	9.81	3.06	1.94	17.62	9.4	14.86
Unemployed	%	9.55	5.4	2.87	33.9	8.22	21.5
Foreign-born	%	2.46	3.28	0.22	36.72	1.62	7.07
Primary edu 25-34y	%	3.9	1.89	0.77	18.88	3.43	7.12
Higher edu 25-34y	%	38.5	9.49	15.89	73.96	36.94	55.96
Income/decl.	€	24967	2721.93	14998	34229	25386	28744.4
Pop Density	Inh/km^2	767.8	2134.82	24.8	24027.9	305.9	1972.97
Ave near GS	m	137.4	115.6	21.8	1132	110.8	285.8
Total GS	% land	72.56	19.7	0	98.54	76.8	95.36

Statistics have been computed based on all Belgian municipalities, with the exception of Herstappe (N=588).

Table A.4. Loadings of the Principal Component Analysis on the five groups of medications, socio-economic and environmental variables used in chapter 5

	Factor 1	Factor 2	Factor 3	Factor 4
$PM_{2.5}$	**0.90**	-0.26	-0.03	-0.15
PM_{10}	**0.89**	-0.22	0.05	-0.11
NO_2	**0.88**	0.18	-0.24	-0.07
BC	**0.91**	0.13	-0.11	0.25
Manual workers	0.06	-0.22	0.03	**-0.93**
Unemployed	-0.08	**0.82**	0.35	0.24
Foreign-born	0.48	**0.68**	-0.27	0.32
Primary edu 25-34y	0.07	**0.90**	-0.03	-0.13
Higher edu 25-34y	0.24	-0.45	-0.28	**0.76**
Income/decl.	-0.07	**-0.89**	-0.21	0.11
Pop Density	**0.59**	**0.58**	-0.27	0.18
Ave near GS	**0.5**	0.47	-0.3	0.13
Total GS	**-0.83**	-0.29	0.11	-0.21
RS	-0.29	0.11	**0.81**	-0.04
CVS	-0.27	0.13	**0.90**	-0.01
LIP	-0.03	-0.07	**0.79**	-0.14
IS	0.20	-0.18	**0.67**	-0.08
CNS	-0.20	0.34	**0.80**	0.01

Loadings of the PCA on the environmental and socio-economic and ASMP variables. In bold are highlighted the loadings > 0.5; loadings with values <0.1 are in light grey as not significant. The proportion of variation explained by the four factors is 81.

Table A.5. Composition of the 3 factors resulting from the PCA on the environmental and SE variables used in chapter 5

	Factor 1	Factor 2	Factor 3
$PM_{2.5}$	**0.87**	-0.32	-0.19
PM_{10}	**0.85**	-0.27	-0.18
NO_2	**0.92**	0.1	-0.03
BC	**0.92**	0.05	0.23
Manual workers	0.05	-0.19	**-0.91**
Unemployed	-0.09	**0.85**	0.19
Foreign-born	0.57	**0.61**	0.4
Primary edu 25-34y	0.15	**0.89**	-0.08
Higher edu 25-34y	0.24	**-0.52**	**0.78**
Income/decl.	0.03	**-0.91**	0.12
Pop Density	**0.67**	**0.5**	0.26
Ave near GS	**0.58**	0.39	0.23
Total GS	**-0.85**	-0.22	-0.21

Loadings of the PCA on the environmental and socio-economic variables only. In bold are highlighted the loadings > 0.5; loadings with values < 0.1 are in light grey as not significant. The proportion of variation explained by the three factors is 82.

Table A.6. Description and statistics of the variables employed in chapter 6

	Units	Mean	Std dev	Min	Max	Median	P95
Prevalence of Purchase	%	36.4	7.01	18.17	53.75	36.21	48.71
Cost/case/year	€	42.72	13.15	16.77	99.44	38.53	70.22
Pop Density	Inh/km^2	767.78	2134.82	24.8	24027.9	305.9	1972.97
Income/decl.	€	24967	2721.93	14998	34229	25386	28744.4
PM_{10}	(μ/m^3)	16.74	2.96	5.81	22.55	17.23	20.4
PM_{10} P95	(μ/m^3)	37.51	5.21	18.7	48	38.8	44.46
Density of pigs	Pigs/km^2	74.29	57.14	0	471.17	64.34	175.76
Total GS	% land	72.56	19.7	0	98.54	76.8	95.36

Statistics have been computed based on all Belgian municipalities, with the exception of Herstappe (N=588).

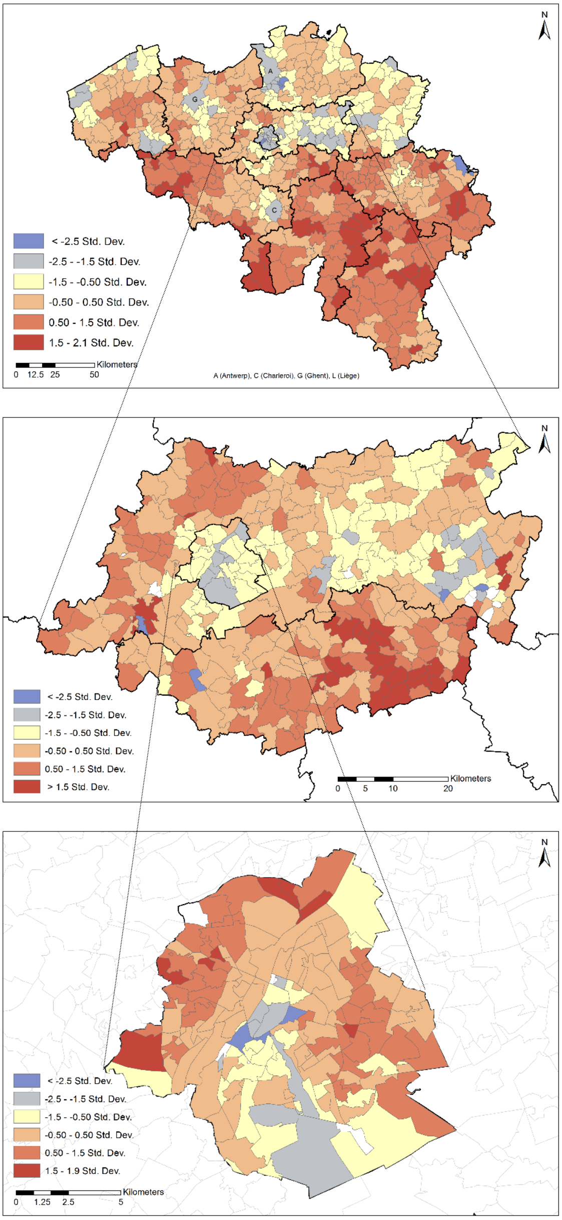


Figure A.4. Prevalence of purchase for asthma medication for the three nested studied areas by means of the discretisation method based on average value and standard deviation (chapter 6)

Bibliography

Adli, M., Berger, M., Brakemeier, E.L., Engel, L., Fingerhut, J., Gomez-Carrillo, A., Hehl, R., Heinz, A., Mayer, J., Mehran, N., Tolaas, S., Walter, H., Weiland, U., and Stollmann, J. (2017). "Neurourbanism: towards a new discipline". *The Lancet Psychiatry*, 4(3): 183–185.

Administration du Cadastre de l'Enregistrement et des Domaines (2005). "La matrice cadastrale". Technical report.

Aerts, R., Honnay, O., and Van Nieuwenhuyse, A. (2018). "Biodiversity and human health: Mechanisms and evidence of the positive health effects of diversity in nature and green spaces". *British Medical Bulletin*, 127(1): 5–22.

Aerts, R., Nemery, B., Bauwelinck, M., Trabelsi, S., Deboosere, P., Van Nieuwenhuyse, A., Nawrot, T.S., and Casas, L. (2020). "Residential green space, air pollution, socioeconomic deprivation and cardiovascular medication sales in Belgium: A nationwide ecological study". *Science of The Total Environment*, 712.

Alcock, I., White, M., Cherrie, M., Wheeler, B., Taylor, J., McInnes, R., Otte im Kampe, E., Vardoulakis, S., Sarran, C., Soyiri, I., and Fleming, L. (2017). "Land cover and air pollution are associated with asthma hospitalisations: A cross-sectional study". *Environment International*, 109: 29–41.

Allen, J. and Balfour, R. (2012). "Natural Solutions for Tackling Health Inequalities: Realising the Opportunity". Technical report, UCL Institute of Health Equity.

Almanza, E., Jerrett, M., Dunton, G., Seto, E., and Ann Pentz, M. (2012). "A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data". *Health and Place*, 18(1): 46–54.

Amoly, E., Dadvand, P., Forns, J., López-Vicente, M., Basagaña, X., Julvez, J., Alvarez-Pedrerol, M., Nieuwenhuijsen, M.J., and Sunyer, J. (2015). "Green and blue spaces and behavioral development in barcelona schoolchildren: The BREATHE project". *Environmental Health Perspectives*, 122(12): 1351–1358.

Annemans, L., Crott, R., De Clerq, H., Hyuybrechts, M., Peys, F., Robays, H., Streens, I., Vansschoubroek, K., and Winderickx, P. (1997). "Pricing and reimbursement of Pharmaceuticals in Belgium". *PharmacoEconomics*, 11(3): 203–209.

Annerstedt, M., Ostergren, P.O., Bjork, J., Grahn, P., Skarback, E., and Wahrborg, P. (2012). "Green qualities in the neighbourhood and mental health - results from a longitudinal cohort study in Southern Sweden." *BMC Public Health*, 12(337).

Arentze, T.A., van der Waerden, P.J.H.J., Bergen, J.W., and Timmermans, H.J.P. (2009). "Measuring the quality of urban environments: A need-based micro-simulation approach". *Applied Spatial Analysis and Policy*, 2(3): 195–209.

Astell-Burt, T., Feng, X., and Kolt, G.S. (2014a). "Green space is associated with walking and moderate-to-vigorous physical activity (MVPA) in middle-to-older-aged adults: Findings from 203 883 Australians in the 45 and Up Study". *British Journal of Sports Medicine*, 48(5): 404–406.

Astell-Burt, T., Feng, X., and Kolt, G.S. (2014b). "Greener neighborhoods, slimmer people evidence from 246 920 Australians". *International Journal of Obesity*, 38(1): 156–159.

Attwell, K. (2000). "Urban land resources and urban planting - Case studies from Denmark". *Landscape and Urban Planning*, 52(2-3): 145–163.

Aujoulat, I., Jacquemin, P., Hermans, M.P., Rietzschel, E., Scheen, A., Tréfois, P., Darras, E., and Wens, J. (2015). "Clinical inertia in general practice, a matter of debate: A qualitative study with 114 general practitioners in Belgium". *BMC Family Practice*, 16(13).

Bajwa, S.G. and Vories, E.D. (2007). "Spatial analysis of cotton (Gossypium hirsutum L.) canopy responses to irrigation in a moderately humid area". *Irrigation Science*, 25(4): 429–441.

Bauer, J., Müller, R., Brüggmann, D., and Groneberg, D.A. (2018). "Spatial Accessibility of Primary Care in England: A Cross-Sectional Study Using a Floating Catchment Area Method". *Health Services Research*, 53(3): 1957–1978.

Baycan-Levent, T. and Nijkamp, P. (2009). "Planning and Management of Urban Green Spaces in Europe: Comparative Analysis". *Journal of Urban Planning and Devekopment*, 135(1).

Becker, D.A., Wood, P.B., and Keyser, P.D. (2012). "Canada Warbler use of harvested stands following timber management in the southern portion of their range". *Forest Ecology and Management*, 276: 1–9.

Belgian Interregional Environment Agency (2014). "Rapport annuel 2014 de la qualité de l'air en Belgique". Technical report, IRCEL.

Berman, M.G., Jonides, J., and Kaplan, S. (2008). "The Cognitive Benefits of Interacting With Nature | The Center for Health Design". *Psychological Science Research*, 19(12): 1207–1212.

Bianchi, M., Clavenna, A., Labate, L., Bortolotti, A., Fortino, I., Merlino, L., Locatelli, G.W., Giuliani, G., and Bonati, M. (2009). "Anti-asthmatic drug prescriptions to an Italian paedriatic population". *Pediatric Allergy and Immunology*, 20(6): 585–591.

Bisgaard, H. and Szefler, S. (2007). "Prevalence of asthma-like symptoms in young children". *Pediatric Pulmonology*, 42(8): 723–728.

Björnberg, A. and Phang, A.Y. (2018). "Euro Health Consumer Index 2018 Report". Technical report. URL https://healthpowerhouse.com/media/EHCI-2018/EHCI-2018-report.pdf.

Bogar, S. and Beyer, K.M. (2015). "Green Space, Violence, and Crime: A Systematic Review". *Trauma, Violence, and Abuse*, 17(2): 160–171.

Boone, C.G., Buckley, G.L., Grove, J.M., and Sister, C. (2009). "Parks and People: An Environmental Justice Inquiry in Baltimore, Maryland". *Annals of the Association of American Geographers*, 99(4): 767–787.

Boone-Heinonen, J., Gordon-Larsen, P., Guilkey, D.K., Jacobs, D.R., and Popkin, B.M. (2011). "Environment and physical activity dynamics: The role of residential self-selection". *Psychology of Sport and Exercise*, 12(1): 54–60.

Borlee, F., Joris Yzermans, C., Aalders, B., Rooijackers, J., Krop, E., Maassen, C.B., Schellevis, F., Brunekreef, B., Heederik, D., and Smit, L.A. (2017). "Air pollution from livestock farms is associated with airway obstruction in neighboring residents". *American Journal of Respiratory and Critical Care Medicine*, 196(9): 1152–1161.

Bossuyt, N., Gadeyne, S., Deboosere, P., and Van Oyen, H. (2004). "Socio-economic inequalities in health expectancy in Belgium". *Public Health*, 118: 3–10.

Bowler, D.E., Buyung-Ali, L., Knight, T.M., and Pullin, A.S. (2010). "Urban greening to cool towns and cities: A systematic review of the empirical evidence". *Landscape and Urban Planning*, 97(3): 147–155.

Bowling, A. (2002). *Research methods in health: investigating health and health services.* Open University Press, Buckingham, 2nd edition.

Brand, P.L.P., Baraldi, E., Bisgaard, H., Boner, A.L., Castro-Rodriguez, J.A., Custovic, A., de Blic, J., de Jongste, J.C., Eber, E., Everard, M.L., Frey, U., Gappa, M., Garcia-Marcos, L., Grigg, J., Lenney, W., Le Souëf, P., McKenzie, S., Merkus, P.J.F.M., Midulla, F., Paton, J.Y., Piacentini, G., Pohunek, P., Rossi, G.A., Seddon, P., Silverman, M., Sly, P.D., Stick, S., Valiulis, A., Van Aalderen, W.M.C., Wildhaber, J.H., Wennergren, G., Wilson, N., Zivkovic, Z., and Bush, A. (2008). "Definition, assessment and treatment of wheezing disorders in preschool children: An evidence-based approach". *European Respiratory Journal*, 32(4): 1096–1110.

Breeze, E., Jones, D.A., Wilkinson, P., Bulpitt, C.J., Grundy, C., Latif, A.M., and Fletcher, A.E. (2005). "Area deprivation, social class, and quality of life among people aged 75 years and over in Britain". *International Journal of Epidemiology*, 34(2): 276–283.

Broberg, A., Kyttä, M., and Fagerholm, N. (2013). "Child-friendly urban structures: Bullerby revisited". *Journal of Environmental Psychology*, 35: 110–120.

Brueckner, J.K., Thisse, J.F., and Zenou, Y. (1999). "Why is central Paris rich and downtown Detroit poor?" *European Economic Review*, 43(1): 91–107.

Burgoine, T., Jones, A.P., Namenek Brouwer, R.J., and Benjamin Neelon, S.E. (2015). "Associations between BMI and home, school and route environmental exposures estimated using GPS and GIS: Do we see evidence of selective daily mobility bias in children?" *International Journal of Health Geographics*, 14(8).

Buyadi, S.N.A., Mohd, W.M.N.W., and Misni, A. (2015). "Vegetation's Role on Modifying Microclimate of Urban Resident". *Procedia - Social and Behavioral Sciences*, 202: 400–407.

Byomkesh, T., Nakagoshi, N., and Dewan, A.M. (2012). "Urbanization and green space dynamics in Greater Dhaka, Bangladesh". *Landscape and Ecological Engineering*, 8(1): 45–58.

Cariñanos, P. and Casares-Porcel, M. (2011). "Urban green zones and related pollen allergy: A review. Some guidelines for designing spaces with low allergy impact". *Landscape and Urban Planning*, 101(3): 205–214.

Carter, M. and Horwitz, P. (2014). "Beyond proximity: The importance of green space useability to self-reported health". *EcoHealth*, 11(3): 322–332.

Casas, L., Simons, K., Nawrot, T.S., Brasseur, O., Declerck, P., Buyl, R., Coomans, D., Nemery, B., and Van Nieuwenhuyse, A. (2016). "Respiratory medication sales and urban air pollution in Brussels (2005 to 2011)". *Environment International*, 94: 576–582.

Chaix, B., Kestens, Y., Perchoux, C., Karusisi, N., Merlo, J., and Labadi, K. (2012). "An interactive mapping tool to assess individual mobility patterns in neighborhood studies". *American Journal of Preventive Medicine*, 43(4): 440–450.

Chaix, B., Simon, C., Charreire, H., Thomas, F., Kestens, Y., Karusisi, N., Vallée, J., Oppert, J.M., Weber, C., and Pannier, B. (2014). "The environmental correlates of overall and neighborhood based recreational walking (a cross-sectional analysis of the RECORD Study)". *International Journal of Behavioral Nutrition and Physical Activity*, 11(1): 20.

Charreire, H., Weber, C., Chaix, B., Salze, P., Casey, R., Banos, A., Badariotti, D., Kesse-Guyot, E., Hercberg, S., Simon, C., and Oppert, J.M. (2012). "Identifying built environmental patterns using cluster analysis and GIS: relationships with walking, cycling and body mass index in French adults." *International Journal of Behavioral Nutrition and Physical Activity*, 9(59).

Chen, E., Shalowitz, M.U., Story, R.E., Ehrlich, K.B., Manczak, E.M., Ham, P.J., Le, V., and Miller, G.E. (2017). "Parents' childhood socioeconomic circumstances are associated with their children's asthma outcomes". *Journal of Allergy and Clinical Immunology*, 140(3): 828–835.e2.

Cheng, C.L., Chen, Y.C., Liu, T.M., and Kao Yang, Y.H. (2011). "Using spatial analysis to demonstrate the heterogeneity of the cardiovascular drug-prescribing pattern in Taiwan". *BMC Public Health*, 11(380).

Cheng, W.C., Chang, J.C., Chang, C.P., Su, Y., and Tu, T.M. (2008). "A fixed-threshold approach to generate high-resolution vegetation maps for IKONOS imagery". *Sensors*, 8(7): 4308–4317.

Chi, W.W.b., Shi, W.W., and Kuang, W.W. (2015). "Spatio-temporal characteristics of intra-urban land cover in the cities of China and USA from 1978 to 2010". *Journal of Geographical Sciences*, 25(1): 3–18.

Clark, N.E., Lovell, R., Wheeler, B.W., Higgins, S.L., Depledge, M.H., and Norris, K. (2014). "Biodiversity, cultural pathways, and human health: A framework". *Trends in Ecology and Evolution*, 29(4): 198–204.

Cohen, B.E., Edmondson, D., and Kronish, I.M. (2015). "State of the art review: Depression, stress, anxiety, and cardiovascular disease". *American Journal of Hypertension*, 28(11).

Cohen, J., Van Wesemael, Y., Smets, T., Bilsen, J., and Deliens, L. (2012). "Cultural differences affecting euthanasia practice in Belgium: One law but different attitudes and practices in Flanders and Wallonia". *Social Science and Medicine*, 75(5): 845–853.

Cong, N., Piao, S., Chen, A., Wang, X., Lin, X., Chen, S., Han, S., Zhou, G., and Zhang, X. (2012). "Spring vegetation green-up date in China inferred from SPOT NDVI data: A multiple model analysis". *Agricultural and Forest Meteorology*, 165: 104–113.

Conway, D., Li, C.Q., Wolch, J., Kahle, C., and Jerrett, M. (2010). "A spatial autocorrelation approach for examining the effects of urban greenspace on residential property values". *Journal of Real Estate Finance and Economics*, 41(2): 150–169.

Coombes, E., Jones, A.P., and Hillsdon, M. (2010). "The relationship of physical activity and overweight to objectively measured green space accessibility and use". *Social Science and Medicine*, 70(6): 816–822.

Coskun Hepcan, C. (2013). "Quantifying landscape pattern and connectivity in a Mediterranean coastal settlement: the case of the Urla district, Turkey." *Environmental monitoring and assessment*, 185(1): 143–155.

Costa, R., Bauwelinck, M., and Deboosere, P. (2019). "IMA / AIM data report : applications and limitations of Belgian prescription data prescription data". URL http://interfacedemography.be/wp-content/uploads/2019/07/Costa{_}Bauwelinck{_}Deboosere{_}2019{_}ID-WP{_}final.pdf.

Coutts, C., Chapin, T., Horner, M., and Taylor, C. (2013). "County-level effects of green space access on physical activity". *Journal of Physical Activity and Health*, 10: 232–40.

Coutts, C. and Hahn, M. (2015). “Green infrastructure, ecosystem services, and human health”. *International Journal of Environmental Research and Public Health*, 12(8): 9768–9798.

Crouse, D.L., Pinault, L., Balram, A., Hystad, P., Peters, P.A., Chen, H., van Donkelaar, A., Martin, R.V., Ménard, R., Robichaud, A., and Villeneuve, P.J. (2017). “Urban greenness and mortality in Canada’s largest cities: a national cohort study”. *The Lancet Planetary Health*, 1(7): e289–e297.

Dadvand, P., de Nazelle, A., Figueras, F., Basagaña, X., Su, J., Amoly, E., Jerrett, M., Vrijheid, M., Sunyer, J., and Nieuwenhuijsen, M.J. (2012a). “Green space, health inequality and pregnancy”. *Environment International*, 40(1): 110–115.

Dadvand, P., Nieuwenhuijsen, M.J., Esnaola, M., Forns, J., Basagaña, X., Alvarez-Pedrerol, M., Rivas, I., López-Vicente, M., De Castro Pascual, M., Su, J., Jerrett, M., Querol, X., and Sunyer, J. (2015). “Green spaces and cognitive development in primary schoolchildren”. *Proceedings of the National Academy of Sciences*, 112(26): 7937–7942.

Dadvand, P., Sunyer, J., Basagaña, X., Ballester, F., Lertxundi, A., Fernández-Somoano, A., Estarlich, M., García-Esteban, R., Mendez, M.A., and Nieuwenhuijsen, M.J. (2012b). “Surrounding greenness and pregnancy outcomes in four Spanish birth cohorts”. *Environmental Health Perspectives*, 120(10): 1481–1487.

Dadvand, P., Villanueva, C.M., Font-Ribera, L., Martinez, D., Basagaña, X., Belmonte, J., Vrijheid, M., Gražulevičienė, R., Kogevinas, M., and Nieuwenhuijsen, M.J. (2014a). “Risks and benefits of green spaces for children: A cross-sectional study of associations with sedentary behavior, obesity, asthma, and allergy”. *Environmental Health Perspectives*, 122(12): 1329–1335.

Dadvand, P., Wright, J., Martinez, D., Basagaña, X., McEachan, R.R., Cirach, M., Gidlow, C.J., de Hoogh, K., Grazuleviciene, R., and Nieuwenhuijsen, M.J. (2014b). “Inequality, green spaces, and pregnant women: Roles of ethnicity and individual and neighbourhood socioeconomic status”. *Environment International*, 71: 101–108.

Dai, D. (2011). “Racial/ethnic and socioeconomic disparities in urban green space accessibility: Where to intervene?” *Landscape and Urban Planning*, 102(4): 234–244.

de Rooij, M.M., Heederik, D.J., Borlée, F., Hoek, G., and Wouters, I.M. (2017). "Spatial and temporal variation in endotoxin and PM10 concentrations in ambient air in a livestock dense area". *Environmental Research*, 153: 161–170.

de Vries, S., ten Have, M., van Dorsselaer, S., van Wezep, M., Hermans, T., and de Graaf, R. (2016). "Local availability of green and blue space and prevalence of common mental disorders in the Netherlands". *BJPsych Open*, 2(6): 366–372.

de Vries, S., van Dillen, S.M., Groenewegen, P.P., and Spreeuwenberg, P. (2013). "Streetscape greenery and health: Stress, social cohesion and physical activity as mediators". *Social Science and Medicine*, 94: 26–33.

de Vries, S., Verheij, R.A., Groenewegen, P.P., and Spreeuwenberg, P. (2003). "Natural environments - Healthy environments? An exploratory analysis of the relationship between greenspace and health". *Environment and Planning A*, 35(10): 1717–1731.

Delmas, M.C., Marguet, C., Raherison, C., Nicolau, J., and Fuhrman, C. (2013). "Les hospitalisations pour asthme chez l'enfant en France, 2002-2010". *Archives de Pediatrie*, 20(7): 739–747.

Devika, S., Jeyaseelan, L., and Sebastian, G. (2016). "Analysis of sparse data in logistic regression in medical research: A newer approach". *Journal of Postgraduate Medicine*, 62(1): 26–31.

Dewulf, B., Neutens, T., De Weerdt, Y., and Van De Weghe, N. (2013). "Accessibility to primary health care in Belgium: An evaluation of policies awarding financial assistance in shortage areas". *BMC Family Practice*, 14.

Di Nardo, F., Saulle, R., and La Torre, G. (2010). "Green areas and health outcomes: a systematic review of the scientific literature." *Ital J Public Health*, 7(4): 402–413.

Diez Roux, A.V. and Mair, C. (2010). "Neighborhoods and health". *Annals of the New York Academy of Sciences*, 1186(1): 125–145.

Dominici, F., Peng, R.D., Bell, M.L., Pham, L., McDermott, A., Zeger, S.L., and Samet, J.M. (2006). "Fine Particulate Air Pollution and Hospital Admission for Cardiovascular and Respiratory Diseases". *Journal of the American Medical Association*, 295(10).

Droomers, M., Jongeneel-Grimen, B., Kramer, D., De Vries, S., Kremers, S., Bruggink, J.W., Van Oers, H., Kunst, A.E., and Stronks, K. (2015). "The impact of intervening in green space in Dutch deprived neighbourhoods on physical activity and general health: Results from the quasi-experimental URBAN40 study". *Journal of Epidemiology and Community Health*, 70(2): 147–154.

Ducharme, F., Tse, S., and Chauhan, B. (2014). "Diagnosis, management, and prognosis of preschool wheeze". *The Lancet*, 383(9928): 1593–1604.

Duguy, B., Alloza, J.A., Baeza, M.J., De La Riva, J., Echeverría, M., Ibarra, P., Llovet, J., Cabello, F.P., Rovira, P., and Vallejo, R.V. (2012). "Modelling the ecological vulnerability to forest fires in mediterranean ecosystems using geographic information technologies". *Environmental Management*, 50(6): 1012–1026.

Dujardin, C., Selod, H., and Thomas, I. (2008). "Residential segregation and unemployment: The case of Brussels". *Urban Studies*, 45(1): 89–113.

Ege, M.J., Frei, R., Bieli, C., Schram-Bijkerk, D., Waser, M., Benz, M.R., Weiss, G., Nyberg, F., van Hage, M., Pershagen, G., Brunekreef, B., Riedler, J., Lauener, R., Braun-Fahrländer, C., and von Mutius, E. (2007). "Not all farming environments protect against the development of asthma and wheeze in children". *Journal of Allergy and Clinical Immunology*, 119(5): 1140–1147.

Ekedahl, A., Andersson, S.I., Hovelius, B., Mölstad, S., Liedholm, H., and Melander, A. (1995). "Drug prescription attitudes and behaviour of general practitioners - Effects of a problem-oriented educational programme". *European Journal of Clinical Pharmacology*, 47(5): 381–387.

Ellaway, A., Macintyre, S., and Bonnefoy, X. (2005). "Graffiti, greenery, and obesity in adults: secondary analysis of European cross sectional survey". *BMJ*, 331: 606–611.

European Commission (2020). "Investing in green infrastructure". URL https://ec.europa.eu/environment/nature/ecosystems/investing/index{_}en.htm.

European Environmental Agency (2007). "CORINE Land Cover". Technical report. URL https://www.eea.europa.eu/publications/COR0-landcover.

European Parliament and Council (2016). "Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)". Technical report. URL http://data.europa.eu/eli/reg/2016/679/oj.

European Union (2011). "Mapping Guide for a European Urban". URL http://www.eea.europa.eu/data-and-maps/data/urban-atlas{#}tab-methodology.

Eurostat (2019). "Health in the European Union - facts and figures". URL https://ec.europa.eu/eurostat/statistics-explained.

Faryadi, S. and Taheri, S. (2009). "Interconnections of urban green spaces and environmental quality of Tehran". *International Journal of Environmental Research*, 3(2): 199–208.

Federal Public Service (2018). "Social security: Everything you always wanted to know (in Belgium)". Technical Report June, FPS Social Security DG Political Support and Coordination (BeSoc), Brussel.

Feitosa, R.C. and Wilkinson, S. (2018). "Green roofs and green walls and their impact on health promotion". *Cadernos de Saude Publica*, 34(7).

Ferreira de Magalhães, M., Amaral, R., Pereira, A.M., Sá-Sousa, A., Azevedo, I., Azevedo, L.F., and Fonseca, J.A. (2017). "Cost of asthma in children: A nationwide, population-based, cost-of-illness study". *Pediatric Allergy and Immunology*, 28(7): 683–691.

Ferreira-Magalhães, M., Sá-Sousa, A., Morais-Almeida, M., Pité, H., Azevedo, L.F., Azevedo, M.I., Bugalho-Almeida, A., and Fonseca, J.A. (2016). "Asthma-like symptoms, diagnostic tests, and asthma medication use in children and adolescents: A population-based nationwide survey". *Journal of Asthma*, 53(3): 269–276.

Figueiras, A., Caamaño, F., and Gestal-Otero, J.J. (2000). "Influence of physician's education, drug information and medical-care settings on the quality of drugs prescribed". *European Journal of Clinical Pharmacology*, 56(9-10): 747–753.

Fortel, L., Henry, M., Guilbaud, L., Guirao, A.L., Kuhlmann, M., Mouret, H., Rollin, O., and Vaissière, B.E. (2014). "Decreasing abundance, increasing diversity and changing structure of the wild bee community (hymenoptera: anthophila) along an urbanization gradient". *PLoS ONE*, 9(8): e104 679.

Foster, C., Hillsdon, M., Jones, A., Grundy, C., Wilkinson, P., White, M., Sheehan, B., Wareham, N., and Thorogood, M. (2009). "Objective Measures of the Environment and Physical Activity—Results of the Environment and Physical Activity Study in English Adults". *Journal of Physical Activity and Health*, 6(s1): S70–S80.

Fotheringham, A.S. and Wong, D.W. (1991). "The modifiable areal unit problem in multivariate statistical analysis". *Environment and Planning A*, 23(7): 1025–1044.

Frumkin, H. (2001). "Beyond toxicity: Human health and the natural environment". *American Journal of Preventive Medicine*, 20(3): 234–240.

Fuertes, E., Markevych, I., von Berg, A., Bauer, C.P., Berdel, D., Koletzko, S., Sugiri, D., and Heinrich, J. (2014). "Greenness and allergies: Evidence of differential associations in two areas in Germany". *Journal of Epidemiology and Community Health*, 68(8): 787–790.

Fujita, M. and Thisse, J.F. (2011). *Economics of agglomeration: Cities, industrial location, and globalization*. Cambridge University Press, Cambridge.

Fung, T. and Siu, W. (2000). "Environmental quality and its changes, an analysis using NDVI". *International Journal of Remote Sensing*, 21(5): 1011–1024.

Gamble, K.R. and Howard, D.V. (2016). "Not just scenery: viewing nature pictures improves executive attention in older adults". *Experimental Aging Research*, 40(5): 1–16.

Garonna, I., de Jong, R., de Wit, A.J., Mücher, C.A., Schmid, B., and Schaepman, M.E. (2014). "Strong contribution of autumn phenology to changes in satellite-derived growing season length estimates across Europe (1982-2011)". *Global Change Biology*, 20(11): 3457–3470.

Gascon, M., Sánchez-Benavides, G., Dadvand, P., Martínez, D., Gramunt, N., Gotsens, X., Cirach, M., Vert, C., Molinuevo, J.L., Crous-Bou, M., and Nieuwenhuijsen, M. (2018). "Long-term exposure to residential green and blue spaces and anxiety and depression in adults: A cross-sectional study". *Environmental Research*, 162: 231–239.

Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Forns, J., Plasència, A., and Nieuwenhuijsen, M.J. (2015). "Mental health benefits of long-term exposure to residential green and blue spaces: A systematic review". *International Journal of Environmental Research and Public Health*, 12(4): 4354–4379.

Gatersleben, B. and Andrews, M. (2013). "When walking in nature is not restorative-The role of prospect and refuge". *Health and Place*, 20: 91–101.

Gidlow, C.J., Smith, G., Martinez, D., Wilson, R., Trinder, P., Gražulevičiene, R., and Nieuwenhuijsen, M.J. (2016). "Research note: Natural environments and prescribing in England". *Landscape and Urban Planning*, 151: 103–108.

Global Initiative For Asthma (GINA) (2017). "Global Strategy For Asthma Management and Prevention". *Global Initiative for Asthma*. URL www.ginasthma.org.

Gobster, P.H. (2002). "Managing urban parks for a racially and ethnically diverse clientele". *Leisure Sciences*, 24(2): 143–159.

Grahn, P. and Stigsdotter, U.A. (2003). "Landscape planning and stress". *Urban Forestry and Urban Greening*, 2(1): 1–18.

Grazuleviciene, R., Danileviciute, A., Dedele, A., Vencloviene, J., Andrusaityte, S., Uždanaviciute, I., and Nieuwenhuijsen, M.J. (2015). "Surrounding greenness, proximity to city parks and pregnancy outcomes in Kaunas cohort study." *International journal of hygiene and environmental health*, 218(3): 358–365.

Green, M.J. and Benzeval, M. (2013). "The development of socioeconomic inequalities in anxiety and depression symptoms over the lifecourse". *Social Psychiatry and Psychiatric Epidemiology*, 48(12): 1951–1961.

Grigsby-Toussaint, D.S., Chi, S.H., and Fiese, B.H. (2011). "Where they live, how they play: Neighborhood greenness and outdoor physical activity among preschoolers". *International Journal of Health Geographics*, 10(66).

Gruebner, O., Rapp, M., Adli, M., Kluge, U., Galea, S., and Heinz, A. (2017). "Cities and Mental Health". *Deutsches Aerzteblatt International*, 114: 121–127.

Gupta, K., Kumar, P., Pathan, S.K., and Sharma, K.P. (2012). "Urban Neighborhood Green Index - A measure of green spaces in urban areas". *Landscape and Urban Planning*, 105(3): 325–335.

Halasa, N.B., Griffin, M.R., Zhu, Y., and Edwards, K.M. (2002). "Decreased number of antibiotic prescriptions in office-based settings from 1993 to 1999 in children less than five years of age". *Pediatric Infectious Disease Journal*, 21(11): 1023–1028.

Halonen, J.I., Kivimaki, M., Pentti, J., Stenholm, S., Kawachi, I., Subramanian, S.V., and Vahtera, J. (2014). "Green and blue areas as predictors of overweight and obesity in an 8-year follow-up study." *Obesity (Silver Spring, Md.)*, 22(8): 1910–1917.

Hanibuchi, T., Kawachi, I., Nakaya, T., Hirai, H., and Kondo, K. (2011). “Neighborhood built environment and physical activity of Japanese older adults: Results from the Aichi Gerontological Evaluation Study (AGES)”. *BMC Public Health*, 11(657).

Haq, S.M.A. (2011). “Urban Green Spaces and an Integrative Approach to Sustainable Environment”. *Journal of Environmental Protection*, 02(05): 601–608.

Hartig, T., Mitchell, R., de Vries, S., and Frumkin, H. (2014). “Nature and Health”. *Annual Review of Public Health*, 35: 207–228.

Helbich, M., de Beurs, D., Kwan, M.P., O’Connor, R.C., and Groenewegen, P.P. (2018a). “Natural environments and suicide mortality in the Netherlands: a cross-sectional, ecological study”. *The Lancet Planetary Health*, 2(3): e134–e139.

Helbich, M., Klein, N., Roberts, H., Hagedoorn, P., and Groenewegen, P.P. (2018b). “More green space is related to less antidepressant prescription rates in the Netherlands: A Bayesian geoadditive quantile regression approach”. *Environmental Research*, 166: 290–297.

Henau, K., Van Eycken, E., Silversmit, G., and Pukkala, E. (2015). “Regional variation in incidence for smoking and alcohol related cancers in Belgium”. *Cancer Epidemiology*, 39(1): 55–65.

Hill, M.J., FitzSimons, J., and Pearson, C.J. (2009). “Creating Land use Scenarios for City Greenbelts Using A Spatial Multi-Criteria Analysis Shell: Two Case Studies”. *Physical Geography*, 30(4): 353–382.

Hillsdon, M., Panter, J., Foster, C., and Jones, A. (2006). “The relationship between access and quality of urban green space with population physical activity”. *Public Health*, 120(12): 1127–1132.

Hoffimann, E., Barros, H., and Ribeiro, A.I. (2017). “Socioeconomic inequalities in green space quality and Accessibility—Evidence from a Southern European city”. *International Journal of Environmental Research and Public Health*, 14(916).

Honold, J., Lakes, T., Beyer, R., and van der Meer, E. (2014). “Restoration in Urban Spaces: Nature Views From Home, Greenways, and Public Parks”. *Environment and Behavior*, 48(6): 796–825.

Houston, D. (2014). “Implications of the modifiable areal unit problem for assessing built environment correlates of moderate and vigorous physical activity”. *Applied Geography*, 50: 40–47.

Huang, H. and Yin, L. (2015). "Creating sustainable urban built environments: An application of hedonic house price models in Wuhan, China". *Journal of Housing and the Built Environment*, 30(2): 219–235.

Huynh, Q., Craig, W., Janssen, I., and Pickett, W. (2013). "Exposure to public natural space as a protective factor for emotional well-being among young people in Canada." *BMC Public Health*, 13(407).

Ibes, D.C. (2015). "A multi-dimensional classification and equity analysis of an urban park system: A novel methodology and case study application". *Landscape and Urban Planning*, 137: 122–137.

Irceline (2018). "Cellule Interrégionale de l'Environnement". URL http://www.irceline.be.

Jamagne, P. (2012). "Secteurs statistiques. Vade-mecum". Technical report, StatBel.

Janssen, I. and Rosu, A. (2015). "Undeveloped green space and free-time physical activity in 11 to 13-year-old children". *International Journal of Behavioral Nutrition and Physical Activity*, 12(26).

Janssen, S., Gerwin, D., Fierens, F., and Mensink, C. (2008). "Spatial interpolation of air pollution measurements using CORINE land cover data". *Atmospheric Environment*, 42(20): 4884–4903.

Jennings, V. and Bamkole, O. (2019). "The relationship between social cohesion and urban green space: An avenue for health promotion". *International Journal of Environmental Research and Public Health*, 16(452).

Jepson, G., Butler, T., Gregory, D., and Jones, K. (2000). "Prescribing patterns for asthma by general practitioners in six European countries". *Respiratory Medicine*, 94(6): 578–583.

Jim, C.Y. and Chen, W.Y. (2010). "External effects of neighbourhood parks and landscape elements on high-rise residential value". *Land Use Policy*, 27(2): 662–670.

Kang, S. and Choi, W. (2014). "Forest cover changes in North Korea since the 1980s". *Regional Environmental Change*, 14(1): 347–354.

Kang, S. and Kim, J.O. (2015). "Morphological analysis of green infrastructure in the Seoul metropolitan area, South Korea". *Landscape and Ecological Engineering*, 11(2): 259–268.

Kaplan, S. (1995). "The restorative benefits of nature: Toward an integrative framework". *Journal of Environmental Psychology*, 15: 169–182.

Kemperman, A. and Timmermans, H. (2014). "Green spaces in the direct living environment and social contacts of the aging population". *Landscape and Urban Planning*, 129: 44–54.

Kihal-Talantikite, W., Padilla, C.M., Lalloué, B., Gelormini, M., Zmirou-Navier, D., and Deguen, S. (2013). "Green space, social inequalities and neonatal mortality in France". *BMC Pregnancy and Childbirth*, 13(1): 191.

Köhler, M., Schmidt, M., Wilhelm Grimme, F., Laar, M., Lúcia de Assunção Paiva, V., and Tavares, S. (2002). "Green roofs in temperate climates and in the hot-humid tropics – far beyond the aesthetics". *Environmental Management and Health*, 13(4): 382–391.

Kong, F., Yin, H., and Nakagoshi, N. (2007). "Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: A case study in Jinan City, China". *Landscape and Urban Planning*, 79: 240–252.

Kong, F., Yin, H., Wang, C., Cavan, G., and James, P. (2014). "A satellite image-based analysis of factors contributing to the green-space cool island intensity on a city scale". *Urban Forestry and Urban Greening*, 13(4): 846–853.

Koopman, L.P., Neijens, H.J., Wijga, A., Smit, H.A., De Jongste, J.C., Kerkhof, M., Gerritsen, J., Vos, A.P., Van Strien, R.T., and Brunekreef, B. (2002). "Early respiratory and skin symptoms in relation to ethnic background: The importance of socioeconomic status; The PIAMA study". *Archives of Disease in Childhood*, 87(6): 482–488.

Koppen, G., Sang, Å.O., and Tveit, M.S. (2014). "Managing the potential for outdoor recreation: Adequate mapping and measuring of accessibility to urban recreational landscapes". *Urban Forestry and Urban Greening*, 13(1): 71–83.

Kruize, H., Driessen, P.P., Glasbergen, P., and Van Egmond, K. (2007). "Environmental equity and the role of public policy: Experiences in the Rijnmond region". *Environmental Management*, 40(4): 578–595.

Kume, T., Umetsu, C., and Palanisami, K. (2009). "Impact of the December 2004 tsunami on soil, groundwater and vegetation in the Nagapattinam district, India". *Journal of Environmental Management*, 90(10): 3147–3154.

Kutzora, S., Weber, A., Heinze, S., Hendrowarsito, L., Nennstiel-Ratzel, U., von Mutius, E., Fuchs, N., and Herr, C. (2018).

"Asthmatic/wheezing phenotypes in preschool children: Influential factors, health care and urban-rural differences". *International Journal of Hygiene and Environmental Health*, 221(2): 293–299.

La Greca, P., La Rosa, D., Martinico, F., and Privitera, R. (2011). "Agricultural and green infrastructures: The role of non-urbanised areas for eco-sustainable planning in a metropolitan region". *Environmental Pollution*, 159: 2193–2202.

La Rosa, D., Barbarossa, L., Privitera, R., and Martinico, F. (2014). "Agriculture and the city: A method for sustainable planning of new forms of agriculture in urban contexts". *Land Use Policy*, 41: 290–303.

Lachowycz, K. and Jones, A.P. (2014). "Does walking explain associations between access to greenspace andlower mortality?" *Social Science and Medicine*, 107: 9–17.

Lambert, K.A., Bowatte, G., Tham, R., Lodge, C., Prendergast, L., Heinrich, J., Abramson, M.J., Dharmage, S.C., and Erbas, B. (2017). "Residential greenness and allergic respiratory diseases in children and adolescents – A systematic review and meta-analysis". *Environmental Research*, 159: 212–221.

Landrigan, P.J., Fuller, R., Acosta, N.J., Adeyi, O., Arnold, R., Basu, N.N., Baldé, A.B., Bertollini, R., Bose-O'Reilly, S., Boufford, J.I., Breysse, P.N., Chiles, T., Mahidol, C., Coll-Seck, A.M., Cropper, M.L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., Hanrahan, D., Hunter, D., Khare, M., Krupnick, A., Lanphear, B., Lohani, B., Martin, K., Mathiasen, K.V., McTeer, M.A., Murray, C.J., Ndahimananjara, J.D., Perera, F., Potočnik, J., Preker, A.S., Ramesh, J., Rockström, J., Salinas, C., Samson, L.D., Sandilya, K., Sly, P.D., Smith, K.R., Steiner, A., Stewart, R.B., Suk, W.A., van Schayck, O.C., Yadama, G.N., Yumkella, K., and Zhong, M. (2017). "The Lancet Commission on pollution and health". *The Lancet Commissions*, 391(10119): 462–512.

Laurent, O., Wu, J., Li, L., and Milesi, C. (2013). "Green spaces and pregnancy outcomes in Southern California". *Health and Place*, 24: 190–195.

Lauwers, L., Trabelsi, S., Pelgrims, I., Bastiaens, H., De Clercq, E., Guilbert, A., Guyot, M., Leone, M., Mortelmans, D., Nawrot, T.S., Van Nieuwenhuyse, A., Remmen, R., Saenen, N., Thomas, I., and Keune, H. (2020). "Urban environment and mental health: the NAMED project, a mixed method approach". *BMJ Open*, 10(e031963).

Le Texier, M., Schiel, K., and Caruso, G. (2018). “The provision of urban green space and its accessibility: Spatial data effects in Brussels”. *PLoS ONE*, 13(10): 1–17.

Lee, A.C. and Maheswaran, R. (2011). “The health benefits of urban green spaces: A review of the evidence”. *Journal of Public Health*, 33(2): 212–222.

Lee, G. and Hong, I. (2013). “Measuring spatial accessibility in the context of spatial disparity between demand and supply of urban park service”. *Landscape and Urban Planning*, 119: 85–90.

Lepczyk, C.A., Aronson, M.F., Evans, K.L., Goddard, M.A., Lerman, S.B., and Macivor, J.S. (2017). “Biodiversity in the City: Fundamental Questions for Understanding the Ecology of Urban Green Spaces for Biodiversity Conservation”. *BioScience*, 67(9): 799–807.

Leprieur, C., Kerr, Y.H., Mastorchio, S., and Meunier, J.C. (2000). “Monitoring Vegetation Cover across Semi-arid Regions: A Comparative Multi-Scale approach of Remote Observations”. *International Journal of Remote Sensing*, 21(2): 281–300.

Leslie, E., Sugiyama, T., Ierodiaconou, D., and Kremer, P. (2010). “Perceived and objectively measured greenness of neighbourhoods: Are they measuring the same thing?” *Landscape and Urban Planning*, 95(1-2): 28–33.

Levy, M.L., Godfrey, S., Irving, C.S., Sheikh, A., Hanekom, W., Bush, A., and Lachman, P. (2004). “Wheeze detection: recordings vs. assessment of physician and parent.” *The Journal of asthma : official journal of the Association for the Care of Asthma*, 41(8): 845–53.

Li, F., Harmer, P.A., Cardinal, B.J., Bosworth, M., Acock, A., Johnson-Shelton, D., and Moore, J.M. (2008). “Built Environment, Adiposity, and Physical Activity in Adults Aged 50-75”. *American Journal of Preventive Medicine*, 35(1): 38–46.

Li, W.C. and Yeung, K.K. (2014). “A comprehensive study of green roof performance from environmental perspective”. *International Journal of Sustainable Built Environment*, 3(1): 127–134.

Li, X., Meng, Q., Li, W., Zhang, C., Jancso, T., and Mavromatis, S. (2014). “An explorative study on the proximity of buildings to green spaces in urban areas using remotely sensed imagery”. *Annals of GIS*, 20(3): 193–203.

Liang, L., Schwartz, M.D., and Fei, S. (2012). “Photographic assessment of temperate forest understory phenology in relation to springtime meteorological drivers”. *International Journal of Biometeorology*, 56(2): 343–355.

Lin, B.B., Fuller, R.A., Bush, R., Gaston, K.J., and Shanahan, D.F. (2014). “Opportunity or orientation? Who uses urban parks and why”. *PLoS ONE*, 9(1): e87 422.

Liu, H.L. and Shen, Y.S. (2014). “The impact of green space changes on air pollution and microclimates: A case study of the taipei metropolitan area”. *Sustainability*, 6: 8827–8855.

Lõhmus, M. and Balbus, J. (2015). “Making green infrastructure healthier infrastructure”. *Infection Ecology and Epidemiology*, 5(1): 30 082.

Lorant, V., Deliège, D., Eaton, W., Robert, A., Philippot, P., and Ansseau, M. (2003). “Socioeconomic inequalities in depression: A meta-analysis”. *American Journal of Epidemiology*, 157(2): 98–112.

Lovasi, G.S., Jacobson, J.S., Quinn, J.W., Neckerman, K.M., Ashby-Thompson, M.N., and Rundle, A. (2011). “Is the environment near home and school associated with physical activity and adiposity of urban preschool children?” *Journal of urban health : bulletin of the New York Academy of Medicine*, 88(6): 1143–1157.

Maantay, J. (2002). “Mapping environmental injustices: Pitfalls and potential of geographic information systems in assessing environmental health and equity”. *Environmental Health Perspectives*, 110(suppl 2): 161–171.

Maas, J., Verheij, R.A., De Vries, S., Spreeuwenberg, P., Schellevis, F.G., and Groenewegen, P.P. (2009). “Morbidity is related to a green living environment”. *Journal of Epidemiology and Community Health*, 63: 967–973.

Maas, J., Verheij, R.A., Groenewegen, P.P., De Vries, S., and Spreeuwenberg, P. (2006). “Green space, urbanity, and health: How strong is the relation?” *Journal of Epidemiology and Community Health*, 60: 587–592.

Macintyre, S., Macdonald, L., and Ellaway, A. (2008). “Do poorer people have poorer access to local resources and facilities? The distribution of local resources by area deprivation in Glasgow, Scotland”. *Social Science and Medicine*, 67(6): 900–914.

Maimaitiyiming, M., Ghulam, A., Tiyip, T., Pla, F., Latorre-Carmona, P., Halik, Ü., Sawut, M., and Caetano, M. (2014). “Effects of green space spatial pattern on land surface temperature: Implications for sustainable urban planning and climate change adaptation”. *ISPRS Journal of Photogrammetry and Remote Sensing*, 89: 59–66.

Mamane, A., Raherison, C., Tessier, J.F., Baldi, I., and Bouvier, G. (2015). "Environmental exposure to pesticides and respiratory health". *European Respiratory Review*, 24(137): 462–473.

Manea, L., Gilbody, S., and McMillan, D. (2012). "Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): A meta-analysis". *CMAJ*, 184(3): E191–E196.

Marfortt, D.A., Josviack, D., Lozano, A., Cuestas, E., Agüero, L., and Castro-Rodriguez, J.A. (2018). "Differences between preschoolers with asthma and allergies in urban and rural environments". *Journal of Asthma*, 55(5): 470–476.

Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A.M., de Vries, S., Triguero-Mas, M., Brauer, M., Nieuwenhuijsen, M.J., Lupp, G., Richardson, E.A., Astell-Burt, T., Dimitrova, D., Feng, X., Sadeh, M., Standl, M., Heinrich, J., and Fuertes, E. (2017). "Exploring pathways linking greenspace to health: Theoretical and methodological guidance". *Environmental Research*, 158: 301–317.

Markevych, I., Thiering, E., Fuertes, E., Sugiri, D., Berdel, D., Koletzko, S., Von Berg, A., Bauer, C.P., and Heinrich, J. (2014a). "A cross-sectional analysis of the effects of residential greenness on blood pressure in 10-year old children: Results from the GINIplus and LISAplus studies". *BMC Public Health*, 14(1): 477.

Markevych, I., Tiesler, C.M.T., Fuertes, E., Romanos, M., Dadvand, P., Nieuwenhuijsen, M.J., Berdel, D., Koletzko, S., and Heinrich, J. (2014b). "Access to urban green spaces and behavioural problems in children: Results from the GINIplus and LISAplus studies". *Environment International*, 71: 29–35.

Martinez, F.D., Wright, A.L., Taussig, L.M., Holberg, C.J., Halonen, M., and Morgan, W.J. (1995). "Asthma and wheezing in the first six years of life. The Group Health Medical Associates." *The New England Journal of Medicine*, 332(3): 133–138.

Maruani, T. and Amit-Cohen, I. (2007). "Open space planning models: A review of approaches and methods". *Landscape and Urban Planning*, 81: 1–13.

McDonnell, M. and Hahs, A. (2009). "Comparative ecology of cities and towns: Past, present and future". In M. McDonnell, A. Hahs, and J. Breuste, editors, *Ecology of Cities and Towns: A Comparative Approach*, chapter 5, pp. 71–89. Cambridge University Press, Cambridge.

McHugh, M.L. (2012). "Interrater reliability: The kappa statistic". *Biochemia Medica*, 22(3): 276–282.

McKinney, M.L. (2008). "Effects of urbanization on species richness: A review of plants and animals". *Urban Ecosystems*, 11(2): 161–176.

McMorris, O., Villeneuve, P.J., Su, J., and Jerrett, M. (2015). "Urban greenness and physical activity in a national survey of Canadians." *Environmental Research*, 137: 94–100.

Melis, G., Gelormino, E., Marra, G., Ferracin, E., and Costa, G. (2015). "The effects of the urban built environment on mental health: A cohort study in a large northern Italian city". *International Journal of Environmental Research and Public Health*, 12(11): 14 898–14 915.

Mendola, P., Nobles, C., Williams, A., Sherman, S., Kanner, J., Seeni, I., and Grantz, K. (2019). "Air pollution and preterm birth: Do air pollution changes over time influence risk in consecutive pregnancies among low-risk women?" *International Journal of Environmental Research and Public Health*, 16(18).

Mérenne-Schoumaker, B., Vandermotten, ChristianVan Hecke, E., Decroly, J.M., Vanneste, D., and Verhetsel, A. (2015). *Atlas de Belgique 5 - Activités économiques.* Academia Press, Ghent.

Michael, S.E., Hull, R.B., and Zahm, D.L. (2001). "Environmental factors influencing auto burglary a case study". *Environment and Behavior*, 33(3): 368–388.

Michael, Y.L., Nagel, C.L., Gold, R., and Hillier, T.A. (2014). "Does change in the neighborhood environment prevent obesity in older women?" *Social Science and Medicine*, 102: 129–137.

Michel, G., Silverman, M., Strippoli, M.P.F.P., Zwahlen, M., Brooke, A.M., Grigg, J., and Kuehni, C.E. (2006). "Parental understanding of wheeze and its impact on asthma prevalence estimates". *European Respiratory Journal*, 28(6): 1124–1130.

Moher, D., Liberati, A., Tetzlaff, J., and Altman, D.G. (2009). "Preferred Reporting Items for Systematic Reviews and Meta-Analyses". *Annulas of Internal Medicine*, 151(4): 264–269.

Morancho, A.B. (2003). "A hedonic valuation of urban green areas". *Landscape and Urban Planning*, 66(1): 35–41.

Mukherjee, D., Safraj, S., Tayyab, M., Shivashankar, R., Patel, S.A., Narayanan, G., Ajay, V.S., Ali, M.K., Narayan, K.V., Tandon, N., and Prabhakaran, D. (2017). "Park availability and major depression in individuals with chronic conditions: Is there an association in urban India?" *Health and Place*, 47: 54–62.

Müller-Rompa, S.E., Markevych, I., Hose, A.J., Loss, G., Wouters, I.M., Genuneit, J., Braun-Fahrländer, C., Horak, E., Boznanski, A., Heederik, D., von Mutius, E., Heinrich, J., and Ege, M.J. (2018). "An approach to the asthma-protective farm effect by geocoding: Good farms and better farms". *Pediatric Allergy and Immunology*, 29(3): 275–282.

Muntaner, C., Eaton, W.W., Miech, R., and O'Campo, P. (2004). "Socioeconomic position and major mental disorders". *Epidemiologic Reviews*, 26: 53–62.

Natural England (2010). "Nature Nearby, Accessible Natural Greenspace Guidance". Technical Report March. URL www.naturalengland.org.uk/publications.

Naughton, C., Bennett, K., and Feely, J. (2006). "Regional variation in prescribing for chronic conditions among an elderly population using a pharmacy claims database". *Irish Journal of Medical Science*.

Nicol, C. and Blake, R. (2000). "Classification and use of open space in the context of increasing urban capacity". *Planning Practice and Research*, 15(3): 193–210.

Nielsen, T.S. and Hansen, K.B. (2007). "Do green areas affect health? Results from a Danish survey on the use of green areas and health indicators". *Health and Place*, 13(4): 839–850.

Nowak, D.J., Crane, D.E., and Stevens, J.C. (2006). "Air pollution removal by urban trees and shrubs in the United States". *Urban Forestry and Urban Greening*, 4(3-4): 115–123.

Nutsford, D., Pearson, A.L., and Kingham, S. (2013). "An ecological study investigating the association between access to urban green space and mental health". *Public Health*, 127(11): 1005–1011.

Nutsford, D., Pearson, A.L., Kingham, S., and Reitsma, F. (2016). "Residential exposure to visible blue space (but not green space) associated with lower psychological distress in a capital city". *Health and Place*, 39: 70–78.

OECD (2017). "Education Policy Outlook: Belgium". Technical Report September, Organisation for Economic Co-operation and Development. URL http://www.oecd.org/edu/profiles.htm.

OECD/European Observatory on Health System and Policies (2019). "Belgium: Country Health Profile 2019, State of Health in the EU". Technical report, European Observatory on Health Systems and Policies, Brussels.

Olad Ghaffari, P. and Monavari, S.M. (2013). "Physical development trend and green space destruction in developing cities: A GIS approach". *Environment, Development and Sustainability*, 15(1): 167–175.

Open Street Map Community (2020). "Open Street Map". URL www.openstreetmap.org.

Openshaw, S. (2009). *Modifiable Areal Unit Problem*. Geo Books, Norwich, geo books edition.

Orioli, R., Antonucci, C., Scortichini, M., Cerza, F., Marando, F., Ancona, C., Manes, F., Davoli, M., Michelozzi, P., Forastiere, F., and Cesaroni, G. (2019). "Exposure to residential greenness as a predictor of cause-specific mortality and stroke incidence in the Rome longitudinal study". *Environmental Health Perspectives*, 127(2): 1–11.

Paoletti, E., Bardelli, T., Giovannini, G., and Pecchioli, L. (2011). "Air quality impact of an urban park over time". *Procedia Environmental Sciences*, 4: 10–16.

Paulot, F. and Jacob, D.J. (2014). "Hidden cost of U.S. agricultural exports: Particulate matter from ammonia emissions". *Environmental Science and Technology*, 48(2): 903–908.

Perchoux, C. (2016). "Accounting for residential and non-residential environments to measure contextual effects on health behavior : the case of recreational walking behavior".

Pereira, G., Foster, S., Martin, K., Christian, H., Boruff, B.J., Knuiman, M., and Giles-Corti, B. (2012). "The association between neighborhood greenness and cardiovascular disease: an observational study." *BMC Public Health*, 12(466).

Peters, A., Liu, E., Verrier, R.L., Schwartz, J., Gold, D.R., Mittleman, M., Baliff, J., Oh, J.A., Allen, G., Monahan, K., and Dockery, D.W. (2000). "Air pollution and incidence of cardiac arrhythmia". *Epidemiology*, 11(1): 11–17.

Philpott, S.M., Cotton, J., Bichier, P., Friedrich, R.L., Moorhead, L.C., Uno, S., and Valdez, M. (2014). “Local and landscape drivers of arthropod abundance, richness, and trophic composition in urban habitats”. *Urban Ecosystems*, 17(2): 513–532.

Piantadosi, S., Byar, D.P., and Green, S.B. (1988). “The ecological fallacy”. *American Journal of Epidemiology*, 127(5): 893–904.

Pietilä, M., Neuvonen, M., Borodulin, K., Korpela, K., Sievänen, T., and Tyrväinen, L. (2015). “Relationships between exposure to urban green spaces, physical activity and self-rated health”. *Journal of Outdoor Recreation and Tourism*, 10: 44–54.

Porta, M. (2014). *A dictionary of epidemiology*. Oxford University Press, New York, 6th edition.

Potestio, M.L., Patel, A.B., Powell, C.D., McNeil, D.A., Jacobson, R.D., and McLaren, L. (2009). “Is there an association between spatial access to parks/green space and childhood overweight/obesity in Calgary, Canada?” *International Journal of Behavioral Nutrition and Physical Activity*, 6.

Querriau, X., Thomas, I., and Vanneste, D. (2011). *Atlas de Belgique 4 - Habitat*. Academia Press, Ghent.

Raaschou-Nielsen, O., Andersen, Z.J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G., Hoffmann, B., Fischer, P., Nieuwenhuijsen, M.J., Brunekreef, B., Xun, W.W., Katsouyanni, K., Dimakopoulou, K., Sommar, J., Forsberg, B., Modig, L., Oudin, A., Oftedal, B., Schwarze, P.E., Nafstad, P., De Faire, U., Pedersen, N.L., Östenson, C.G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K.T., Sørensen, M., Tjønneland, A., Ellermann, T., Eeftens, M., Peeters, P.H., Meliefste, K., Wang, M., Bueno-de Mesquita, B., Key, T.J., de Hoogh, K., Concin, H., Nagel, G., Vilier, A., Grioni, S., Krogh, V., Tsai, M.Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Trichopoulou, A., Bamia, C., Vineis, P., and Hoek, G. (2013). “Air pollution and lung cancer incidence in 17 European cohorts: Prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE)”. *The Lancet Oncology*, 14(9): 813–822.

Rafiee, R., Salman Mahiny, A., and Khorasani, N. (2009). “Assessment of changes in urban green spaces of Mashad city using satellite data”. *International Journal of Applied Earth Observation and Geoinformation*, 11(6): 431–438.

RAHMAN, A., AGGARWAL, S.P., NETZBAND, M., and FAZAL, S. (2011). "Monitoring Urban Sprawl Using Remote Sensing and GIS Techniques of a Fast Growing Urban Centre, India". *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(1): 56–64.

REKLAITIENE, R., VIRVICIUTE, D., TAMOSIUNAS, A., BACEVICIENE, M., LUKSIENE, D., SAPRANAVICIUTE ZABAZLAJEVA, L., RADISAUSKAS, R., BERNOTIENE, G., GRAZULEVICIENE, R., DEDELE, A., VENSLOVIENE, J., BOBAK, M., and NIEUWENHUIJSEN, M.J. (2014). "The relationship of green space, depressive symptoms and perceived general health in urban population". *Scandinavian Journal of Public Health*, 42(7): 669–676.

ROCHA, V., RIBEIRO, A.I., SEVERO, M., BARROS, H., and FRAGA, S. (2017). "Neighbourhood socioeconomic deprivation and health-related quality of life: A multilevel analysis". *PLoS ONE*, 12(12): 1–16.

ROJAS-RUEDA, D., NIEUWENHUIJSEN, M.J., GASCON, M., PEREZ-LEON, D., and MUDU, P. (2019). "Green spaces and mortality: a systematic review and meta-analysis of cohort studies". *The Lancet Planetary Health*, 3(11): e469–e477.

RÖSSLER, S. (2008). "Green space development in shrinking cities – opportunities and constraints". *Urbanistični inštitut Republike Slovenije*, 19(2): 147–152.

RUPPRECHT, C.D. and BYRNE, J.A. (2014). "Informal urban green-space: Comparison of quantity and characteristics in Brisbane, Australia and Sapporo, Japan". *PLoS ONE*, 9(6): e99 784.

RUPPRECHT, C.D.D. and BYRNE, J.A. (2018). "Informal Urban Green Space as Anti-Gentrification Strategy?" In T.H. WINIFRED CURRAN, editor, *Just Green Enough: Urban development and environmental gentrification*, pp. 209–226. Routledge.

SACKER, A. and CABLE, N. (2005). "Do adolescent leisure-time physical activities foster health and well-being in adulthood? Evidence from two British birth cohorts". *European Journal of Public Health*, 16(3): 331–335.

SAMET, J. and KREWSKI, D. (2007). "Health Effects Associated With Exposure to Ambient Air Pollution". *Journal of Toxicology and Environmental Health, Part A*, 70(3-4): 227–242.

SARKAR, C., WEBSTER, C., and GALLACHER, J. (2018). "Residential greenness and prevalence of major depressive disorders: a cross-sectional, observational, associational study of 94 879 adult UK Biobank participants". *The Lancet Planetary Health*, 2(4): e162–e173.

Satellite Imaging Corporation (2020). "Satellite Sensors". URL https://www.satimagingcorp.com/.

Saw, L.E., Lim, F.K.S., and Carrasco, L.R. (2015). "The relationship between natural park usage and happiness does not hold in a tropical city-state". *PLoS ONE*, 10(7): e0133 781.

Schipperijn, J., Stigsdotter, U.K., Randrup, T.B., and Troelsen, J. (2010). "Influences on the use of urban green space - A case study in Odense, Denmark". *Urban Forestry and Urban Greening*, 9(1): 25–32.

Sciensano (2019). "Health Interview Survey". URL https://his.wiv-isp.be/SitePages/Home.aspx.

Slocum, T.A. (1999). *Thematic cartography and visualization.* Prentice Hall, Upper Saddle River, NJ.

Sofianopoulou, E., Rushton, S.P., Diggle, P.J., and Pless-Mulloli, T. (2014). "Association between respiratory prescribing, air pollution and deprivation, in primary health care". *Journal of Public Health*, 36(1): 502–509.

Song, H., Lane, K.J., Kim, H., Kim, H., Byun, G., Le, M., Choi, Y., Park, C.R., and Lee, J.T. (2019). "Association between urban greenness and depressive symptoms: Evaluation of greenness using various indicators". *International Journal of Environmental Research and Public Health*, 16(2).

Sonnenschein-van der Voort, A.M.M., Jaddoe, V.W.V., Raat, H., Moll, H.A., Hofman, A., de Jongste, J.C., and Duijts, L. (2012). "Fetal and infant growth and asthma symptoms in preschool children: the Generation R Study." *American Journal of Respiratory and Critical Care Medicine*, 185(7): 731–737.

StatBel (2018). "Structure of the population". URL https://statbel.fgov.be/en/themes/population/structure-population.

Stevens, C.A., Turner, D., Kuehni, C.E., Couriel, J.M., and Silverman, M. (2003). "The economic impact of preschool asthma and wheeze". *European Respiratory Journal*, 21(6): 1000–1006.

Strachan, D.P. (1989). "Hay fever, hygiene, and household size." *BMJ*, 299(6710): 1259–60.

Strosnider, H., Kennedy, C., Monti, M., and Yip, F. (2017). "Rural and urban differences in air quality, 2008-2012, and community drinking water quality, 2010-2015 - United States". *MMWR Surveillance Summaries*, 66(13).

Takano, T., Nakamura, K., and Watanabe, M. (2002). "Urban residential environments and senior citizens' longevity in megacity areas: The importance of walkable green spaces". *Journal of Epidemiology and Community Health*, 56(12): 913–918.

Talen, E. (2010). "The spatial logic of parks". *Journal of Urban Design*, 15(4): 473–491.

Tamura, K., Wilson, J.S., Goldfeld, K., Puett, R.C., Klenosky, D.B., Harper, W.A., and Troped, P.J. (2019). "Accelerometer and GPS Data to Analyze Built Environments and Physical Activity". *Research Quarterly for Exercise and Sport*, 90(3): 395–402.

Taylor, B.T., Fernando, P., Bauman, A.E., Williamson, A., Craig, J.C., and Redman, S. (2011). "Measuring the quality of public open space using google earth". *American Journal of Preventive Medicine*, 40(2): 105–112.

Taylor, L. and Hochuli, D.F. (2017). "Defining greenspace: Multiple uses across multiple disciplines". *Landscape and Urban Planning*, 158: 25–38.

Taylor, M.S., Wheeler, B.W., White, M.P., Economou, T., and Osborne, N.J. (2015). "Research note: Urban street tree density and antidepressant prescription rates-A cross-sectional study in London, UK". *Landscape and Urban Planning*, 136: 174–179.

Thomas, I., Adam, A., and Verhetsel, A. (2017). "Migration and commuting interactions fields: A new geography with community detection algorithm?" *Belgeo*, 4: 1–17.

Thurston, G.D., Kipen, H., Annesi-Maesano, I., Balmes, J., Brook, R.D., Cromar, K., De Matteis, S., Forastiere, F., Forsberg, B., Frampton, M.W., Grigg, J., Heederik, D., Kelly, F.J., Kuenzli, N., Laumbach, R., Peters, A., Rajagopalan, S.T., Rich, D., Ritz, B., Samet, J.M., Sandstrom, T., Sigsgaard, T., Sunyer, J., and Brunekreef, B. (2017). "A joint ERS/ATS policy statement: What constitutes an adverse health effect of air pollution? An analytical framework". *European Respiratory Journal*, 49(1).

Tischer, C., Gascon, M., Fernández-Somoano, A., Tardón, A., Materola, A.L., Ibarluzea, J., Ferrero, A., Estarlich, M., Cirach, M., Vrijheid, M., Fuertes, E., Dalmau-Bueno, A., Nieuwenhuijsen, M.J., Antó, J.M., Sunyer, J., and Dadvand, P. (2017). "Urban green and grey space in relation to respiratory health in children". *European Respiratory Journal*, 49(6).

Tobler, W. (1970). “A computer movie simulating urban growth in the Detroit Region”. *Economic Geography*, 46: 234–240.

Trabelsi, S., Casas, L., Nemery, B., Nawrot, T.S., and Thomas, I. (2019). “Geographies of asthma medication purchase for pre-schoolers in Belgium”. *Respiratory Research*, 20(1).

Triguero-Mas, M., Dadvand, P., Cirach, M., Martínez, D., Medina, A., Mompart, A., Basagaña, X., Grazuleviciene, R., and Nieuwenhuijsen, M.J. (2015). “Natural outdoor environments and mental and physical health: relationships and mechanisms.” *Environment International*, 77: 35–41.

Troy, A., Morgan Grove, J., and O’Neil-Dunne, J. (2012). “The relationship between tree canopy and crime rates across an urban-rural gradient in the greater Baltimore region”. *Landscape and Urban Planning*, 106(3): 262–270.

Ulrich, R. (1984). “View through a window may influence recovery from surgery”. *Science*, 224(4647): 420–421.

UN/DESA/Population Division (2019). “World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)”. Technical report, United Nations, New York.

van den Berg, A.E., Jorgensen, A., and Wilson, E.R. (2014). “Evaluating restoration in urban green spaces: Does setting type make a difference?” *Landscape and Urban Planning*, 127: 173–181.

van den Berg, A.E., Maas, J., Verheij, R.A., and Groenewegen, P.P. (2010). “Green space as a buffer between stressful life events and health”. *Social Science and Medicine*, 70(8): 1203–1210.

Van den Berg, M., Wendel-Vos, W., van Poppel, M., Kemper, H., van Mechelen, W., and Maas, J. (2015). “Health benefits of green spaces in the living environment: A systematic review of epidemiological studies”. *Urban Forestry and Urban Greening*, 14(4): 806–816.

Van Herzele, A. and de Vries, S. (2012). “Linking green space to health: A comparative study of two urban neighbourhoods in Ghent, Belgium”. *Population and Environment*, 34(2): 171–193.

Van Herzele, A. and Wiedemann, T. (2003). “A monitoring tool for the provision of accessible and attractive urban green spaces”. *Landscape and Urban Planning*, 63(2): 109–126.

van Lenthe, F.J., Brug, J., and MacKenbach, J.P. (2005). "Neighbourhood inequalities in physical inactivity: The role of neighbourhood attractiveness, proximity to local facilities and safety in the Netherlands". *Social Science and Medicine*, 60(4): 763–775.

Venn, S.J. and Niemelä, J.K. (2004). "Ecology in a multidisciplinary study of urban green space: The URGE project". *Boreal Environment Research*, 9(6): 479–489.

Villeneuve, P.J., Jerrett, M., G. Su, J., Burnett, R.T., Chen, H., Wheeler, A.J., and Goldberg, M.S. (2012). "A cohort study relating urban green space with mortality in Ontario, Canada". *Environmental Research*, 115: 51–58.

Visser, H. and De Nijs, T. (2006). "The map comparison kit". *Environmental Modelling and Software*, 21(3): 346–358.

Waldhoer, T., Wald, M., and Heinzl, H. (2008). "Analysis of the spatial distribution of infant mortality by cause of death in Austria in 1984 to 2006". *International Journal of Health Geographics*, 7: 1–9.

Wall, M.M., Larson, N.I., Forsyth, A., Van Riper, D.C., Graham, D.J., Story, M.T., and Neumark-Sztainer, D. (2012). "Patterns of obesogenic neighborhood features and adolescent weight: A comparison of statistical approaches". *American Journal of Preventive Medicine*, 42(5): e65—-e75.

Wangia, V. and Shireman, T.I. (2013). "A review of geographic variation and Geographic Information Systems (GIS) applications in prescription drug use research." *Research in Social and Administrative Pharmacy*, 9(6): 666–87.

Weber, N., Haase, D., and Franck, U. (2014). "Zooming into temperature conditions in the city of Leipzig: How do urban built and green structures influence earth surface temperatures in the city?" *Science of the Total Environment*, 496: 289–298.

Weier, J. and Herring, D. (2000). "Measuring Vegetation (NDVI & EVI) Normalized Difference Vegetation Index (NDVI)". URL https://earthobservatory.nasa.gov/features/MeasuringVegetation.

Wendel, H.E., Downs, J.A., and Mihelcic, J.R. (2011). "Assessing equitable access to urban green space: The role of engineered water infrastructure". *Environmental Science and Technology*, 45(16): 6728–6734.

WHO - Regional Office for Europe (2016). "Urban green spaces and health". Technical report, World Health Organisation.

WHO Collaborating Centre for Drug Statistics Methodology (2018). “ATC/DDD Index 2019”. URL https://www.whocc.no/atc{_}ddd{_}index/.

Wilker, E.H., Wu, C.D., McNeely, E., Mostofsky, E., Spengler, J., Wellenius, G.a., and Mittleman, M.a. (2014). “Green space and mortality following ischemic stroke.” *Environmental Research*, 133: 42–48.

Wolch, J.R., Byrne, J., and Newell, J.P. (2014). “Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'”. *Landscape and Urban Planning*, 125: 234–244.

World Health Organization (1948). “Constitution of WHO: principles”. URL https://www.who.int/about/who-we-are/constitution.

World Health Organization (2017a). “Depression and Other Common Mental Disorders. Global Health Estimates”. Technical report, WHO, Geneva.

World Health Organization (2017b). “Global health estimates 2015: deaths by cause, age, sex, by country and by region, 2000-2015”. URL http://www.who.int/healthinfo/global{_}burden{_}disease/estimates/en/index1.html.

World Health Organization (2020). “Determinants of Health”. URL https://www.who.int/hia/evidence/doh/en/.

World Health Organization and Norwegian Institute of Public Health (2013). “WHO Collaborating Centre for Drug Statistics Methodology”.

World Health Organization United Nations Environment Programme (1992). “Urban air pollution in megacities of the world”. Technical report, WHO/UNEP, Oxford.

Wu, C.D., McNeely, E., Cedeño-Laurent, J.G., Pan, W.C., Adamkiewicz, G., Dominici, F., Lung, S.C.C., Su, H.J., and Spengler, J.D. (2014). “Linking student performance in Massachusetts elementary schools with the "greenness" of school surroundings using remote sensing”. *PLoS ONE*, 9(10): e108 548.

Xu, K., Soucat, A., Kutzin, J., Brindley, C., Maele, N.V., Touré, H., Garcia, M.A., Li, D., Barroy, H., Flores, G., Roubal, T., Indikadahena, C., Cherilova, V., and Siroka, A. (2018). “Public Spending on Health: A Closer Look at Global Trends”. Technical report, World Health Organization, Geneva.

Xu, X., Duan, X., Sun, H., and Sun, Q. (2011). “Green space changes and planning in the capital region of China”. *Environmental Management*, 47(3): 456–467.

Yang, B.Y., Qian, Z.M., Li, S., Chen, G., Bloom, M.S., Elliott, M., Syberg, K.W., Heinrich, J., Markevych, I., Wang, S.Q., Chen, D., Ma, H., Chen, D.H., Liu, Y., Komppula, M., Leskinen, A., Liu, K.K., Zeng, X.W., Hu, L.W., Guo, Y., and Dong, G.H. (2018). “Ambient air pollution in relation to diabetes and glucose-homoeostasis markers in China: a cross-sectional study with findings from the 33 Communities Chinese Health Study”. *The Lancet Planetary Health*, 2(2): e64–e73.

Zhang, C., Chen, Y., and Lu, D. (2015). “Mapping the land-cover distribution in arid and semiarid urban landscapes with Landsat Thematic Mapper imagery”. *International Journal of Remote Sensing*, 36(17).

Zhao, C., Fu, G., Liu, X., and Fu, F. (2011). “Urban planning indicators, morphology and climate indicators: A case study for a north-south transect of Beijing, China”. *Building and Environment*, 46(5): 1174–1183.

List of Figures

List of Tables