Development and analysis of land-use/land-cover spatio-temporal metrics in urban environments: Exploring urban growth patterns and linkages to socio-economic factors

Marta Sapena Moll

Advisor: Prof. Dr. Luis Ángel Ruiz Fernández

PhD in Geomatics Engineering

UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

October 2020

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Development and analysis of land-use/land-cover spatio-temporal metrics in urban environments: Exploring urban growth patterns and linkages to socio-economic factors

Marta Sapena Moll

Advisor: Prof. Dr. Luis Ángel Ruiz Fernández

Geo-Environmental Cartography and Remote Sensing Group

Department of Cartographic engineering, Geodesy and Photogrammetry

Universitat Politècnica de València, Spain

PhD in Geomatics Engineering

Valencia, October 2020

DEPARTMENT OF CARTOGRAPHIC ENGINEERING, GEODESY AND
PHOTOGRAMMETRY, UNIVERSITAT POLITÈCNICA DE VALÈNCIA

UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

**Development and analysis of land-use/land-cover spatio-temporal
metrics in urban environments: Exploring urban growth patterns and
linkages to socio-economic factors**

PhD thesis presented by Marta Sapena Moll and directed by Dr. Luis A. Ruiz,
Professor at the Department of Cartographic Engineering, Geodesy and
Photogrammetry (UPV). The thesis was developed within the Geo-
Environmental Cartography and Remote Sensing research group (CGAT) to
obtain the title of Doctor in Geomatics Engineering. It has been conducted
within the interuniversity program between Universidad Politécnica de Madrid
and Universitat Politècnica de València.

The thesis has been completed in five years, completing all requirements from
the doctoral program. In total five publications in scientific journals, six
international and national conference participations and publications, and
attendance to several seminars, courses and transverse activities with more
than 2000 hours recognized as specific and transverse activities, exceeding the
requirement of 600 hours.

During her research, the candidate was twice a guest scientist at the German
Aerospace Center (DLR) within the Geo-Risks and Civil Security Department at
the German Remote Sensing Data Center. Two research stays were completed
that recognize the international mention of this thesis. The first stay lasted
three months, where the research collaboration between the two research
centers was stablished. During the last year and a half of this thesis the
candidate was an employee of DLR and completed her research with the close
collaboration of the two institutes.

:CGAT
Grupo de Cartografía
GeoAmbiental y Teledetección

DLR **Deutsches Zentrum
für Luft- und Raumfahrt**
German Aerospace Center

Abstract

Urbanization is a multi-dimensional process that involves economic, social and environmental changes, alters urban form and size, population density and distribution, and land-use/land-cover (LULC), ultimately affecting global sustainability. Urbanization has been rapid in recent years and this is expected to continue in the coming decades at unprecedented rates. Therefore, promoting sustainable urbanization, limiting natural land consumption and ensuring the well-being of population have become policy targets in urban areas at various levels, from local to global. The spatial structure of urban areas and their growth patterns determine how the physical, socio-economic and environmental characteristics of urban areas change over time. These interrelationships play a major role in the daily life of urban dwellers and leads decision-makers to seek better-informed choices for the sustainable planning of urban areas. Thus, a better understanding of the relationships between the spatial structure of urban areas and their socio-economic performance is of crucial relevance.

Monitoring, quantifying and characterizing the deveoplment of urban areas enriches our understanding of past and present trends, provides evidence-based information and supports decision-making processes, which allows anticipating to unsustainable patterns and their potential consequences. This is possible due to Earth observation (EO) and Geographic Information Systems (GIS) techniques. EO provides accurate and frequent data and the tools to monitor urban growth spatially and temporally at multiple levels, which can be further analyzed by means of GIS. Spatial metrics have been increasingly used to monitor and quantify urban growth. However, the ever-growing availability of EO-derived data and the emergence of LULC and geospatial databases require the development of new methods and tools that leverage the growing availability of geo-data, which is becoming a valuable source of data for urban studies, enabling the comprehensive analysis of urban areas spatially, temporally and at multiple levels. In this context, the main objective of this dissertation is the development and analysis of new tools and methods for monitoring and characterizing urban growth using geo-data and LULC databases, as well as exploring their relationships with socio-economic factors, providing new evidences regarding the use of LULC data for urban characterization at different levels by means of spatial and statistical methods.

First, we reviewed and compiled the most common spatio-temporal metrics and more than fifty metrics and indicators were collected and tested in urban environments. These metrics were implemented within a software tool, IndiFrag, which computes the metrics using LULC vector databases as input.

Then, we present a methodology based on spatio-temporal metrics and propose a new index (Population and Urban Growing Imbalance, PUGI), that quantifies the inequality of growth between population and built-up areas, to analyze and compare urban growth patterns at different levels. We computed spatio-temporal metrics at local level in a testing sample of six urban areas in Europe from the Urban Atlas LULC database, then uncorrelated metrics were selected and the data were interpreted at various levels. This allowed for a differentiation of growing patterns, discriminating between compact and sprawl trends. The index proposed complemented the analysis by including demographic dynamics, being also useful for assessing the growing imbalance between the progression on residential areas and the population change at multiple levels. The analysis at various levels (i.e., functional urban areas, urban and peri-urban sectors and administrative units) contributed to a better understanding of urban growth patterns and their relation to sustainable policies.

We quantified the two-way relationship between the urban structure in cities and their socio-economic status. Therefore, we measured spatial patterns of 31 cities in North Rhine-Westphalia, Germany, using spatial metrics derived from a Local Climate Zone classification obtained by classifying remote sensing and open geo-data with machine learning techniques. Based on these data, we quantified relationship between spatial metrics and socio-economic indicators related to education, health, living conditions, labor, and transport by means of multiple linear regression models, explaining the variability of the socio-economic variables from 43% up to 82%. Then, we grouped cities according to their level of quality of life using the socio-economic variables, and found that the spatial pattern of low-dense built-up types was different among socio-economic groups. The proposed method is transferable to other datasets, levels, and regions. The limitations and needed considerations when conducting such studies were discussed.

We assessed the use of spatio-temporal metrics derived from LULC maps to identify urban growth spatial patterns. For this, we applied LULC change models to simulate different long-term scenarios of urban growth following various spatial patterns (i.e., expansion, compact, dispersed, road-based and leapfrog) on diverse baseline urban forms (i.e., monocentric, polycentric, sprawl and linear). Then, we computed spatio-temporal metrics for the simulated scenarios, selected the most explanatory by applying a discriminant analysis and classified the growth patterns using clustering methods. Two metrics, which account for the densification, compactness and concentration of urban growth, were the most significant for classifying the five growth patterns, despite the influence of the baseline urban form. These metrics have

the potential to identify growth patterns to be used for monitoring the development and evaluating the management of urban areas.

Finally, we identified empirical relationships between income, inequality, GDP, air pollution and employment indicators and their change over time with the spatial structure of the built and natural elements in up to 600 urban areas from 32 countries. We measured the spatial structure by means of spatio-temporal metrics extracted from geo-information available at the global level. We employed random forest regression models and these metrics were able to explain between 32% and 68% of the variability of socio-economic variables. This confirms that spatial patterns and their change are linked to socio-economic indicators. We identified the most relevant spatio-temporal metrics in the models: we found that urban compactness, concentration degree, the dispersion index, the densification of built-up growth, accessibility, and LULC density and change, could be used as proxies for some socio-economic indicators. This is a fundamental step for the identification of such relationships on a global scale. The proposed methodology is highly versatile, the inclusion of new datasets is straightforward, and the increasing availability of multi-temporal geospatial and socio-economic databases is expected to boost the study of these relationships from a multi-temporal perspective in the near future.

This work contributes to a better understanding of urban growth patterns and improves knowledge about the relationships between urban spatial structure and socio-economic factors, providing new tools and methods for monitoring and assessing urban sustainability by means of LULC databases, which could be used by researchers, urban planners and decision-makers to ensure the sustainable future of urban environments pursued by the Sustainable Development Goals (SDGs).

Resumen

La urbanización es un proceso multidimensional que implica cambios económicos, sociales y ambientales y además altera la forma y el tamaño de las ciudades, la densidad y distribución de la población y los usos y coberturas del suelo (LULC, por sus siglas en inglés), afectando a la sostenibilidad global. La urbanización ha aumentado considerablemente en los últimos años y se espera que continúe a un ritmo sin precedentes en las próximas décadas. Por lo tanto, la promoción de pautas de urbanización sostenible, la limitación de consumo de suelos naturales y el bienestar de la población en las áreas urbanas se han convertido en objetivos políticos a distintos niveles, desde el nivel local hasta el nivel mundial. La estructura espacial de las áreas urbanas y sus patrones de crecimiento determinan la forma en que sus características físicas, socioeconómicas y ambientales cambian con el tiempo. Estas interrelaciones desempeñan un papel importante en la vida diaria de los ciudadanos y su conocimiento ayuda a los responsables políticos a tomar decisiones sensatas e informadas para la planificación sostenible de las áreas urbanas. Así pues, es de crucial importancia entender las relaciones que existen entre la estructura espacial de las áreas urbanas y su comportamiento socioeconómico.

Monitorizar, cuantificar y caracterizar el desarrollo de las áreas urbanas mejora la comprensión de los procesos pasados y actuales, además, proporciona información basada en evidencias y respalda los procesos de toma de decisiones, lo cual permite anticiparse a patrones poco sostenibles y a sus posibles consecuencias. Esto es posible gracias a las técnicas de observación de la Tierra y de los Sistemas de Información Geográfica (SIG). La observación de la Tierra proporciona datos precisos y frecuentes, así como las herramientas necesarias para monitorizar el crecimiento urbano, espacial y temporalmente a múltiples niveles, lo cual puede analizarse más a fondo mediante los SIG. Las métricas espaciales se utilizan cada vez más para monitorizar y cuantificar el crecimiento urbano. Sin embargo, la disponibilidad cada vez mayor de datos derivados de la observación de la Tierra y la aparición de bases de datos LULC y geoespaciales, requieren la elaboración de nuevos métodos y herramientas que aprovechen esta creciente disponibilidad de datos geográficos, los cuales se están convirtiendo en una valiosa fuente de información para estudios urbanos, ya que permiten el análisis exhaustivo de las áreas urbanas desde un punto de vista espacio-temporal en múltiples niveles. En este contexto, el principal objetivo de esta tesis es el desarrollo y análisis de nuevas herramientas y métodos para monitorizar y caracterizar el crecimiento urbano utilizando datos geográficos y bases de datos LULC, así como explorar sus relaciones con factores socioeconómicos, aportando nuevas evidencias sobre la utilización de los datos LULC para la caracterización urbana en diferentes niveles mediante

métodos espaciales y estadísticos.

En primer lugar, se revisaron y compilaron las métricas espacio-temporales más comunes, en total se evaluaron más de cincuenta métricas e indicadores en entornos urbanos. Estas métricas se implementaron dentro de una herramienta software, IndiFrag, que las calcula utilizando bases de datos LULC vectoriales como dato de entrada. A continuación, se presenta una metodología basada en métricas espacio-temporales y se propone un nuevo índice (*Population and Urban Growing Imbalance*, PUGI), que cuantifica la desigualdad entre el crecimiento de la población y las zonas edificadas, y sirve para analizar y comparar los patrones de crecimiento urbano en diferentes niveles. Se calcularon métricas espacio-temporales a nivel local en una muestra de seis áreas urbanas en Europa a partir de la base de datos LULC *Urban Atlas*, a continuación se seleccionaron las métricas no correlacionadas y se interpretaron los datos a varios niveles. Esto permitió una distinción de los patrones de crecimiento, discriminando entre tendencias de compactación y dispersión. El índice propuesto complementa el análisis incluyendo la evolución demográfica, que resulta útil para evaluar la desigualdad entre el aumento de las áreas residenciales y el cambio poblacional en múltiples niveles. El análisis a diversos niveles (es decir, en zonas urbanas funcionales, sectores urbanos y periurbanos y unidades administrativas) contribuyó a una mejor comprensión de los patrones de crecimiento urbano y su relación con políticas sostenibles.

En segundo lugar, se cuantificaron relaciones bidireccionales entre la estructura urbana de las ciudades y su situación socioeconómica. Para ello, se midieron los patrones espaciales de 31 ciudades en Renania del Norte-Westfalia (Alemania), utilizando métricas espaciales derivadas de un mapa de zonas climáticas locales, obtenido de una clasificación generada a partir de datos de teledetección y datos geográficos abiertos, con técnicas de aprendizaje automático. Posteriormente, se cuantificó la relación entre las métricas espaciales y los indicadores socioeconómicos relacionados con la educación, salud, condiciones de vida, trabajo y transporte, mediante modelos de regresión lineal múltiple, explicando la variabilidad de los indicadores socioeconómicos desde el 43% hasta el 82%. Seguidamente, se agruparon las ciudades según su nivel de calidad de vida utilizando las variables socioeconómicas y se observó que los patrones espaciales de las tipologías con edificaciones de baja densidad eran diferentes según los distintos grupos socioeconómicos. El método propuesto es transferible a otros conjuntos de datos, niveles y regiones. Además, se examinaron las limitaciones y las consideraciones necesarias para realizar este tipo de estudios.

Posteriormente, se evaluó el uso de las métricas espacio-temporales derivadas de los mapas LULC para identificar patrones espaciales de crecimiento urbano.

Para ello, se utilizaron modelos de cambio de usos del suelo para simular diferentes escenarios de crecimiento urbano a largo plazo siguiendo varios patrones espaciales (expansivo, compacto, disperso, basado en carreteras y *leapfrog*) en diversas formas urbanas de partida (monocéntrica, policéntrica, dispersa y lineal). A continuación, se calcularon las métricas espacio-temporales para los escenarios simulados, se seleccionaron las más explicativas aplicando análisis discriminante y se clasificaron los patrones de crecimiento utilizando métodos estadísticos de agrupación. Dos métricas, que miden la densidad, la compacidad y la concentración del crecimiento urbano, fueron las más significativas para clasificar los cinco patrones de crecimiento, a pesar de la influencia de la forma urbana de partida. Estos parámetros permiten determinar los patrones de crecimiento para ser utilizados como información para monitorizar el desarrollo y evaluar la gestión de las áreas urbanas.

Por último, se identificaron relaciones empíricas entre indicadores de ingresos, desigualdad, PIB, contaminación atmosférica y empleo y su evolución a lo largo del tiempo, con la estructura espacial de los elementos construidos y naturales en hasta 600 áreas urbanas de 32 países. Se midió la estructura espacial mediante métricas espacio-temporales extraídas de información geográfica disponible a nivel mundial. Se aplicaron modelos de regresión *random forest* y las métricas fueron capaces de explicar entre el 32% y el 68% de la variabilidad de los indicadores socioeconómicos. Esto confirma que los patrones espaciales y sus cambios están vinculados a los indicadores socioeconómicos. Además, se identificaron las métricas espacio-temporales más relevantes en los modelos: se observó que la compacidad urbana, el grado de concentración, el índice de dispersión, la densificación del crecimiento del suelo edificado, la accesibilidad, la densidad LULC y su variación, podrían ser utilizados como *proxies* para algunos indicadores socioeconómicos. Este es un paso fundamental para la identificación de relaciones a escala global. La metodología propuesta es altamente versátil, la incorporación de nuevos conjuntos de datos es sencilla y se espera que la creciente disponibilidad de bases de datos geoespaciales y socioeconómicas multitemporales promueva el estudio de tales relaciones desde una perspectiva multitemporal en un futuro próximo.

Este trabajo contribuye a una mayor comprensión de los patrones de crecimiento urbano y amplía el conocimiento sobre las relaciones entre la estructura espacial urbana y los factores socioeconómicos. Se proporcionan nuevas herramientas y métodos para monitorizar y evaluar la sostenibilidad urbana a partir de bases de datos LULC, que podrían ser utilizadas por los investigadores, planificadores urbanos y responsables políticos para garantizar un futuro sostenible en los entornos urbanos, lo cual se contempla en los

objetivos de desarrollo sostenible (SDGs).

Resum

La urbanització és un procés multidimensional que implica canvis econòmics, socials i ambientals i a més altera la forma i la grandària de les ciutats, la densitat i distribució de la població i els usos i cobertures del sòl (LULC, per les seues sigles en anglés), cosa que afecta a la sostenibilitat global. La urbanització ha augmentat considerablement en els últims anys i s'espera que continue a un ritme sense precedents en les pròximes dècades. Per tant, la promoció de pautes d'urbanització sostenible, la limitació de consum de sòl natural i el benestar de la població en les àrees urbanes s'han convertit en objectius polítics a diferents nivells, des del nivell local fins al nivell mundial. L'estructura espacial de les àrees urbanes i els seus patrons de creixement determinen la forma en què les seues característiques físiques, socioeconòmiques i ambientals canvien amb el temps. Aquestes interrelacions exerceixen un paper important en la vida diària dels ciutadans i el seu coneixement ajuda els responsables polítics a prendre decisions sensates i informades per a la planificació sostenible de les àrees urbanes. Així doncs, és de crucial importància entendre les relacions que existeixen entre l'estructura espacial de les àrees urbanes i el seu comportament socioeconòmic.

Monitorar, quantificar i caracteritzar el desenvolupament de les àrees urbanes enriqueix la nostra comprensió dels processos passats i actuals, a més, proporciona informació basada en evidències i recolza els processos de presa de decisions, això permet anticipar-se a patrons poc sostenibles i a les seues possibles conseqüències. Això és possible gràcies a les tècniques d'observació de la Terra i dels Sistemes d'Informació Geogràfica (SIG). L'observació de la Terra proporciona dades precises i freqüents, així com les ferramentes necessàries per a monitorar el creixement urbà, des d'un punt de vista espacial i temporal a múltiples nivells, els quals poden analitzar-se més a fons mitjançant els SIG. Les mètriques espacials s'utilitzen cada vegada més per a monitorar i quantificar el creixement urbà. No obstant això, la disponibilitat cada vegada major de dades derivades de l'observació de la Terra i l'aparició de bases de dades LULC i geoespacials, requereixen l'elaboració de nous mètodes i de ferramentes que aprofiten aquesta creixent disponibilitat de dades geogràfiques, les quals s'estan convertint en una valuosa font d'informació per a estudis urbans, ja que permeten l'anàlisi exhaustiva de les àrees urbanes des d'un punt de vista espaciotemporal en múltiples nivells. En aquest context, el principal objectiu d'aquesta tesi és el desenvolupament i l'anàlisi de noves ferramentes i mètodes per a monitorar i caracteritzar el creixement urbà utilitzant dades geogràfiques i bases de dades LULC, així com explorar les seues relacions amb factors socioeconòmics, i aportar noves evidències sobre la utilització de les dades LULC per a la caracterització urbana

en diferents nivells mitjançant mètodes espacials i estadístics.

En primer lloc, es van revisar i compilar les mètriques espaciotemporals més comunes, en total es van avaluar més de cinquanta mètriques i indicadors en entorns urbans. Aquestes mètriques es van implementar dins d'un programari, IndiFrag, que les calcula utilitzant bases de dades LULC vectorials com a dada d'entrada. A continuació, es presenta una metodologia basada en mètriques espaciotemporals i es proposa un nou índex (*Population and Urban Growing Imbalance*, PUGI), que quantifica la desigualtat entre el creixement de la població i les zones edificades, i serveix per a analitzar i comparar els patrons de creixement urbà a diferents nivells. Es van calcular mètriques espaciotemporals a nivell local en una mostra de sis àrees urbanes a Europa a partir de la base de dades LULC *Urban Atlas*, a continuació es van seleccionar les mètriques no correlacionades i es van interpretar les dades a diversos nivells. Això va permetre una distinció dels patrons de creixement i es va discriminar entre tendències de compactació i dispersió. L'índex proposat va complementar l'anàlisi incloent l'evolució demogràfica, que resulta útil per a avaluar la desigualtat entre l'augment de les àrees residencials i el canvi poblacional en múltiples nivells. L'anàlisi a diversos nivells (és a dir, en àrees urbanes funcionals, sectors urbans i periurbans i unitats administratives) va contribuir a una millor comprensió dels patrons de creixement urbà i la seua relació amb polítiques sostenibles.

En segon lloc, es van quantificar relacions bidireccionals entre l'estructura urbana de les ciutats i la seua situació socioeconòmica. Per a això, es van mesurar els patrons espacials de 31 ciutats a Renània del Nord-Westfàlia (Alemanya), utilitzant mètriques espacials derivades d'un mapa de zones climàtiques locals, obtingut d'una classificació generada a partir de dades de teledetecció i dades geogràfiques obertes, amb tècniques d'aprenentatge automàtic. Posteriorment, es va quantificar la relació entre les mètriques espacials i els indicadors socioeconòmics relacionats amb l'educació, salut, condicions de vida, treball i transport mitjançant models de regressió lineal múltiple, la qual va explicar la variabilitat dels indicadors socioeconòmiques des del 43% fins al 82%. Seguidament, es van agrupar les ciutats segons el seu nivell de qualitat de vida utilitzant les variables socioeconòmiques i es va observar que els patrons espacials de les tipologies amb edificacions de baixa densitat variaven segons els diferents grups socioeconòmics. El mètode proposat és transferible a altres conjunts de dades, nivells i regions. A més, es van examinar les limitacions i les consideracions necessàries per a realitzar aquest tipus d'estudis.

Seguidament, es va avaluar l'ús de les mètriques espaciotemporals derivades dels mapes LULC per a identificar patrons espacials de creixement urbà. Per a

això, es van utilitzar models de canvi d'usos del sòl per a simular diferents escenaris de creixement urbà a llarg termini seguint diversos patrons espacials (expansiu, compacte, dispers, basat en carreteres i *leapfrog*) en diverses formes urbanes de partida (monocèntrica, policèntrica, dispersa i lineal). A continuació, es van calcular les mètriques espaciotemporals per als escenaris simulats, es van seleccionar les més explicatives aplicant anàlisi discriminant i es van classificar els patrons de creixement utilitzant mètodes estadístics d'agrupació. Dues mètriques, que mesuren la densitat, la compactació i la concentració del creixement urbà, van ser les més significatives per a classificar els cinc patrons de creixement, malgrat la influència de la forma urbana de partida. Aquests paràmetres permeten determinar els patrons de creixement per a ser utilitzats com a informació per a monitorar el desenvolupament i avaluar la gestió de les àrees urbanes.

Finalment, es van identificar relacions empíriques entre indicadors d'ingressos, desigualtat, PIB, contaminació atmosfèrica i ocupació i la seua evolució al llarg del temps, amb l'estructura espacial dels elements construïts i naturals en fins a 600 àrees urbanes de 32 països. Es va mesurar l'estructura espacial mitjançant mètriques espaciotemporals obtingudes d'informació geogràfica disponible a nivell mundial. Es van aplicar models de regressió *random forest* i les mètriques van poder explicar entre el 32% i el 68% de la variabilitat dels indicadors socioeconòmics. Això confirma que els patrons espacials i els seus canvis estan vinculats als indicadors socioeconòmics. A més, es van identificar les mètriques espaciotemporals més rellevants en els models: es va observar que la compactació urbana, el grau de concentració, l'índex de dispersió, la densificació del creixement del sòl edificat, l'accessibilitat, la densitat LULC i la seua variació, podrien ser utilitzats com *proxies* per a alguns indicadors socioeconòmics. Aquest és un pas fonamental per a la identificació de relacions a escala global. La metodologia proposada és altament versàtil, la incorporació de nous conjunts de dades és senzilla i s'espera que la creixent disponibilitat de bases de dades geoespacials i socioeconòmiques multitemporals promoga l'estudi d'aquestes relacions des d'una perspectiva multitemporal en un futur pròxim.

Aquest treball contribueix a una major comprensió dels patrons de creixement urbà i amplia el coneixement sobre les relacions entre l'estructura espacial urbana i els factors socioeconòmics. Es proporcionen noves ferramentes i mètodes per a monitorar i avaluar la sostenibilitat urbana a partir de bases de dades LULC, que podrien ser utilitzades pels investigadors, planificadors urbans i responsables polítics per a garantir un futur sostenible en els entorns urbans, cosa que es contempla en els objectius de desenvolupament sostenible (SDGs).

Acknowledgments

I would like to briefly show my gratitude to everyone who helped me, supported me or simply was there during these five intense years of my research. A time that was not always easy, but from which I take great personal and professional experiences and learning.

En primer lugar, quiero agradecer a mi director y tutor de tesis Luis A. Ruiz por todo. En especial por haberme enseñado, orientado y ayudado durante todos estos años, así como por proponer el tema que desencadenó esta investigación y con ello brindarme la oportunidad de realizar mi tesis doctoral. Sobre todo, gracias por introducirme en el mundo de la investigación y permitirme descubrir mi vocación.

A todas/os mis compañeras/os del CGAT por cada momento vivido estos años, vuestra amistad y apoyo. Gracias por hacer que trabajar fuera tan fácil y divertido. En especial, quiero agradecer a Elena por estar siempre ahí, entenderme tanto, ser mi ejemplo e inestimable amiga.

Thanks to DLR for having me as a guest scientist, it was a very enriching experience from which I learned a lot. I would like to show my gratitude to my colleagues of the City and Society Team, thank you for making me feel so comfortable and part of the team from the very beginning. I feel very fortunate. I want to give special thanks to Hannes, for your trust, understanding and invaluable support to finish my PhD.

Por último, a mi familia y amigas/os, estoy segura de que no ha sido fácil aguantar mis estados de ánimo en esta montaña rusa, pero al final mereció la pena. Os quiero por vuestra comprensión y paciencia. A Pau, no hay palabras para agradecer todo lo que has hecho por mí y lo que has aguantado. Gracias por tu apoyo infinito y creer en mí de manera incondicional.

Table of contents

of life

Spatio-temporal metrics for urban growth spatial pattern categorization

Linking spatio-temporal metrics of built-up areas to socio-economic indicators on a semi-global scale

List of figures

List of tables

Chapter 1

Introduction

1.1. Background

1.1.1. The relevance of monitoring urban growth

Since the beginning of the twenty-first century, urban areas are home to the majority of population. The United Nations (2018) foretold that the imbalance between urban and rural population is expected to soar in the coming years. However, urban population will grow at different levels and paces across countries, regions, and especially across income groups. Low-income and lower-middle-income countries are expected to experience most of the growth, whereas high-income countries will also experience an increase in urban population but at slower pace, since they already have relatively high levels of urbanization. In general, urban environments will face new challenges in the wake of an unprecedented growing number of urban dwellers.

Not only population is expected to grow in cites, but also the number and size of cities. This process is known as urbanization and the United Nations (2019) defined it as follows: "*Urbanization is a complex socio-economic process that transforms the built environment, converting formerly rural into urban settlements, while also shifting the spatial distribution of a population from rural to urban areas. It includes changes in dominant occupations, lifestyle, culture and behavior, and thus alters the demographic and social structure of both urban and rural areas […]. Urbanization refers both to the increase in the percentage of population residing in urban areas […] and the total area occupied by urban settlements*". This means that urbanization is a multi-dimensional process linked to socio-economic factors. It entails changes in the environment, urban form and size, population density and distribution, and triggers land-use/land-cover changes, which in turn have social and economic effects.

The rapid urbanization expected in the coming years will bring opportunities and challenges to urban areas. On the one hand, well-planned and well-managed urbanization is a powerful tool for sustainable development both, in developing and developed countries. Urban areas and cities bring new opportunities and economic growth, are more resource efficient, provide better access to employment, education, services and amenities, etc. (Seto et al., 2017; Ahlfeldt et al., 2018). Therefore, cities can be the solution for a better future, since the agglomeration of people, energy and resources creates hotspots for innovation and governments usually get involved in global and local affairs towards sustainable future, such as climate, resilience, health, cultural tolerance, and race and gender inequality (Acuto et al., 2018), ultimately improving the well-being of urban dwellers.

On the other hand, cities occupy less than 3% of the Earth, but they consume the majority of the energy and produce most of the carbon emissions (UN, 2019). Land is a limited resource and it is shrinking due to urbanization pressure. Heretofore, population growth has been generally accompanied by an outpacing increase of the urban layout, turning in lower population densities, higher land take and causing, inter alia, ecosystem degradation, socio-economic consequences, and loss of quality of life (Taubenböck et al., 2012; Haase et al., 2013; Ribeiro-Barranco et al., 2014; Kompil et al., 2015; Güneralp et al., 2020). For example, urban land in European cities grew from 1990 to 2006 even in those cities with declining population, raising the natural land consumption in this period and decreasing population density, measured as the population per km^2 of urban area (Haase et al., 2013). The authors suggested that the increased economic capacity fostered the affordability of higher square footage and the preference to live alone, which increased per capita living space and may have been a reason for the unequal growth between urban land and population in Europe. Not only in Europe, but also in India, China and North America population densities declined between 1970 and 2010, exhibiting low levels of land-use efficiency (Güneralp et al., 2020). This is especially relevant in rapid expanding urban areas of Africa and Asia where there are particular challenges; population growth puts a high pressure on local governments responding to issues of job, housing, provision of infrastructure, social and health services (Acuto et al., 2018).

Inequality and urban sprawl are only two of the manifold consequences that urbanization triggers when inefficiently managed. Almost 1 billion of people are living in deprived conditions nowadays (UN, 2020), meanwhile many cities in higher-income countries experienced urban sprawl in the last 30 years (OECD, 2018). In the last years, these among many other issues have been under the spotlight for scholars, urban planners and policymakers (Wei and Ewing, 2018). For instance, Arribas-Bel et al. (2011) measured urban sprawl in Europe, and identified the most sprawled regions in central Europe, around Germany; they proposed a multi-dimensional index to embrace the complexity of this phenomenon. Later on, the European Environmental Agency (EEA) in response of the increasing concern of sprawl in Europe developed a detailed report highlighting the importance of monitoring urban sprawl, its effects, drivers, how to measure it, and its policy relevance and implications, with the ultimate aim of increasing policy awareness and ensuring a healthy environment and well-being (EEA, 2016a). Gielen et al. (2019) quantified the added cost of sprawl at the municipal level, showing a higher expenditure in security, public transport and community well-being (i.e., providing services of waste collection, sanitation, water and light supply, etc.). Regarding inequality, this is higher in urban areas than in rural areas and it affects both developed

and developing countries. In the last thirty years, economic inequalities increased in developed countries, while declined in Latin American, Asian and African countries. However, these regions still show high levels of inequality. In most cities, wealth areas coexist with areas of severe deprivation (UN, 2020). Thus, deriving reliable spatial information on the size and location of deprived areas has gained much interest to locate settlements with fewer resources, inadequate housing or worse living conditions, which is crucial information for integrating these areas into governing structures (e.g., Kuffer et al., 2018; Wurm et al., 2019a).

Overall, population growth accompanied by inequitable, inefficient, and unsustainable nature's resources consumption and technological development have driven several effects in the past decades and are likely to increase in the future, including climate change, ocean acidification, land degradation, water scarcity, fishery over-exploitation, biodiversity loss, poverty, sprawl, etc. These effects bring serious challenges to the global health and sustainability (Whitmee et al., 2015; Wei and Ewing, 2018). These trends contradict the global targets of sustainable development in urban areas, where reaching a sustainable urbanization process depends on well-managed urban growth, optimizing agglomeration, and reducing environmental degradation (UN, 2018). In the last ten years, new planning initiatives, international agreements, and programs have been developed to reconsider the urbanization process and promote sustainable land use. In 2015, the United Nations Development Agency defined seventeen Sustainable Development Goals (SDGs) for ensuring a sustainable future for people and the planet to be completed by 2030 (UN, 2015), which aim at ending poverty by means of promoting economic growth, addressing social needs, while protecting the environment and fighting climate change. Several of the goals are related to urban areas, especially the eleventh goal which focuses on the environmental, social and economic sustainability of cities (SDG 11). One year later, the New Urban Agenda from the United Nations developed guidelines and recommendations for a better and more sustainable future considering the SDGs. Three more examples in Europe are the Informal Ministerial Meeting on Urban Development Declaration (2010), the 7th Environment Action Program (EC, 2013) and the European Union Urban Agenda (EC, 2017), which promote urban recycling, compact city planning, improve green infrastructure and soil protection as measures for a more sustainable development of cities. These reports and programs exemplify that the sustainability of developing and developed urban areas is an ongoing concern worldwide.

Monitoring the development of urban and peri-urban areas enriches the understanding of how urban areas developed in the past and where are

growing. It provides important insights and evidence-based information that help decision making, anticipate potential adverse effects of unsustainable development trends and, thus, resulting in more effective solutions (Patino and Duque, 2013). Schneider and Woodcock (2008) measured the shape and patterns of urban expansion in a set of global cities and compared the patterns with the unsustainable sprawling urban forms seen in many US cities, then they discussed the possible drivers of these patterns, such as demographic, socio-economic, geographic or land use policies. Besides, other study developed a methodology to monitor urban growth, in particular focused on different types of sprawl, which aimed at monitoring land use, identifying problematic areas to derive sprawl-type specific strategies, and policy measures in line with the sustainability strategy goals to reduce land consumption set by the government (Siedentop and Fina, 2010). Similarly, Salvati et al. (2016) analyzed urban expansion in three compact cities in Europe, finding scattered growth trends with different local patterns influenced by place-specific factors, which affected the uneven socio-economic configuration between urban and rural areas. Therefore, the urban form and urban growth spatial pattern are, directly and indirectly, related to the global sustainability. The knowledge of how urban areas are spatially configured, their size, form, density, spatial pattern, and their variations over time in urban and peri-urban areas is essential to successfully monitor the urbanization impacts on the environment and their socio-economic effects.

1.1.2. Relationships between urban structure and socio-economic and ecological factors

How we organize space in urban areas has a decisive influence on how we live and what effects this has on our closest environment: what kind of mobility we choose, how large our ecological footprint is, how close we are to utilities, or what access we have to jobs or leisure facilities. At the same time, urban areas are continuously changing their form to adapt to socio-economic conditions; they are mainly shaped by our way of life, communications, transport technology, and market forces (Williams et al., 2000). This two-sided link between the structure of urban areas and the environmental, social and economic factors has been discussed by many theorists, urban planners and scholars (Jacobs, 1961; Williams et al., 2000; Taubenböck et al., 2009; Salat, 2011; Tonkiss, 2013; Oliveira, 2016; Lobo et al., 2020).

For example, a recent study observed that urban form influenced economic performance and productivity in cities (Duque et al., 2019), as well as purchasing power (Huang et al., 2007), while land use mixture was related to wealth indicators (Tapiador et al., 2011). A case study in Toronto, Canada,

found that a reinvestment into the inner city, which resulted in the development of high-rise buildings and a significant increase of population, triggered the spatial segregation of the city, concentrating wealth in the inner city, raising poverty in the suburbs, and reducing the number of middle-income neighborhoods (Lehrer and Wieditz, 2009). Another study in cities from United States, Europe and China showed that urban growth patterns determined the sustainability of development processes and, therefore, the environmental impact that urbanization has on our surroundings, such as the amount of land consumption, the choice of means of transport and greenhouse gas emissions (Dong et al., 2019). Angel et al. (2011) found that the spatial structure of cities, specifically fragmented footprints, were explained by means of population size, income, car ownership, topographic constraints, water availability and planning restrictions. Regarding health issues, the prevalence of non-communicable diseases, such as those related to physical health, dietary habits or alcohol consumption, has been related to the socio-economic status of the population (Allen et al., 2017; Belsky et al., 2019); in addition, the availability of accessible green spaces has been associated with a reduction of the risk of cardiovascular and respiratory diseases (Villeneuve et al., 2012), and the percentage and proximity of greenness in the living environment have a positive relationship with physical and mental health (Weigand et al., 2019); meanwhile, habitat loss and fragmented landscapes increase the probability of the emergence of infectious diseases in humans (Patz et al., 2004; Wilkinson et al., 2018; Zohdy et al., 2019). With regards to land use planning, two different reports showed the influence of the British planning system on housing prices, employment, transport systems and access to services, among other socio-economic characteristics (Cheshire et al., 2012; Williams, 2014). These are just a few of the many examples that show the influence spatial design and growth patterns have in socio-economic factors, quality of life and sustainability, and vice versa.

In this context, urban planning and land-use policies have the capacity to rethink and reshape urban areas by means of designing and implementing well-informed decisions and specific policies aimed at the type of development to be planned to accommodate new coming dwellers, when demands for housings, servicies, food and instrastructure grow (UN-Habitat, 2013). How urban growth is planned has long-term impacts that will determine the social, economic and environmental outcomes of our future. Consequently, land-use evolution needs to be monitored and evaluated over time in order to validate urban planning and development policies, allowing to adjust or reorient strategic measures towards sustainable choices when necessary (OECD, 2017a). Besides, identifying social, economic and environmental underlying processes of urbanization and land-use changes improves the understanding

of cause–effect relationships and helps in the development of strategies for sustainable development (Wentz et al., 2014). Additionally, the analysis over time would allow gaining knowledge of how urban growth affects cities and urban areas spatially and socio-economically, and inversely the influence of socio-economic policies and evolution on the urban growth patterns, as well as quantifying their interrelationships.

1.1.3. Development of Earth Observation programs and databases

Earth Observation (EO) and Geographic Information System (GIS) techniques play an important role for monitoring urban areas, analyzing urban form and growth (Liu et al., 2010; Ju et al., 2016; Zhao et al., 2016; Abrantes et al., 2019). EO provides timely, accurate and frequent data, as well as the tools to remotely capture urban expansion and characterizing urban environments spatially across time at different levels. It allows the measurement from coarse to fine patterns of urban form and dynamics in a consistent way (Patino and Duque, 2013; Wentz et al., 2014; Wentz et al., 2018), which is essential to understand urbanization processes. EO data has been recognized from urban planners to policymakers as a fundamental source of geospatial data for urban studies (UN, 2019; Lobo et al., 2020; Zhu et al., 2019). The combined use of EO and GIS allows for a thorough analysis and mapping of geo-information from a multi-temporal and multi-scale perspective. In general, satellite images can be used to categorize the landscape into land-use/land-covers (LULC), while GIS enables the spatial analysis of LULC, LULC changes and urban growth (Schneider and Woodcock, 2008; Jiao, 2015; Wu et al., 2016). Thus, at the global level, using Sentinel-2 satellite imagery and GIS methods to delineate consistenly urban areas, the classification into urban structural types and land covers served to identify mophological types across the world based on urban spatial configurations, which evidenced geographical-cultural similarities (Taubenböck et al., 2020). At the continental level, urban areas in Latin America were classified using Landsat imagery from two dates, then urban expansion and sprawling trends were measured with GIS tools (Inostroza et al., 2013). As an example of the metropolitan level, urban land use change in a large metropolis of the US was monitored by means of remote sensing-based LULC classification in two dates, and the subsequent GIS-based analyses of LULC change, spatial variations and urban expansion were made (Liu and Yang, 2015). At the local scale, using very high resolution remote sensing data, such as satellite images and Airborne Laser Scanning data combined with ancillary geospatial information, allowed creating 3D building models and classifying them into urban land uses and functions, which were used for population distribution models, representing valuable information for multiple applications, such as risk assessment (Aubrecht et al., 2009).

With the proliferation of sensors in the last decades there is an ever-growing availability of satellite images covering spatially and temporally the Earth, which has subsequently boosted the development of new methods and techniques to characterize urban settings, transforming image data into reliable geo-information (Taubenböck, 2019). For instance, the categorical classification of images into LULC maps is valuable information for urban studies (Donnay et al., 2000; Hermosilla et al., 2012), the characterization of cities into urban structural types and land cover has great potential in its relation with urban functions (Bechtel et al., 2015), and the identification of built-up areas at different scales, from single building to urban settlements provides accurate representations of spatial patterns in urban areas (Hermosilla et al., 2011; Stiller et al., 2019; Cao et al., 2020; Qui et al., 2020).

The increasing demand for more accurate, higher resolution and up-to-date LULC datasets from different user communities, and the gained relevance of EO data for urban studies, has resulted in the development of geospatial datasets representing accurately urban and natural environments. Understanding and monitoring LULC spatial distribution and their change is of high relevance in urban and natural environmental studies and planning. Updated LULC data is essential for several organizations, policymakers and stakeholders for quantifying LULC patterns and dynamics in order to develop and implement measures and programs aiming to achieve a sustainable development, support climate change mitigation, aid in national, regional and local planning and in the management of urban areas (Grekousis et al., 2015; Copernicus, 2020a). In this context several local, national, continental and global institutions and administrations, through EO and mapping programs, have developed geospatial datasets for aiding in planning, managing and monitoring urban areas. This is possible thanks to the improved technological and algorithm advancement for LULC characterization, and the increasing availability of multi-temporal and multi-resolution satellite imagery (Yang et al., 2018).

These datasets provide harmonized and consistent geo-information over time and across regions, which are essential to successfully monitor urban areas, enables comparison of urban spatial patterns, improves the understanding of LULC dynamics and their implications, as well as contributes to the development of empirical studies (Ribeiro-Barranco, et al., 2014; Sapena et al., 2016; Pazúr et al., 2017). A collection of available geospatial datasets at global and semi-global levels managed and maintained by institutions and agencies is summarized in Table 1.1, showing the great efforts that are being made in this regard.

Table 1.1. Open geospatial datasets covering and representing urban areas at different levels, scales, spatial and temporal resolutions. Only datasets with spatial resolution better than 1km and covering more than one country are included. (*) means under development. The providers are: European Commission (EC), Join Research Center (JRC), European Spatial Agency (ESA), German Aerospace Center (DLR), National Geomatics Center of China (NGCC), Socioeconomic Data and Applications Center (SEDAC), Lincoln Institute of Land Policy (LILP) and Ney York University (NYU). Pan-EU means Pan-European (EU27, EFTA countries, West Balkans, Turkey, and UK). Minimum Mapping Unit (MMU), Minimum Mapping Width (MMW), Google Earth (GE) and Open Street Maps (OSM).

Dataset	Description	EO-source	Level	Scale	Spatial res.	Year	Provider	Reference
CCI-LC	The Climate Change Initiative Land Cover (CCI-LC) consists of a time series of 23-class land cover maps on an annual basis.	ENVISAT MERIS sensor	Continuous	Global	300 m	1992-2018	ESA	(ESA, 2017)
CGLS-LC100	Copernicus Global Land Cover Layers (CGLS) are complete discrete classifications of 23-class maps and sets of cover fractions (%) for 10 classes on annual basis.	PROBA-V	Continuous	Global	100 m	2015-2019	Copernicus	(Buchhorn et al., 2020a) (Buchhorn et al., 2020b)
GlobeLand30	It is a high-resolution 10-class land cover map.	Landsat and Chinese HJ-1	Continuous	Global	30 m	2000, 2010	NGCC	(Chen et al., 2017)
GHSL	The Global Human Settlement Layer (GHSL) produces global spatial information about the human presence on the planet over time.	Landsat	Continuous	Global	30 m, 250 m, 1 km	1975, 1990, 2000, 2014	EC and JRC	(Corbane et al., 2018a)
HBASE and GMIS	The Global Human Built-up and Settlement Extent (HBASE) and Global Man-made Impervious Surface (GMIS) are estimates of global urban extent and man-made imperviousness, respectively.	Landsat	Continuous	Global	30 m	2010	SEDAC	(Brown de Colstoun et al., 2017) (Wang et al., 2017)
GHSL (S1)	A global map of built-up presence derived from backscattered information of Sentinel1 images.	Sentinel-1	Continuous	Global	20 m	2016	EC and JRC	(Corbane et al., 2018b)
GUF	The Global Urban Footprint (GUF) is a worldwide mapping of settlements.	TerraSAR-X and TanDEM-X	Continuous	Global	12 m	2012	DLR	(Esch et al., 2012)

WSF2015	The World Settlement Footprint (WSF) is a worldwide human settlement layer.	Sentinel-1 and Landsat-8	Continuous	Global	10m	2015	ESA	(Marconcini et al., 2020)
WSF-Evo	Worldwide growth of human settlement on a year-by-year basis.	Sentinel-1 and Landsat-8	Continuous	Global	30m	1985-2015*	ESA	ESA (2019)
Atlas of urban expansion	It provides the urban extension and spatial changes in 200 sampled cities around the world.	Landsat	Cities	Global	30m	1990, 2000, 2014	LILP, UN-Habitat, and NYU	(Angel et al., 2016)
CLC	The CORINE Land Cover (CLC) is a standardized data collection on 44-class land cover in Europe to support environmental policy development.	Landsat, SPOT-4/5, IRS P6 LISS III, RapidEye, Sentinel-2, Landsat-8	Continuous	Pan-EU	100m (MMU 25ha) (MMW 100m)	1990, 2000, 2006, 2012, 2018	Copernicus	(Copernicus, 2020b)
IMD	The Imperviousness degree (IMD) captures the spatial distribution of artificially sealed areas, including the level of sealing of the soil per area unit, in a range from 0-100%.	IRS-P6, Resourcesat-2 LISS-III, SPOT 5, Landsat 8, and VHR imagery	Continuous	Pan-EU	20m, 100m (MMW 20m)	2006, 2009, 2012, 2015, 2018*	Copernicus	(Copernicus, 2020b)
ESM	The European Settlement Map (ESM) represents the percentage of built-up area coverage per spatial unit.	SPOT5 and SPOT6	Continuous	Pan-EU	2.5m, 10m, 100m	2012, 2015	Copernicus	(Corbane and Sabo, 2019)
Urban Atlas	Detailed and comparable land cover and land use information over major European city areas (from 319 to 788 urban areas depending on the year)	SPOT 5, SPOT6, Formosat-2, Pléiades, Planet, KOMPSAT, OSM SuperView, GE,	Functional urban areas (FUAs)	Pan-EU	1/10,000 (MMU 0.25ha urban, 1ha rural)	2006, 2012, 2018*	Copernicus	(Copernicus, 2020b)
NALCMS	The North American Land Change Monitoring System (NALCMS) is a collective effort. A harmonized multi-scale land cover monitoring approach which ensures high accuracy and consistency in monitoring 19-class land cover changes.	MODIS, Landsat, RapidEye	Continuous	North America (Canada, Mexico and US)	30m, 250m	2005, 2010 (250m) 2010, 2015 (30m)	NALCMS	(CEC, 2020)

However, not only public entities, but also scholars are making available their datasets and codes to the scientific community (e.g., Inglada et al., 2017; Leyk and Uhl, 2018; Demuzere et al., 2019; Cao et al., 2020; Gong et al., 2020; Liu et al., 2020; Qui et al., 2020; Weigand et al., 2020), allowing for the comparison of results in order to evaluate them, detect potential weaknesses and agreements (Uhl et al., 2020), as well as to reproduce results in other areas and to update the datasets whenever new input data are available, improving the cross-comparability and multi-temporality of geospatial datasets. Most significantly, in the last few years there have been numerous attempts to create global maps of annual urban land coverage (Zhou et al., 2018; He et al., 2019; Cao et al., 2020; Gong et al., 2020; Liu et al., 2020; Qui et al., 2020), which are fundamental geo-information for climate change mitigation and monitoring urban expansion to support the SDGs.

In this context, the spatial analysis of urban areas and their development using quantitative methods, both, at the intra-urban and inter-urban levels is becoming more common, since more data and analytical methods are available (Venerandi et al., 2018). For example, the comparison of cities from different countries is now more feasible, accordingly the categorization of cities based on their urban form and sprawl is possible (Schwarz, 2010; Arribas-Bel et al., 2011), as well as the analysis of urban development trends and population indicators (Haase et al., 2013; Ribeiro Barranco et al., 2014). Another quantitative approach is based on modeling urban expansion, economic and demographic changes to predict future urban growth or the creation of alternative growth scenarios based on planning strategies (Uuemaa et al., 2013; Kompil et al., 2015; Wissen Hayek et al., 2015; Van de Voorde et al., 2016).

1.1.4. Spatio-temporal metrics

Traditionally, landscape metrics have been used in landscape ecology studies, landscape monitoring and planning, since they are useful in quantifying the configuration, composition and spatial heterogeneity of landscape structures and biodiversity (Herold et al., 2005; Schindler et al., 2008; Plexida et al., 2014). Landscape metrics are static or mono-temporal indicators that summarize the complexity of spatial patterns into quantitative terms from categorical maps at specific scales (Llausàs and Nogué, 2012); therefore, when used in multi-temporal analyses the values of the metrics are compared over time and the changes interpreted (e.g., Lausch and Herzog, 2002; Olsen et al., 2007; Malaviya et al., 2010).

In the last twenty years there has been a growing interest and use of landscape metrics in urban landscapes, since their ability to measure the spatial

configuration and patterns from categorical maps has a great potential in urban studies (Uuemaa et al., 2013; Liu et al., 2015; Zhu et al., 2019). Commonly, when landscape metrics are used in urban environments they are referred as spatial metrics (Herold et al., 2005; Lowry and Lowry, 2014; Reis et al., 2015; Van de Voorde et al., 2016); however, they can also be found as landscape metrics, indices, urban metrics, among others (Schwarz, 2010; Arribas-Bel et al., 2011; Van de Voorde et al., 2011). In this dissertation the term spatial metrics is used hereafter. Like landscape metrics, spatial metrics are also mono-temporal, yet they are commonly used multi-temporally for several applications, as it is for example the case with urban expansion monitoring (e.g., Herold et al., 2005; Araya and Cabral, 2010). Some authors indicated the constraints for quantifying urban dynamics using traditional spatial metrics (Liu et al., 2010), and proposed the use of metrics for quantifying directly urban dynamics in two or more time points (Liu et al., 2010; Wilson et al., 2003). Therefore, there are two types of metrics that can be computed from categorical maps. First, spatial metrics are extracted for one single map, and the difference between the two independent resulting values has to be computed to analyze dynamic patterns between two different maps or dates. Second, multi-temporal metrics are those metrics that use two or more categorical maps to measure spatial patterns over time (e.g., the distance between a new patch in the second map and the closest patch from the first map). In this thesis, we refer to the combination of these metrics as spatio-temporal metrics (i.e., spatial metrics, the differences between spatial metrics over time and multi-temporal metrics). Therefore, not only spatio-temporal metrics measure landscape characteristics, such as the spatial configuration, aggregation properties, diversity of land uses, shape and size, but also describe landscape changes and growth patterns.

Spatio-temporal metrics have been used in many applications in urban environments aiming to support planning policies, comparing patterns among regions, or understanding urbanization patterns (Reis et al., 2015). For example, measuring different dimensions of urban sprawl with spatial metrics, combining their analysis with demographic and socio-economic data, and using clustering methods to identify the most sprawled areas in Europe (Schwarz, 2010; Arribas-Bel et al., 2011). These two studies used available databases at the European level, the CORINE Land Cover (CLC) for the LULC data (Table 1.1) and Urban Audit (Eurostat, 2016a) for population and socio-economic indicators. Dewan et al. (2012) applied spatio-temporal metrics to several dates in order to quantify land use dynamic patterns and measure landscape fragmentation due to urbanization. Spatio-temporal metrics have also been used as goodness-of-fit for calibrating LULC change simulation models, using the values of the metrics to assess the simulated urban growth

pattern scenarios (Van de Voorde et al., 2016). Spatio-temporal metrics based on identifying infill, expansion and outlying growth types have been used to characterize long-term patterns of urban growth (Sun et al., 2013).

1.1.5. Tools for computing spatio-temporal metrics

The available software and tools to compute spatial metrics from categorical maps are diverse and each can be useful for different settings (MacLean and Congalton, 2015). Table 1.2 collects the existing software that measure spatio-temporal metrics from categorical maps and their functionality is compared based on the computation levels, input and output data formats, dependencies from other software and the main application. The most common software tools for extracting spatial metrics from LULC data are FRAGSTATS v4 (McGarigal et al., 2012), Patch Analyst 5.1 (Rempel et al., 2012) and V-LATE (Lang and Tiede, 2003). The first one is the most used software in ecology applications, but the main shortcomings in urban applications are that it only works with raster data and metrics cannot be applied to multiple regions of interest (ROIs). The last two, based on FRAGSTATS metrics and developed as an extension of ArcGis software (ESRI®), include the capability of working with vector data. In particular, Patch Analyst allows working with different ROIs. Similar tools, also working from ArcGis, are PolyFrag (MacLean and Congalton, 2013), ZonalMetrics (Adamczyk and Tiede, 2017) and Arc_LIND (Yu et al., 2019). PolyFrag is based on FRAGSTATS metrics, it proposes also a vector-based analysis and solves some limitations from Patch Analyst and V-LATE, however, the computation on ROI is not included. ZonalMetrics decreases the number of metrics but includes the option to compute several types of ROIs. Other tools relying on open software are r.li (Porta and Spano, 2008), Land-metric DIY (Zaragozí et al., 2012), LecoS (Jung, 2016), GuidosToolbox (Vogt and Riitters, 2017), r.pi (Wegmann, et al., 2018), landscapemetrics (Hesselbarth, et al., 2019), and PyLandStats (Bosch, 2019). r.li and r.pi are plug-ins in the GIS open software GRASS (GRASS development team, 2017), they measure several patch and class metrics to raster maps, but their application to several classes or ROIs is limited. Land-metric DIY, landscapemetrics and PyLandStats are open-source extensible libraries that allow users to contribute in their development, to include new metrics and their functions can be integrated into advanced computational workflows. They require programming knowledge of the users in specific languages, but are more adaptable. The LecoS is a QGIS plug-in that computes a subset of metrics from FRAGSTATS in an intuitive manner for users. GuidosToolbox is a digital image analysis software with routines to quantify landscape patterns, extract fragmentation indices and conduct distance analyses from binary raster data.

Table 1.2. Description of the existing software and tools to compute spatial metrics. The (*) marks the software that were not available for downloading at the current moment. ROIs stands for Regions of Interest, when the different levels of computation (object, class and landscape metrics) can be applied to several regions delimited by a second raster or vector.

Name	Levels of analysis	Input data format	Output data format	Dependency	Main application	Reference
FRAGSTATS v4	Patch, class, landscape	Raster	Tables	Standalone application	Landscape ecology	(McGarigal et al., 2012)
V-LATE	Class, landscape	Vector	Tables	ArcGis plug-in	Landscape ecology	(Lang and Tiede, 2003)
r.li	landscape, moving window	Raster	Raster, table	GRASS plug-in	Landscape ecology	(Porta and Spano, 2008)
Patch Analyst 5.1*	Class, landscape, ROIs	Raster/ Vector	Tables	ArcGis plug-in	Landscape ecology	(Rempel et al., 2012)
Land-metric DIY	Class, landscape, ROIs	Vector	Tables	standalone application	Landscape ecology	(Zaragozí et al., 2012)
PolyFrag	Patch, class, landscape	Vector	Shapefiles, table	ArcGis plug-in	Landscape ecology	(MacLean and Congalton, 2013)
LecoS	Class, landscape, ROIs	Raster	ROI shapefile, table	QGIS plug-in	Landscape ecology	(Jung, 2016)
USM	Urban, ROIs	Raster	ROI shapefile	Standalone application	Urban sprawl	(Nazarnia et al., 2016)
GuidosToolbox	Landscape	Raster	Tables, maps	Standalone application	Landscape ecology	(Vogt and Riitters, 2017)
ZonalMetrics	Class, ROIs	Vector	ROI shapefile	ArcGis plug-in	Landscape ecology	(Adamczyk and Tiede, 2017)
r.pi	Patch, class	Raster	Raster, tables, matrices	GRASS plug-in	Landscape ecology	(Wegmann, et al., 2018)
Arc_LIND*	Patch, class, landscape	Vector	Shapefile, table	ArcGis plug-in	Landscape ecology	(Yu et al., 2019)
landscapemetrics	Patch, class, landscape	Raster	Tables	R package	Landscape ecology	(Hesselbarth, et al., 2019)
Cloud-based landscape metrics*	Patch, class, landscape, ROIs	Raster	Tables, maps, plots	Amazon cloud, ArcGis and FRAGSTATS	Landscape ecology	(Deng et al., 2019)
PyLandStats	Patch, class, landscape	Raster	Tables	Python library	Landscape ecology	(Bosch, 2019)

From the available tools, only one computes spatial metrics focused on urban environments, the Urban Sprawl Metrics (USM) (Nazarnia et al., 2016), although some tools have been successfully applied to quantify urban sprawl, such as the GuidosToolbox (Sharaf et al., 2018). The USM automatizes the extraction of a few spatial metrics used to obtain the Weighted Urban Proliferation (WUP) indicator presented by Jaeger and Schwick (2014) and used to quantify urban sprawl in Europe (EEA, 2016a). The USM quantifies the degree of urban sprawl combining the size of the built-up areas, their spatial configuration and the land uptake. However, the tool has two main limitations. First, it does not work with categorical data and relies simply on a built-up binary mask; therefore, the interaction of urban sprawl with other land uses is disregarded. Second, it has very few metrics that are then combined to compute a sprawl indicator, limiting the applicability to other urban analyses.

Overall, some limitations from the tools presented in Table 1.2, especially regarding urban applications are: (i) All the available tools are mono-temporal, which means that they only compute spatial metrics for a given time, none of the tools includes multi-temporal metrics. Though PyLandStats has a spatio-temporal module, it basically computes spatial metrics on LULC maps from several dates; therefore, using these tools multi-temporal analyses can only be conducted comparing the change from spatial metrics over time; (ii) only USM offers specific metrics for urban applications, but the number of metrics is scarce. Besides, any of the tools includes metrics specifically dealing with road networks, commonly applied in urban studies (e.g., Venerandi et al., 2018); additionally, (iii) the outputs are limited, since the majority provide the results in a tabular format, which is useful for further statistical analyses but it constraints the creation of maps based on the results per ROIs, which is helpful in cross-comparative analyses; and (iv) only the cloud-based landscape metrics tool proposes the creation of plots to analyze the values of the metrics; however, to the author's knowledge, this tool is not available for users yet. The rest of the tools do not generate plots and graphs as outputs, which are useful and intuitive to interpret spatial patterns in urban areas and LULC changes.

As shown in Table 1.2, the number of available tools to compute spatial metrics has increased in the last few years. Most them were developed for landscape ecology applications and their use can be limited for monitoring urban environments, especially for cross-comparative and multi-temporal analyses with high-resolution categorical data. Consequently, besides the ones included in the available tools, some scholars opted to compute additional spatio-temporal metrics more suitable for urban studies, by means of spatial operations within GIS software (e.g.: Arribas-Bel et al., 2011; Schwarz, 2010;

Goerlich et al., 2017; Gielen et al., 2018). Therefore, the creation of a tool which embraces several metrics suitable for urban studies, includes multi-temporal metrics, deals with vector data, applies the metrics to different regions in the same landscape, and generates maps and plots to describe in an intuitive way the spatio-temporal patterns, would be beneficial to help researches and planners in the quantitative analysis of urban environments.

1.2. Research justification

The management of urban areas has become one of the most important development challenges of this century for its impact on the global sustainability. Consequently, there is a need to generate more robust scientific knowledge that allows improving the understanding of urbanization processes and their consequences in order to orientate urban growth towards more sustainable trajectories. The complexity of urban systems requires multi-dimensional approaches that should be tackled at different levels and across regions, nevertheless, multi-level and cross-comparative empirical work are still sparse (Seto et al., 2017; Wentz et al., 2018). Despite the growing availability and development of new geo-information and databases in urban environments, there is still a massive lack of knowledge on urban issues across the globe (Taubenböck, 2019).

The development of urban areas is conditioned by local and regional diverse processes, but it also shares drivers and consequences. A global view is necessary to identify these relationships (Esch et al., 2018). EO-based geo-information provides a current and comprehensive image of the built environment and its change over time. It has become a powerful source of data for urban monitoring studies, especially when combined with ancillary data such as socio-economic variables. The combination of spatial data on urban areas with demographic, economic, social and environmental factors allows unravelling the linkages that exist within these processes on different levels, and over time (Abrantes et al., 2019; Venerandi et al., 2018; Schwarz, 2010). However, methods that quantify such relationships depend on the availability of spatial and socio-economic data at different levels. On the one hand, more and better comparative and consistent data collections are necessary to advance in the understanding of urbanization (Lobo et al., 2020). On the other hand, the development of systematic and quantitative methods able to identify relationships between multiple aspects in urban areas is required to provide conclusions supported by evidence.

The use of spatio-temporal metrics to quantify physical characteristic in urban environments such as urban form, size, growth patterns, and LULC changes is

widespread (Reis et al., 2015). Nevertheless, there is a vast number of spatial metrics, which are often correlated since they represent similar spatial attributes, scholars are proposing more metrics, and there is not an agreement on the best subset of metrics that best quantifies certain aspects of urban form and growth patterns. Spatio-temporal metrics are able to capture the spatial attributes of urban areas and their change, therefore, identifying subsets of relevant metrics describing urban form and growth processes will provide quantitative measures from EO-base data that can be used further to identify relationships to the sustainability of these areas, for example, to environmental degradation, economic growth, or social inequality. Moreover, the relationship between the spatial configurations of urban areas to socio-economic factors might provide evidences of two-sided linkages between different patterns of urbanization and economic development. These linkages, in addition to improve the understanding of urban areas and their processes, will eventually help to create indicators extracted from EO-based data that could be used as proxies of socio-economic parameters.

In this context, the development of methods to efficiently characterize and monitor urban areas with geospatial data will allow the analysis of urban areas on a massive scale, their comparison across different levels, regions and over time. This will be used to evaluate development policies, as well as exploring the relationships with other factors, such as geographical, political, social, economic, and environmental, thus improving the understanding of urban environments and serving as basis for decision making tasks, aiming at planning better future cities.

1.3. Thesis outline

This document is organized in seven chapters. First, this introductory chapter describes the state of the art, introducing the research problem and its background. Then, the structure of the document is outlined. Second, the hypotheses and objectives of the dissertation are presented in chapter 2. Chapters 3 to 6 are edited versions of international scientific publications and compose the core of this dissertation. These chapters are complemented with methods and analyses published in national journals and conference archives. Finally, chapter 7 presents the general conclusions of this research. Appendices supporting the chapters outcomes and bibliographic references are at the end of the document.

Figure 1.1 illustrates the structure of the document. It summarizes for each chapter the datasets used, the level and scale of analysis, the methodological approach, and the main findings. The research problem and state of the art in
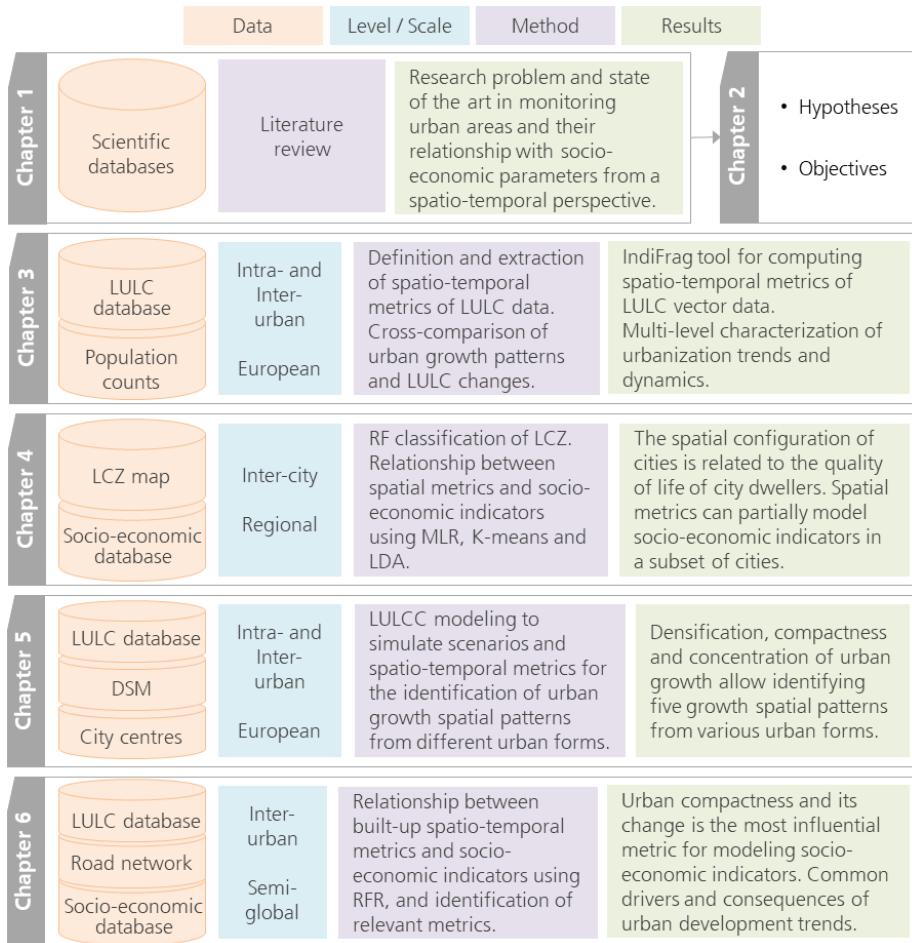
Figure 1.1. Document structure. For each chapter the employed datasets, level and scale of analysis, method, and main results are summarized. The abbreviations used are: Land-use/land-cover (LULC), local climate zones (LCZ), multiple linear regression (MLR), linear discriminant analysis (LDA), digital surface model (DSM), land-use/land-cover change (LULCC), and random forest regression (RFR).

chapter 1 are used as a basis for defining the main hypotheses and objectives and as a general background for each chapter.

Chapter 3 presents a methodology based on the use of spatio-temporal metrics from LULC data and a new index that quantifies the inequality of growth between population and the urban layout, to analyze and compare urban growth patterns at different levels. The analysis was carried out in a testing sample of six urban areas from the Urban Atlas database. Statistical

methods were applied for data reduction and optimization, and the results were interpreted at various levels (inter-urban and intra-urban). The index proposed complements the spatial analysis by including demographic dynamics, being also useful for assessing the growing imbalance between the progression on residential areas and the population change at the local level. The analysis at various levels contributes to a better understanding of urban growth patterns.

Additionally, in this chapter, the computing tool IndiFrag is introduced. This tool was created in a first stage of this thesis to ease and automatize the extraction of spatio-temporal metrics from LULC maps of urban environments used in this thesis. The tool is registered in the Patents and Software Catalog of the Universitat Politècnica de València since 2016, under the reference code R-18020-2016 (available from: https://aplicat.upv.es/exploraupv/ficha-tecnologia/patente_software/15271). The formulas and description of the spatio-temporal metrics are explained in the Appendix A, while a detailed description and case study are published in Sapena and Ruiz (2015). The tool is free downloadable at http://cgat.webs.upv.es/software/ and currently it counts with 118 downloads from 32 countries.

Chapter 4 studies the two-sided relationship between the urban spatial structure of cities and their socio-economic levels. Spatial metrics were derived from a Local Climate Zone (LCZ) classification in North Rhine-Westphalia, Germany. A reduced subset of uncorrelated metrics was used to quantify the relationships between spatial patterns and socio-economic variables in a sample of 31 cities by means of multiple linear regression models. The variability of 'education', 'health', 'living conditions', 'labor', and 'transport' variables was partially explained by means of spatial metrics from LCZ data. Cities were also grouped according to their level of quality of life and similarities in the spatial organization of urban structural types within groups were evidenced. The methodology proposed in this chapter is transferable to other datasets, levels, and regions. The findings support the hypothesis that cities with a similar urban structure will feature similar socio-economic characteristics, always considering other geographical and cultural factors.

This chapter is the result of collaboration with the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt, DLR), which started as part of an ERASMUS + grant (10/2018-01/2019), financed by the European Union, that promotes internationalization and student mobility. The candidate was a guest scientist for 3 months at the Earth Observation Center (EOC) in the Geo-Risks and Civil Security Department, supervised by Dr. Hannes Taubenböck and mentored by Dr. Michael Wurm within the Research Team City and Society.

Chapter 5 assesses the potential use of spatio-temporal metrics from multi-temporal LULC maps to identify urban growth spatial patterns in developing urban areas. Land-use/land-cover change (LULCC) models were used to simulate long-term growth and different scenarios of urban growth spatial patterns (i.e. expansion, compact, dispersed, road-based and leapfrog) on various baseline urban forms (i.e. monocentric, polycentric, sprawl, and linear). By means of discriminant analysis and clustering methods of the spatio-temporal metrics, two metrics were identified to better differentiate growth patterns despite the influence of the baseline urban form. These metrics account for the densification, compactness and concentration of urban growth, based on the chapter findings they could be used to identify growth patterns for monitoring and evaluating the management of developing urban areas towards a sustainable development.

Chapter 6 presents a methodology to explore the relationships between the urban structure and growth patterns with socio-economic indicators from a multi-temporal perspective in a semi-global scale using worldwide available datasets. Open multi-source and multi-temporal datasets are leveraged from six hundred functional urban areas in 32 countries, for 2000 and 2014. Spatio-temporal metrics are derived from the built-up areas, land covers and road networks. Then, relationships between metrics and socio-economic indicators and their change are measured by means of regression models. The analysis allows identifying the spatio-temporal metrics that are most significant for modeling 'income', 'GDP', 'Gini', 'air quality', and 'employment'. This is a first and fundamental step that shows the high potential of open datasets to identify global development trends in urban areas and their socio-economic and environmental relationships from a multi-temporal perspective.

This chapter is the result of a second research stay and continued collaboration with the Team City and Society from DLR. This second stay aimed at conducting a multi-temporal analysis in urban areas. It was defined based on the methods and findings from previous chapters, so that it encompasses everything learned from the comparative spatial analysis in urban areas, relationships between urban structure and quality of life, and the identification of urban forms and growth patterns.

1.4. Publications

This document is a compilation of the edited version of international scientific publications produced while developing this thesis. The publications are listed below with the approval of all co-authors. This compilation satisfies the requirements of the PhD in Geomatics Engineering, an interuniversity program

between Universidad Politécnica de Madrid and Universitat Politècnica de València, Spain.

Chapter 3:

- **Sapena, M.**, Ruiz, L.A., 2019. Analysis of land use/land cover spatio-temporal metrics and population dynamics for urban growth characterization. *Computers, Environment and Urban Systems*, 73, 27-39. Doi:10.1016/j.compenvurbsys.2018.08.001 (IF 2019: 4.655).

Chapter 4:

- **Sapena, M.**, Wurm, M., Taubenböck, H., Tuia, D., Ruiz, L.A., 2021. Estimating quality of life dimensions from urban spatial pattern metrics. *Computers, Environment and Urban Systems*, 85, 101549. Doi:10.1016/j.compenvurbsys.2020.101549 (IF 2019: 4.655).

Chapter 5:

- **Sapena, M.**, Ruiz, L.A., 2020 (on-line). Identifying urban growth patterns through land-use/land-cover spatio-temporal metrics: Simulation and analysis. *International Journal of Geographical Information Science*. Doi:10.1080/13658816.2020.1817463 (IF 2019: 3.733).

Chapter 6:

- **Sapena, M.**, Ruiz, L.A., Taubenböck, H., 2020. Analyzing links between spatio-temporal metrics of built-up areas and socio-economic indicators on a semi-global scale. *ISPRS International Journal of Geo-information*, 9(7), 436. Doi:10.3390/ijgi9070436 (IF 2019: 2.239).

Regarding the copy rights for scholarly purposes, chapters 3 to 6 are edited versions of the abovementioned scientific papers, with full acknowledgement of the original publications and without any kind of commercial use. Other publications produced during this research that complement this dissertation:

- **Sapena, M.**, Ruiz, L.A., 2015. Description and extraction of urban fragmentation indices: The Indifrag tool. *Revista de Teledetección*, 43, 77-90. Doi:10.4995/raet.2015.3476.

- **Sapena, M.**, Ruiz, L.A. 2015. Analysis of urban development by means of multi-temporal fragmentation metrics from LULC data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-7/W3, 1411-1418. Doi:10.5194/isprsarchives-XL-7-W3-1411-2015.

- **Sapena, M.**, Ruiz, L.A., Palomar, J.M., 2015. Estudio evolutivo de los usos del suelo urbano mediante índices de distribución espacio-temporal. *Actas del XVI Congreso Nacional de la Asociación Española de Teledetección*, 21-23 October, Sevilla, Spain.

- **Sapena, M.**, Ruiz, L.A., Goerlich, F.J., 2016. Analysing relationships between urban land use fragmentation metrics and socio-economic variables. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B8, 1029-1036. Doi:10.5194/isprs-archives-XLI-B8-1029-2016.

- **Sapena, M.**, Ruiz, L.A., 2017. Aplicaciones de los índices de fragmentación de los usos del suelo para caracterizar la expansión urbana. I *Congreso en Ingeniería Geomática*, 136-143, 5-6 July, València, Spain. Doi:10.4995/CIGeo2017.2017.6621.

- **Sapena, M.**, Ruiz, L.A., Joó, K., 2017. Modelos de simulación de expansión urbana a partir de imágenes de satélite: Adecuación al análisis temporal de la fragmentación de los usos del suelo. *Nuevas plataformas y sensores de teledetección. XVII Congreso de la Asociación Española de Teledetección*, 311-314, 3-7 October, Murcia, Spain.

- **Sapena, M.**, Ruiz, L.A., 2018. Caracterización de los patrones espaciales del crecimiento urbano aplicando índices espacio-temporales de los usos del suelo. *Perspectivas multidisciplinares en la sociedad del conocimiento. XVIII Congreso de Tecnologías de la Información Geográfica*, 580-590, 20-22 June, València, Spain.

Chapter 2

Hypotheses and objectives

Urban areas are growing at unprecedented rates causing significant socio-economic and environmental impacts on developing and developed urban environments. In any case, the sustainable development of urban areas is based on the monitoring and analysis of their growth by urban planners and decision-makers. As described in the introductory section, there is a link between the structure of urban areas, their spatial change, and the environmental, social and economic factors, affecting the well-being of urban dwellers. Investigating these underlying processes will help in disentangling and understanding these cause-effect relationships in order to reorient urban development patterns to more sustainable ones.

In the last decades the ever-increasing Earth observation programs have provided with large amount of satellite imagery, which in many cases has been systematically processed to produce a variety of geo-data and LULC databases that are essential for the multi-temporal and spatial study of urban areas at different levels, which is possible by means of GIS-based methods. The use of spatial metrics and their change over time have recently shown a high potential for monitoring urban form and growth spatial patterns. However, further steps must be done in this sense to advance in the development of new methods and tools to measure, quantify and monitor urban environments. This could be done by means of LULC datasets from a multi-temporal perspective and for different levels and regions—both from the physical and socio-economic points of view and also their interrelations—in order to improve the understanding of these highly dynamic spaces with the ultimate aim to facilitate their economic, social and environmental sustainability. This is of high relevance in the current political context, where global targets of sustainable development have been stablished and urban environments play a significant role in their implementation.

In this context, this research aims to develop effective tools and methods for urban monitoring and characterizing growth spatial patterns using data derived from Earth observation, as well as to study their relationships with socio-economic factors, providing new evidences about the utility of geo-data and LULC databases for urban analyses at different levels (i.e., from local to broad levels) by means of GIS and statistical methods. Thus, the scientific community, planners and land managers could benefit from the proposed methods and tools generated in this thesis to extract and quantify information in order to analyze dynamic urban environments from available geospatial databases using a spatio-temporal perspective. To this end, a series of specific hypotheses and objectives have been proposed and defined as follows:

Hypothesis 1: Spatial and temporal metrics obtained from LULC databases are suitable to capture and quantify the spatial configuration and structure of

urban environments and their structural changes. Population data combined with multi-temporal metrics may enhance the capabilities for the characterization of urban growth. The creation of a supporting tool to compute spatio-temporal metrics and interpretation outputs that could be applied on different levels may provide new insights for urban growth characterization.

> **Objective 1:** To review and compile available spatio-temporal metrics suitable for urban analyses and urban growth characterization, integrating them in a software tool allowing us to extract these metrics from LULC databases at spatial and temporal levels and to create graphical and analytical tools as outputs, applying and evaluating them for the analysis of urban growth at intra- and inter-urban levels.

> **Objective 2:** To propose a new index that integrates multi-temporal LULC and population data and assess its suitability for the characterization of urban growth patterns in combination with spatio-temporal metrics at intra- and inter-urban levels, in a diverse sample of urban areas, allowing to interpret the relation of these growth patterns with sustainability policies.

Hypothesis 2: Urban areas with similar physical appearance are more likely to feature similar demographic, socio-economic and environmental characteristics. The spatial organization of urban structural types, as defined by the Local Climate Zones (LCZ) in urban environments, may reflect the distribution of activities and infrastructure, as well as the social and economic characteristics of their inhabitants, which influence the quality of life and well-being of population.

> **Objective 3:** To identify and interpret the two-sided relationships between the spatial structure of urban areas and the socio-economic factors of their inhabitants, exploring the suitability of the structural information implicit in the Local Climate Zone (LCZ) classification framework to quantify urban spatial patterns at the inter-city level, and to identify groups of cities according to their similarities in terms of quality of life as defined by socio-economic indicators.

Hypothesis 3: The identification of growth spatial patterns in urban areas improves the knowledge of urban planners for designing new policies and applying sustainable development measures. Urban structure and its evolution can be characterized to a certain level by a selected subset of significant spatio-temporal metrics. Their identification could be used to effectively categorize standard growth spatial patterns in urban areas with different urban forms and geographical contexts.

Objective 4: To compile and describe the most standard urban forms and growth spatial patterns, as defined in the literature, to evaluate the use of spatio-temporal metrics derived from LULC databases and alternative scenarios created from land use change simulation models for the categorization of these urban growth spatial patterns, identifying the most significant spatio-temporal metrics for this task and analyzing the influence of baseline urban forms in the identification of such patterns.

Hypothesis 4: There is a global two-sided relationship between socio-economic and ecological factors in urban environments and the urban structure from a spatio-temporal perspective. Therefore, demographic, economic or social changes imply an alteration in the spatial structure of cities and urban areas. Similarly, the transformation of the urban structure alters demographic distribution, has an economic, social and environmental impact. Hence, monitoring the development of urban areas, analyzing their patterns, and exploring their relationships with socio-economic factors at a broad scale could provide more empirical evidence-based information and improve the understanding of urbanization processes from physical and socio-economic points of view.

Objective 5: To quantify the relationships at a broad-scale level between socio-economic and environmental variables, such as income, inequality, GDP, air quality and employment, and spatio-temporal metrics extracted from geospatial databases, both on a specific date and in terms of their variation over time, identifying those spatio-temporal metrics that are most related to socio-economic and environmental variables and could be extracted from global geospatial databases, allowing to obtain preliminary conclusions on the use of spatial patterns and their development over time as proxies of socio-economic parameters at the global level.

Chapter 3

Spatio-temporal analysis of LULC and population in urban areas

3.1. Introduction

Quantifying urban growth and its characterization in different spatial patterns is crucial for evaluating its environmental, economic, and social impacts, since the degree of compact or sprawl growth differs both, in causes and in consequences (Bhatta, 2010). Urban growth can be categorized as sprawl or compact according to the spatial arrangement of built-up areas, land uptake per inhabitant and the amount of built-up area in the landscape (EEA, 2016a).

The characterization of the spatial configuration and change patterns of LULC is based on methods that allow for multi-temporal assessment. Spatio-temporal metrics contribute to characterize the urban growth process (Herold et al., 2005; Uuemaa et al., 2013). However, spatio-temporal metrics do not account for the land uptake per inhabitant, which has been mentioned as a relevant variable to characterize the growth process. A joint analysis of urban growth and population distribution provides an overview of the human use of the landscape and its tendency to sprawl (EEA, 2016a; Martinuzzi et al., 2007). Recent studies have combined spatial metrics with population data to categorize urban patterns. For instance, Arribas-Bel et al. (2011) used population density and distribution indices for an inter-city comparison and combined them with spatial metrics for clustering European cities according to their level of sprawl at a single date. Jaeger and Schwick (2014) introduced a metric that integrates urban expansion, dispersion, and the land uptake per inhabitant at intra-city level for a single date. Afterward, it was applied to the built-up area in Europe at various levels: national, regional and 1-km^2-grid (EEA, 2016a; Hennig et al., 2015). They found that the application at local scale eased the detection of changes; however, it was hardly comparable with socio-economic data at this level. Other studies revealed that population density combined with other drivers (i.e. spatial characteristics, socio-economic, policies, among others) is suitable for predicting urban growth and its type (Dubovyk et al., 2011; EEA, 2016a).

Besides the potential of their combined study, several studies have pointed out the large inequality between the growth pace of built-up areas and population in Europe. Kasanko et al. (2006) analyzed the difference between built-up and population growth rates from the fifties to the nineties at inter-city level, and built-up grew faster in almost all of the 15 cities studied, presenting different growth patterns according to their geographical location. However, they did not propose a way to quantify this inequality. More recent studies obtained similar conclusions studying samples of 29 (Ribeiro-Barranco et al., 2014) and 188 European cities (Haase et al., 2013). They observed that even when population decreased built-up change was positive. This mainly occurred in Southern cities where a faster built-up growth was experienced in the studied

periods, while lower rates were found in Eastern cities. However, these results obtained at broader scale (city level) cannot be assumed at local level (intra-city level). The dynamics of urban areas are not homogeneous and they should be quantified independently to characterize the inherent heterogeneity of urban areas, but interpreted and analyzed together at various levels to obtain more accurate conclusions.

Analysis at multiple levels is essential for different reasons. On the one hand, the analysis at broad level shows an overall value of the actual trends, while detailed levels are more informative (EEA, 2016a). On the other hand, policies are applied at national, regional and local levels causing different growth trends (DG REGIO, 2011). Previous studies have reported that the degree of compactness or sprawl of the urban layout and its interpretation differs widely depending on the levels employed (Altieri et al., 2014; Hennig et al., 2015). The imbalanced development of population and built-up areas previously detected in European cities may also vary if analyzed at various levels. A concurrent multi-level analysis of the population and urban growth rates combined with LULC spatio-temporal metrics may help to the characterization of the urban growth process, moreover, the use of several land uses and metrics would be useful for the selection of the most suitable ones to identify growing patterns. In this framework, the main objectives of this chapter are: (1) To integrate an exhaustive set of spatio-temporal metrics reviewed in the literature in a software tool adding visualization and interpretation outputs, including new metrics focused on urban and population spatio-temporal characterization and assessing their added value for interpretation of urban growth; and (2) to present a methodology based on spatio-temporal metrics that allows the analysis and comparison of urban growth at inter-city and intra-city levels and to interpret its relation with urban sustainability policies.

3.2. Data

3.2.1. Description of datasets

The study was performed using the Urban Atlas database, which is part of the local component of the Copernicus Land Monitoring Services (EEA, 2016b). It provides harmonized, inter-comparable and high-resolution LULC vector maps (scale 1: 10,000) from 319 Functional Urban Areas (FUAs) with more than 100,000 inhabitants for the year 2006 (UA2006), and 785 FUAs above 50,000 inhabitants for 2012 (UA2012) as of September 2020. Currently, the year 2018 is under development and non-validated products are partially available. The minimum mapping unit is 0.25 ha for urban and 1 ha for rural areas, and the minimum overall accuracy is 85% in urban and 80% in rural areas. The

OECD and the European Commission have jointly developed a harmonized definition of FUAs in a consistent way across countries, it represents the city and its commuting zone, with a population greater than 50,000 (Poelman and Dijkstra, 2015). FUAs represent the economic and functional spatial extent of the city (using population density and travel-to-work flow data).

Since our purpose was to assess a methodology rather than the in-depth analysis of specific urban areas, a sample testing dataset composed of six FUAs was selected attending to the following criteria: the availability of population data and administrative unit boundary datasets to calculate the metrics, the existence of high LULC change to test temporal indices, and the geographical diversity to cope with different urbanization contexts. As a result, the FUAs selected were Berlin, Paris, Rome, Krakow, Lisbon and Valencia (Figure 3.1).
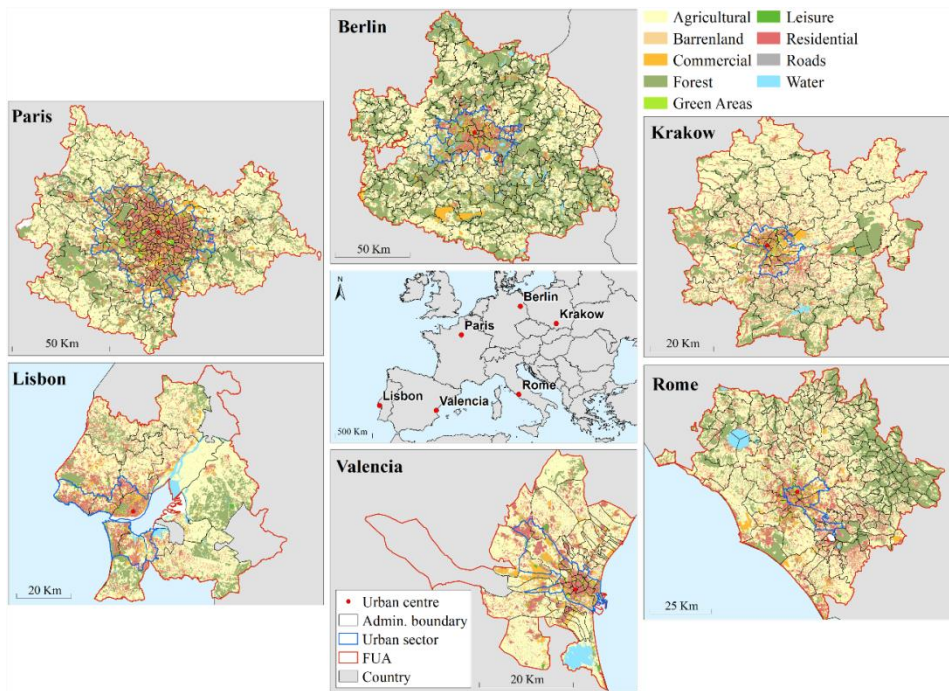
Figure 3.1. Testing sample areas. Location in Europe (center); UA2012 maps, FUA, urban sector and administrative unit boundaries (municipalities or equivalent local administrative units and city districts), and urban centers.

The UA2006 was initially focused on urban and peri-urban areas represented by twenty classes, seventeen urban and three rural. The UA2012 was extended from three to ten rural classes to allow for a better understanding of the urban fringe. This led us to a legend adaptation before comparing UA2006 and

UA2012, harmonizing and simplifying the legend for our urban analysis purpose. We reclassified the legend to nine aggregated land use classes following the criteria of class similarity, thematic coherence and simplification of processing and interpretation tasks (Table 3.1).

Table 3.1. Legend adaptation. Reclassification of the Urban Atlas for 2006 (UA2006) and 2012 (UA2012) legends into nine aggregated land use classes, where SL means sealing level. Deeper insight into the original thematic classes can be found in Copernicus (2016).

Decision rules		UA2006	UA2012	Legend
Artificial surfaces	Urban fabric	Continuous (SL > 80%) Discontinuous (SL 50-80%) Discontinuous (SL 30-50%) Discontinuous (SL 10-30%) Discontinuous (SL < 10%) Isolated structures	Continuous (SL > 80%) Discontinuous (SL 50-80%) Discontinuous (SL 30-50%) Discontinuous (SL 10-30%) Discontinuous (SL < 10%) Isolated structures	Residential
	Industrial, commercial, public, military, private and transport	Industrial, commercial, etc. Port areas Airports	Industrial, commercial, etc. Port areas Airports	Commercial
		Fast transit roads Other roads Railways	Fast transit roads Other roads Railways	Roads
	Mine, dumps, construction sites	Mine and dumps Construction sites Land without current use	Mine and dumps Construction sites Land without current use	Barren land
	Non-agricultural vegetation areas	Green urban areas Sport and leisure facilities	Green urban areas Sport and leisure facilities	Green areas Leisure areas
Little/ no human influence	Agricultural	Agricultural, semi-natural and wetland areas	Arable land (annual crops) Permanent crops Pastures Mixed cultivation patterns Orchards Wetlands	Agricultural
	Natural/semi-natural areas		Herbaceous vegetation associations Open spaces with little vegetation	
		Forest	Forest	Forest
Water		Water	Water	Water

Since disparities in urbanization trends within FUAs and cities are expected, according to the European Environment Agency (EEA) (2016) report, more detailed levels were also considered in our analysis. Thus, the FUA level was subdivided into Local Administrative Units (LAU), dividing the territory into municipalities or equivalent units. According to Salvati and De Rosa (2014), this territorial unit is relevant for the purpose of planning and statistical analyses at local level. Cities were also subdivided into districts (SCD), which are zones defined according to population criteria (EU, 2016). Both levels are referred henceforth to as administrative units. Administrative unit boundaries were obtained from official institutions, as well as population data from 2006 and 2012 (Table 3.2).

Table 3.2. Data sources. Administrative units and population data for 2006 and 2012 of the sample FUAs.

FUA	Administrative units	Population (2006/2012)
Berlin	Federal Agency for Cartography and Geodesy (BKG) Dienstleistungszentrum (www.geodatenzentrum.de/geodaten/gdz_rahmen.gdz_div)	Statistical Office for Berlin-Brandenburg (*www.statistik-berlin-brandenburg.de*)
Paris	National Institute of Geographic and Forest Information (IGN) (*http://professionnels.ign.fr/geofla#tab-3*)	National Institute of Statistics and Economic Studies (INSEE) (*www.insee.fr/fr/statistiques*)
Rome	LAU: Italian National Institute of Statistics (ISTAT) (*www.istat.it/it/archivio/104317*) SCD: IPTSAT S.R.L. (www.iptsat.com/index.php/it/download)	LAU: Italian National Institute of Statistics (ISTAT) (*http://demo.istat.it/*) SCD: Roma Capitale open data (https://dati.comune.roma.it/)
Krakow	LAU: GIS Support Sp. z o.o. (*http://gis-support.pl/*) SCD: delineated using the Municipal Spatial Information System of Krakow (*http://krakow.pl/plan*)	LAU: Local Data Bank (*https://bdl.stat.gov.pl*) SCD: Public Information Bulletin (*www.bip.krakow.pl*)
Lisbon	General Direction of the Territory (DGTerritório) (www.dgterritorio.pt/cartografia_e_geodesia/cartografia/)	Statistics Portugal (INE) (*www.ine.pt/*)
Valencia	National Institute of Statistics (INE) (*www.ine.es/*)	National Institute of Statistics (INE)

Since different growth patterns are expected in urban and peri-urban areas the FUA level was further subdivided into sub-areas or sectors: (i) Urban, and (ii) peri-urban areas, defined as those areas around urban settlements which blend into the rural landscape, where usually low-density urban growth is present (EC, 2012). These sectors were delimited following a dominant land use density criteria in the administrative units of classes forest, agricultural and urban (artificial surfaces, in Table 3.1). Thus, the urban sector corresponds to those areas where the urban density overpasses agricultural and forest densities, and the peri-urban sector comprises the rest.

3.3. Methods

The methodology consisted on the compilation and implementation of spatio-temporal metrics in a software tool and the development of a new index that quantifies the inequality of growth between population and urban areas. Then the metrics were computed at local level in a testing sample of six urban areas, uncorrelated metrics were selected and the data were interpreted to analyze and compare urban growth patterns at different levels.

3.3.1. Creation of IndiFrag: An object-based spatio-temporal metrics extraction tool

As introduced in chapter 1, different software and tools are currently available to compute spatial metrics (Table 1.2). However, they are unsuitable for a multi-level, multi-temporal, and high-resolution analysis with LULC vector data for multiple regions. In order to be able to consistently use and extract spatio-temporal metrics needed to accomplish the different objectives of this thesis, and after an exhaustive literature review, a set of spatio-temporal metrics focused on urban environments was compiled in a software tool (IndiFrag) to automatize the spatial and GIS operations required to compute the metrics. Besides, this tool was designed to provide visualization outputs and plots to represent and support the interpretation of the results of some metrics.

Thus, IndiFrag is a processing tool created for the extraction of a set of spatio-temporal indices and metrics to quantitatively describe the level of fragmentation and spatial distribution of LULC in response to morphological, spatial and typological properties of cartographic objects. The tool also quantifies spatial changes over a period of time using a particular set of multi-temporal metrics. It works with categorical maps in vector format. Thus, topological relationships are used, easing the interpretation and analysis of the spatio-temporal distribution of LULC data in different areas and regions of a particular area of study (Sapena and Ruiz, 2015a, 2015b).

The tool consists on geoprocessing scripts in Python 2.7 and it was implemented as a ToolBox within the software ArcGis (a new version for Python 3 is under development, which will be compatible with ArcGis Pro). It calculates the indices and metrics according to different levels of hierarchy and to the type of attributes or properties they describe (Sapena and Ruiz, 2015a). The three hierarchical levels based on the calculation scale are: super-object (describes relationships between objects from various classes in a broader context, e.g. territorial unit), class (describes relations between a set of objects from the same class, e.g. land use) and object (describes characteristics of each object, e.g. plot). Besides, metrics were divided into five semantic groups

based on the described spatial attributes: area and perimeter, shape, aggregation, diversity and contrast. Both characteristics are illustrated in Figure 3.2.

In view of the foregoing, IndiFrag was conceived for comparative urban studies as it allows working with different LULC independently (class level) and for each administrative unit (super-object level) in the same process. An extensive description of the tool and a complete list of the spatio-temporal metrics currently included in IndiFrag with their formula, description, units, level and reference is included in the Appendix A, and it is summarized in Table A.1.
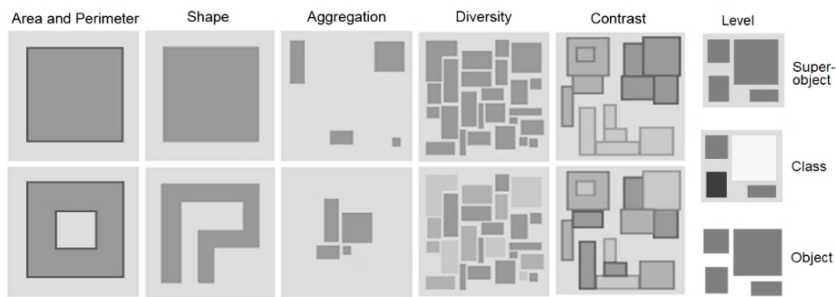
Figure 3.2. Graphical representation of semantic groups of metrics quantifying different spatial attributes in IndiFrag (i.e. area and perimeter, shape, aggregation, diversity and contrast). Metrics can be computed at different hierarchical levels (i.e. super-object, class, and object levels).

3.3.2. Extraction of land use spatio-temporal metrics

In order to analyze and compare LULC changes and to highlight growth patterns in FUAs, administrative units and land use classes, two types of metrics were calculated at the administrative unit level: those that consider all land uses within the administrative unit (administrative unit metrics), and metrics referred to one land use within an administrative unit (class metrics). Hence, we computed: (i) spatial metrics for two dates (years 2006 and 2012) and their derived changes, and (ii) multi-temporal metrics. As a result, a collection of spatio-temporal metrics was obtained for each administrative unit and class (Figure 3.3).

Duplicity and redundant information are usually present when working with such a large set of spatial metrics (Cushman et al., 2008), therefore a selection of metrics was applied to avoid redundancies and increase the efficiency of the process. We computed 167 single-date spatial metrics (23 per administrative unit, plus 18 per class, except for roads) and 248 two-date metrics (167 changes from the spatial metrics, plus one per administrative unit and 10
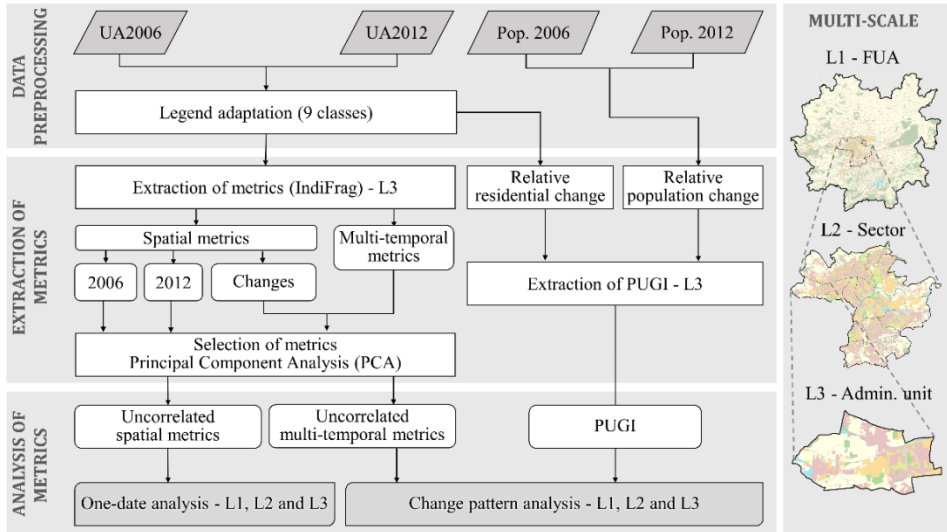
Figure 3.3. Workflow. Legend adaptation of Urban Atlas; population data and residential areas for 2006 and 2012 and their changes are extracted; spatial metrics for 2006 and 2012, their derived changes, multi-temporal metrics and PUGI index are computed at the administrative unit level; uncorrelated metrics are selected using PCA; One-date and change pattern analyses are interpreted at three levels: FUA (L1), sectors (L2) and administrative unit (L3).

multi-temporal metrics per class) for 833 administrative units. The objective selection of the most relevant metrics was achieved by applying the Principal Component Analysis (PCA) method using R statistical software (R Team Core, 2019). The selection of class metrics was divided into two processes according to the sector. In the urban sector analysis, we focused on the residential class for its particular interest, but also on the most dynamic classes in this sector: commercial and industrial, referred henceforth to as commercial, leisure and green urban areas. The peri-urban sector was focused on forest class and its modification in response to urban growth. Metrics at administrative unit level were included in both sectors.

PCA is a multivariate statistical method allowing for the transformation of a large number of correlated variables into uncorrelated variables (Jolliffe, 2002). Four different PCAs were performed: in urban and peri-urban sectors, and using single-date and two-date metrics. The indices were grouped according to the weights of the first and second components discarding those indices with similar weights in both components and preserving only one per group, ensuring non-correlation between the selected indices. Figure 3.4 shows the final subsets of indices selected for the analyses.
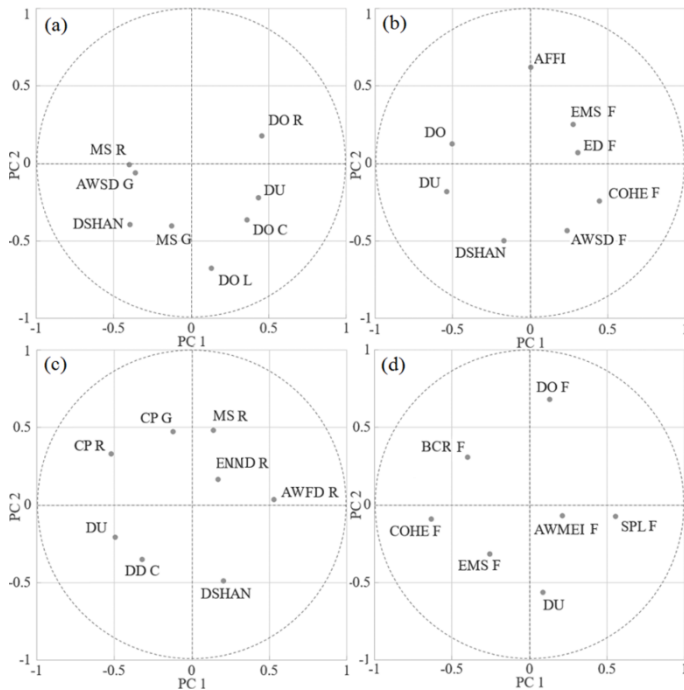
Figure 3.4. Graphs of spatial distribution of the final uncorrelated metrics selected in the space defined by the first and second principal component weights. Four independent PCAs, where: (a) Single-date metrics for urban and (b) peri-urban sectors, and two-date based metrics for (c) urban and (d) peri-urban sectors. See Table 3.1 for abbreviation meanings.

Table 3.3 shows and describes the final set of indices selected for analysis. The results per administrative unit can be found in the supplementary material available from: https://doi.org/10.1016/j.compenvurbs-ys.2018.08.001.

In order to compare overall results among FUAs, we conducted two sub-analyses. For inter-city analysis and once metrics were calculated for each administrative unit, we computed their mean and coefficient of variation for each FUA and sector (urban and peri-urban) within FUAs. This allows for the comparison of metrics and their homogeneity between different FUAs, which provides useful information when comparing values at broad levels. In addition, we used global growth graphs, concentric circle and sector analysis extracted also from IndiFrag. These graphs are useful to quantify changes and analyze their spatial distribution at different distances and orientations from a central point. We used central points defined by Urban Audit and based on GISCO settlement layer dataset (Data source: GISCO - Eurostat, European Commission).

Table 3.3. Description of the spatio-temporal metrics extracted from IndiFrag and selected using PCA. The name, abbreviation, description, time: single-date (1t) and two-date (2t), level of metric: administrative unit (LAU) or class: residential (R); green areas (G); commercial (C); leisure areas (L); and forest (F), and equation reference (Eq.) are reported. More detailed information about the metrics can be consulted in Appendix A.

Name	Definition	Time	Level	Eq.
Spatial metrics				
Urban density (DU)	Ratio between urban area and the total LAU area.	1t 2t	LAU LAU	(A.1)
Mean object size (MS)	Average of the size of the patches from a class.	1t 2t	G,R R	(A.4)
Edge density (ED)	Sum of lengths of patches from a class divided by its area.	1t	F	(A.5)
Area-weighted mean fractal dimension (AWFD)	Average of fractal dimension of patches in a class, weighted by patch's area.	2t	R	(A.13)
Object density (DO)	Number of patches divided by the area of the LAU.	1t 2t	LAU,C,L,R F	(A.17)
Area-weighted standard distance (AWSD)	Average of distances from patches to the centroid of the class.	1t	G, F	(A.19)
Euclidean nearest neighbor mean distance (ENND)	Average of the distances between nearest patches of a class	2t	R	(A.21)
Effective mesh size (EMS)	Size of patches dividing the LAU into n areas with the same degree of division.	1t 2t	F F	(A.22)
Cohesion (COHE)	Connectedness of the patches from a class. It increases as the class becomes more aggregated.	1t 2t	F F	(A.23)
Splitting index (SLP)	No. of patches dividing LAU into equal parts with the same division degree.	2t	F	(A.24)
Shannon diversity (DSHAN)	Minus the sum of proportional abundance of each class multiplied by its proportion.	1t 2t	LAU LAU	(A.31)
Density-diversity (DD)	Sum of the amount of a class as proportion of the largest class.	2t	C	(A.34)
Absolute functional fragmentation index (AFFI)	Ratio between the LAU and the sum of every class perimeter.	1t	LAU	(A.36)
Multi-temporal metrics				
Change proportion (CP)	Ratio between the change area of a class and the area of the LAU.	2t	G, R	(A.40)
Landscape expansion index (LEI)	Categorizes new patches in: infilling (≥50% adjacent to its class), edge-expansion (0>50%), and outlying (=0%) types by comparing perimeters between new and old patches.	2t	R,C,L	(A.41)
Area-weighted mean expansion index (AWMEI)	Sum across all new patches of the percentages of adjacencies weighted by the area of the new patch.	2t	F	(A.43)
Change rate (RC)	Annual rate of class change using the compound interest formula.	2t	Forest	(A.44)

3.3.3. Population and urban growing imbalance index (PUGI)

Inequality of urban dynamics regarding the increase of built-up area with respect to population is related to the type of evolution experimented by urban areas over time and it can be especially relevant to monitor the sustainability of urban development (Ribeiro-Barranco et al., 2014). In order to quantify how urban growth outpaces population increase or vice versa and based on the assumption that the distance of the population and urban growth rates—if they are plotted on two axes—to the line of equal growth is related to the imbalance of both rates (Kasanko et al., 2006), we propose a multi-temporal index for a better understanding of the balance in urban growing and population increase in urban dynamic areas: The Population and Urban Growing Imbalance index (PUGI). This index quantifies the inequality between two variables, population and residential land use relative growths extracted at two different dates. We used the area of residential land use, since this is more related and comparable to the actual increase of population, as suggested by Kasanko et al. (2006).

In order to define the index, the increase/decrease of population and the increase of residential area are converted to relative terms as relative change to the first year:

$$rcr = \frac{(r_{t2} - r_{t1})}{r_{t1}} \cdot 100 \qquad (3.1)$$

$$rcp = \frac{(p_{t2} - p_{t1})}{p_{t1}} \cdot 100 \qquad (3.2)$$

where, r_{t1} and r_{t2} represent the areas of residential class, and p_{t1} and p_{t2} the population at the beginning and end of the studied period.

Administrative units are plotted in a four-quadrant scatterplot with (3.1) and (3.2) in the axes (Figure 3.5). Similar scatterplots have been previously used to represent urban sprawl by plotting the compactness degree against urban proportion at a single date (Altieri et al., 2014), to analyze the relation between economic development and urban growth (Chen et al., 2013), to compare urbanization and population growth rates (Kasanko et al., 2006), and to classify the development of cities according to their position in the plot (Ribeiro-Barranco et al., 2014). Here, we propose the quantification of the mentioned distance as a measure of the disproportion between rates.

Having the proportion of population change in the abscissas, the proportion of residential increase in the ordinates, and considering the quadrants delineated
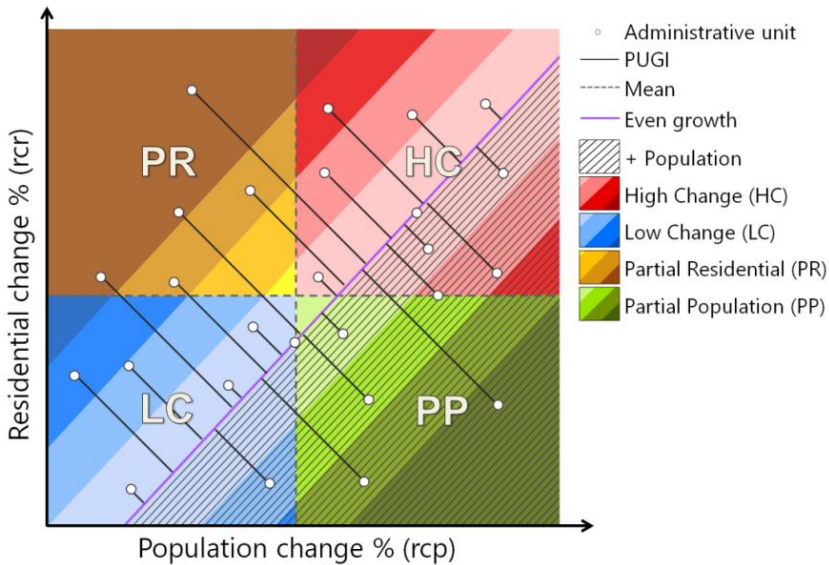
Figure 3.5. Example of four-quadrant scatterplot. Calculation of the minimum distance from a point to the even growth line (PUGI) and classification of administrative units according to the quadrant delineated by means: high change (HC), low change (LC), partial population (PP), and partial residential (PR). A gradient color is given according to the distance to the equal growth line, lighter colors show more balanced growth, while darker colors show more imbalanced growth. The stripe background shows when population grows faster than residential areas.

by the mean values of the two variables, administrative units can be classified into four groups according to the type of change experimented (Ribeiro-Barranco et al., 2014):

> The upper right quadrant indicates a high change (HC) in both variables.
> The lower left quadrant represents a more stable and low change (LC).
> The upper left quadrant corresponds to a high residential growth complemented by a low or negative population change (Partial residential change, PR).
> The lower right quadrant corresponds to a high increase in population followed by a low or null residential growth (Partial population change, PP).

The even growth line represents the same pace of growth rate in both variables, as an ideal or balanced development situation (Figure 3.5). Administrative units above this line have undergone faster growth of residential areas with respect to population, and in those below the line, the

population has exceeded residential growth. The farther the administrative unit is from this line, the larger the difference between the two growth rates. This magnitude is represented by the PUGI index, shown in equation (3.3), defined as the minimum distance between the location in this bi-variate space and the even growth line. It is computed as the Euclidean distance from a point to a line and measured along a perpendicular line to the even growth line (Figure 3.5). The sign of the index represents whether the administrative unit is located above or below the line. Thus, a negative value means that the point is below, and the population growth is higher than residential increase. A positive value indicates that the residential area grows faster than population. The administrative unit coordinates are: relative change of population (rcp as x-coordinate) and relative change of residential (rcr as y-coordinate). Considering that the equation of an even growth line is an identity function, and knowing the formula of the Euclidean distance from a point to a line, the PUGI index is obtained as:

$$PUGI = \frac{(rcr\text{-}rcp)}{\sqrt{2}} \tag{3.3}$$

3.4. Results

The results of metrics computed in 2012 and from 2006 to 2012 are interpreted at three levels (i.e. FUA, sector and administrative unit). First, we analyze them at FUA and sector levels, and then we focus on each FUA at administrative unit level.

3.4.1. Analysis at inter-city level

Attending to the spatial metrics from 2012 at FUA level, Paris and Valencia present the highest values of mean Urban density (DU) and the lowest coefficients of variation (CV), showing a compact and homogeneous spatial distribution of built-up areas (Table 3.4). By contrast, Berlin and Rome present lower mean values and the highest CV, showing a more heterogeneous distribution of urban density than the rest of the FUAs. However, focusing at sector level, the DU in the urban sector is consistently more uniform than in the peri-urban, which presents a higher CV and, as unlike at FUA level, Valencia doubles the density of Paris in the peri-urban sector and has lower CV, while in the urban sectors the values are quite similar.

Analyzing the mean values of Shannon diversity (DSHAN) at FUA level, Lisbon and Paris are significantly more diverse than the rest of the FUAs and present an even distribution (Table 3.4), while Rome presents low mean DSHAN and

Table 3.4. Examples of mean values and coefficients of variation (in parentheses) of some spatial metrics for 2012 at FUA and sector levels (urban and peri-urban), where: urban density (DU), Shannon diversity (DSHAN), density of commercial (DO$_{Commercial}$), object mean size of residential (MS$_{Residential}$), and effective mesh size index of forest (EMS$_{Forest}$).

	DU			DSHAN			DO Commercial		MS Resid.	EMS Forest
	FUA	Urban	Peri-urban	FUA	Urban	Peri-urban	FUA	Urban	Urban	Peri-urban
Berlin	0.16 (1.23)	0.67 (0.29)	0.1 (0.82)	0.98 (0.3)	1.43 (0.18)	0.93 (0.27)	1.12 (1.26)	4.38 (0.56)	2.13 (0.14)	1.51 (1.11)
Krakow	0.31 (0.92)	0.77 (0.23)	0.17 (0.53)	1.03 (0.34)	1.47 (0.1)	0.9 (0.32)	2.63 (1.29)	7.94 (0.43)	1.64 (0.16)	0.39 (1.83)
Lisbon	0.36 (0.69)	0.63 (0.16)	0.18 (0.5)	1.2 (0.24)	1.48 (0.11)	1.02 (0.17)	3.44 (0.64)	5.7 (0.2)	1 (0.12)	0.82 (1.94)
Paris	0.71 (0.44)	0.84 (0.22)	0.22 (0.7)	1.22 (0.26)	1.25 (0.26)	1.14 (0.23)	7.28 (0.68)	8.78 (0.51)	1.26 (0.29)	0.53 (1.2)
Rome	0.19 (0.94)	0.75 (0.27)	0.16 (0.79)	0.96 (0.25)	1.32 (0.09)	0.94 (0.24)	1.47 (1.38)	7.99 (0.55)	1.25 (0.31)	1.73 (1.67)
Valencia	0.56 (0.51)	0.82 (0.2)	0.39 (0.53)	1.03 (0.28)	1.25 (0.21)	0.89 (0.24)	10 (1.09)	10.7 (0.49)	0.64 (0.26)	0.002 (4.1)

CV values. In contrast, when analyzed at sector level, Rome is not the least diverse FUA. Instead, Valencia presents less diversity in both sectors, having an intermediate CV. Berlin and Krakow have similar responses in both sectors.

Class metrics show that Object density of commercial (DO$_{Commercial}$) is variable among FUAs. For instance, Valencia and Paris present high mean values and they are significantly denser than Berlin, Krakow and Rome. However, in the urban sector the differences and CV are much lower, showing uniformity in the distribution of commercial use, especially in Lisbon. Mean object size of residential (MS$_{Residential}$) in urban sectors shows significant differences in buildings size between Berlin and Valencia. Another example of discrepancies among FUAs is the mean values of the effective mesh size index of forest (EMS$_{Forest}$) in the peri-urban sector, with lower fragmentation values of forest in Berlin and Rome (larger patches and less fragmented) compared to Paris and Krakow, that present more fragmentation (Table 3.4).

According to the evolution of DU from 2006 to 2012 (Table 3.5), the FUAs of Krakow, Lisbon and Valencia are very dynamic and homogeneous in terms of built-up surface. Moreover, the population and urban growing imbalance index (PUGI) shows high positive values in Krakow and Lisbon, especially Lisbon in the peri-urban sector and Krakow in the urban sector, evidencing the rapid increase of residential areas with respect to the population growth, probably related with a sprawl development (Table 3.5). Valencia presents a more balanced development with a negative PUGI value at the FUA level, while

Table 3.5. Examples of mean values and coefficients of variation (in parentheses) of two spatio-temporal metrics and the PUGI index for the period 2006-2012 at FUA and sector levels (urban and peri-urban), where: urban density change (ΔDU), Shannon diversity change (ΔDSHAN), and population and urban growing imbalance (PUGI).

	ΔDU			ΔDSHAN			PUGI		
	FUA	**Urban**	**Peri-urban**	**FUA**	**Urban**	**Peri-urban**	**FUA**	**Urban**	**Peri-urban**
Berlin	0.002 (3.34)	0.004 (1.18)	0.002 (3.87)	0.003 (3.71)	-0.008 (1.58)	0.005 (2.5)	-0.982	-1.629	0.993
Krakow	0.01 (0.74)	0.011 (0.95)	0.01 (0.65)	0.013 (1.86)	-0.019 (0.98)	0.022 (0.7)	2.355	3.132	0.678
Lisbon	0.016 (0.72)	0.026 (0.3)	0.009 (0.88)	-0.003 (18.7)	-0.004 (9.33)	-0.002 (30)	3.331	1.050	5.763
Paris	0.003 (1.93)	0.003 (2.33)	0.005 (1.06)	0.001 (38.1)	-0.002 (13.6)	0.01 (1.37)	-1.115	-1.211	-1.883
Rome	0.007 (1.24)	0.005 (1.07)	0.007 (1.25)	0.016 (1.54)	-0.001 (8.28)	0.017 (1.48)	-1.728	0.263	-3.085
Valencia	0.012 (1.47)	0.012 (1.74)	0.013 (1.32)	0.004 (9.57)	-0.013 (2.42)	0.015 (2.86)	-0.697	1.099	-5.083

the peri-urban sector has a high negative PUGI value, evidencing a densification process in this sector. Berlin and Paris experimented less DU changes but with more spatial variability, accompanied by low and negative PUGI values, meaning that population grew slightly faster than residential land use. Berlin, where the variability of ΔDU is particularly high, increases its CV and has a positive PUGI value in the peri-urban sector. Rome presents an intermediate ΔDU and CV compared to the rest of the FUAs and sectors, with a global negative PUGI that is higher in the peri-urban sector, meaning higher inequality of growth in this sector. Regarding the changes in DSHAN (Table 3.5), all FUAs increase their diversity except Lisbon. Paris shows a low change in diversity, but this is heterogeneously distributed (high CV value) along the FUA. However, at sector level all, except Lisbon, present two different patterns: Urban areas reduce their diversity, whereas peri-urban interfaces increase it, showing a high variety of land uses with a homogeneous distribution in the peri-urban sectors.

The global growth graphs of the residential land use close to the city centers present a compact built-up area with permanent land use in some FUAs (Figure 3.6). In Berlin, Rome, Krakow, Lisbon and Valencia there is a peak in residential land use growth at approximately 16 km away from the urban center, and Rome has a second peak farther from the center. Paris reaches its maxima in the development of residential land use around 35 km away from the center and focused in the West, East and less in the South area. The growth directions in Berlin, Lisbon, Rome and Valencia are different, mainly
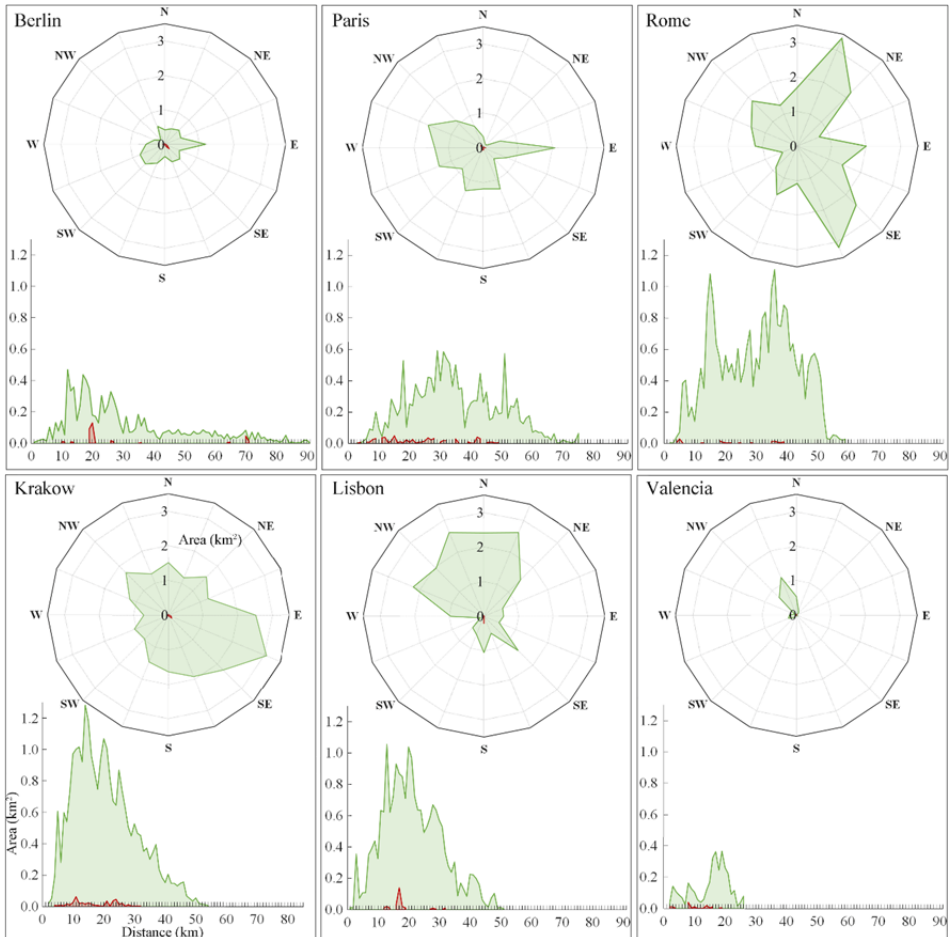
Figure 3.6. Global growth graphs. (Radar chart above) The sector analysis represents the spatial orientation of residential class changes in the six FUAs, the radius means the change in residential area in square kilometers by orientation, and (area chart below) the concentric circles analysis show the variation of residential area with respect to their central point. Green color means residential growth, while red shows lost patches.

due to physical and topographic constraints (e.g. the sea or rivers). Furthermore, partial losses of residential areas are present, for example, in Lisbon due to the extension of the road network; or the construction of an airport in Berlin.

Analyzing the results of the landscape expansion index (LEI) for residential, commercial and leisure land uses, in general, the expansion process has been mainly edge-expansive and outlying in the six FUAs (Figure 3.7). Considering

the compact growth as a combination of infilling and edge-expansive growths and the dispersed growth as outlying, the urban growth at FUA level in Berlin, Paris, Rome and Krakow tend to be mainly compact, resulting in a more continuous urban cover. However, Lisbon and Valencia have a more disperse growth.

Figure 3.7 shows the loss of natural and semi-natural vegetation in each FUA as a consequence of urban growth. Despite the double loss of forest in Berlin with respect to Rome, the mean change of the Splitting index in the peri-urban sector in Rome ($\triangle SPL_{Forest}$=185, CV=12) is much higher than in the rest of the FUAs (e.g. Berlin, with $\triangle SPL_{Forest}$=1, CV=9), showing a stronger trend of forest fragmentation in Rome.
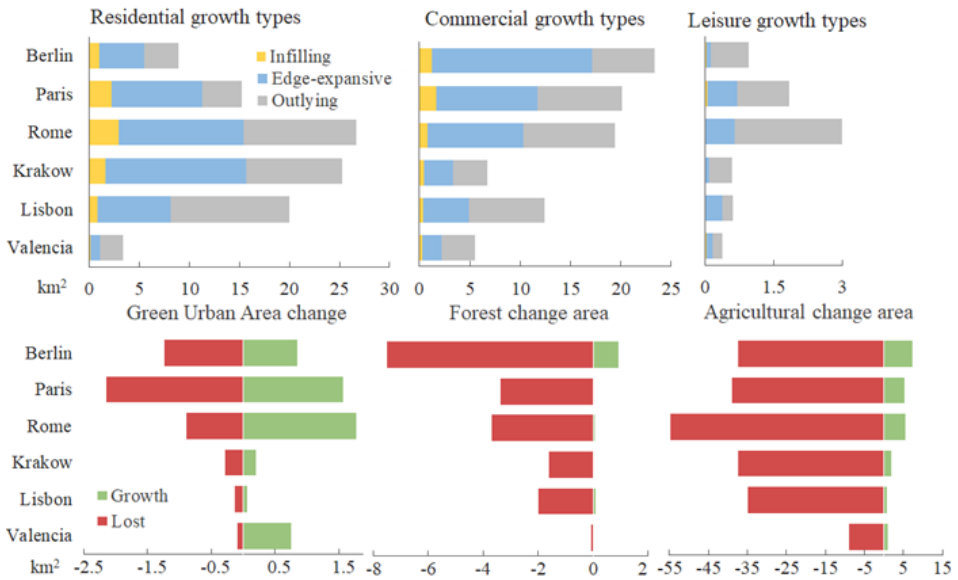
Figure 3.7. Growth and loss per land use at FUA level. Above, area of growth type in square kilometers (infilling, edge-expansive and outlying) of each FUA by class: residential, commercial and leisure. Below, gain and loss in square kilometers, of each FUA by class: green urban areas, forest and agricultural.

3.4.2. Analysis at intra-city level

As previously commented in the sector analysis, in 2012 high values of urban density (DU) are mainly located in the urban centers of the FUAs, however, there are variations within FUAs and sectors (Figure 3.8). For instance, in Berlin, there are some isolated units with high-density values located in the southern half of the FUA. Paris, Rome and Valencia also present scattered administrative
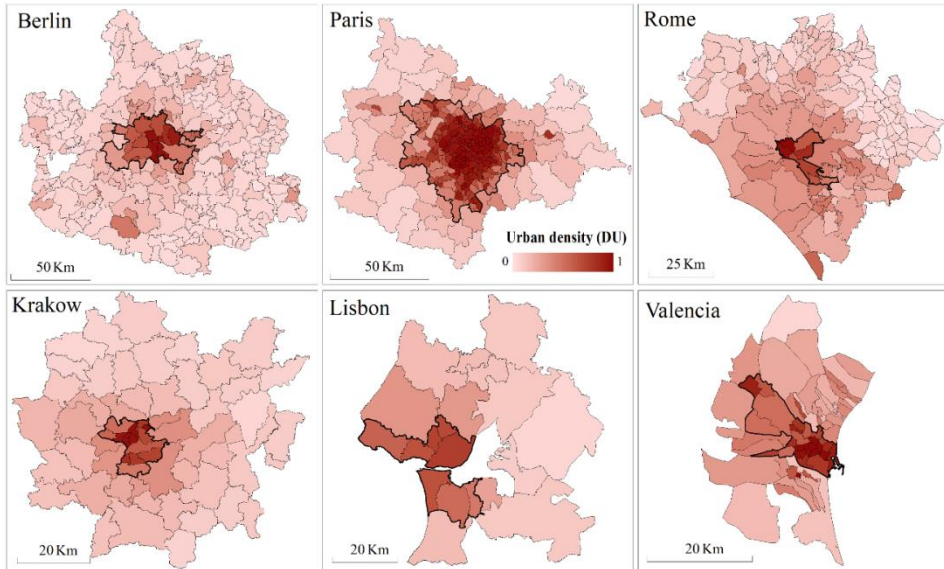
Figure 3.8. Urban density (DU) in 2012. DU quantitative maps of the administrative units in 2012 for the six FUAs. Bold lines separate urban from peri-urban sectors.

units with high DU out of the urban centers in different directions. Krakow and Lisbon show a gradual degradation of DU from the urban sector towards the peri-urban reaching their lowest values in the boundary of the FUA. With regard to Shannon diversity (DSHAN), high and medium values are located not only in the urban sector, but also in the contiguous administrative units, as the mix of land uses is usually higher along the boundary of the urban and peri-urban areas. The lowest values of DSHAN are found in the North-East of several FUAs: Berlin, Rome, Krakow and Valencia (Figure B.1).

When interpreted together, the object density ($DO_{Residential}$) and mean object size ($MS_{Residential}$) of residential class inform about the quantity and type of the residential patches in each administrative unit (Figure B.2). Results show that Berlin, Krakow and Lisbon present more uniform values of $DO_{Residential}$ and $MS_{Residential}$ than the rest of the FUAs. In Paris, Rome and Valencia a more variable response is observed, $DO_{Residential}$ varies along the urban sector, as well as their $MS_{Residential}$. On the other hand, regarding the area-weighted standard distance of green areas ($AWSD_{Green}$), that shows the aggregation of these elements, different compactness degrees are observed in the urban sector of Berlin, where administrative units differ widely (Figure B.3).

The analysis of temporal metrics at administrative unit level revealed significant changes during the analyzed period (2006-2012). In Berlin and Paris, slight

increases of DU took place at transition areas between urban and peri-urban sectors. A few administrative units present a slight loss of urban areas, but this effect is mainly due to the transition from barren land (included in artificial land uses in UA legend) to non-urban land uses. Berlin presents also the highest value of ΔDU in the southern part of the urban sector (ΔDU=0.1). Rome and Valencia, in general, increase their artificial surface in specific administrative units, while main changes are located in the peri-urban sector in different directions. DU in Lisbon and Krakow follows a gradient growth pattern from the urban center, reducing its intensity in the periphery, while in the rest of the FUAs presents a more random and scattered distribution (Figure 3.9).
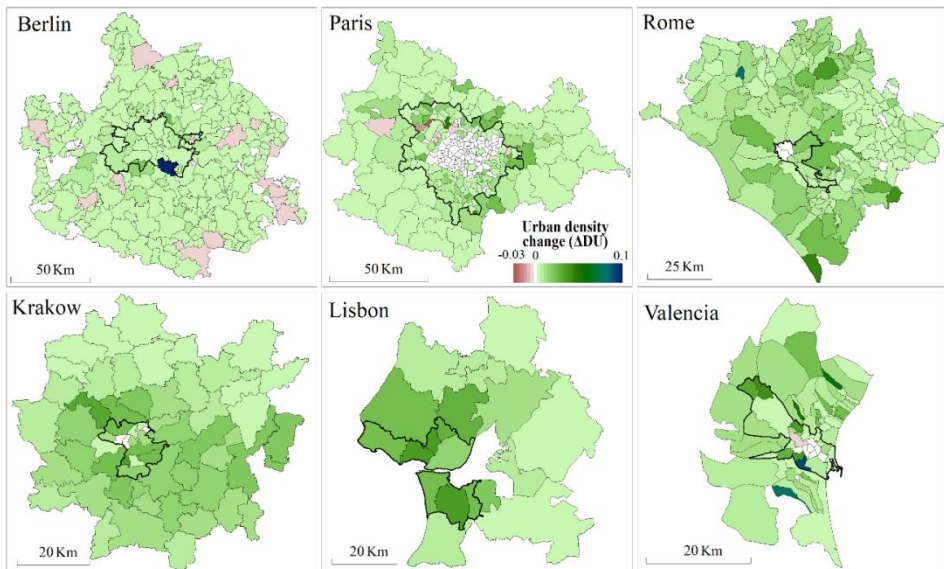
Figure 3.9. Urban density change (ΔDU). DU change quantitative maps of the administrative units from the six FUAs. Green values mean urban growth in this period, while maroon values show a partial loss of urban areas.

Negative variations of DSHAN are mostly located in the urban sectors of these FUAs, with some exceptions. However, in those areas where there has been an urban growth process, there is an increase of DSHAN and the diversity of land uses (Figure B.4). Analyzing the change of per-class indices, we observed a greater occurrence of density-diversity of commercial ($DD_{Commercial}$) in those administrative units along the border between sectors, reaching a maximum in Berlin ($\Delta DD_{Commercial}$= 0.19). The administrative units with an increase of the $DD_{Commercial}$ in Valencia presented a scattered spatial distribution, while in Rome were concentrated along the coast (Figure B.5). In addition, the tendency of

most FUAs in green areas growth is negative (Figure B.6), as in the inter-city analysis, except for Rome, where only one has negative change proportion of green areas (CP_{Green}), and Valencia, with null or positive values (there is a maximum of CP_{Green}=4.35). With respect to residential areas, $MS_{Residential}$ variation shows a tendency to smaller patches, except in some administrative units located on West Rome, South-East Lisbon, and inside and around the urban center of Valencia, where the overall increase of $MS_{Residential}$ implies larger new patches (Figure B.7). The residential class in the peri-urban sector of Rome has a compact growth pattern according to the changes of the Euclidean nearest neighbor mean distance of residential class ($ENND_{Residential}$), that reaches the maximum negative change value ($\Delta ENND_{Residential}$ =-51.48 m), meaning that the residential class is more clustered than others, especially in the North. In Krakow, residential patches are more aggregated, mainly in the South-East (maximum negative value of $\Delta ENND_{Residential}$=-9.17m). High positive values of $\Delta ENND_{Residential}$ may evidence that previous residential class is suffering a sprawl process since the mean value of the distances between patches is increasing (Figure B.8). Regarding forest class, variations in the EMS_{Forest} show a general reduction of forest patches, decreasing in peri-urban sectors mainly due to the general urban growth dynamics. The most affected FUAs are those with more presence of forest. Berlin, for example, presents a maximum (ΔEMS_{Forest} = -0.9 km^2) but also has a general decrease in the North and South. In Rome, fragmentation increases in the administrative units from the North and North-West (maximum ΔEMS_{Forest} = -0.3 km^2). Paris and Krakow show reduced forest patch sizes in South-West and South-East, respectively (Figure B.9).

Administrative units were also classified based on population and residential paces of growth, providing a quantitative measure of their imbalance (PUGI). As an example, administrative units inside and along the border of the urban sector in Berlin are characterized by high change and partial population change, with low and negative PUGI values (Figure 3.10 and Figure B.10), accompanied by low and partial residential changes in the limit of the FUA, along with positive PUGI values. There is an exception in the peri-urban sector, where an administrative unit presents a high negative value (PUGI = -30.17), where population grew a 43% and the residential class remained unchanged, meanwhile, spatio-temporal metrics showed a unique slightly positive value of $\Delta DD_{Commercial}$. As opposite, Paris presents low change and partial population changes in the administrative units of the urban sector, with almost no residential increase but a significant population increase. In the peri-urban sector, population increase exceeds residential growth, accompanied mostly by negative $\Delta ENND_{Residential}$, which evidences the densification and transition to more compact administrative units. Rome has a more random distribution of
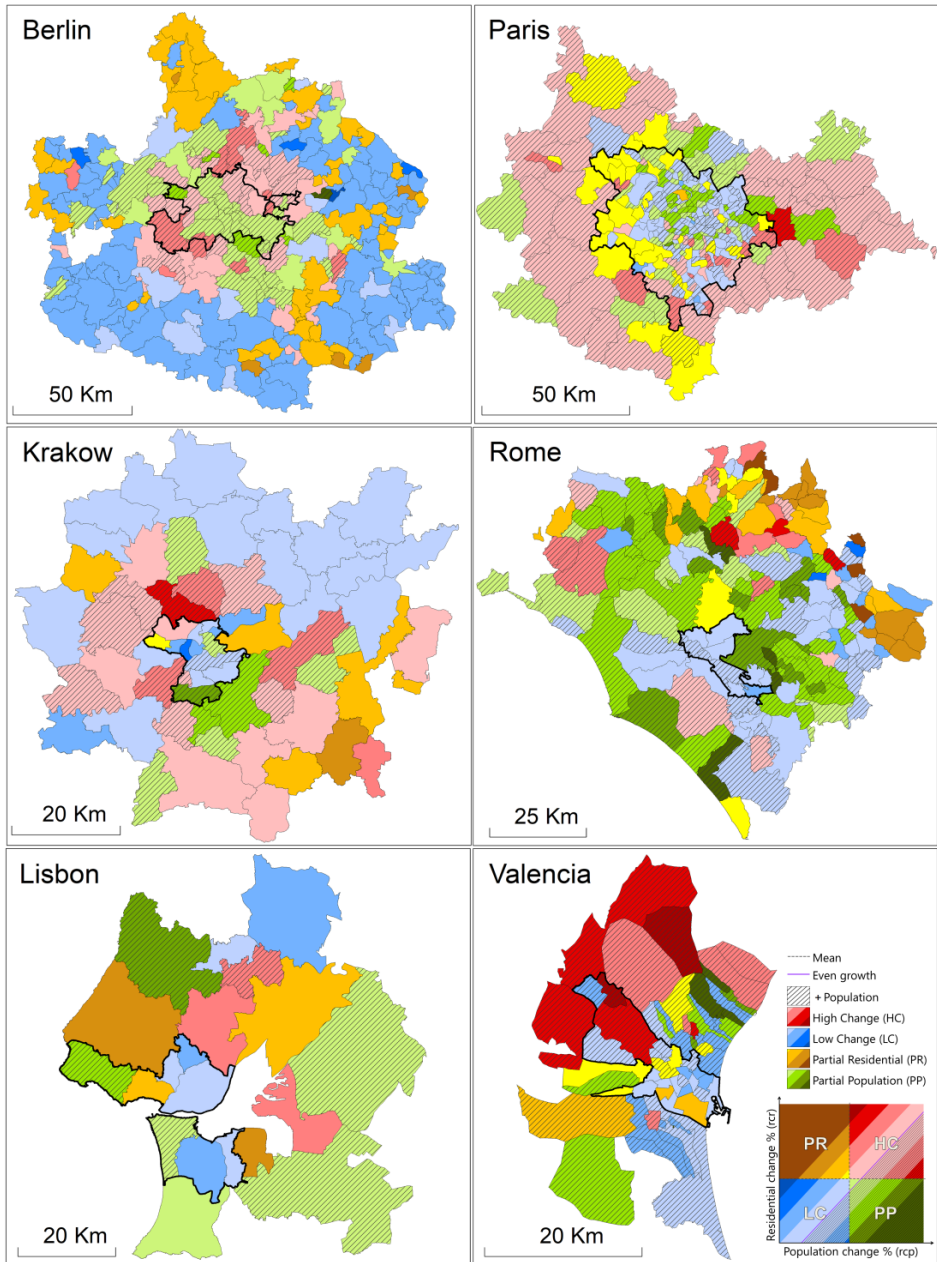
Figure 3.10. Graphical representation of the administrative units classified in change types by color, the hue depends on the PUGI distance, where darker colors means more inequality between residential and population growths. The stripe background shows when the population grows at a faster pace than residential areas.

growth classes. Small and balanced changes are located not only in the urban sector but also in the South and East of the peri-urban sector. Partial population change is located at the interface of peri-urban and urban sectors and near the coast, with high negative PUGI values showing a prominent population increase, while spatio-temporal metrics show a slight increase in $DD_{Commercial}$ and CP_{Green}, and a reduction in $ENND_{Residential}$. Partial residential change occurs in the North and North-East, with high and positive PUGI, showing an increase in residential class despite the loss of population in these areas, along with a general decrease of $\Delta ENND_{Residential}$, meaning a more compact distribution. However, there is also an increase of forest fragmentation (maximum $\Delta SPL_{Forest}=12$). Lisbon presents a significant residential increase, with positive PUGI and $\Delta ENND_{Residential}$ values in the urban sector and surroundings, evidencing a sprawl trend as previously detected in $LEI_{Residential}$. However, the westernmost administrative unit in the urban sector presents not only negative PUGI and $\Delta ENND_{Residential}$, but also positive $DD_{Commercial}$ and CP_{Green}. Krakow has low change but positive PUGI values in the North, due to the loss of population in these areas. Higher change is focused on the interface of urban and peri-urban sectors, presenting more population increase, while in three administrative units of the urban sector there is a general decrease of $ENND_{Residential}$ and a reduction of CP_{Green}. In general, Valencia has a prominent population increase, particularly in the North. Most of the administrative units in the urban sector suffered low changes (low positive PUGI values, slight or null residential increase accompanied by population loss). However, spatio-temporal metrics in Valencia reveal that negative PUGI values are generally together with a more compact residential growth and the increase of green urban areas (negative $ENND_{Residential}$ and positive CP_{Green}).

3.5. Discussion

The proposed methodology and the metrics analyzed provide useful information of the multi-temporal processes of urban growth between and within FUAs (inter- and intra-city). However, the extrapolation of these tendencies to other urban areas or periods should be taken carefully since only a 6-year interval of a reduced sample of urban areas was considered.

The interpretation of results at FUA level provides an overview of the state of urban areas and their evolution, allowing for the comparison of different FUAs. The analysis of sectors, urban and peri-urban, increases the level of detail and allows for a better differentiation of the type of urban expansion, compact or scattered, and the intra-city analysis complements the spatial distribution of

the growth patterns and allows for a local analysis of the evolution of cities. This information is complementary. In some of the examples presented, the analysis provided a uniform response of metrics in a sector, but a variable response at the different administrative units within that sector, reflecting different behavior at different levels of analysis. This is useful for the comparison of FUAs and the analysis of their internal spatial variability. The definition of urban and peri-urban sectors has an evident influence on the results obtained, and this should be properly defined attending to the final aim of each particular study.

The LEI index allows for the classification of the new patches in three growth types, which is useful in order to assign the compactness and sprawl degree of each FUA and land use. Our results are in consonance with a previous report (EEA, 2016a) that quantified urban sprawl from 2006 to 2009 in similar urban areas, showing a decrease of the degree of urban sprawl for NUTS-2 (i.e. basic regions for the application of regional policies) of Berlin and Paris, remaining the same in Rome, rising slightly in Krakow and Lisbon regions, and increasing sharply in Valencia. The LEI index might reveal the effect of the compact growth policies supported by the European Communities (1999), encouraging regional authorities to seek the development of sustainable, polycentric, balanced and compact urban systems. When applied at FUA level, this index provides an overview of the growth process, but at the administrative unit level, it allows for the detection of isolated sprawled areas.

On the one hand, in this period only two FUAs presented an increase of green areas in the FUA and urban sector levels. This seems to contradict the current idea of green cities in Europe (DG REGIO, 2011), and the Green Infrastructure Strategy and policies developed by the European Commission (EC, 2016). In this sense, monitoring the change proportion of green areas (CP_{Green}) would allow for the evaluation of the effectiveness of past and present policies. On the other hand, the variation in size of residential patches suggests a change in the typology of new buildings, such as detached houses or large buildings. The Euclidean nearest neighbor mean distance of residential class ($ENND_{Residential}$) represents the restructuring of the class into less or more dispersed, its alteration through the time emphasizes potential areas where residential growth process is being sprawled (a positive variation). This metric may detect, for instance, the variation of distances between residential areas and services. In this sense, the Urban Agenda reports that a compact city model benefits from the reduced distances between services (EC, 2017), and this can be quantified and monitored using this metric.

The classification of administrative units based on population and residential paces of growth, and the values of the PUGI index, provide additional

information for the study of growth patterns in the dynamics of urban areas. Similar classification methods have been applied without using population data (Altieri et al., 2014; Chen et al., 2013) and including this variable (Kasanko et al., 2006; Ribeiro-Barranco et al., 2014), but inequality of both variables had not been quantified. The increase in residential class and urban areas do not necessarily have a linear relation with the increase of population at different levels, and the proposed PUGI index quantifies this potential asymmetry. Some authors (EEA, 2011; Haase et al., 2013; Kabisch and Haase, 2013; Ribeiro-Barranco et al., 2014) have stated that, in general, European cities tend to grow faster in built-up than in population when studied at broad levels. However, when this phenomenon is analyzed at local level, results may vary. According to our results, population relative change outpaced residential relative increase from 2006 to 2012 at FUA level in Berlin, Paris, Rome and Valencia, and higher disparities were found at the intra-city level. In this sense, the PUGI index proposed quantifies the growing imbalance between the progress of the new residential areas and the population, allowing for the identification of differences of growth patterns and such behaviors may reflect differences in local policies or economic models. The PUGI index adds demographic information to the spatial metrics traditionally used in landscape ecology. The high land consumption per inhabitant is considered one of the contributing drivers of urban sprawl (EEA, 2016a; Jaeger et al., 2010a; Martinuzzi et al., 2007), thus the use of this metric may assist in the categorization of the urban growth as compact or sprawl, and even estimate the degree of both, being relevant in the context of urban sustainability. Moreover, the combination of this index with changes of spatio-temporal metrics, such as urban density, commercial density-diversity, Euclidean nearest neighbor mean distance of residential, proportion of green areas, and splitting index of forest, allows to identify the type of growth pattern and may help to assess the effect of past or current policies in the development of land uses and the subsequent impact in the quality of life of urban areas. Furthermore, with detailed information about the urban area and its background, this metric combination may assist in the interpretation of drivers of the urban growth process. For instance, in Valencia, the collapse of the construction and real estate sectors that took place during the studied period had economic consequences. Concurrently, the migration of rural population to coastal and inland municipalities close to urban areas, due to the extension of residential areas as a mean of decongesting the urban core, harmed the territorial and social cohesion (IVIE, 2013). These processes were revealed with local values of PUGI in Valencia (mostly negative in coastal and peri-urban administrative units and low positive in the urban sector), quantifying population movements and a deceleration of housing construction.

The interpretation of the PUGI index is quite intuitive, as the combination of class and magnitude outlines if the change process is balanced at the level of the administrative units. Positive values mean low-dense growth, while negative values reflect the reduction in the land consumption, and hence a densification process. A constraint of this index is the possibility to get a high positive value when there is not relative residential change but population has deeply decreased (since land consumption per inhabitant increases, this case is also a low-dense growth). However, the identification of these cases is straightforward, since the class assigned is usually low change. Another possible limitation is related to the definition of the index. Since the variables involved have relative values its interpretation may lead to confusion, i.e. a slight increase in a small administrative unit will show a great relative change, affecting the mean value used as classification threshold. In this case, different statistics (median, mode, etc.) should be used to avoid possible outliers. The integrated analysis approach based on the use of PUGI, its class and the spatio-temporal metrics is useful to overcome these limitations.

In addition to the potential of the PUGI index itself, analogous indices, obtained by simply modifying its variables, may be applied with different goals and scenarios. For instance, Kabisch and Haase (2013) did not find correlation between population change and the development of new green urban areas, but the application of a modified version of the PUGI index, using the relative population change and the relative green areas change as variables, could provide deeper insight and more specific conclusions at local level. Nowadays, variables related to the dynamics of the landscape (residential areas, green areas, etc.) can be updated using remote sensing techniques (Gil-Yepes et al., 2016).

Finally, some limitations related to the data and methods proposed in this study should be pointed out. The first is related to the scale effect, some spatial metrics vary in response to changes in the spatial extent and scale of the analysis (Šímová and Gdulová, 2012), and hence the conducted metrics might be affected by the minimum mapping unit and the administrative unit size. This constraint could be reduced by including a parameter that specifies the scale, as previously seen in Jaeger et al. (2010b) or by conducting a grid cell based analysis to improve comparability. Another limitation is the quality and thematic accuracy of the dataset, as discussed by Šímová and Gdulová (2012). In our particular test, the overall accuracy of Urban Atlas database was 85% in urban and 80% in rural land uses. However, according to the validation report of the UA2006-2012 change map, the overall accuracy of the transition from artificial to agricultural land uses is 50% in the selected testing sample. Therefore, the decrease of the urban density at local level found in a

few administrative units may be related to the poor classification accuracy of these particular classes. Moreover, classification errors are not balanced when working with temporal datasets. For this reason, the interpretation of changes should be done cautiously when working with LULC databases.

3.6. Conclusions

In this chapter we explored the application of spatio-temporal metrics and the PUGI index extracted from the Urban Atlas and demographic databases at two dates to compare and analyze urban growth patterns from a testing sample of six FUAs across Europe.

From an objective selection of spatio-temporal metrics quantifying land use variations, we performed a three-fold analysis: an inter-city comparison at FUA level, a sector level analysis between FUAs, and an intra-city analysis at administrative unit level. Discrepancies between patterns observed in the urban and peri-urban sectors were evidenced. Working at administrative unit level presented advantages over the FUA and sector levels since a more specific and spatially explicit identification of urban growth type is feasible. Moreover, it is closer to the boundaries employed by local authorities responsible for spatial planning, and it may be potentially used for monitoring the effect of local and regional policies implemented.

The results showed that the spatio-temporal metrics are useful for comparison of growth patterns at different levels. Nevertheless, a single metric is not sufficient to properly describe the urban growth process, but the combined analysis of a selection of spatio-temporal metrics and the proposed PUGI index, a qualitative and quantitative metric that relates built-up areas and population dynamics, enables a deeper analysis of urban growth patterns. Its integration into the analysis emphasizes the imbalance between residential land use and population growth rates, providing complementary information related to the per-person land consumption and supporting the characterization of the degree of sprawl in the urbanization process, a relevant issue in the context of urban sustainability. The input data for the PUGI index are affordable and frequently made available by local agencies, and its representation allows for the straightforward interpretation of population and residential dynamics and its balance.

LULC multi-temporal databases allow for more precise urban dynamic studies. Currently, the Urban Atlas dataset has only one period of time available (2006-2012), which is still insufficient for detecting reliable growth trends. Longer and more frequent time-series would allow for more accurate and comprehensive urban dynamic studies. In this sense, Urban Atlas is currently

publishing LULC maps for 2018, and it is expected to be updated every six years, progressively increasing possibilities of analysis in the near future. This chapter highlights the suitability of LULC databases for urban growth studies and their potential for analyzing urbanization trends.

Chapter 4

Relationships between spatial patterns of urban structure and quality of life

4.1. Introduction

The quality of life and sustainable development of urban and peri-urban areas depend on the successful management of their growth. Both are common goals in cities around the world. They are described in multiple dimensions: 'quality of life' is a broad concept assessed on various factors ranging from living conditions and employment to experience of life. It is usually represented by a multiple set of indicators such as income, deprivation rate, education attainment, employment rate, life expectancy, air quality, etc. (Eurostat, 2017; OECD, 2017). Also 'sustainable development' is addressed by the United Nations in their 2030 Agenda for Sustainable Development, which aims at ending poverty by means of promoting economic growth, addressing social needs, while protecting the environment and fighting climate change.

Urban form is constituted by spatial and socio-economic processes developed over time and space (Salat, 2011; Abrantes et al., 2019). It is accepted in scientific literature to wield a powerful influence on shaping societies (Salat, 2011; Tonkiss, 2013; Oliveira, 2016). Urban form is a key element for understanding urban systems as it drives where people live and work and how the interaction is spatially structured (e.g. Grimm et al., 2015; Taubenböck, 2019). However, it is not self-evident to establish a universal link between the urban spatial structure, here considered as the organization of urban areas in terms of the distribution of physical structures and human activities (Krehl and Siedentop, 2019), and quality of life. Accordingly, in this study we want to explicitly investigate this relation between urban structural features and socio-economic parameters, and whether quality of life can be interpreted based on spatial and statistical methods.

As introduced in previous chapters, urban areas with similar physical appearance tend to feature similar social, economic, and environmental characteristics. Consequently, several authors have described qualitatively and quantitatively these influences. Concerning social factors, many relevant concerns such as crime, public safety, gentrification, health, and poverty, have been linked to diversity and configuration of land uses, road network patterns, or remote sensing derived variables (e.g. Jacobs, 1961; Lehrer and Wieditz, 2009; Patino et al., 2014; Sandborn and Engstrom, 2016; Hankey and Marshall, 2017; Wurm et al., 2019b). In terms of economic issues, wealth indicators were positively related to the diversity of land uses, and productivity and innovation were influenced by density, centricity, and urban size (e.g. Tapiador et al., 2011; UN-Habitat, 2015). For the environmental dimension, the identification of land cover and urban structural types allowed for instance determining urban heat islands or green area facilities, which contributed to climate change studies (Stewart and Oke, 2012; Bechtel et al., 2019), while

pollution, energy use, and transport means have also been related to different properties of urban form, such as density, diversity or centrality of land use (e.g.: Anderson et al., 1996; Hankey and Marshall, 2017). However, there are few studies that measure this widely agreed linkage between the spatial structure of cities and their socio-economic status in a quantitative manner. These studies mostly rely on EO data to extract the physical information, such as buildings, roads, LULC and their spatial distribution, or structure and texture features. This approach has been applied so far to model neighborhood deprivation (Venerandi et al., 2018), poverty (Duque et al., 2015; Jean et al., 2016; Wurm and Taubenböck, 2018), income and property value (Taubenböck et al., 2009), and demographic, living conditions, labor and transport factors (Sapena et al., 2016). These examples present previous attempts to identify links between urban spatial structure and socio-economic parameters.

Insofar, the investigation of these relations has been possible due to the increasing accessibility of open databases and EO products. On the one hand, satellite images allow for increasing capabilities to provide high-resolution geo-information. In this context, LULC data have been an important source of information for urban studies; however, it lacks three-dimensional information of urban structures, considered a fundamental aspect in such studies (Wentz et al., 2018). Therefore, the characterization of cities into urban structural types and land cover, with Local Climate Zones (LCZ) (Stewart and Oke, 2012) as one concept, has great potential in its relation with socio-economic functions (Bechtel et al., 2015). LCZ have additional inherent information on the physical composition of cities compared to other LULC legends by their density, building types, heights, greenness and their land cover that are worthwhile to explore. Besides, it is a conceptually consistent, generic, and culturally-neutral description and thus a replicable classification system. On the other hand, global, national and local institutes provide more and more statistical data for different dates and spatial levels. Notwithstanding all the urban theories relating these two components, and the growing availability of both, spatial and socio-economic databases, studies aiming to quantify the relations between the spatial structure of urbanized areas and the quality of life of inhabitants or the well-known SDGs are still scarce. Methods based on the quantification of spatial patterns by means of spatial metrics and the clustering of urban areas based on their socio-economic performance (e.g.: Schwarz, 2010; Sapena et al., 2016; Abrantes et al., 2019) have shown to be suitable for the combined analysis of spatial and socio-economic variables in urban areas as well as for their relation.

In this framework, the general objective of this chapter is to understand better the relationship between the spatial structure of cities and the socio-economic

level of city dwellers. For this reason, we explore the value of the LCZ, as urban structural types, in relation to quality of life indicators at the city level. First, we quantify the relationships between socio-economic variables and the spatial distribution of LCZ. Then, we group cities according to their similar levels of quality of life and describe their spatial structure.

4.2. Data

4.2.1. Study area

We selected North Rhine-Westphalia (NRW) as a study case for its socio-economic relevance in Europe, reinforced by the availability of statistical data. The historical and political background of many cities located in this Federal State is similar, which diminishes external influences in our analysis. We base our study on a sample of 31 cities in NRW as consistent spatial and socio-economic databases are available there. The location and identification of cities is presented below in Figure 4.2.

Regarding the Federal State of NRW, it is the most populous of the sixteen German states, accounting for 21.7% of the total population in Germany (Eurostat, 2019). The Ruhr industrial region, in NRW, is a competitive industrial region of Germany. NRW is an economic center in Europe, with a regional GDP of € 672 billion in 2016 (21.4% of the German GDP). However, the per capita level is slightly below the national level. Nowadays, the economy of NRW is based on small and medium-sized enterprises, hosting more than 20% of companies in Germany, and providing work to near 80% of the active population (EC, 2019).

4.2.2. Socio-economic variables

For the socio-economic analysis we used the City statistics database (Eurostat, 2016a). This database was originally created with the purpose to provide information that supports more evidence-based decisions in planning and managing tasks (Eurostat, 2016b). The City statistics project covers several aspects of quality of life—i.e., demography, housing, health, economic activity, labor market, income disparity, educational qualifications, environment, climate, travel patterns, tourism, and cultural infrastructure—for cities and their commuting zones in Europe (Eurostat, 2018). At the city level, it contains 171 variables and 62 indicators for more than one thousand cities that have an urban core of at least 50,000 inhabitants. The data are available at different dates from 1990 onwards. In this study the city level is the basic spatial unit. At this level a rich source of data for comparative studies in Europe is provided.

For the purpose of this study, we selected a set of socio-economic variables and indicators for 31 cities in NRW for the year 2009 (Table 4.1), to coincide with the date of the satellite images used for LCZ classification. When data from 2009 were not available, the previous or subsequent year was used instead. Subject to the availability of data, we selected indicators of five dimensions of 'quality of life' covered in the database: education, health, living condition, labor and transport. We linked the dimensions to the SDGs policy commitments, as was previously done by the OECD (2017) to evidence the global efforts that are being made to reduce inequalities in the socio-economic level of citizens (Table 4.1).

Table 4.1. Description of the selected socio-economic variables (dependent variables in the models from Table 4.5) representing five dimensions of quality of life and their link to the Sustainable Development Goals (SDGs).

Dimension	Name	Description	SDGs
Education	education	The proportion of population (aged 25-64) with lower secondary as the highest level of education	SDG 4 (education)
Health	health	Crude death rate per 1000 inhabitants	SDG 3 (health)
Living conditions	housing	Average price for buying an apartment in euros	SDG 1 and 11 (poverty and sustainable cities)
	income	Median disposable annual household income in euros	
	affordability	Ratio reflecting the ability of a city to pay for housing. Housing price compared to income.	
Labor	employment	Number of employments per 1000 inhabitants (work place-based)	SDG 8 (decent work and economy)
Transport	transport	The share of journeys to work by car or motorcycle (%)	SDG 9 and 11 (Infrastructure and sustainable cities)
	commuting	People commuting out of the city per 1000 residents	

4.2.3. Earth observation and ancillary data

For classification of the physical structures describing the cities' spatial structure we rely on remotely sensed and geospatial data extracted from three data sources:

> High-resolution remote sensing imaging: a Rapid-Eye mosaic for the year 2009 was constructed for the whole area. This satellite provides images at 6.5 meters resolution (orthorectified and resampled to 5 meters) with five spectral bands (red, green, blue, near infrared and red edge).

> 3D model: A normalized digital surface model (nDSM) was derived

from 135 individual Cartosat-1 stereo images (collected between 2009-2013) and processed according to Wurm et al. (2014) to retrieve above ground heights.

> GIS layers from OpenStreetMap: the amenities and road layers from the open repository of geospatial data was used (downloaded in 2014, openstreetmap.org).

4.3. Methods

The methodology consisted on measuring spatial patterns of cities based on spatial metrics derived from a LCZ classification obtained by fusing remote sensing and open GIS data with a machine learning approach. Then, we quantified the relationship between the spatial metrics and socio-economic variables by means of multiple linear regression models. Clustering methods were applied to group cities according to their level of 'quality of life' using the socio-economic variables and similarities in the spatial patterns were analyzed.

4.3.1. Patterns describing the spatial structure of cities

For the derivation of the spatial patterns describing the spatial structure of cities, we applied the LCZ framework that allows characterizing the morphologic appearance of cities in a conceptually consistent manner. It comprises several urban structural and land covers types with uniform surface cover, structure, material and use (Stewart and Oke, 2012). Out of the 17 original LCZ classes, 12 were present in the region (Figure 4.1). The spatial pattern describes the distribution of phenomena across space, e.g., concentration, dispersion, clustered patterns, etc. (Getis and Paelinck, 2004). In particular, we refer to the arrangement of urban structural types and land covers within cities.
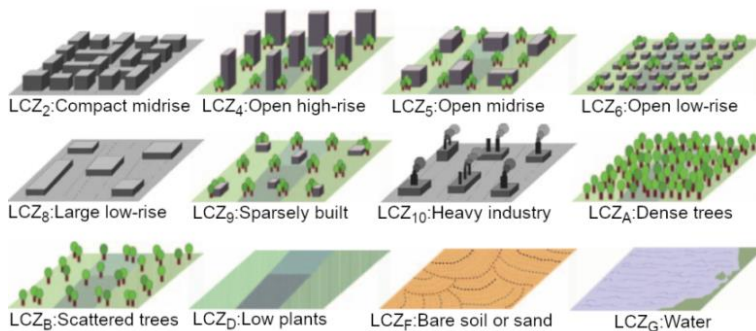
Figure 4.1. Summary of the LCZ classes present in NRW (from Stewart and Oke, 2012). LCZs 2 to 10 are built-up classes, LCZs A to G land cover types.

For the classification of LCZ, we followed the protocol presented in Tuia et al., (2017). We modeled LCZs on a grid composed of cells of size 200×200 m. In total, 89 variables were extracted for each cell (Table 4.2) to train the classifier. A ground truth of 2,658 cells was defined by photointerpretation, where the cognitive perception of an interpreter was used to define the predominant LCZ.

Table 4.2. Description of the geospatial variables for each cell from remote sensing and GIS data.

Source	Type	Description	No. of variables
Remote sensing	Bands	Mean and standard deviation of the pixel values. Data were atmospherically corrected and haze removed.	10
	Texture	Co- and occurrence features (local standard deviation, average, homogeneity, entropy, dissimilarity, correlation, contrast and angular moments).	60
	3D	Mean and standard deviation for both nDSM and buildings only. The number of buildings was also added as a feature.	5
	Land cover	Percentage of the area occupied by buildings, trees, grassland and impervious surfaces. The land cover is issued from an object-based classification on a Rapid-Eye mosaic (Montanges et al., 2015).	4
GIS	Roads	Total line length for highway, primary, secondary, tertiary roads, residential streets, streams and rivers, smoothed with a Gaussian kernel at the cell level.	7
	POIs	Counts for POIs cafes, restaurants and rail stations, smoothed with a Gaussian kernel at the cell level.	3

The classifier was based on random forests, a method building several decision trees with heavy randomization of features (Breiman, 2001a). The initial result was then further improved by making the model aware of two spatial relationships between cells using a Markovian Random Field formulation (see Tuia et al., (2017)): (1) By predicting with higher probability co-occurrence of neighboring LCZ that attract or repel each other spatially; (2) By favoring a map respecting a rank-size distribution of urban settlements, according to Zipf's law.

For testing the relationships between socio-economic and spatial structure of cities, we extracted all the spatial metrics included in IndiFrag. We used the LCZ classification as a base to characterize the spatial structure of cities. The level of analysis to extract the spatial metrics was the city level that corresponds to the level in which the socio-economic variables are provided. Therefore, we obtained one set of metrics per LCZ class for the spatial level of the city, and another set of metrics for the city, regardless of the LCZ classes. Then, we standardized the values of the metrics as the mean divided by the

standard deviation, in order to obtain comparable regression coefficients and avoid influence of measurement units.

We applied feature selection methods to avoid potential noise introduced by highly-correlated features, affecting the accuracy of results. In particular, regression models need the independence of predictors to minimize the multicollinearity, which makes the model unstable. We followed three consecutive approaches for the objective selection of metrics. First, we discarded the non-discriminative spatial metrics, those with a coefficient of variation lower than 5%. Second, we conducted a correlation analysis to identify redundancies in the spatial information. We omitted those metrics showing strong correlations to others (Pearson correlation coefficient > 0.8), keeping one metric per group of correlated metrics. Third, we applied a recent method proposed by Genuer et al. (2015), called Variable Selection Using Random Forests (VSURF) that selects a specific subset of metrics adapted to each socio-economic variable. This is based on measuring the relevance of every metric in relation to each socio-economic variable using a random forest regression. We kept one subset of metrics for each socio-economic variable (the different subsets of selected metrics are reported in Table 4.4).

4.3.2. Estimating socio-economic and spatial pattern links

A model was obtained for each socio-economic variable from Table 4.1 applying stepwise multiple linear regression analysis, using the subset of spatial pattern metrics previously selected as independent variables. We applied a min-max normalization transforming the socio-economic variables in a range between zero and one as follows:

$$z_i=(x_i-min(x))/(max(x)-min(x))$$

where $x=(x_1,...,x_n)$, x_i is the ith original value and z_i is the normalized value.

For education, health, housing, affordability, transport, and commuting the normalization was inversed and thus, higher values mean better conditions for all variables. The number of independent variables was restricted to a maximum of four spatial metrics to avoid overfitting, considering the limited number of observations (cities) in our dataset. The residuals were tested for normality using the Shapiro-Wilk test (Shapiro and Wilk, 1965), and for statistical significance by requiring *p-values* to be lower than 0.05. Leave-one-out cross-validation was employed to evaluate the models. We estimated the root mean squared error (RMSE) and the coefficient of determination (R^2) to summarize the proportion of variance explained by the model, and thus the goodness-of-fit.

To verify whether the level of 'quality of life' in cities is reflected in their urban spatial structure we conducted a two-step analysis: (1) we used the *k*-Means clustering method to group cities according to their values of socio-economic variables, representing variables from five dimensions of quality of life considered in our study (Table 4.1), out of nine (Eurostat, 2017). Using the Elbow method (Ketchen and Shook, 1996) we found an appropriate number of groups. Consequently, we created and described four clusters that group cities based on their socio-economic similarities. Moreover, we represented the 'quality of life' for each city and group using star plots, as well as the average of the region, which facilitates the interpretation of the different groups of cities; (2) we applied a stepwise discriminant analysis for selecting a relevant and reduced set of spatial metrics—based on their significance—that better separates the cities into these groups. Afterwards, the values of the spatial metrics, and thus the spatial structure of cities, were interpreted for each group.

4.4. Results

4.4.1. Spatial analysis of cities

In Table 4.3 we present the composition of the training/test sets and the per-class and overall accuracies obtained for the LCZ classification. Per-class

Table 4.3. Numerical results of the LCZ classification and number of samples used for train/test steps. User's and producer's accuracy and global statistics.

Code	LCZ	No. samples (train/test)	User's Accuracy	Producer's Accuracy
Land cover				
LCZ A	Dense trees	186 / 114	76.99%	76.32%
LCZ B	Scattered trees	118 / 107	74.77%	74.77%
LCZ D	Low plants	169 / 131	81.41%	96.95%
LCZ F	Bare soil or sand	108 / 192	97.33%	94.79%
LCZ G	Water	194 / 106	98.85%	91.13%
Built-up				
LCZ 2	Compact midrise	82 / 39	45.59%	79.49%
LCZ 4	Open high-rise	13 / 26	100%	3.85%
LCZ 5	Open midrise	124 / 72	67.95%	73.61%
LCZ 6	Open low-rise	152 / 48	88.10%	77.08%
LCZ 8	Large low-rise	100 / 101	91.01%	80.20%
LCZ 9	Sparsely built	153 / 55	81.67%	89.10%
LCZ 10	Heavy industry	148 / 192	88.62%	90.83%
	Average accuracy	-	82.69%	76.51%
	Overall accuracy	-	83.08%	
	kappa	-	0.81	

accuracy is given by the user's and producer's accuracy, where the number of correct classified cells in a class divided by the total number of cells classified as that class is the user's accuracy (commission error), and if divided by total number of cells of a class in the ground truth is the producer's accuracy (omission error) (Congalton, 1991). The region was split in two parts (North and South) and training was performed on the Northern region, while testing was performed on the Southern to avoid positive biases related to spatial co-location of cells. We obtained an overall accuracy of 83%, which is slightly lower than in Tuia et al. (2017), most probably due to the larger amount of testing samples used in this study. With the exception of the LCZ_2 "Compact midrise" class and the LCZ_5 "Open midrise" all the other classes are classified with more than 70% accuracy. Moreover, the average accuracies of 76.5% and 82.7% also show that the errors are not systematic on the small classes.

In Figure 4.2 we illustrate the LCZ classification for the 31 sample cities in NRW. The detailed example of the city of Münster reveals how the structural variety of the built and natural landscape is captured by the LCZ classification.
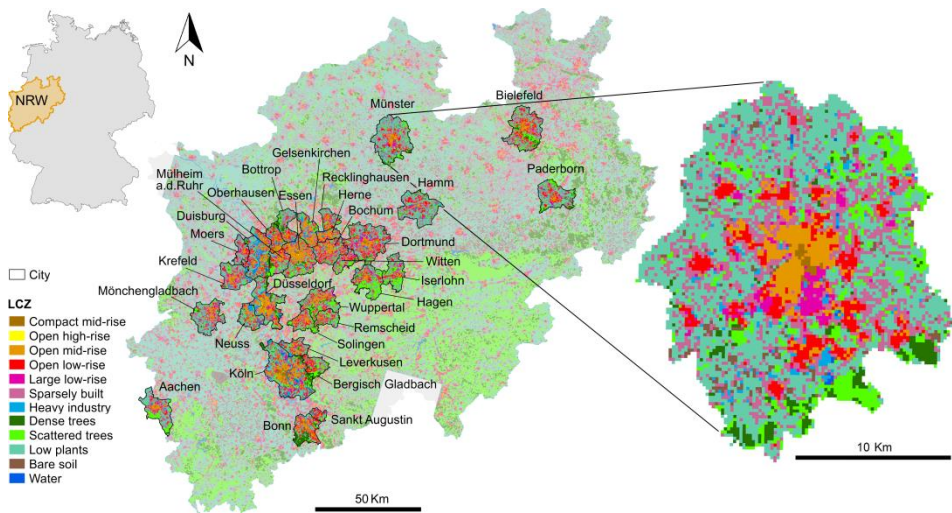
Figure 4.2. Result of the classification for the NRW region highlighting the classification in the analyzed cities (left). Detailed example of the classification for Münster (right). Maps are north-oriented.

Concerning the spatial metrics, in total 22 global metrics per city and 24 class metrics per LCZ and city were calculated. Since our classification map had 12 LCZ classes, 310 metrics were obtained at city level. After the correlation analysis, a reduced subset of 88 uncorrelated metrics remained. This subset was the input in VSURF for each socio-economic variable, obtaining one group

Table 4.4. Description of the selected spatial metrics (independent variables in the models from Table 4.5. The significant relations between metrics and socio-economic variables according to VSURF, are shown in the intersection of the rows and columns. The characters show whether metrics computed at the class level (with their LCZ short codes, see Table 4.3), at the city level (X), or lack of relation (-). Formulas (Eq.) can be consulted in Appendix A. Patch means a group of contiguous pixels with the same LCZ class.

Spatial metric	Description	Eq.	education	health	housing	income	affordab.	employ.	transport	commut.
Compactness (C)	Shape complexity of the class.	(A.27)	-	F	A,F,5,9	-	A,F,5,9	A,F,5,9	A,5,9	A,F,5,9
Class density (DC)	The ratio between the class and city area.	(A.2)	2,4,10	D,F,5,9	A,F,G,2,5,6,9	6	A,F,G,2,5,6	9	6,9	6
Density-diversity (DD)	Informs about the richness and heterogeneity.	(A.34)	-	-	X	-	-	X	X	X
Euclidean nearest neighbor mean distance (ENND)	Mean distance between the nearest patches from the same class.	(A.21)	B,5	A,B,8,9	10	9	10	5,9	5,9	9
Pixel Euclidean nearest neighbor mean distance (ENNDp)	Mean distance between the nearest pixels from the same class.	(A.21)	A,5,6	D,6	8	A	A	D	-	A
Area-weighted mean standard distance (AWSD)	Informs about the concentration degree.	(A.19)	-	A,5	-	5	A,6	A	D,6	8
Object density (DO)	The number of patches of the same class per km².	(A.17)	6,9	B,5,6	D,5	-	D	D,5	D,8	D,5
Urban density (DU)	The ratio between the built-up (LCZ2-10) and the city area.	(A.1)	-	X	-	-	-	X	X	-
Radius dimension (Dim$_R$)	The centrality of the class with respect to the city center.	(A.26)	A,6,9	A,B,G	A,6,9	G,8,9	6,9	6,9	A,6,9	6,9
Coherence degree (CD)	The probability that two random points are in the same patch in a city.	(A.25)	X	-	X	-	X	X	X	X
Shape index (SI)	The complexity of patch shape is compared to a square.	(A.14)	2	-	2	-	2	-	2	2
Splitting index (SPL)	The number of patches when dividing the class into equal	(A.24)	X,G	F,10	X,6	A,6	X,6	X,6	X,6	X,6,10

	size parts with the same division.								
Leapfrog (LPF)	The proportion of isolated pixels with respect to the entire class.	(A.18)	6	5,8	A,B	A	5,6	-	5,6	5,6
Urban-/porosity (PU, P)	The ratio of open space compared to the city or class area.	(A.8)	X	-	5,9	X,9	5	-	5	5
Boundary contrast ratio (BCR)	Measures the contrast between classes based on pixel's neighbors.	(A.37)	D	D	A	X	-	X,D	D	-
Effective mesh size (EMS)	Measures the connectivity. Low values mean fragmentation.	(A.22)	2	6,8	2,8	2,6,8	2,6,8	X,2,8	X,2,8	X,6
Mean object size (MS)	Average size of the patches from a class.	(A.4)	2,4,9	F,G,9	2	G,2,4	F,2	-	2,9	2

of metrics per socio-economic variable with sizes between 19 to 31 metrics (Table 4.4). This output was part of the input in the following section as explained below.

4.4.2. Models of socio-economic variables

In Table 4.5 we show the results of the eight fitted models, one for each socio-economic variable. The numerical goodness-of-fit indicators show that the models are statistically significant (*p-value* < 0.05) and explain from 43% to 82% of the variability (R^2) of the socio-economic variables by means of the spatial structure of cities, with RMSEs ranging from 0.10 to 0.17. The values of the model of housing were normalized with a logarithmic transformation to obtain a normal distribution of the residuals and improve the adjustment (Table 4.5). The spatial metrics included in each model and their associated coefficients allow interpreting which and to what extent spatial patterns explain the modelled variable (Table 4.5). As they are all standardized to z-scores prior to the analysis, their direct contribution is represented by the regression coefficients.

Table 4.5. Multiple linear regression models for the normalized socio-economic variables (SE), where higher values mean better conditions for all variables (dependent variables), using the spatial metrics (independent variables in bold, with the LCZ class in the subscript). The leave-one-out cross-validation coefficient of determination (R^2), the root mean square error (RMSE), the p-value of the model, and the number of observations or cities (Ob) are shown. The acronyms of LCZ and spatial metrics can be found in Table 4.3 and Table 4.4, respectively.

SE	Model	R^2	RMSE	p-value	Ob
education	$0.407+0.159 \cdot \textbf{ENNDp}_5+0.119 \cdot \textbf{Dim}_{R\,A}-0.11 \cdot \textbf{PU}+0.077 \cdot \textbf{DC}_4$	54.87	0.174	$7.7 \cdot 10^{-6}$	31
health	$0.389+0.13 \cdot \textbf{DC}_F+0.102 \cdot \textbf{ENND}_B+0.081 \cdot \textbf{ENND}_9+0.081 \cdot \textbf{SPL}_{10}$	49.78	0.175	$9.9 \cdot 10^{-6}$	31
log(housing)	$0.536+0.138 \cdot \textbf{C}_5+0.093 \cdot \textbf{P}_9-0.06 \cdot \textbf{MS}_2+0.057 \cdot \textbf{AWSD}_5$	50.61	0.151	$1.2 \cdot 10^{-5}$	31
income	$0.560+0.112 \cdot \textbf{EMS}_6-0.092 \cdot \textbf{SPL}_A-0.057 \cdot \textbf{P}_9$	43.40	0.165	$2.7 \cdot 10^{-5}$	31
affordability	$0.580-0.088 \cdot \textbf{MS}_2+0.076 \cdot \textbf{C}_5-0.067 \cdot \textbf{SPL}_6$	53.98	0.153	$3.9 \cdot 10^{-6}$	31
employment	$0.281+0.108 \cdot \textbf{Dim}_{R\,6}+0.108 \cdot \textbf{ENND}_9+0.093 \cdot \textbf{EMS}$	56.84	0.157	$2.0 \cdot 10^{-7}$	31
transport	$0.457+0.236 \cdot \textbf{EMS}-0.149 \cdot \textbf{CD}+0.114 \cdot \textbf{DO}_8-0.057 \cdot \textbf{DC}_9$	51.22	0.166	$5.1 \cdot 10^{-6}$	29
commuting	$0.637+-0.222 \cdot \textbf{C}_9+0.1 \cdot \textbf{LPF}_6+0.078 \cdot \textbf{ENND}_9-0.059 \cdot \textbf{SPL}_6$	82.29	0.101	$7.4 \cdot 10^{-11}$	31

The relationships we found between the spatial structure of the cities in this region and the socio-economic variables are as follows: cities with a better *level of education* have less open, and thus more continuous, built-up (PU), however, the distribution of open midrise is more scattered ($ENNDp_5$), dense tree patches are furthest away from the city center ($Dim_{R\,A}$), and there is a higher density of open high-rise buildings (DC_4). In terms of *health*, the model relates a lower death rate in cities with a fragmented and distant distribution of sparsely built ($ENND_9$) and scattered tree ($ENND_B$) patches. Conversely,

larger and less fragmented areas of heavy industry (SPL_{10}) are usually present in cities with higher levels of death rates. On the one hand, a compact shape of open midrise (C_5), scattered from city centers towards the suburban areas ($AWSD_5$) with a compact midrise core (MS_2) is related to lower prices of *housing*. On the other hand, *income* is higher in cities with bigger extensions of open low-rise (EMS_6), clustered dense trees (SPL_A), and contiguous areas of sparsely built with very few open areas (P_9). Regarding the *ability to pay* for housing based on income, the model is similar to housing model (MS_2 and C_5), except that the affordability is inversely proportional to the fragmentation of open low-rise (SPL_6). Therefore, the ability to pay is lower in bigger cities with a compact midrise core surrounded by fragmented clusters of open low-rise structures.

In relation to the economic aspects (*employment*), open low-rise located towards the periphery of the city ($Dim_{R\ 6}$) and a fragmented and distant distribution of sparsely built ($ENND_9$) are characteristic of cities with higher employment rate, moreover, LCZ patches are bigger, which means more continuous LCZ classes and, in general, less isolated small patches (EMS). Concerning *transport*, fragmented cities (EMS) in small continuous clusters (CD), with higher proportion of sparsely built areas (DC_9) and a lower number of large low-rise areas (DO_8) commute more by car or motor cycle. Meanwhile, citizens living in cities associated with more compact areas of sparsely built structural type (C_9) commute more out of the city (commuting). The way in which open low-rise is allocated affects commuting patterns. The higher the number of compact clusters (LPF_6 and IS_6) the more the commuting proportion.

4.4.3. Categorization of cities

Figure 4.3 shows the clustering of cities according to their socio-economic similarities using the normalized values of six socio-economic variables (we excluded the share of journeys to work by car or motor cycle since statistics were available only from 29 cities and the ability to pay for housing since housing and income were included instead). The individual plots show the location of cities by means of the bi-dimensional spaces defined by each pair of socio-economic variables. Cities are identified by means of a number and color. The map depicts how cities and groups are distributed in the region. It can be seen that the first group (green) is easily identified by means of income and commuting levels (income and commuting plots in Figure 4.3). While education discerns the second group (blue, education plots in Figure 4.3), the identification of the third group (orange) is not straightforward. However, the fourth (red) can be identified by means of death rate, price of buying an apartment and employment rates (health, housing, and employment plots in

Figure 4.3).
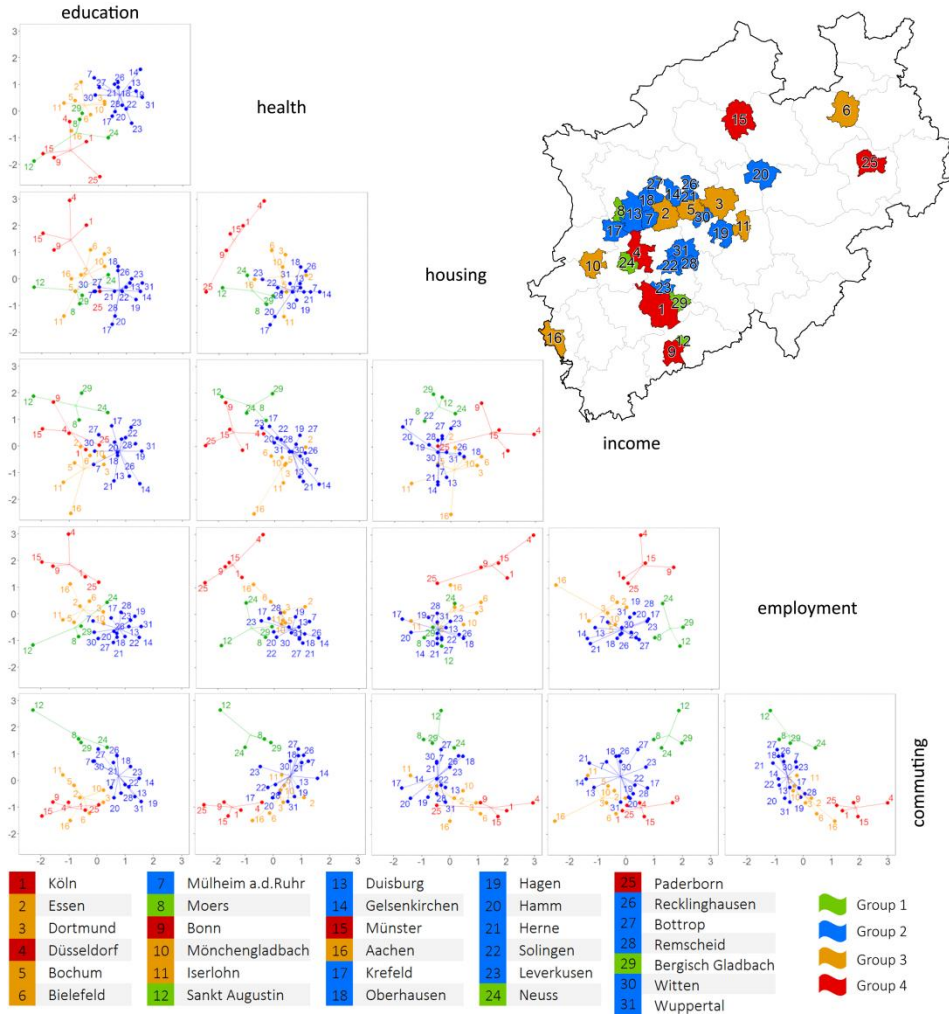
Figure 4.3. Clustering of cities into four groups using the scaled socio-economic variables. The individual scatter plots show: the location of the cities according to each pair of socio-economic variables (row and column, e.g.: the top-left plot corresponds to 'education' and 'health'), the centroid of each group, and the distance of cities to their centroid. The map locates spatially the clusters and combined with the table identifies the cities (identification number, group and name of the city). It compares cities relatively based on to their socio-economic performance and groups them according to their similarities.

The interpretation of groups by their mean values (i.e., group centroids using the non-scaled socio-economic variables, Table 4.6), shows that *group 1* is

formed by four cities with medium and low rates of mortality and low-education, the prices of buying an apartment are the lowest in contrast to the highest income levels (i.e., the capacity to pay for housing is higher), however, the low employment is balanced by the highest commuting level to work out of the city. *Group 2* accounts for the majority of cities (15 out of 31). This group is characterized for having lower education and employment together with higher death rates, the prices for buying an apartment and the income are medium-low in comparison with the rest of the groups, and a close to 15% commute out of the city. *Group 3*, which most closely approximates to the mean values of the region (Table 4.6), clusters seven cities; in this group the education level and health are medium, the price of buying an apartment is high in contrast with the lower income levels (i.e., low capacity to pay for housing), however, the employment rate is medium-high and there are low commuting rates. Finally, *group 4* gathers five cities with the lowest proportion of the low-educated population, lower rates of mortality, and the highest prices for buying an apartment accompanied by high-income values; however, the huge discrepancy suggests housing prices are less affordable, the employment rate is the highest of the region and the level of commuting out of the city is the lowest.

Table 4.6. Mean of non-scaled socio-economic variables (centroids) and number of cities per group. The last row shows the mean values of the NRW region.

Group	cities	education (%)	health (n°/1000)	housing (€)	income (€)	employment (n°/1000)	commuting (n°/1000)
1	4	27.59	9.92	91 875	23 975	426.01	215.27
2	15	37.24	11.89	91 933	20 900	421.49	148.49
3	7	29.44	11.36	104 000	19 700	498.79	112.94
4	5	26.37	9.01	134 400	22 200	686.05	97.82
NRW	31	32.48	11.05	101 500	21 235	482.20	140.91

By representing the cities multi-dimensionally using the socio-economic values by means of star plots (Figure 4.4), the shape of each city becomes an indicator of its 'quality of life'—here based on five dimensions—, the more complete (i.e., the area of the gray circle is covered) the better. *Group 1* shows high levels of commuting out of the city, education, health, house affordability, and income, but very few employments (work place-based). This shape can be related to satellite cities with good quality of life (regarding education, health and living conditions) but a less desirable situation in terms of sustainability due to the high commuting shares, to balance against the low employment rate. In *group 2* we find the lowest values of education and health in the region, commuting is medium-high and housing is affordable compared to income levels, however employment is quite low. There are similarities with the first group in the values of employment and housing,
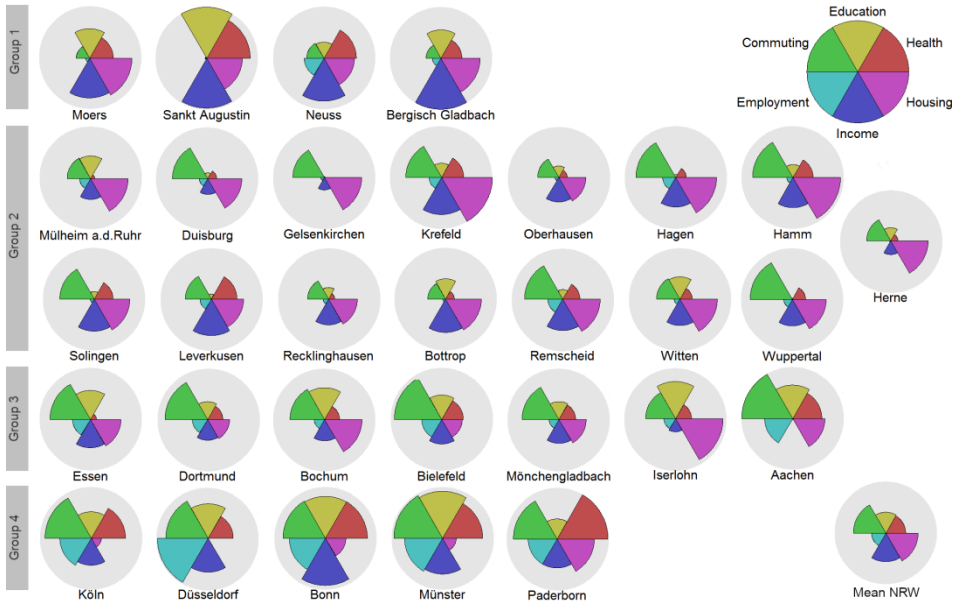
Figure 4.4. Multi-dimensional quality of life star plots of cities by group. Values are relative, as the socio-economic variables were min-max normalized between zero and one. For education, health, housing, and commuting the normalization was inversed and thus, higher values mean better conditions. The legend (top right) shows the maximum value of each socio-economic variable, equal to one, and its name related to the position and color. The mean values of the NRW region are represented in the bottom right. The gray background shows the maximum reachable value.

however, the analysis of the rest of variables suggests that this group has the lowest quality of life relative to the entire region. *Group 3* presents the lowest values of income in the region, and health is slightly lower than that of the mean NRW value. However, the remaining socio-economic variables are quite close to the mean values, which may suggest a quality of life close to the NRW average. Finally, group 4 has the lowest values of commuting out of the city, while education, health, employment and income are considerably high and, as a counterpart, housing is less affordable. Additionally, this can be considered the most sustainable group in terms of commuting shares. Thus, according to the analyzed dimensions, it could be objectively said that it shows the highest quality of life in the region.

The spatial structure of urban spaces, as mentioned in the introduction, is related to this measured 'quality of life'. To explore such relationships, we selected the spatial metrics that best identify these groups. We started from the subset of metrics selected with the VSURF method. Five spatial metrics for three structural types were the most influential in terms of grouping cities into

different levels of quality of life. Those metrics were: the distance between sparsely built and open midrise structures patches within the city (ENND$_9$ and ENND$_5$), the number of open areas within the sparsely built patches (P$_9$), the connectivity and size of open low-rise patches (EMS$_6$), the compactness of open midrise (C$_5$), and the centrality (proximity to the city center) of sparsely built and open low-rise (Dim$_{R\,9}$ and Dim$_{R\,6}$). The spatial patterns that better differentiate between the derived levels of quality of life can be analyzed by representing the values of these metrics for each socio-economic group in box-and-whiskers plots (Figure 4.5). The spatial patterns that better represent the cities in *group 1* are the presence of the biggest continuous areas of open low-rise, the highest compact shapes of open midrise patches but spatially scattered, and the compact distribution of sparsely built close to the city centers. For *group 2*, the metrics portray an even distribution of the sparsely built areas through the city, with fragmented and centralized open low-rise. *Group 3* shows open midrise structures scattered across the city, plus high values of open areas in the sparsely built environment, close to each other but farther from the urban cores, pointing that these urban structures are located in the surrounding areas of the city centers that are mainly occupied by high and medium rise types. Finally, cities in *group 4* are especially characterized by a compact nucleus of open midrise structures, with irregular shape, combined with fragmented distribution of sparsely built far from the urban cores, probably as they are located in the outskirts of the city, as well as the fragmented and decentralized distribution of open low-rise (Figure 4.5). That is, cities in *group 4* have a compact urban core becoming gradually less compact as the distance to the core increases, eventually with low-dense
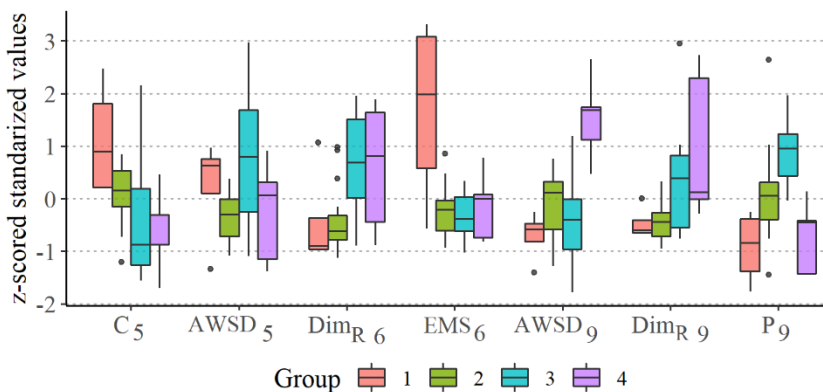
Figure 4.5. Box-and-whiskers plot illustrating the standardized values of the spatial metrics for each socio-economic group of cities, where: C=compactness, ENND= Euclidean nearest neighbor mean distance, Dim$_R$=radius dimension, EMS=effective mesh size, and P=porosity. The subscript shows the LCZ: 5=open midrise, 6 =open low-rise and 9=sparsely built.

structures located in the outskirts. For example, Münster (detailed example from Figure 4.2) present this spatial pattern, with a compact midrise core (orange), with decentralized fragmented clusters of open low-rise (red) intermixed with a scattered and isolated distribution of sparsely built (pink).

4.5. Discussion

This study in the cities of North Rhine-Westphalia in Germany shows the interrelation of urban spatial structure with quality of life dimensions. Our findings show that the education, mortality, income, employment, and other quality of life indicators can be partially explained by urban spatial pattern metrics extracted from urban structural types and land covers. No more than four metrics were needed to explain more than 40% of the variability of the socio-economic levels in cities with a similar economic and historical background for a given time. For example, the level of education tended to be better in more compact cities but also in cities with low-dense structures (i.e., open low-rise and sparsely built), which correspond to major cities and their satellite cities in NRW, respectively. This link can be related to higher-educated people moving to bigger cities seeking better job opportunities, and eventually moving to satellite cities. This seems to differ with a study where higher education levels were found in low-dense urban areas against high-dense areas in North America (Batchis, 2010). Cities with distant agglomerations of sparsely built areas and vegetation, combined with fewer and more scattered industry areas showed fewer death rates. In this sense, Oliveira (2016) compiled case studies that related walkability, diversity of land uses, and urban form with an improvement in health habits. The positive relation between death rate and bigger areas of heavy industry, besides higher shares of death in cities from groups 2 and 3, could be related to the fact that most of these cities are located in the highly industrialized Ruhr region, where death rates are high (Kibele, 2012). Apartments in cities of NRW with midrise structures (compact in the core and open towards the suburbs) and patches of dense tree are prone to be more expensive. We also found that income is measured higher in cities with a larger share of continuous and homogeneous areas of very low-dense built-up areas (i.e. sparsely built and open low-rise structural types), a spatial pattern that is especially seen in the satellite cities in NRW. Similarly, we found that commuting out of the city is higher in cities with more clusters and more compact areas of these low-dense built-up structural types, and the share of people choosing to commute by car or motor cycle is higher in less diverse and low-dense cities, which could be related to more monofunctional and dispersed cities. This tendency is widely discussed in the literature, for example, Travisi et al. (2010) found higher automobile

dependency in low-dense Italian cities. The positive relation of low-dense cities with higher incomes and commuting shares, especially by car or motor cycle, is likely to be linked to preferences of high-income households to live in less dense areas despite the higher travel cost. Additionally, the proportion of employment showed a positive relation to the homogeneity of structural types, they seem to be more organized, that may suggest that cities with more jobs are planned in a more uniform spatial distribution, with the exception of the sparsely built that tends to be more fragmented in these cities. Other authors also found relationships between spatial metrics and percentages of land uses with employment sector statistics (Ghafouri et al., 2016).

The socio-economic variables used in this study cover several dimensions of quality of life (Eurostat, 2017). Therefore, grouping cities according to the socio-economic variables allowed us identifying various levels of quality of life within the analyzed cities. One group presented the lowest level in the region, but it does not necessarily mean that the quality of life is poor because we are comparing relative values. On the contrary, two groups stood out for having better levels of quality of life. These groups differ in commuting patterns, housing affordability, and employment rates, and coincide with major cities and satellite cities. Despite having a good quality of life, satellite cities here identified with low-dense built structures, are unsustainable in terms of commuting and transport choices, besides low-dense cities are more inefficient in the use of land, energy and resources (Bhatta, 2010). We also found common spatial patterns related to the built-up structural types in cities that had similar levels of quality of life, which again suggests the two-sided impact of spatial structure of cities on their socio-economic levels. We should note here that this specific morphology found for cities in NRW for a given date do not necessarily have the same relations in other areas. Context is—as Tonkiss (2013) argues—all in this debate. However, similar correlations between urban spatial structures and economical functions have been previously discussed in the literature. For instance, Mouratidis (2018) found a positive relation between social well-being and high density, short distances to the center, and land use diversity. Venerandi et al. (2018) modelled deprivation to population density, higher proportion of bare soil and regular street patterns, while other authors predicted indicators of wealth, poverty and crime through satellite data (Irvine et al., 2017). Studies relating urban spatial structures with socio-economic values are usually focused in single dimensions, such as education, poverty, transport, air pollution, health, energy consumption, etc. (e.g.: Batchis, 2010; Duque et al., 2015; Sandborn and Engstrom, 2016; Hankey and Marshall, 2017, Wurm et al. 2019a). Only a few of them analyze various indicators (e.g.: Irvine et al., 2017; Sapena et al., 2016) or combine them (Tapiador et al., 2011). In contrast, we tackled several aspects individually

related to the quality of life and also in a combined way. We found that cities with a higher socio-economic status in NRW have a core with spatially compact midrise structures, while on the periphery there are small groups of low-rise structures and sparsely structures disaggregated with a high proportion of open green spaces. However, we are fully aware that the urban spatial structure of cities does not define or fully explain their success, since there are many other factors that play an important role. However, in the region analyzed, cities with similar spatial appearance also had similar socio-economic levels; this is a strong indication that the spatial structures of the cities do influence socio-economic performance. It should also be recalled that correlation does not imply causation, and thus the variation of a spatial pattern does not necessarily improve the socio-economic level of a certain area, although, it will certainly alter its state (Lehre and Wieditz, 2009; Williams, 2014).

We also faced some limitations. We based the relations on a large and consistent set of variables. However, a comprehensive set of variables is inexistent which means that we are not able to depict the manifold interrelationships holistically. Moreover, it is worth noting that we did not include external influences in the models, such as policies, individual historical background, etc. While of course, every city is unique, the overall urban spatial structure of the analyzed cities is comparable to a certain extent, thus reducing these influences. However, when conducting a global analysis, externalities should be considered, as well as measuring the spatial stratified heterogeneity (Wang et al., 2016) to test whether the variables are distributed unevenly across different parts of the study area, in which case it would be convenient to perform different models. Besides, it was not possible to model some socio-economic factors such as the 'proportion of economically active population' or the 'share of persons at risk of poverty after social transfers' with the spatial distribution of LCZs, and thus they were not included in the analysis. Regarding the data used in this study, it is important to mention the following: on the one hand, the urban spatial patterns of cities were extracted by means of spatial pattern metrics; therefore, selecting the most significant and non-redundant metrics is important. Another consideration is that the accuracy and spatial resolution of the image classification (here measured with 83% overall accuracy) affects to the spatial metrics, this fact needs to be considered when extracting conclusions of such studies. On the other hand, the socio-economic data used in this study have a great potential for comparative studies in Europe, but as mentioned, the average values at city level disregard internal socio-economic variations assuming that cities are homogeneous. Another limitation is the use of only five aspects of quality of life based on eight indicators, instead of a larger subset of socio-economic variables to enrich the

analysis. Although these variables were able to represent a significant part of the different levels of quality of life in the region, we were subject to the availability of data. Whereas remote sensing and GIS derived products, such as the LCZs classification, have no boundary or time-scale limitations, socio-economic statistics are usually restricted by administrative boundaries and census dates. Nevertheless, there is a recent tendency to provide these data in a different format, such as gridded datasets that swap irregularly shaped census boundaries to a regular surface (EFGS, 2019). These new datasets will serve as an opportunity to conduct studies that are not restricted by administrative boundaries.

The quality of life of the population and the sustainable development of urban areas are in the spotlight (OECD, 2017). Improving the understanding of the spatial structure of urban areas, the demographic, social and economic levels of these areas and their interrelations contributes to planning the development of cities with a view to meeting the global policy objectives set out in the New Urban Agenda (UN-Habitat, 2016). In order to unravel the interactions between the spatial structure of cities and their socio-economic levels, in this paper we quantified their relationships by means of statistical models. This supports the hypotheses that assume that the spatial structure of cities reflects social and economic indicators of their inhabitants, and eventually influence their quality of life. The applied methodology can be used as a tool to obtain empirical evidences as well as learning from past trends and understanding the present to design a better future.

4.6. Conclusions

The spatial structure of urban spaces is related to the quality of live and sustainability of our cities. This is clearly confirmed by this analysis of cities in NRW. We extracted the spatial structure of cities using spatial pattern metrics from a LCZ classification based on machine learning algorithms applied to multimodal geospatial data. These attributes explained the variability of quality of life related indicators, which are linked to six out of seventeen SDGs. Moreover, grouping cities into different levels of quality of life showed common spatial patterns within the groups. We ascertained that the spatial structure of cities has a strong influence on their socio-economical functions, but does not fully determine them.

In times of increasing availability of socio-economic and spatial data (e.g., from remote sensing) in ever-increasing spatial resolutions, there is a huge demand for systematic research in this direction. Of particular interest is research that systemizes these relations in dependence of context, that is policies, culture,

demography, etc., for a more general and quantifiable knowledge of the influence of the urban spatial structure on socio-economic parameters of cities and their people—this chapter testifies to this.

Although this study accounts for cities in NRW in a specific period, and thus is not globally representative, results show a trend that is worthwhile investigating further. This is feasible due to the growing availability of data for both local and global levels. Moreover, the methods applied in this study are directly transferable to other regions and datasets, which would broaden the analysis and derived conclusions.

Chapter 5

Spatio-temporal metrics for urban growth spatial pattern categorization

5.1. Introduction

Urban form refers to the spatial configuration of the physical built environment and human activities (Georg et al., 2016; Abrantes et al., 2019). Meanwhile, the urban growth spatial pattern is a dynamic process of urban change that, in some cases, modifies the initial urban form. Not only is there a wide diversity of urban forms (e.g.: monocentric, polycentric, sprawl or linear) (Marshall, 2005; Nabielek, et al., 2016), but also the spatial growth patterns are manifold (e.g.: expansive, compact, fragmented, ribbon, dispersed, etc.) (Camagni et al., 2002; Wilson et al., 2003; Marshall, 2005). They are conditioned by the history of territorial development, form, topography, geography, economic and social development, land use policies, etc. (Schneider and Woodcock, 2008; European Union, 2016). Both urban form and growth patterns affect population well-being in many aspects. Up until now it has been demonstrated that they influence transport systems, commuting choices (Song et al., 2017), energy consumption (Chen et al., 2011), air quality, and health (Hankey and Marshall, 2017), among many other factors. In this context, land-use planning plays an important role in the growth of urban areas by reshaping their form and function (UN-Habitat, 2015).

Urban growth patterns are of particular interest since they have a diverse impact on environmental, social, and economical aspects (Williams, 2014; European Union, 2016; Zhao et al., 2016; Wei and Ewing, 2018). Some examples in this regard are how compact growth usually improves accessibility to employment and services but increases housing prices; unlike the creation of new satellite agglomerations that normally provide affordable housing but increase car commuting due to their mono-functional use (Williams, 2014). Infectious disease outbreaks have been linked to anthropogenic land use changes. For example, fragmented growth and road building were related to phenomena such as deforestation and habitat fragmentation in some areas, which increase the risk of disease through a higher interaction among pathogens, vectors, and hosts (Patz et al., 2004). Therefore, the more informed decision-makers and planners are, the more efficient their actions will be and the better the sustainability of urban areas will be addressed. Thus, the identification of patterns accommodating growth is the first step to analyze their suitability (UN-Habitat, 2013).

Urban growth has been characterized using a diversity of GIS methods. Reis et al. (2015) compiled from the literature an extensive list of spatial metrics used for characterizing and quantifying urban growth, outlying that some of them may vary with growth context and spatial scales. Tian et al. (2011) described the spatial growth patterns of five urban areas by means of urban growth rate, size distribution and spatial metrics, using these values to describe growth

patterns as a diffuse or coalescent growth process. Other studies used a straightforward index to quantify the adjacencies between urban and newly urban patches, categorizing them into infill, edge-expansion and outlying growth types (Liu et al., 2010; Shi et al., 2012), which may serve as a basis for a growth pattern classification. Jiao (2015) proposed different indicators to characterize urban growth, measuring urban land density decline, urban compactness, expansion rate and degree of sprawl. Two recent studies proposed new methods for characterizing urban growth. The first combines spatial expansion dynamics with urban forms (Shi et al., 2017). The second study combines spatio-temporal metrics with the imbalance between population and urban growth (Sapena and Ruiz, 2019). Although there are several spatial patterns of urban growth, most studies focus on various degrees of compact-sprawl growth (e.g.: Tian et al. (2011), and Jiao (2015) classified urban growth processes as compact, sprawl or intermediate phase). In order to optimize urban development management plans, the establishment of a methodology for the identification of additional urban growth patterns, beyond compact-sprawl, is recommended to forecast mid- and long-term urban growth patterns, and the potential impacts on the surrounding ecosystems (Shi et al., 2012). Spatio-temporal metrics have been largely used to characterize urban growth. However, studies analyzing the relevance of spatio-temporal metrics to classify a diverse and well-defined set of urban growth patterns have not yet been conducted.

Despite the growing number of available satellites providing images worldwide, there is still scarce availability of long time-series and fine scale LULC databases, which are the basis for monitoring urban growth and evaluating their spatial patterns. Hence, the use of simulation strategies, such as land-use/land-cover change (LULCC) simulation models, may provide alternative data sources, creating synthetic and diverse urban scenarios based on different priorities and policies (Van de Voorde et al., 2016; Liang et al., 2018). These LULCC statistical models are spatial and location-based computational approaches that reproduce the dynamics of geographical features, considering a wide range of factors as change drivers (Tong and Feng, 2019). LULCC models have been used for evaluating planning policies based on multiple growth scenarios. For instance, Van de Voorde et al. (2016) and Ustaoglu et al. (2018) simulated alternative scenarios under different planning strategies to foresee their implications and to use as a tool for planning cities accordingly, whereas Dorning et al. (2015) and Sun et al. (2018) simulated different growth scenarios based on various planning strategies to assess the effectiveness of regional natural resource conservation plans and to explore optimal strategies for improving ecosystem services. Additionally, Hoymann and Geotzke (2016) evaluated the effect of policy measures to

mitigate climate change and developed new strategies based on simulated urban development scenarios. Overall, scholars have relied on LULCC models to propose, test and validate methodologies that aim to reverse unsustainable trends in cities (Musa et al., 2017; Pickard et al., 2017).

Within this context, we compiled the urban forms and growth patterns described in the literature in order to explore whether spatio-temporal metrics derived from LULC databases can be used to identify urban growth patterns. Therefore, the main purpose of this study is to identify a significant subset of spatio-temporal metrics for the identification of different urban growth spatial patterns, evaluate them in a diversity of baseline urban forms, and assess the influence of the initial urban form in the identification of such patterns.

5.2. Materials and methods

The methodology followed is intended to be adapted to data availability, particularly in the near future, as well as the repeatability and suitability of the classification approach. Thus, LULC databases are being increasingly available locally and globally due to international EO and mapping programs (Table 1.1). Spatio-temporal metrics can be systematically extracted from these databases and used as input into clustering methods, allowing for a straightforward classification of patterns. Figure 5.1 summarizes the overall methodology followed. First, we describe the urban forms and growth spatial patterns used in this study. Then, we select four urban areas that represent these urban forms (Figure 5.1.1), and apply a LULCC model for simulating five long-term urban growth patterns from the baseline forms (Figure 5.1.2 and 5.1.3). This provides a wider range of possible scenarios to evaluate the metrics and overcomes the lack of long-term LULC series. Afterwards, the extraction and selection of spatio-temporal metrics for every simulated scenario are described. We focus on metrics that quantify aggregation, spatial relations, and their variations, extracting an exhaustive set of spatio-temporal metrics followed by the objective selection of the most significant based on Pearson's correlation and discriminant analysis (Figure 5.1.4 and 5.1.5A). Then, growth patterns are classified using the spatio-temporal metrics by means of an unsupervised clustering method (Figure 5.1.5B), and the results are interpreted and described, including the influence of the initial urban forms in identifying growth classes (Figure 5.1.5C).

5.2.1. Definition of urban forms and growth spatial patterns

In this paper, we consider the urban form to be the static physical configuration of the urban cover. We consider four theoretical urban forms
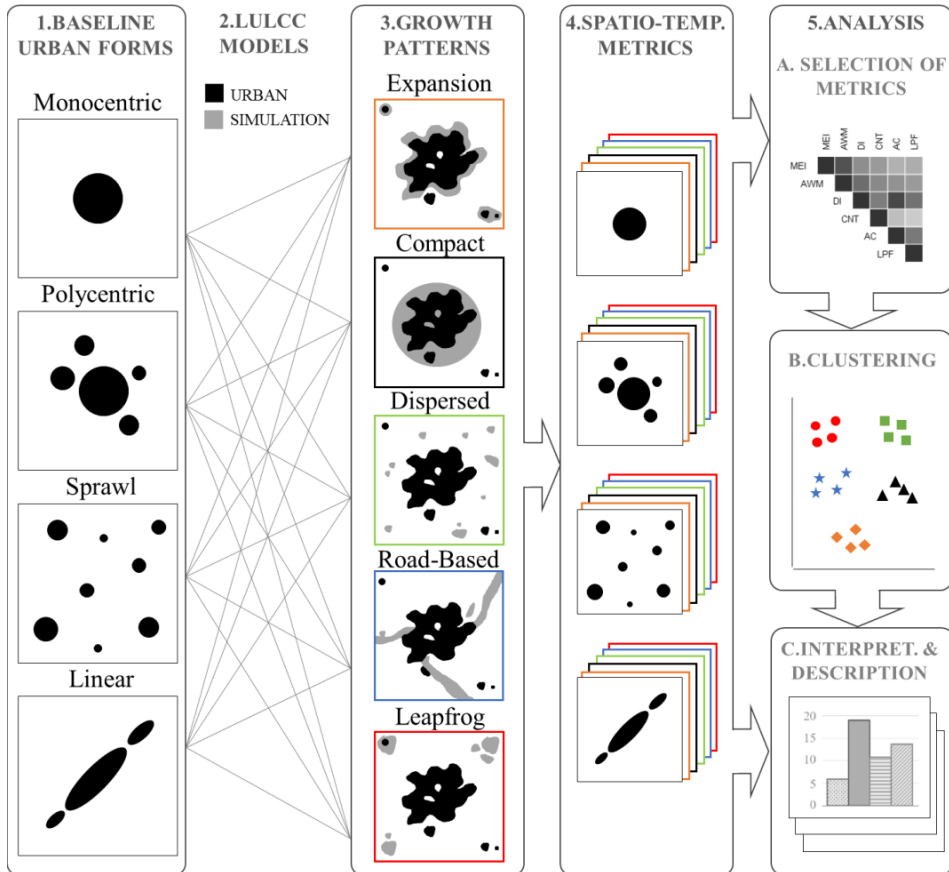
Figure 5.1. Workflow of the methodology: (1) Definition and selection of four initial urban areas having four different urban forms. (2) Application of the land-use/land-cover change (LULCC) model for the simulation of (3) five urban growth spatial patterns. (4) Computation of spatio-temporal metrics for the twenty pairs of baseline-growth simulated scenarios. (5A) Selection of a subset of significant metrics, (5B) classification of growth patterns using the metrics, and (5C) interpretation of results.

described in the literature (ESPON, 2005; Marshall, 2005; Taubenböck et al., 2014; Georg et al., 2016; Nabielek et al., 2016; Salvati et al., 2016; Wei and Ewin, 2018): Monocentric, polycentric, sprawl, and linear (Table 5.1). Regarding the different urban growth patterns defined in the literature (Camagni et al., 2002; Chin, 2002; Wilson et al., 2003; Marshall, 2005; Schneider and Woodcock, 2008; Terando et al., 2014; Georg et al., 2016; Salvati et al., 2016; Wu et al., 2016) we categorized them into five patterns: expansion, compact, dispersed, road-based, and leapfrog (Table 5.1). It must be considered that both, form and growth pattern defined, are pure

theoretical prototypes and they are often combined in real urban areas.

Table 5.1. Name and description of urban forms and growth spatial patterns that are combined by means of a LULCC model.

Name	Description	References
Urban form		
Monocentric	A highly-dense urban settlement spreads over a wide area, density decreases as the distance to the city center increases. Consists of a dominant city and several dependent cities or towns.	(ESPON 2005, Marshall 2005, Georg *et al.* 2016, Nabielek *et al.* 2016, Salvati *et al.* 2016)
Polycentric	It consists of a single functional unit formed by compact subcenters that are well connected, close to each other and consolidated around the main city.	(Marshall 2005, Georg *et al.* 2016, Nabielek *et al.* 2016, Salvati *et al.* 2016)
Sprawl	It is formed by a few relatively small settlements scattered and separated by long distances with low urban densities. Usually characterized by mono-functional land uses.	(ESPON 2005, Marshall 2005, Nabielek *et al.* 2016, Georg *et al.* 2016, Wei and Ewin 2018)
Linear	An elongated urban agglomeration. Usually follows the shape of physical restrictions such as transport routes, rivers, coastlines or valleys. It may not have an obvious center.	(Marshall 2005, Georg *et al.* 2016, Nabielek *et al.* 2016)
Urban growth spatial pattern		
Expansion	It increases the built-up area from the boundaries of the urbanized area, fostering a greater extension of the urban cover. Some authors named it edge-expansion, edge or fringe growth.	(Camagni *et al.* 2002, Wilson *et al.* 2003, Marshall 2005, Terando *et al.* 2014, Wu *et al.* 2016)
Compact	This pattern fosters a more compact urban form by processes such as densification, coalescence, intensification or infilling among disconnected urban patches. Also called land recycling or re-used land, such as barren land growth.	(Camagni *et al.* 2002, Wilson *et al.* 2003, Marshall 2005, Schneider and Woodcock 2008)
Dispersed	When low-density urban growth occurs out of the city boundaries in a scattered form, it is a process of decentralization and suburbanization; some authors relate it to unplanned or spontaneous urban growth. It is also known isolated, outlying, discontinuous, diffuse, sprawl, fragmented or scattered growth, among other terms.	(Camagni et al. 2002, Wilson et al. 2003, Marshall 2005, Schneider and Woodcock 2008, Terando et al. 2014, Salvati et al. 2016)
Road-based	The urban growth takes place along linear structures such as highway or railway axes, also called ribbon, strip, and linear branch growth.	(Camagni *et al.* 2002, Wilson *et al.* 2003, Marshall 2005, Terando *et al.* 2014, Georg *et al.* 2016, Salvati *et al.* 2016)
Leapfrog	When secondary new centers emerge at different distances from the inner city with vacant land interspersed. It can be found as cluster or new satellite agglomerations. It is usually large, compact and dense growth.	(Camagni *et al.* 2002, Chin 2002, Wilson *et al.* 2003, Marshall 2005, Salvati *et a*l. 2016)

5.2.2. Data

Four functional urban areas were selected as working data. The selection criteria were: (i) diversity: they represented different urban forms; (ii) extent: they had similar areas; and (iii) availability: they were available in the Urban Atlas database (EEA, 2016b). After a thorough visual review of the database and based on analyses of external studies, as referenced below, we selected the following urban areas (Figure 5.2): (a) Dijon, France, as an example of monocentric agglomeration, according to Baumont et al. (2004). (b) Manchester, United Kingdom, as a conglomeration formed by the coalition of several cities originally separated (polycentric), fused later to form a continuous urban area (ESPON, 2005). (c) The region of Passau, Germany, identified as ex-urban sprawl growing in non-protected semi-rural areas in a discontinuous way (Siedentop and Fina, 2010). (d) Innsbruck, Austria, shows a linear pattern following the topography of the main valleys (Krakover and Borsdorf, 2000). These areas were selected not as study cases, but as a representation of the four different spatial urban forms defined, providing the baseline for the analysis of potential growth scenarios.
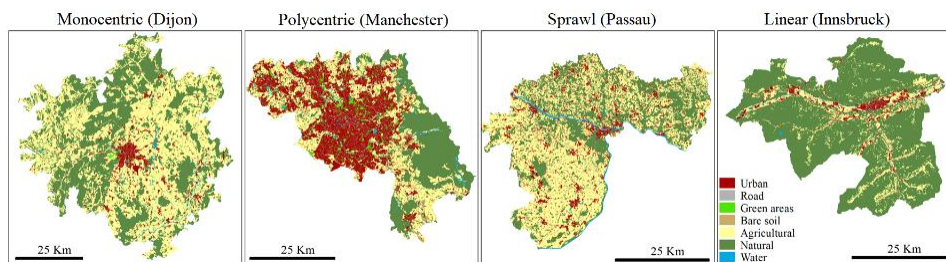
Figure 5.2. The four urban areas representing different baseline urban forms (the name of the functional urban areas is in parenthesis). Source: Urban Atlas 2012 (EEA 2016b), with an aggregated legend.

For simulating growth scenarios, we used LULC data from the Urban Atlas dataset for the year 2012. Digital elevation models (EU-DEM) from the land-monitoring services of Copernicus (https://land.copernicus.eu/imagery-in-situ/eu-dem) (25 meter/pixel), and location of city centers from Eurostat (https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data) were also used.

5.2.3. Land-use/land-cover change models

Even though image classification techniques in remote sensing are continuously improving, we still lack high-resolution long-term time series of LULC datasets. A promising initiative in Europe is the Urban Atlas dataset. However, the period of six years is still insufficient for detecting reliable growth

patterns. Therefore, we aim to create longer LULC time-series using the LULCC simulation models.

We reviewed the most widely used LULCC models and analyzed their suitability for evaluating spatio-temporal metrics under different growth scenarios on urban environments. In general, LULCC models are based on known past LULC change trends and a number of triggering factors, which are used to estimate the probability of change from one land use to another for each pixel or object. This probability is used in the allocation process, where limited by the demand or amount of expected change between land uses, those pixels with higher probability of change are converted to a new land use until the demand is reached (Sapena et al., 2017). There are different types of LULCC models: those that consider political and population decisions are agent-based (Jokar Arsanjani et al., 2013); those that use the neighborhood of a cell to determine whether a change will happen are based on cellular automata (Torrens, 2009); those based on the extrapolation from the past are statistical; as well as models that use other algorithms or combine them (Musa et al., 2016). With regards to the triggering factors, they can be geographic, economic, social or biological, and static or dynamic, if updated in each iteration (Rosa et al., 2014). Regarding the allocation process, this can be deterministic, such as the top-down ranking where the pixels are sorted from highest to lowest probability of change and the land use is changed according to this order having a fixed result (Engelen et al., 2007), or stochastic, as the Monte Carlo method, which compares the transition probability with a random number, from 0 to 1, and if the probability is greater than the number then the pixel changes, having multiple possible results (Meentemeyer et al., 2013). Table 5.2 shows a summary of the most used and innovative models, with particular emphasis on the type of input data and the methods applied.

Table 5.2. Description of LULCC simulation models focused on the input data, triggering factors, algorithm and allocation methods. For the input data (T) is the minimum number of dates required. Digital terrain model (DTM) and land-use/land-cover (LULC)

LULCC Model	Input data and factors	Algorithm	Allocation
SLEUTH (Clarke et al., 1997)	Urban mask (4T) LULC map (2T) Road network (2T) Slope Hillshade Constraints	Cellular automata. Transition rules: suitability map, accessibility, Moore neighborhood, stochastic perturbation, and coefficients (diffusion:breed:spread:road:slope)	Monte Carlo iterations (frequency of change). Threshold to set the number of pixels to change to urban.
DINAMICA (Soares-Filho et al., 2002)	LULC map (2T) Slope and DTM Soil Road network	Cellular automata and logistic regression. Transition rules: probability map, Moore neighborhood, stochastic	Demand of growth. Monte Carlo and top-down ranking iterations.

		perturbation.	Allocates seed pixel and then the patch.
CLUE-S (Verburg et al., 2002)	Distance to elements Dynamic variables LULC map (2T) Population Distance to elements Slope and DTM Constraints	Patch-based: set size. Constrained cellular automata and stepwise logistic regression. Transition rules: constraints, conversion matrices, probability map.	Demand of growth. Dynamic ranking allocating the LULC with higher probability.
FORE-SCE (Sohl et al., 2007)	LULC map (1T) Slope and DTM Soil index, temperature, precipitation. Distance to elements Population density	Stepwise logistic regression Patch-based: set size. Transition rules: transition matrices, probability map, fragmentation trends.	Demand of growth. Monte Carlo by LULC to avoid repetition. Allocates seed pixel and then the patch.
MOLAND (Engelen et al., 2007)	LULC maps (1T) Socio-economic variables Distance to roads Slope Constraints Regions	Constrained cellular automata. Transition rules: suitability, accessibility, zoning, von Neumann neighborhood, constraints, stochastic perturbation.	External LULC demand per region. Top-down ranking.
StocModLCC (Rosa et al., 2013)	Binary masks (2T) Historical maps Soil Protected areas Distance to elements Socio-economic variables	Stepwise logistic regression	Monte Carlo iterations (frequency of change).
FUTURES (Meentemeyer et al., 2013)	Binary mask (1T) LULC map (1T) Distance to elements Slope and DTM Road network Other predictors Administrative units Dynamic variables Constraints Population projection	Cellular automata and logistic regression. Transition rules: probability, accessibility, urban pressure (kernel, occurrence or gravity), incentive power (set infilling – sprawl degree). Patch-based: calibrate size and shape.	Demand based on population projection or set by user. Monte Carlo or top-down ranking. Allocates seed pixel and then the patch.
Lulcc (Moulds et al., 2015)	Binary mask (1T) Distance to elements Slope Constraints	Logistic regression, random forest, recursive partitioning and regression trees. Transition rules: suitability, different neighborhood, constraints.	Demand of growth. Dynamic ranking allocating the use with higher probability, top-down ranking and Monte Carlo.
MachCA (Feng et al., 2016)	Binary masks (2T) Distance to elements Constraints	Cellular automata and least squares support vector machines (LS-SVM) Transition rules (updated in each iteration): Moore neighborhood, constraints, stochastic perturbation.	Without demand. The allocation depends on the binary classifier of the LS-SVM.

CA-MAS (Tian et al., 2016)	Binary masks (2T) Distance to elements Slope and DTM Road network Constraints	Cellular automata, logistic regression, and multi-agent system. Transition rules: neighborhood, constraints, stochastic perturbation, probability based on landscape and human systems.	Demand of growth. Stablished demand of change. Monte Carlo.
GeoSOS-FLUS (Liu et al., 2017)	LULC maps (1T) Socio-economic, Climate and population variables Distance to elements Road network Slope and DTM	Cellular automata and artificial neural networks. Transition rules: neighborhood, probability of occurrence of each land use.	Demand using top-down system dynamics and population, economy, climate and land use variables. Allocation using roulette selection, self-adaptive inertia coefficient and conversion cost for each LULC.

To evaluate the suitability of the models for the creation of alternative scenarios, the following characteristics were considered: input data; simplicity of implementation; possibility of modeling several land uses; capability to alter dispersion or compact degree; whether it is patch-oriented; and the adaptation of the result to reality (Sapena et al., 2017). In this context, the FUTURES model (FUTure Urban-Regional Environment Simulation) was the most suitable model to simulate long-term urban growth spatial patterns from different baseline forms. Among other facts, creating alternative synthetic growth scenarios by altering a few factors is feasible using this model (i.e.: density, infilling and constraints) (Dorning et al., 2015; Sapena and Ruiz, 2018).

5.2.3.1. FUTURES LULCC model and factors of urban growth

The model FUTURES is a cellular automata, stochastic and patch-based LULCC model based on the logistic regression method, and was implemented in GRASS GIS (Petrasova et al., 2016; GRASS development team, 2017). It requires an urban mask and geographic, economic, and social factors that determine where growth is likely to occur. We used FUTURES because it allows the variation of the sprawl degree in the simulation, as well as the modification of several factors, constraints (limiting growth in specific areas, e.g.: subject to political decisions) and stimulus (boosting growth in specific areas, e.g.: subject to land-use planning). This high adaptability facilitates the creation of alternative growing scenarios (Sapena and Ruiz, 2018).

The LULC data were rasterized to 10-meter pixel size for the simulation. Accordingly, the EU-DEM was resampled using bilinear interpolation. From these datasets, we calculated several factors as possible predictors of urban

growth (Figure 5.3). The proximity to specific geographical elements may contribute to the development of new buildings, for instance, due to resident preferences to live in residential areas, close to the business district, with a good accessibility, nearby gardens or leisure areas, etc. These social and economic factors are included using the Euclidean distance to residential, commercial and industrial buildings, city center, road network, green urban areas, leisure areas, agricultural plots, or natural areas, all extracted from the legend of the Urban Atlas and Eurostat (see the legend in Figure 5.3). Similarly, under the assumption that development stimulates more development in near proximity, we computed three different types of development pressure based on the distance-decay effect, following the equation (5.1) from Meentemeyer et al. (2013):

$$\text{Pressure}_j = \sum_{i=1}^{n_j} \left(\frac{\text{pixel}_i}{d_{ij}^{0.5}} \right) \tag{5.1}$$

where j is the evaluated pixel, n_j is the number of pixels in the neighborhood from pixel_j within a given radius, pixel_i is a binary variable that indicates if the ith pixel from the neighborhood is urban (1) or non-urban (0), and d_{ij} is the distance between pixels i and j.

We calculated the urban pressure within a radius of 1 km, the road network pressure within 0.5 km, and the urban nucleus pressure within 5 km. The urban nucleus was defined as the biggest urban cluster when combining all urban plots within a distance of 200 m, based on the concept of urban morphological zones defined by the EEA (2014). Since topographic conditions may limit or ease urban growth, we included elevation and slope factors extracted from the EU-DEM. Finally, two additional factors were included: the constraint and the stimulus. The constraint limits the urban growth in specific areas, in our case roads, water bodies and green urban areas, since they may be protected or have low probability of change. The stimulus encourages urban growth in specific areas, such as boosting centralized growth and land-recycling from barren land patches (compact growth) or stimulating growth around the emergence of new centers (leapfrog growth). All the aforementioned factors were scaled to a range from zero to one to avoid the influence of the measurement units (Figure 5.3).

5.2.3.2. Simulated urban growth patterns

Based on the growth patterns described above, five growth scenarios were simulated:

> Expansion growth represents an expansion of the existing urban cover
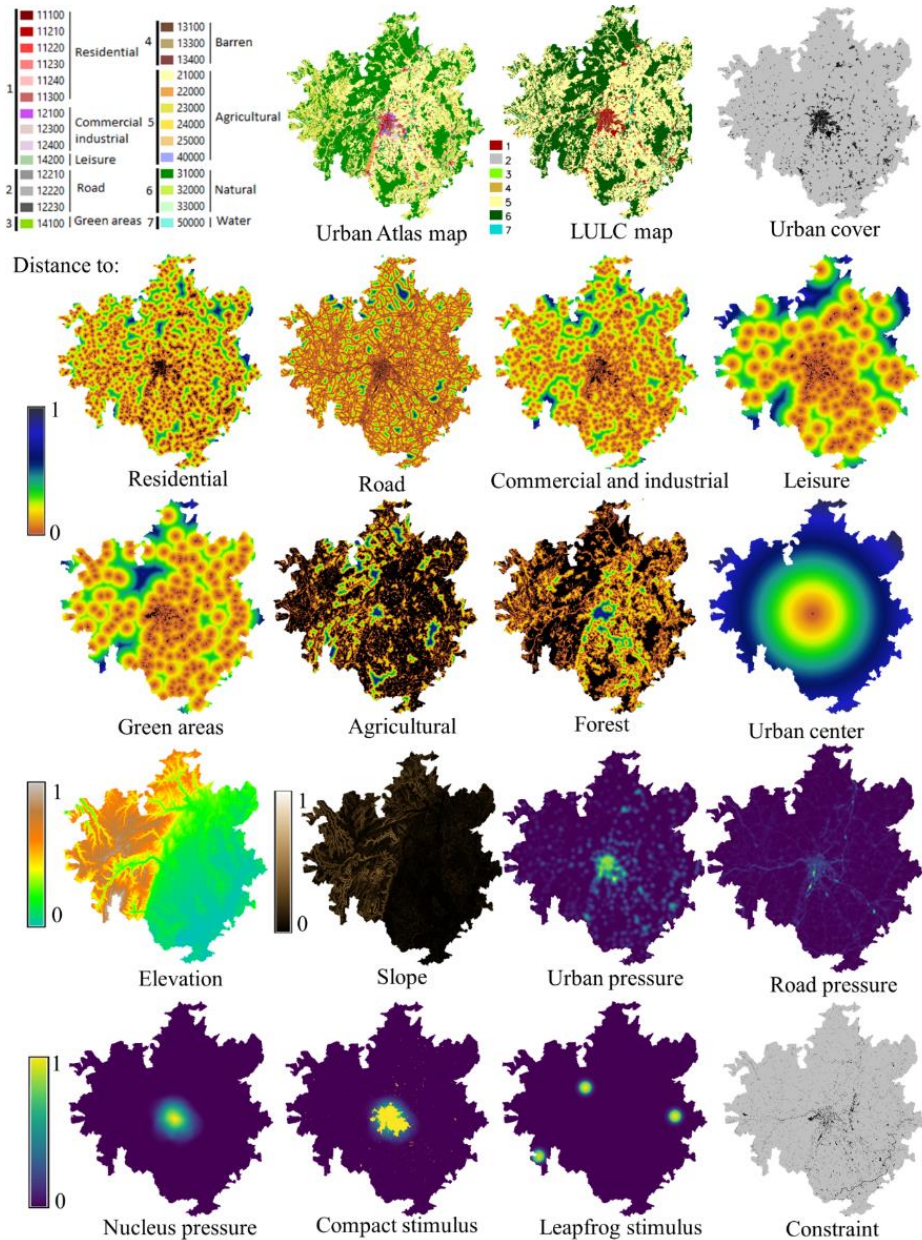
Figure 5.3. Example of factors computed for the monocentric form (Dijon). On the top left, reclassification of Urban Atlas legend (five digits, see the interpretation in https://land.copernicus.eu/user-corner/technical-library/urban-atlas-2012-mapping-guide-new/) into nine classes for computing factors (distances to, pressures, stimulus and constraint), and into seven classes for creating the reference LULC map.

from the urban fringe.

> *Compact growth* encourages infill growth and land use recycling, prioritizing the urbanization of open land inside urban areas, near the urban nucleus, and bare soil.

> *Dispersed growth* follows a scattered, isolated and uncontrolled urban growth beyond developed areas.

> *Road-based growth* occurs when the urban growth takes place near the road network.

> *Leapfrog growth* creates new urban centers at a considerable distance from the developed area.

In order to have different scenarios simulating urban growth in different pathways, we computed twenty different models, one for each combination of baseline forms (monocentric, polycentric, sprawl and linear) and simulated growth (expansion, compact, dispersed, road-based and leapfrog). The simulation steps were:

(1) Training the logistic regression model with five percent of the study area, using the urban cover as a dependent variable and the factors as independent variables (5.2). The set of factors included in the model was adapted to the simulated growth pattern. In this first step, factors were pre-selected depending on the targeted simulated pattern (see Figure 5.4).

$$\log\left(\frac{P}{1-P}\right) = \beta_0 + \sum_{i=1}^{n} \beta_i \cdot x_i \tag{5.2}$$

where P is the potential or probability of a pixel to be or become urban, β_0 is the constant of the model, β_i is the coefficient of factor x_i, and *n* is the total number of factors.

(2) Retraining the model by discarding those factors not statistically significant (*p-values* > 0.05) (see Figure 5.4).

(3) Applying the trained model to the total study area, predicting the probability of becoming urban for each pixel. This output is called potential (P) based on (5.2).

(4) The potential can be modified by the incentive parameter (IP) that applies a power function to transform the probability gradient into the new potential (P^{IP}). This transformation increases or decreases the probability of urban growth by altering site suitability, allowing the model to encourage compact or dispersed growth patterns (Figure 5.5).

(5) Calibration of urban patches, a list of sizes and shapes is stored and will be used in the patch allocation process.

Factors			Scenarios																				Potential
	Urban form		Monocentric					Polycentric					Sprawl					Linear					
	Urban growth pattern		E	C	D	R	L	E	C	D	R	L	E	C	D	R	L	E	C	D	R	L	
	Distance to	Residencial	■																				
		Road		·	■																		
		Commercial		·																			
		Leisure	·																				
		Green areas	·										·	·				·	·				
		Agricultural	·					·	·														
		Natural	·					·					·					·					
		Urban center						·	·				·	·				·	·				
	Topography	Elevation											·	·				·					
		Slope																					
	Development pressure	Urban	■										■										
		Road		·	■									·					·				
		Nucleus		■										■					■				
	Stimulus	Compact		■																			
		Leapfrog									■												
	Constraints	Excluded																					
	Incentive parameter (IP)		1	3.5	0.1	1	1	1	4.0	0.5	1	1	1	2.5	0.5	1	1	1	2.5	0.5	1	1	
	Demand (n° of pixels)		50% (992,620)					25% (1,476,156)					50% (785,215)					50% (411,259)					

Figure 5.4. (Left) Factors used in the LULCC model are in gray (rows) for the twenty simulated scenarios (columns). White grids were not included in the model, while the dot (·) means that the factor was not statistically significant. Growth patterns: (E) expansion, (C) compact, (D) dispersed, (R) road-based, and (L) leapfrog. The incentive parameter and the demand are shown on the last two rows. (Right) Example of the patch allocation steps.

(6) Defining the demand based on the spatial area of growth instead of the population growth or time span. We established that fifty percent of the total urban area is developed, except for the polycentric urban area where twenty-five percent was established; due to the fact that this urban area was initially highly developed (the demand of growth in number of pixels is shown in Figure 5.4).

(7) Iterative allocation of growth using the Monte Carlo method until the demand is achieved. First a potential seed is located. Then, based on suitability of contiguous pixels and a random size and shape from the calibration step list, the patch is finally allocated (Figure 5.4, right).

5.2.4. Computing spatio-temporal metrics

We computed twenty-four spatio-temporal metrics related to the aggregation and spatial distribution of LULC change for each simulated growth scenario, using IndiFrag. We conducted a correlation analysis of the spatio-temporal metrics to discard those metrics with a strong correlation (Pearson's $r \geq 0.8$) and avoid redundancies in the spatial information. As a result, only the eleven metrics described in Table 5.3 were used. They were computed using the initial LULC maps and the final simulated scenarios.
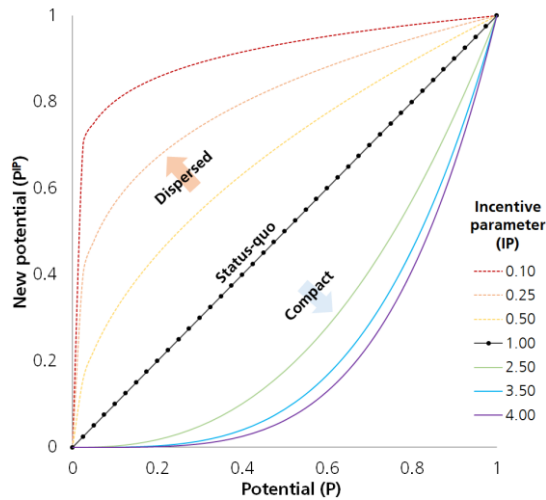
Figure 5.5. Transformation of the potential (P) by varying the incentive parameter (IP). Values (IP < 1) increase the potential for change by making the model more flexible and more dispersed. Values (IP > 1) reduce the potential by restricting the model and resulting in a more compact model. When IP = 1, the model follows the trend calculated according to the statistical model (Sapena and Ruiz, 2018).

Table 5.3. Description of the selected spatio-temporal metrics. Formulas (Eq.) can be found in the Appendix A.

Metric	Description	Eq.
Leapfrog (LPF)	Proportion of isolated urban patches.	(A.18)
Porosity (P)	The ratio of open space compared to the total land cover area.	(A.8)
Weighted Euclidean Distance (AWSD)	Concentration degree relative to the centroid.	(A.19)
Mean nearest neighborhood distance (ENND)	Aggregation degree. Mean distance between nearest patches.	(A.21)
Compactness (C)	Shape complexity of the urban cover.	(A.27)
Radius dimension (Dim_R)	Centrality of the urban cover with respect to the urban center.	(A.26)
Effective mesh size (EMS)	Measures connectivity. Lower values mean more fragmentation.	(A.22)
Splitting index (SPL)	The number of patches when dividing the cover into equal size parts with the same division.	(A.24)
Area-weighted mean expansion index (AWMEI)	Weighted growth compactness. Adjacencies between new urban patches and the urban cover.	(A.43)
Disaggregation index (DIS)	Mean distance from new urban patches to the closest patch of the urban cover.	(A.49)
Centroid displacement (CNT)	The distance between the geometrical centroid of the urban cover at two different times.	(A.50)

5.2.5. Urban growth spatial pattern classification

In order to harmonize the differences in units, the values of the metrics were standardized to mean zero and standard deviation one. From the pre-selected metrics (Table 5.3), a supervised stepwise linear discriminant analysis was applied to select the best combination of metrics for classification. In this method, all variables are progressively reviewed and evaluated at each step to determine which will contribute most to the discrimination between classes, that variable is included in the model and the process is iterated (Hermosilla et al., 2012). As a result, the most discriminant metrics were: the weighted mean expansion index (AWMEI$_{urban}$), the variation of the weighted Euclidean distance (AWSD$_{urban}$), the disaggregation index (DIS$_{urban}$), and the change in the compactness degree (C$_{urban}$), all referred to the urban cover. Starting from these metrics, the classification of urban growth patterns was performed by means of the unsupervised k-Means Clustering method. This is an iterative algorithm that divides the m observations (twenty scenarios) in n dimensions (four spatio-temporal metrics) into k groups (five growth patterns) until the within-group sum of squares is minimized (Hartigan and Wong, 1979). Therefore, data were classified into five clusters that were interpreted and assigned a growth pattern class. The result was evaluated using the confusion matrix and its derived indices: the overall accuracy, and the omission and commission errors of the classification.

Finally, in order to assess how the baseline urban form influences the classification of growth patterns, three outputs were compared: the classification error rates per urban form, the behavior of metrics per urban form using graphs and the interpretation of concentric and sector growth graphs. These graphs provide a description of the urban growth according to the distance and direction to the city center. For the polycentric city the centroid of the ten city centers was used. In the distance analysis, the interval was set to 2.5 kilometers and an angle of 22.5 degrees to define the sectors of the graphs.

5.3. Results

5.3.1. Categorization of urban growth spatial patterns

As a result of the LULCC modeling we created twenty growth scenarios, whose distinctive features are shown in Figure 5.6. These scenarios recreate the behavior of five growth patterns that may happen in a developing area, and how these patterns will progress in different urban areas with specific baseline urban forms. The spatio-temporal metrics were extracted from these
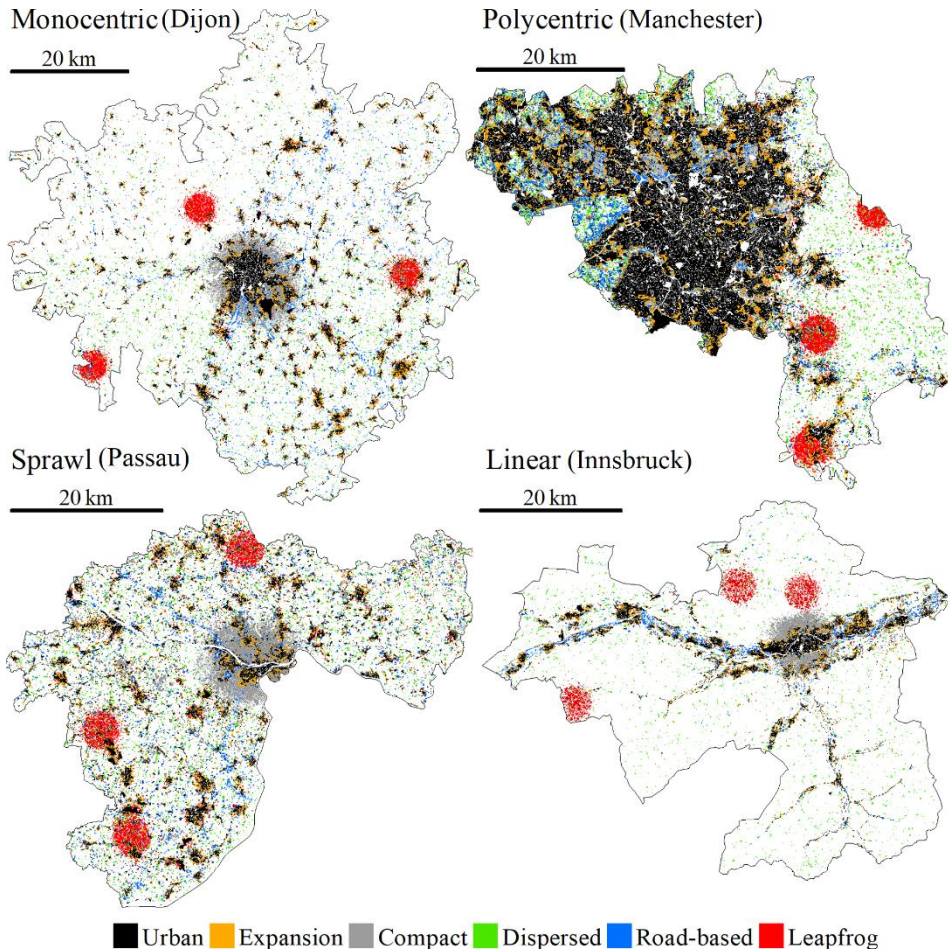
Figure 5.6. Simulated growth scenarios from the four baseline urban forms following five urban growth spatial patterns. The baseline urban covers are shown in black, while growth patterns are shown in different colors.

scenarios, graphically represented in Appendix B (Figure B.11 to Figure B.15).

Figure 5.7 shows the distribution of scenarios by means of the standardized values of the selected spatio-temporal metrics ($AWMEI_{urban}$, DIS_{urban}, C_{urban}, and $AWSD_{urban}$), where the baseline urban forms and growth patterns are represented with a different shape and color, respectively. The distances between growth pattern scenarios in the space represented by each pair of metrics are inversely related to their similarities. Observing the combination by pairs of metrics, the contribution of metrics for the identification of growth patterns can be analyzed from Figure 5.7. $AWSD_{urban}$ adequately discriminates
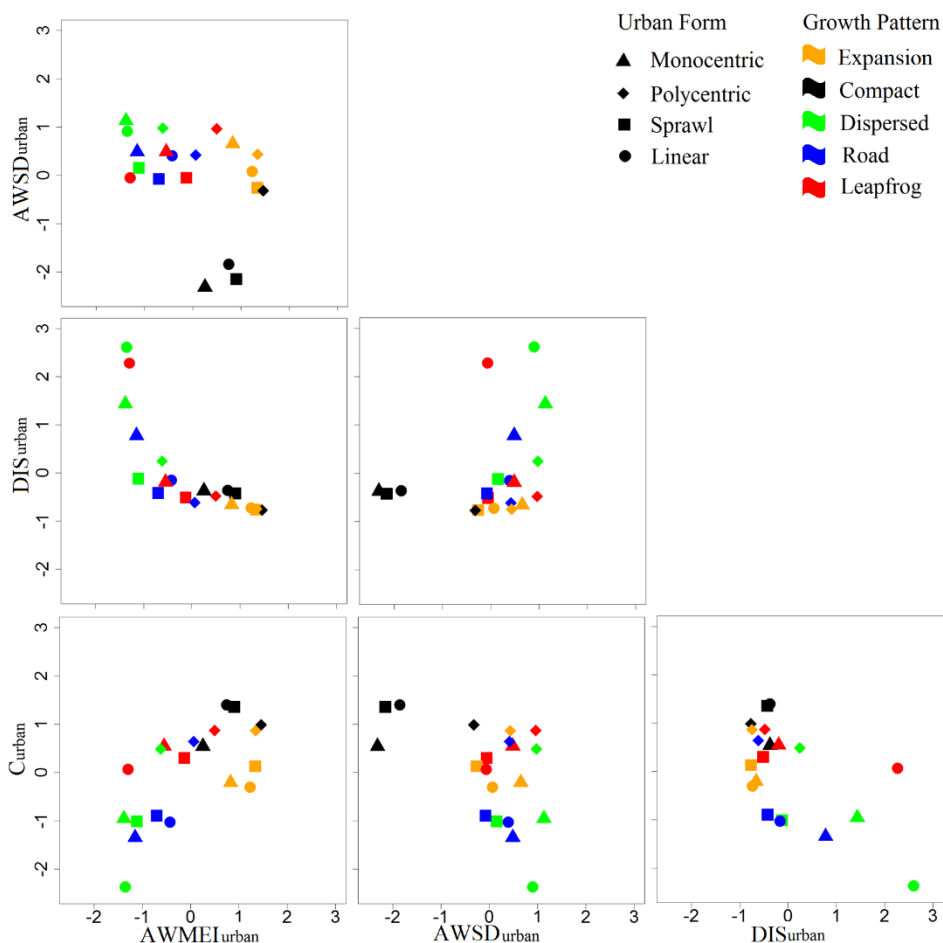
Figure 5.7. The distribution of simulated growth scenarios according to the combination of the standardized values of $AWMEI_{urban}$, DIS_{urban}, $AWSD_{urban}$, and C_{urban} metrics. The color represents the simulated growth pattern, while the symbol is the initial urban form.

the compact pattern and $AWMEI_{urban}$ the expansion, with some exception. The DIS_{urban} separates the dispersed pattern from the rest, even if it is sometimes mixed with other patterns. Finally, C_{urban} helps to discriminate the road-based and disperse patterns from the rest but groups them together. The leapfrog pattern seems to be the most difficult to identify using this subset of metrics.

The classification of urban growth spatial patterns was conducted applying iterative cluster analyses, one for each combination of metrics from one to four. Overall accuracies in the identification of growth scenarios using a single metric ranged from 50% to 60% (with C_{urban} and $AWSD_{urban}$), they quantify

the variation in compactness of the urban cover and its concentration degree. By combining two metrics we reached the highest accuracy in classifying the five growth patterns, with a value of 75%, using $AWMEI_{urban}$ and $AWSD_{urban}$. $AWMEI_{urban}$ enriches $AWSD_{urban}$ with adjacency properties of new urban patches. The addition of the third and fourth metrics did not improve the classification results.

Table 5.4 shows the classification errors of the clustering method for each scenario, using $AWMEI_{urban}$ and $AWSD_{urban}$. The omission error (OE) gives the proportion of underclassification of a pattern, while the commission error (CE) informs about the overclassification. Accordingly, the expansion pattern is the one with highest accuracy, followed by the compact and dispersed that were underclassified in one case. The road-based scenario presents the lowest accuracy, followed by the leapfrog growth, which are intermixed, as seen in Figure 5.7 (upper-left). This response owes to the strong influence that the shape of the road network and the location of the new nuclei have on these patterns and both are related to the baseline form.

Table 5.4. Classification of scenarios into five clusters (color and shapes) using $AWMEI_{urban}$ and $AWSD_{urban}$. Omission (OE) and Commission Errors (CE) are shown per pattern. The Urban Form derived Error (UFE) is the error rate per baseline form. The centroids of the classified clusters are compared against the actual pattern centroids by means of the Euclidean distance in the space defined by $AWMEI_{urban}$ and $AWSD_{urban}$.

	MONO	POLY	SPRA	LINE	OE	CE	Cluster centroid		Pattern centroid		Euclidean distance
							AWMEI	AWSD	AWMEI	AWSD	
EXPA	▲	▲	▲	▲	0	0.2	31.65	0.373	31.07	0.505	0.595
COMP	●	▲	●	●	0.3	0	25.16	-2.310	27.37	-1.772	2.275
DISP	◆	◆	—	◆	0.3	0	6.27	1.450	6.28	1.192	0.258
ROAD	—	■	—	■	0.5	0.5	6.84	0.388	12.35	0.604	5.514
LEAP	■	■	■	—	0.3	0.4	17.13	0.769	14.33	0.640	2.803
UFE	0	0.4	0.2	0.4							

The comparison of the centroids of the classified clusters against the actual patterns shows the highest difference in the growth adjacency ($AWMEI_{urban}$) of the road-based pattern. This is because even if the road-based growth patterns are quite clustered by means of $AWMEI_{urban}$ and $AWSD_{urban}$ they are overlapped by the leapfrog growth (Figure 5.7). Consequently, only two scenarios in this cluster were identified correctly, modifying the $AWMEI_{urban}$ values and displacing the centroid of the cluster to the left. However, as $AWSD_{urban}$ centroids are quite similar for road-based and leapfrog patterns, the differences are the least in this metric (Table 5.4). With regards to the rest, the

centroids are quite similar (Table 5.4).

5.3.2. Influence of the urban form on growth spatial pattern categorization

According to the error rate per urban form (UFE) from Table 5.4, growth patterns derived from the monocentric form were successfully identified. When the sprawl form grew in a dispersed way the algorithm was unable to identify it, as $AWSD_{urban}$ has a different behavior than the other forms and was classified as road-based with a lower mean value of $AWSD_{urban}$. This occurs because the sprawl form is highly dispersed and more dispersion does not substantially influence the concentration degree. Finally, the polycentric and linear forms add uncertainties in the proper identification of growth patterns (Table 5.4).

Figure 5.8 shows the spatial distribution of growth, based on the distance and orientation from the center, according to the baseline urban form of each pattern. As expected, since this type of growth keeps the shape of the urban form, the expansion growth differs widely among baseline forms (also seen in Figure 5.6 in orange color). On the contrary, the compact growth shows similar behavior between urban forms, where the peak of growth is between distances of 5 and 10 kilometers from the city center, with some isolated growth further from the center due to the urban recycling of open and bare land plots. Nevertheless, the polycentric does not have a strong peak, as the urban nuclei is already densely urbanized, and instead it fills the open spaces in a gradual manner, this is why the direction is also different, as the open land is mainly located on the west side of the center, while the rest are located homogenously around the city center. Regarding the dispersed growth, the first distances from the center experiment very slow growth, while the rest are distributed uniformly across the study areas. Besides, considering the boundaries, the directions are quite similar among baseline forms. The direction of the road-based growth depends on the road network, especially noted in the linear urban form, while the other networks are denser and more uniformly distributed. In this case, the distances of growth to the center evidence where the road network is denser. The distances of the growth patches are similar to the dispersed pattern. Lastly, the leapfrog growth has a very clear pattern, since the three new urban nuclei can be easily detected by the distances and orientations to the center. However, the peaks are dissimilar between the different baseline scenarios, as the new nuclei are located randomly throughout the study areas.

When interpreted individually, the adjacency and concentration degrees of urban growth ($AWMEI_{urban}$ and $AWSD_{urban}$) present different responses
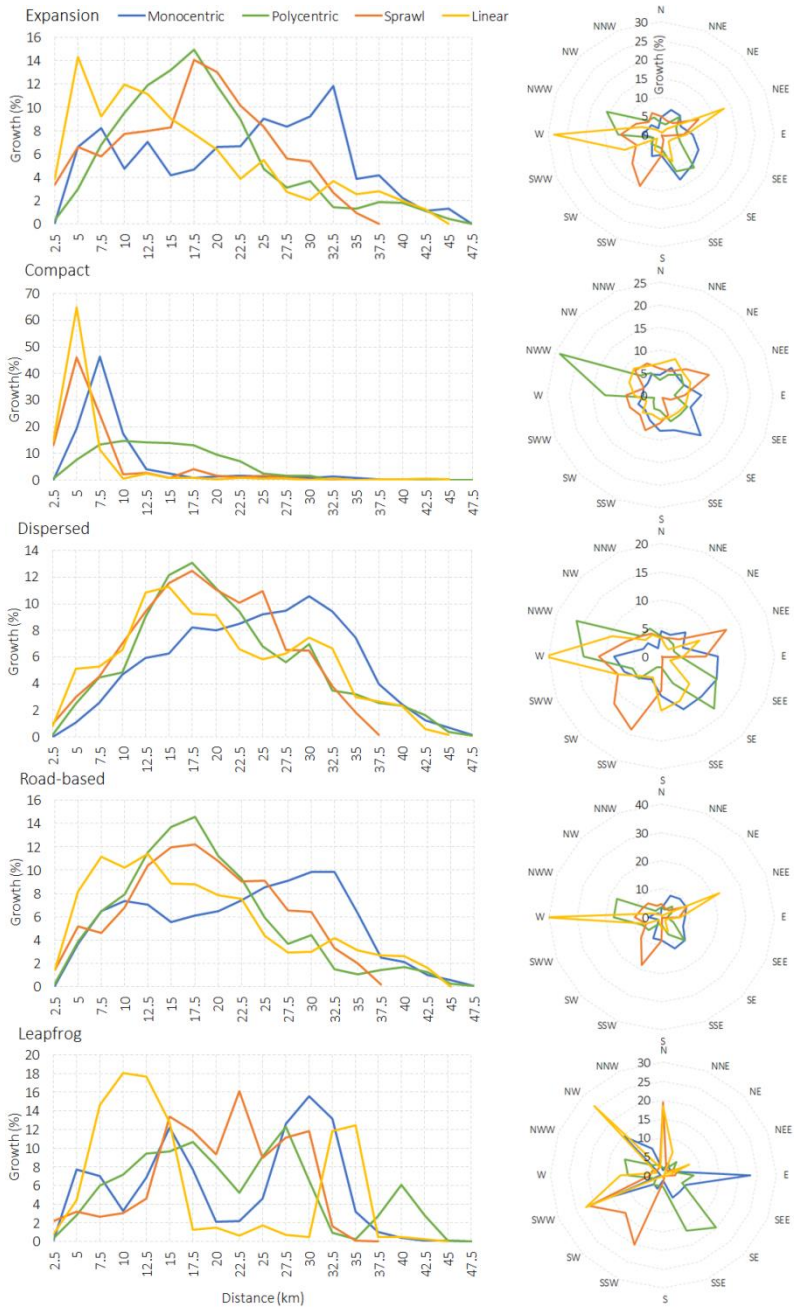
Figure 5.8. Concentric (left) and sector (right) growth graphs. Distribution of growth according to the simulated patterns. The growth is shown as the proportion of the total growth. The distance and orientation are referred to the city center points.

depending on baseline forms (Figure 5.9). The expansion and compact growths have similar values of AWMEI$_{urban}$, but different values of the changes of AWSD$_{urban}$. On the other hand, the leapfrog growth from the linear form had an unexpected value of AWMEI$_{urban}$ compared to the rest of the simulations, as also seen in Figure 5.7. This scenario has the particularity that the hilly areas are not urbanized along with the fact that the simulated leapfrog pattern projected the new urban clusters randomly—in the hilly areas—and, consequently, the adjacency of new urban elements to previous urban areas are much lower compared to the rest of the scenarios. The polycentric form is characterized for being highly urbanized and compact. Therefore, it not only has higher values of AWMEI$_{urban}$ in all patterns, due to higher probabilities of growth adjacent to the urban elements, but also AWSD$_{urban}$ of the compact pattern increases weakly, since there are not open lands within the nucleus, which influences the identification of patterns. Regarding sprawl form, as stated above the already spread urban cover together with new isolated urban patches slightly increase the distance to the centroid. In fact, the changes in AWSD$_{urban}$ in all patterns are quite low with the exception of the compact growth that has a strong impact in this form. These irregular responses of metrics for the scenarios depending on the baseline forms highlight the notable influence that urban form has on the identification of spatial patterns.
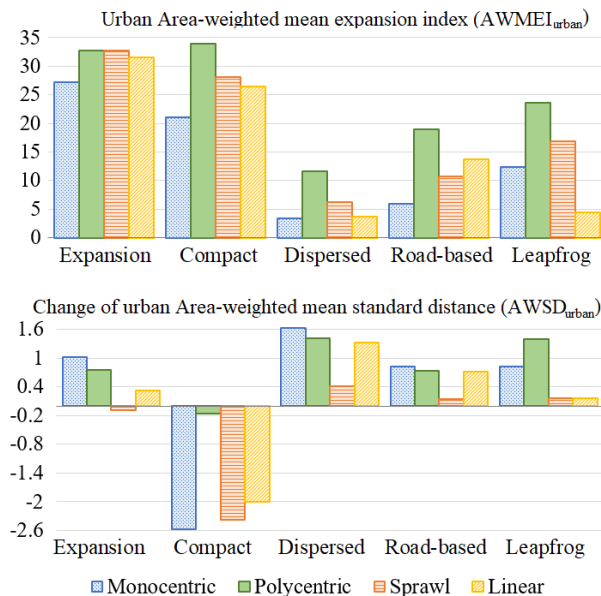
Figure 5.9. Values of the spatio-temporal metrics used in the cluster analysis. Metrics are grouped by growth pattern, and colors represent the baseline forms.

5.4. Discussion

This chapter proposes a methodology for the early identification of five different growth patterns in urban areas based on a significant subset of spatio-temporal metrics derived from LULC maps. To the authors' knowledge, it differs from other studies as it attempts to identify growth categories rather than degrees between compactness and dispersion (e.g.: Tian et al., 2011; Liu et al., 2010; Jiao, 2015).

Our results illustrate that four spatio-temporal metrics are the most explanatory for discriminating between expansive, compact, dispersed, road-based and leapfrog urban growth spatial patterns. Thus, $AWMEI_{urban}$ is particularly helpful to discriminate the expansion growth as it quantifies adjacencies, DIS_{urban} separates disperse growth as it measures distance to old urban patches, C_{urban} detects road-based and dispersed growths as they tend to be less compact in shape, and $AWSD_{urban}$ identifies compact growth since it measures the concentration degree. However, we found that the use of only two spatio-temporal metrics is sufficient to accurately identify the five growth spatial patterns analyzed, which has a practical relevance for their use in monitoring urban growth. The change in the concentration degree of the urban cover ($AWSD_{urban}$) is the metric that individually best identifies patterns. It measures the concentration and fragmentation of urban cover. When combined with the degree of adjacency of the urban growth ($AWMEI_{urban}$), which quantifies urban densification and growth compactness, the identification of patterns improves (with an overall accuracy of 75%). The complementarity of these two metrics allows for describing the main properties for discrimination of urban growth spatial patterns. While the first accounts for the spatial distribution of the new urban elements from the urban center, the second quantifies the level of aggregation of the new growth.

Despite the successful identification of growth patterns using spatio-temporal metrics from LULC maps, the analyses revealed that the urban form influences the response of spatio-temporal metrics. We simulated the same five growth patterns in four urban areas representing different urban forms and observed dissimilar responses in the spatio-temporal metrics for some scenarios. According to our analysis, the polycentric and linear initial urban forms are those which are adding more uncertainty into the classification of growth patterns. The polycentric form is highly urbanized and compact, which makes urban growth slightly more compact for all growth patterns (Figure 5.9). As such, it is expected that polycentric urban forms have higher values of $AWMEI_{urban}$. On the contrary, the monocentric form facilitated the identification of patterns; this might be related to the greater differences in the values of metrics among patterns (larger distances between monocentric

scenarios in Figure 5.7). Therefore, the identification of growth patterns may be more evident in urban areas with monocentric form. The sprawl form showed a generalized low variation of $AWSD_{urban}$ (Figure 5.6 and Figure 5.7), with the exception of the compact growth. Only the dispersed growth was misclassified. However, $AWSD_{urban}$ should be cautiously interpreted when it concerns urban areas with sprawl form. In general, the most challenging spatial pattern to identify was the road-based growth. Based on this, further analyses should include metrics that quantify the distances of new grown patches to the road network. Therefore, when applying spatio-temporal metrics for growth pattern classification, multi-temporal analyses or the comparison of growth patterns in several urban areas, the influence of the urban form should be considered. Approaches to overcome the influence of the urban form in spatio-temporal metrics and in the classification of growth patterns are still required for the identification of growth patterns at a global scale. In this sense, the inclusion of the baseline urban form as a qualitative variable in the classification procedure would be worthwhile to investigate in future research to improve the discrimination of growth patterns.

Besides the influence of urban form, other aspects should be considered. Classification is always complex and some growth patterns can be actually interpreted as combination of others (Camagni et al., 2002; Wilson et al., 2003; Clark et al., 2009). This is the case, for instance, of the leapfrog growth in its initial stages, which can also be considered a dispersed pattern as remote areas are being urbanized (Wilson et al., 2003), but in a longer term, these areas may trigger the transformation from monocentric to polycentric urban form (Salvati et al., 2016). This may be understood as a consolidation process with a compact growth pattern in the long term. Therefore, this complexity may derive in errors when identifying growth patterns, as their boundary is sometimes undefined, highly dependent on the phase of growth and on the urban baseline form.

The use of simulated scenarios has its benefits and shortcomings. First, it allows synthetic scenarios to be created following pure theoretical growth patterns according to their definitions in the literature (Table 5.1), which facilitates evaluating the spatio-temporal metrics based on known growth patterns under different conditions (i.e.: baseline urban forms). Second, the ability to simulate the same growth on different urban forms facilitates the finding of unexpected behavior in metrics due to the baseline urban form. This is not simple to detect with real data, since the same growth pattern is rarely seen. Third, having a wide range of scenarios for testing the approach and metrics provides transferability to other geographical areas, despite their differences in morphology or growth pattern. Conversely, the main

shortcoming is not related to the simulation itself but to the data source. The scale of the map from which spatial metrics are derived actually influences the metric values, and this is an aspect to be considered when applying this methodology to other areas and using diverse data sets. The MMU affects the spatio-temporal metrics in a way in which a higher MMU normally leads to a decrease in the value of most spatial metrics, since smaller patches disappear and boundaries become more simplified (Šímová and Gdulová, 2012). Therefore, the selected metrics for identifying growth spatial patterns applied on very different datasets should be taken into consideration cautiously. In addition, the chosen cities are prototypes of theoretical urban forms, and urban areas usually are a combination of such forms, as well as the simulated growth patterns. Therefore, when applied to real data, these could present a more blended combination of the five growth patterns than in the simulated scenarios, and some parameters, such as the centroids of the growth patterns based on the spatio-temporal metrics (Table 5.4), could vary slightly.

Since we conducted a data-driven analysis and applied statistical methods for the objective selection of metrics, we can state that the proposed spatio-temporal metrics represent the fundamental spatial properties or attributes to be considered for the classification of the five growth patterns analyzed. In other words, more than the metrics themselves, that may slightly vary from one case to another, what will be essential to identify growth spatial patterns in urban areas are the subjacent spatial attributes they represent. These attributes are described as the change in the concentration degree of the urban cover (i.e.: distances of the urban elements with respect to the urban centroid) and the densification and compactness of the urban growth, given by the spatial adjacency of the new urban elements.

The increasing availability of frequent and updated urban data, in particular those related to LULC, will offer new opportunities in this field, requiring tools and methods, as well as interpretable indicators to efficiently categorize urban growth. Eventually, when databases and LULC data increase, new studies based on real growth cases, instead of simulations, can be conducted. Thus, the selected spatio-temporal metrics could be applied to additional urban areas in different geographical context in order to identify growth patterns and evaluate their sustainability, with the added value of considering the urban form in the classification of growth patterns.

Overall, we validated the use of two spatio-temporal metrics that quantify the densification, compactness and concentration degrees of growth, for identifying growth spatial patterns in different urban areas. These metrics can be further used for monitoring urban growth patterns whenever temporal LULC is available, in order to validate city planning, infrastructures, social

policies and territory management. As a future work, the identification of growth patterns in several cities worldwide using the weighted Euclidean distance (AWSD$_{urban}$) and the weighted mean expansion (AWMEI$_{urban}$) metrics, will allow us to study their impact on environmental, social and economic factors.

5.5. Conclusions

The development of methodologies for the description and quantification of urban growth is useful to monitor urban areas, to diminish the consequences of rapid urban growth and to improve planning and sustainability of urban environments. In the absence of long-term LULC data at high-resolution, we simulated urban growth of different cities and scenarios, and applied spatio-temporal metrics derived from LULC maps to analyze the influence of different initial or baseline urban forms in the classification of urban growth patterns. As a result, it was found that two spatio-temporal metrics that quantify densification, compactness and concentration of growth are sufficient to classify five growth spatial patterns (i.e. expansion, compact, dispersed, road-influenced and leapfrog) with an overall accuracy of 75%. The spatio-temporal metrics demonstrated their usefulness for the categorization of urban growth spatial patterns in diverse urban forms despite the notable influence of the urban form on the growth processes. The monocentric and sprawl forms eased the identification of patterns in comparison to the polycentric and linear forms that added uncertainties in the classification. Our results show the potential of spatio-temporal urban distribution metrics for monitoring dynamic urban areas. The early detection of growth patterns and thus, the ability to foresee their consequences will be valuable for land-use planning in urban and peri-urban areas.

Chapter 6

Linking spatio-temporal metrics of built-up areas to socio-economic indicators on a semi-global scale

6.1. Introduction

Urban form organizes people, space and flows. As such, urban areas are simultaneously shaped by economic and demographic processes; social relations; legal and political systems; and historical, cultural and climate contexts; etc. (Tonkiss, 2013; Zhu et al., 2019). The urbanization process affects dwellers in many dimensions. For example, one impact concerns cities, where air pollution and its impact on health, inequality and environmental degradation are increasing threats as a consequence of rapid growth (UN, 2019). The development of urban areas is not only conditioned by manifold local and regional factors but also by global trends that contain drivers and consequences.

In recent years, the number of studies quantifying the relationships between EO-derived data and socio-economic variables has risen. Consequently, various elements of the built and natural environment, as well as atmospheric parameters derived from EO, have been related to different socio-economic indicators: For example, image-derived metrics and features have been used to model poverty levels. For example, severe poverty was associated with the travel time to major market towns, and the percentage of woodland and winter crop cover (Watmough et al., 2016). Duque et al. (2015) developed a composite poverty index based upon a wide set of variables related to land cover composition and urban spatial patterns. Poverty was found to be higher in areas with less impervious surfaces with the absence of clay roofs, a higher complexity of the urban fabric, and a lower diversity of landscapes. Similarly, deprived living conditions in major UK cities were related with population density, vast portions of unbuilt land, regular street patterns and cul-de-sacs (Venerandi et al., 2018). Meanwhile, a local study in Liverpool, UK found that the percentage of vegetation and water, and the variability and homogeneity of the image intensity values were the best predictors of deprivation (Arribas-Bel et al., 2017). GDP exhibits a high correlation not only with built-up density in a set of Canadian cities (Faisal et al., 2016) but also with the intensity and density of night-time lights in a city of China (Liang et al., 2020). Regarding air quality, it has been related to both the built and natural environments. Continuous urban development was associated with better air quality in urban areas of the USA, while the presence of proximate forest was significantly related to an improvement in air quality when demographic factors and the degree of urbanization were controlled for (McCarty and Kaza, 2015). Generally, a low centrality of the urban fabric, a low density, worse transport services and limited land diversity are correlated with higher pollutant concentrations (Hankey and Marshall, 2017).

A general finding from these studies is that the built-up structure, night-light

emissions, transport network, population distribution and LULC configuration and diversity are related to socio-economic-ecological factors in urban areas. Such relationships have been mainly analyzed based on correlations, multiple regression and random forest methods; they proved to be techniques suitable for modeling statistical variables by means of EO-derived data. However, the majority of studies are intra-urban analyses conducted at the city level, with only few at the regional or national levels. A minority are based on global inter-urban analyses, which provide a more comprehensive, but less detailed, picture of development patterns. Examples of inter-urban studies demonstrated that, in European cities, an equal distribution of LULC is associated with lower inequality in life satisfaction (Olsen et al., 2019) and that quality-of-life-related indicators can be modeled by means of LULC spatial metrics (Sapena et al., 2016). In urban areas of the USA, similarities in the structures of urban landscapes were linked to transport behaviors (Stokes and Seto, 2019).

On balance, relationships between the built and natural environments and socio-economic-ecological factors have been proven, but large area and multi-temporal analyses remain rare. These analyses bring the opportunity to create, based on predetermined relationships, spatial indicators of social, economic and environmental parameters among and across countries. In this direction, geospatial data have been used as proxies of income inequality (De Leeuw et al., 2010; Mveyange, 2015), unsustainable urban growth (Stokes and Seto, 2019), economic disparities (Taubenböck et al., 2017), and GDP, especially useful in countries with low-quality statistical systems (Chen and Nordhau, 2011). Hence, unraveling the links between urban form and LULC and statistical variables, both at a particular moment in time and in terms of their evolution over time, aids in mapping and assessing the temporal evolution of socio-economic and ecological processes. Some examples in this regard are foreseeing the loss of farmland and food security issues (Rimal et al., 2017), predict the risk of and exposure to diseases (Wilkinson et al., 2018) or comparing the evolution of socio-economic factors, such as employment and poverty, in response to specific policies (Oldekop et al., 2018; Sims et al., 2019).

There has been a recent call regarding the need for cross-comparative empirical analyses across different regions that reveal the consistency of these relationships and that allow the drawing of reliable conclusions on the sustainability of urban development (Seto et al., 2017; Zhu et al., 2019; Lobo et al., 2020). However, these analyses are usually limited by the scarce or inconsistent availability of data at a global scale. For the needed socio-economic datasets, currently, the availability of global and still comparable

data at resolutions of intra-urban level is still limited. On the one hand, some institutions are delivering socio-economic and environmental statistics for cities and functional urban areas. Two examples are the City Statistics from Eurostat (Eurostat, 2016b) and the Organization for Economic Co-operation and Development (OECD) (OECD, 2019a). They provide comparable statistics associated with territorial units with large-scale coverage for multiple time periods. On the other hand, there has been a growing interest in integrating statistical and spatial information to produce spatially explicit socio-economic data, swapping from irregularly shaped boundaries to a regular surface, easing comparisons within and across regions at lower levels. Two of these initiatives are GEOSTAT (GEOSTAT, 2020) and the Socioeconomic Data and Applications Center (SEDAC) (SEDAC, 2020). Although the variables and the time coverage are still limited, they are promising data sources that are under development. For the needed spatial datasets, concurrently, recent EO-based efforts have been made in the global mapping and characterization of human settlements and land covers over time. Some examples are shown in Table 1.1. Furthermore, the development of methods and algorithms to automatically classify urban environments across the globe is progressing rapidly (e.g.: Bechtel et al., 2015; Cao et al, 2020; Qui et al., 2020). The global coverage and high spatial and temporal resolutions of EO-derived products combined with the high capacity to automatize processes allows the frequent updating of geospatial datasets. This, however, is still an issue in socio-economic databases, since they depend on surveys and censuses with low temporal frequency, and they are limited or even inexistent in some geographical areas.

Accordingly, our aim is to use spatial patterns and their development over time as proxies of socio-economic parameters at the global level. With the help of easily quantifiable spatial metrics extracted from openly available EO-derived and ancillary data, we aim to prove the feasibility. With the growing availability of spatial and socio-economic datasets, this is an opportunity in terms of methodological fine-tuning for defining empirical methods that could be applied globally in the near future, when higher-resolution data with a global reach will be available. In this context, a semi-global analysis will bring the opportunity to obtain first fundamental conclusions and foresee potential subsequent analyses when more and higher resolved (i.e., spatially, temporally, thematically, and better quality) data become available. Therefore, the purpose of this study is to quantify the relationships between socio-economic and environmental variables, such as income, inequality, GDP, air quality and employment, and spatio-temporal metrics issued from geospatial databases, both on a specific date and in terms of their variation over time. Subsequently, the purpose is to identify the spatio-temporal metrics that are most related to socio-economic and environmental variables and can be extracted on a

massive scale from current global geospatial databases.

6.2. Material and methods

In this study, we leveraged multi-temporal open datasets with global and semi-global geographic and socio-economic data. Figure 6.1 outlines the general workflow of the study. The manifold datasets are described in Section 6.2.1, while Section 6.2.2 defines the preprocessing steps to ensure the harmonization of datasets that are necessary for the subsequent extraction of spatio-temporal metrics (Section 6.2.3). Then, the spatio-temporal metrics are related to the socio-economic variables from a multi-temporal perspective by means of regression models, and the relevant metrics are identified (Section 6.2.4).
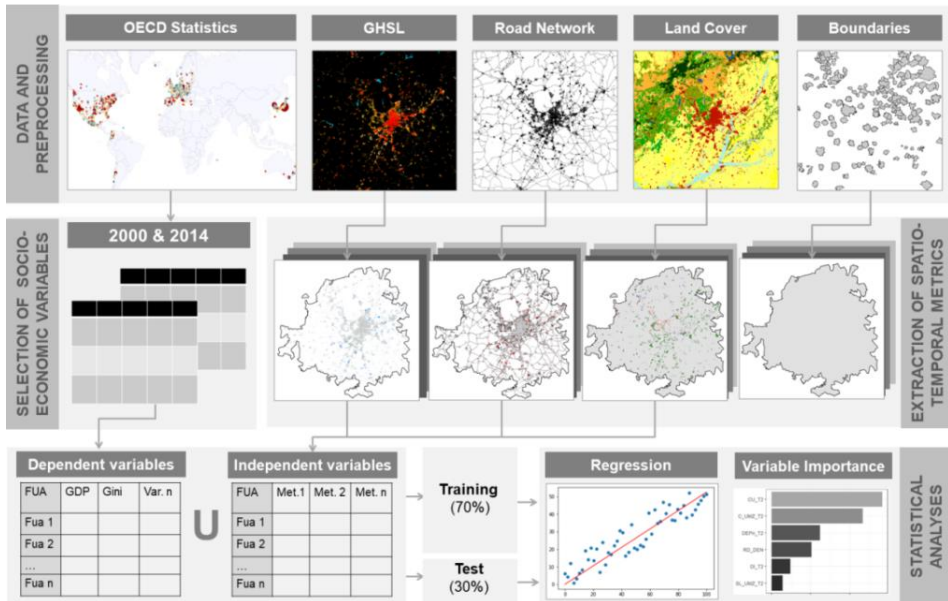
Figure 6.1. Workflow. First, data are downloaded and prepared for further analysis. Second, socio-economic variables are selected from the OECD database for the years 2000 and 2014. Third, spatio-temporal metrics are extracted from geographic data for each boundary individually (functional urban areas, FUAs) corresponding to the same two years. Last, socio-economic variables (dependent variables) and spatio-temporal metrics (independent variables) are combined and split into training and test samples to build regression models and rank the contribution of metrics.

6.2.1. Socio-economic, EO-derived and ancillary datasets

6.2.1.1. Global Human Settlement Layer (GHSL)

The GHSL consists of a global multi-temporal classification of built-up areas created by the Joint Research Centre from the European Commission. The GHSL considers "built-up" as building footprint areas (i.e., roofed constructions above ground). It is derived from Landsat imagery collections at a 30 m spatial resolution in four-time steps: 1975, 1990, 2000 and 2014. We used the latest version of the product released at the end of 2019, which has considerable improvements over the previous version, the *GHS_BUILT_LDSM T_GLOBE_R2018A_3857_30_V2_0* (Corbane et al., 2018a), for the years 2000 and 2014 to coincide with the socio-economic data. The dataset is a categorical raster in VRT format, with GeoTIFF tiles, where different categories represent built-up land at each epoch, water, non-built land and no data, in the coordinate reference system (CRS) Pseudo Mercator (EPSG: 3857).

6.2.1.2. OECD Regional Statistics

The OECD offers the regional statistical database in which the metropolitan areas dataset is the lowest level (OECD, 2019a). This dataset contains data on demographic, economic, income distribution, environmental and labor statistics. For February 2020, 649 functional urban areas (FUAs) with over 250,000 inhabitants in 33 OECD member countries and Colombia from the year 2000 onwards were available. The variables presented in the database are calculated using different methods. The majority are modeled based on the aggregation of local administrative data, and others, using geospatial data sources (e.g., air quality) or by downscaling variables available from larger regions through the use of population grids (e.g.: GDP) (OECD, 2019b).

We gathered statistics for 32 countries for the years 2000 and 2014, or the previous or following year when data were not available; for example, the Gini and income variables were only available for the years 2013, 2015 and 2016; we used them as an approximation for the year 2014. The availability differs widely between years and countries, and from variable to variable; therefore, the number of FUAs we applied varied between variables. We selected socio-economic variables related to economic, income, labor and environmental topics for 2014 and for change between 2000 and 2014 (Table 6.1).The statistical data used in this study refer to data available in the metropolitan areas dataset as of February 2020. The list of FUAs available for each socio-economic and environmental variable and their values are presented in detail in www.mdpi.com/2220-9964/9/7/436/s1. FUAs with over 250,000 inhabitants not listed are due to their unavailability for our study years. Since the OECD Regional Statistics are updated from time to time, changes in the available FUAs and socio-economic variables may occur. For this reason, the original downloaded dataset is included in the supplementary material.

Table 6.1. Description of socio-economic and environmental variables modeled for 2014 or their change between 2000 and 2014.

Variable	Description	Year/s
GDP	Gross domestic product per capita (GDP) is the value added created through the production of goods and services during a certain period per capita. It is expressed in United State dollars (USD) constant prices and constant Purchasing Power Parities (PPPs) with the base year 2010 (i.e., differences in price levels between countries are eliminated based on PPP rates). The GDP is less suitable for comparisons over time, as growth is affected by changes in prices and dollars per capita (OECD, 2020a).	2014
Gini	It is an indicator of income inequality among individuals. The Gini coefficient is based on the comparison of the cumulative proportions of the population against the cumulative proportions of income they receive; this ratio ranges from 0 in the case of perfect equality to 1 in the case of perfect inequality (OECD, 2020b).	2014
Income	It is defined as household disposable income in a particular year measured in USD. It consists of earnings, self-employment and capital income and public cash transfers; taxes and contributions are deducted (OECD, 2020b).	2014
Air quality	Fine particulate matter (PM2.5) is the air pollutant that poses the greatest risk to health, affecting more people than any other pollutant. Chronic exposure to PM2.5 increases the risk of respiratory and cardiovascular diseases. Average level in μg/m³ (OECD, 2020c).	2014 2000/ 2014
Employment rate	Employment rate measures the extent to which available labor resources (people available to work) are being used, calculated as the ratio of the employed to the working age population (aged 15 or over) (OECD, 2020d).	2000/ 2014
Population	Population, all ages. It is used to derive a spatio-temporal metric.	2000/ 2014

6.2.1.3. Boundaries of EU-OECD FUAs

The EU-OECD FUA is the unit of analysis. FUAs were defined to maximize international comparability, to overcome the limitations of using purely administrative approaches, and for policy analyses on topics related to urban development (OECD, 2012). This dataset was used for two different reasons: (i) to provide a spatial dimension to the socio-economic data (Figure 6.2), and (ii) to delimit the geographic datasets with the same boundary in order to extract metrics and statistics at the same level as for the socio-economic variables. The boundaries of the FUAs can be downloaded by country in shapefile format in the CRS WGS84 (EPSG:4326).

6.2.1.4. Climate Change Initiative Land Cover

Land cover data from the Land Cover project of the European Space Agency Climate Change Initiative (CCI-LC) were used to obtain land cover densities and dynamics due to urban growth and development. The CCI-LC project
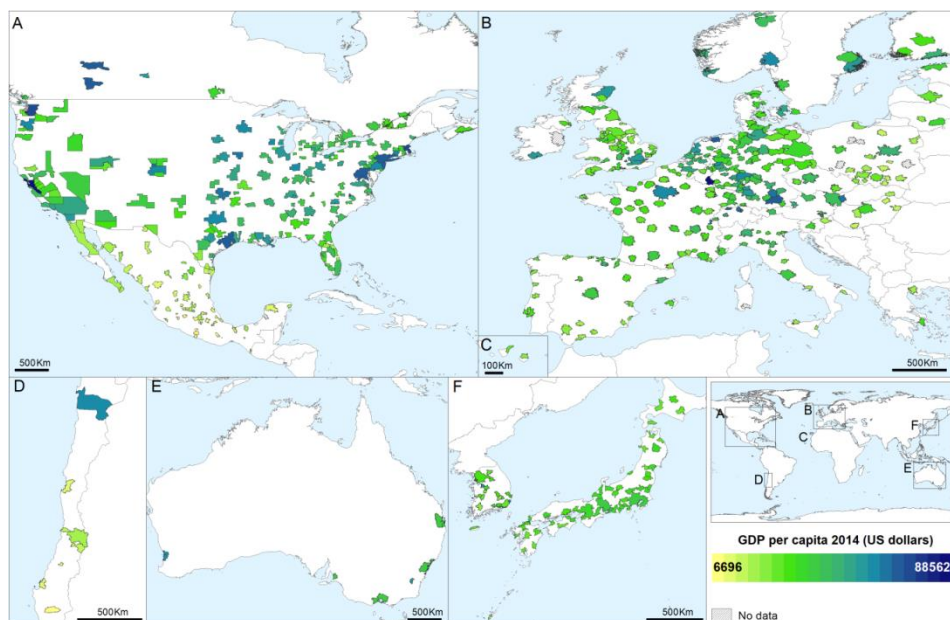
Figure 6.2. Example of the gross domestic product per capita (GDP) in USD for the year 2014 combined with the FUA boundaries in 32 OECD countries: (A) Canada, the USA and Mexico; (B) and (C) European countries; (D) Chile; (E) Australia; (F) South Korea and Japan. "No data" FUAs are included in the OECD metropolitan area dataset but do not have a GDP value for the year 2014.

delivers consistent global land cover maps at a 300 m spatial resolution on an annual basis for 1992 to 2018 (ESA, 2017). We used the *ESACCI-LC-L4-LCCS-Map-300m-P1Y-1992_2015-v2.0.7* dataset for the years 2000 and 2014 to coincide with the rest of the datasets. The land cover map is a categorical multiband raster, in GeoTIFF format in the CRS WGS84, where each band represents one year.

6.2.1.5. Road Network

The Global Roads Inventory Project (GRIP) dataset was developed to provide a recent and consistent global road dataset for use in global environmental models (Meijer et al., 2018). We used five different datasets to cover the regions included in OECD FUAs (North America, Central and South America, Europe, South and East Asia and Oceania). The datasets are in shapefile format in the CRS WGS84.

6.2.2. Preprocessing and harmonization of datasets

The data came in different formats, resolutions and coordinate reference systems; therefore, some preliminary steps were necessary before integrating the data from different sources. The required data and codes to reproduce this work have been made available in the supplementary material. The preprocessing steps were as follows:

> The boundaries of the EU-OECD FUAs from each country were merged in a shapefile, and only those FUAs with statistical information in the metropolitan area dataset were kept. Colombian FUAs were not included in the analysis due to GHSL underclassification, cloud presence or a lack of socio-economic variables.

> The European region of the GRIP dataset was georeferenced using control points from OpenStreetMaps, as it was originally displaced (about 100 m).

> Then, two built-up epochs were extracted from the GHSL. Categories 4 to 6 represent the built-up area in 2000, and categories 3 to 6, that in 2014. This generated two built-up maps.

> Regarding the CCI-LC, two bands corresponding to the years 2000 and 2014 were extracted (bands 9 and 23). The legend of the CCI-LC was grouped into seven major land cover types, as follows: agricultural areas (categories from 10 to 30, both included), high semi-/natural vegetation (40–100 and 160–180), low semi-natural/natural vegetation (110–153), urban areas (190), bare areas (200–202), water bodies (210) and permanent snow (220). To see the original legend and the link between the categories and land covers, refer to the European Spatial Agency (ESA) (2011). This process generated two land cover maps.

> The resulting global built-up and land cover maps and road network dataset were clipped using the boundaries of the FUAs in the CRS of the dataset to be clipped, transforming the FUA boundaries when necessary.

> After that, the built-up and land cover maps in their original spatial resolutions were vectorized to shapefile format, since the tool used for the extraction of the metrics works with vector data.

> Finally, the data were transformed to a local projected CRS to allow the measurement of areas and distances, which are basic attributes in most of the spatial metrics. To do so, the centroid of the FUA was used to determine the EPSG code to project the data to their Universal Transverse Mercator (UTM) zone (e.g., Madrid has the EPSG code 32630, which corresponds to the CRS WGS84/UTM zone 30N). Thus, all the FUAs have similar adapted and local CRSs in the same units, meters.

As a result, for each individual FUA, there were two built-up maps and two land cover maps for 2000 and 2014, the road network, and the boundary delimiting the area of analysis, with a common format and CRS prepared for further analysis.

6.2.3. Extraction of spatio-temporal metrics

In order to quantify the urban form and urban growth spatial patterns of the FUAs, we computed spatio-temporal metrics for the built-up, road network and land cover maps using IndiFrag. We applied a set of uncorrelated metrics that allow the measurement of density, aggregation and spatial distribution properties and their variation over time (we discarded metrics with a Pearson's r > 80%, the ones affected by the size of the boundary, and diversity and contrast metrics). Two types of metrics were considered: the spatial metrics extracted for one date and the multi-temporal metrics computed using maps from two different dates. Therefore, a set of spatial metrics was extracted to quantify the urban form in 2014, and another set of spatio-temporal metrics was extracted for the years 2000 and 2014 to measure the urban growth spatial patterns and land cover dynamics.

Some metrics are applied specifically to the largest and second largest urban cores of the FUA, instead of to the entire built-up area. The cores are based on the urban morphological zone definition (EEA, 2014). From the largest built-up patch, the core is measured by including all the built-up patches within a distance of 200 m, and the same applies for the second largest built-up patch. In this manner, one core split by a feature such as a river, or two built-up pixels connected by a corner were included in the urban core.

6.2.3.1. Spatial metrics (2000 and 2014)

We calculated the following spatial metrics individually for the two time points 2000 and 2014:

> The urban compactness (CU) measures the complexity and fragmentation of the built-up area; it is for both the FUAs and for the largest urban core (C_{UC}). High values show a more compact shape and aggregated distribution (A.28).
> The dispersion index (DI) is the ratio between the normalized number of patches and the proportion of built-up area occupied by the largest patch. Low values indicate coalescence, while high values represent dispersion (A.30).
> The normalized area-weighted standard distance (AWSD) measures the centrality of the built-up area, quantifying the degree to which objects are concentrated around their centroid. It is normalized to the

shape and size of the FUA by means of the "maximum distance", measured as the standard distance of a regular grid covering the FUA extension to the centroid. Normalized values range from 0 to 100, where lower distances show a concentrated distribution of built-up patches around the core, and higher values show built-up patches homogeneously distributed across the entire FUA, without a special clustering around the center (A.20).

> The density is the percentage of built-up area (DU) (A.1) and other land covers (DC) (A.2) relative to the total FUA area.

> The percentage of the urban core (L_{UC}) is the percentage of the built-up area that occupies the largest core (A.9). When the value is high, it shows a monocentric form. Since the spatial metric is highly correlated to the DI, only the change was computed and included as a multi-temporal metric.

> The second largest urban core (SL_{UC}) is the percentage of the built-up area that occupies the second largest core (A.10). When the value is close to L_{UC}, it suggests a polycentric form.

> The elongation ratio (ER_{UC}) of the largest urban core quantifies the elongation shape of the urban core. This metric measures the elongation, dividing the diameter of the circumference with the same area as the core by the largest side of the core (A.16). It ranges from 0 to 1. Values closer to zero show elongated shapes, i.e., a linear urban form.

> The density of road network (D_{road}) is the total length of roads per square kilometer (A.3).

6.2.3.2. Multi-temporal metrics (2000-2014)

> We calculated the following metrics as the differences between the spatial metrics for the two different years, 2000 and 2014: the change in urban compactness (CU_{CH}), urban core compactness ($C_{UC\ CH}$), dispersion index (DI_{CH}), normalized area-weighted standard distance ($AWSD_{CH}$), density (DU_{CH}, DC_{CH}), percentage of the urban core ($L_{UC\ CH}$), second largest urban core ($SL_{UC\ CH}$) and elongation ratio ($ER_{UC\ CH}$).

> The urban change rate (UCR) is the percentage of built-up growth relative to the built-up area for the first date (A.47).

> The area-weighted mean expansion index (AWMEI) quantifies the aggregation and densification of growth. It ranges from 0 to 100. A high value indicates a densification (infilling growth) and therefore a more compact growth pattern, and an intermediate value shows expansive growth, while a low value represents scattered growth (A.43).

> The area-weighted mean accessibility index (AWMAI) quantifies the accessibility of new built-up patches to the road network. This is measured with the mean of the inverse distance between the new built-up patches and their closest roads, weighted by the areas of the patches. It ranges from 0 to 100. Higher values show shorter distances to roads and better accessibility (A.43).

> The population and urban growth imbalance index (PUGI) measures the inequality between the increase in the built-up area with respect to population growth or decline (based on population counts from Table 6.1). It provides information related to the land consumption per capita (i.e., the amount of built-up land per population change) and the degree of sprawl in the urbanization process (Sapena and Ruiz, 2019). Positive values show more urban growth, zero means equal growth, and negative values mean higher population growth (3.3).

> The change proportion (CP) of the land cover is the ratio representing the change in a particular land cover with respect to the total area of the FUA, and it measures the relative area of change (A.40).

6.2.4. Regression models and identifying spatio-temporal metrics' relevance

We used random forest regression models to quantify how much urban spatial patterns and their change over time are related to socio-economic indicators and their multi-temporal variations. The use of random forest over linear and non-linear regression models has been discussed in the recent literature. Many studies have compared different algorithms, and random forest performed the best in most of the cases (e.g.: Breiman, 2001b; Arribas-Bel et al., 2017; Gonzalez and Leboulluec, 2019; Liang et al., 2020; Paul, et al., 2020). Random forest is a supervised learning algorithm that uses an ensemble learning method for classification and regression (Breiman, 2001a). For building the models, we trained 500 decision trees with random splits of two thirds of the data, leaving one third for testing, which is the out-of-bag (OOB) sample. The predictions and accuracies of the models are calculated with the OOB samples. This method builds the model by minimizing the mean square error (MSE). In order to evaluate the model's performance, we applied the following accuracy indices to the OOB sample: (i) The coefficient of determination (R^2) measures the proportion of the total variability explained by the model; (ii) the MSE measures the average squared difference between the observed value and its prediction; and (iii) the root mean squared error (RMSE) is the standard deviation of the differences between the observed values and their predictions; the RMSE estimates the concentration of predictions around the 1:1 line (when the prediction equals the observation), and it is measured in the same units as

the observed variable, which limits the comparison of models of different units. Therefore, we also included (iv) the normalized RMSE with the standard deviation (sd-NRMSE). It represents the ratio between the variations not explained by the model against the overall variation in the observed variable. The sd-NRMSE will be close to zero if the model explains the variation well and around one when it explains it partially, and bigger values indicate a weak performance (Otto, 2019). (v) The normalized range-based RMSE (range-NRMSE) gives the error as a percentage of the total range of the observed variable (Otto, 2019).

In order to explore the relevance of the spatio-temporal metrics in terms of their relationships with socio-economic and environmental variables, we ranked the metrics according to the variable importance measure. This is a widely used and robust index that captures nonlinear and interaction effects (Breiman, 2001b; Arribas-Bel et al., 2017; Probst et al., 2019; Liang et al., 2020). It reflects the increase in the MSE when a metric is randomly permuted in a tree, averaged over all trees. Metrics with larger differences were ranked first in terms of importance. Additionally, to test the significance of the metrics' importance in the model, the MSE was compared against a null distribution of the MSE. We did this by running the model 100 times and permuting the dependent variable randomly, reporting the significantly important metrics (*p-value* < 0.05). We built different models for the same variable combining subsets of spatio-temporal metrics (for example, using only spatial metrics for 2014, adding the road density, the PUGI, the land cover change, and/or the multi-temporal metrics). In this manner, the final model only keeps the combination that performs best, removing the metrics with a negative influence based on the importance measure. Finally, since the random forest regression and the variable importance measure do not report the positive or negative relationships between variables and metrics, we divided the socio-economic variables into five quantiles with the same number of FUAs from lower to higher levels. Thereby, we represented the standardized values (*z-score*) of the significantly important metrics for each quantile. This eases the interpretation of found relationships.

6.3. Results

6.3.1. Estimation of socio-economic variables

Table 6.2 reports the accuracy indices of the regression models for the socio-economic and environmental variables for 2014 using the best subsets of metrics. For the metrics included in each model subsection 6.3.3, Figure 6.5.

Table 6.2. Accuracy indices for mono-temporal models. GDP per capita, Gini, income and air quality regression models for 2014 using spatio-temporal metrics. FUAs represents the total number of FUAs in the model; R^2, the coefficient of determination; MSE, the mean square error; RMSE, the root mean square error in the same units as the variable; and sd-NRMSE and range-NRSME, the standard deviation and range-based normalized RMSE, respectively.

Variable (unit)	FUAs	R^2	MSE	RMSE	sd-NRMSE	range-NRMSE
GDP (USD)	597	43.97	102,028,574	10,101	0.7479	0.1234
Gini (ratio)	142	52.2	0.0011	0.0326	0.689	0.1577
Income (USD)	280	68.07	45,985,090	6781	0.564	0.1232
Air quality (µg/m³)	599	52.9	20.8591	4.5672	0.6857	0.1324

The model for the gross domestic product per capita (GDP) explained almost 44% of its variability (R^2), which shows a mid-high relationship with the urban spatial pattern. It has a mean error of 10,101 USD (RMSE), representing 12.3% of the total range of the GDP (range-RMSE). FUAs like Obihiro in Japan, Lane in USA, and Wuppertal in Germany had the lowest errors. However, the errors with respect to the total variability of the GDP are considerably high (sd-NRMSE = 0.75), due to the presence of some outliers when the GDP is high (e.g., San Francisco and Luxembourg, Figure 6.3). The model was not able to capture the spatial attributes, particularly in FUAs with relative high GDP values. Regarding the income inequality of individuals (Gini), the number of available FUAs and countries is limited. Still, 52% of its variability was explained by the model (R^2) with an error of 0.03, which represents 16% of its range. However, the variability between the FUAs is not totally captured by our model (sd-NRMSE = 0.69). In this case, both low and high inequalities, relative to the sampled FUAs, were over- and under-estimated by the model (Figure 6.3). For instance, Bordeaux in France and Oslo in Norway have low Gini values, and the model predicted much higher values. On the contrary, Calgary and Vancouver in Canada, New Haven and Miami in the USA, and Lisbon and Porto in Portugal were underestimated, since much lower inequality values were predicted. Meanwhile, examples of good estimates are Fayette in USA, Winnipeg in Canada and Florence in Italy. The income model is the one with best performance. It shows the highest R^2 and lowest sd-NRSME. It explained 68% of the total variability of the income between the FUAs by means of spatio-temporal metrics. The errors with respect to the total variability of income are considerably low (sd-NRMSE = 0.56), representing 12% of the income range within the FUAs (range-NRMSE), with a mean error of 6,781 USD (RMSE). The model failed in the estimation of low income values, especially seen in Mexican FUAs (Figure 6.3) with the exceptions of Benito Juarez, Hermosillo and Tijuana; one reason might be that they present

Figure 6.3. Observed versus predicted variables in mono-temporal models. The more the FUAs, represented as points, that are closer to the 1:1 line, the better the estimation by the model. The black dashed lines show the 1:1 lines (lines of perfect fit), while the blue dashed lines show the root mean square error of the model (± RMSE). The labels for the FUAs with the highest errors are shown, to identify outliers. The color represents the country. The units are GDP and income, USD; Gini, ratio; and air quality, PM2.5 in µg/m³.

different urban forms and growth patterns but very similar mean income values at the FUA level. Finally, according to the environmental variable, the air quality due to fine particulate matter is also related to the urban spatial patterns. Almost 53% of its variability was explained by means of the spatio-temporal metrics with a mean error of 4.56 µg/m³, representing 13% of the air quality range (range-NRMSE). However, the error relative to the variability of the air quality is considerable (sd-NRMSE = 0.68). When the particular matter was above 30 µg/m³, the model predicted lower values (Figure 6.3). This underestimation is especially seen in the FUAs in Mexico (olive green),

Korea (electric blue), Poland (dark pink) and Santiago in Chile (dark green). As seen in Figure 6.3, the RMSE is highly sensitive to outliers. Even if the majority of the FUAs have a good prediction (they are close to the 1:1 line, e.g., for the Netherlands, Germany, France and the USA), the lack of ability of the model to estimate some of them, creating outliers, widely increases the RMSE and their normalized values.

6.3.2. Estimation of the variation of socio-economic variables

Regarding the temporal variation of the socio-economic and environmental variables, as expected, the performance of the models was lower than that of the mono-temporal models; however, a significant amount of Air quality change and Employment change was explained by spatio-temporal metrics (Table 6.3).

Table 6.3. Accuracy statistics for multi-temporal change models. Air quality and employment rate change regression models for the period between 2000 and 2014 by means of spatio-temporal metrics. FUAs is the total number of FUAs included in the model; R^2, the coefficient of determination; MSE, the mean square error; RMSE, the root mean square error in the same units as the variable; and sd-NRMSE and range-NRSME are the standard deviation and range-based normalized RMSE, respectively.

Variable (unit)	FUAs	R^2	MSE	RMSE	sd-NRMSE	range-NRMSE
Air quality change (µg/m³)	599	41.16	0.6172	0.7856	0.7664	0.1076
Employment change (%)	313	31.56	10.7334	3.2762	0.826	0.1413

First, the change in air quality, measured as the variation in the content of fine particulate matter in the air (µg/m³), was predicted with an R^2 of 41%, which shows that only part of its variability was captured by the model. It has a mean error of 0.78 µg/m³, which represents 11% of the range in the variable (range-NRMSE). However, compared to the total variability of the air quality, this is considerably high (sd-NRMSE = 0.76). According to Figure 6.4, all the FUAs experienced an improvement in air quality between 2000 and 2014. The largest drops in air pollution were measured in some Mexican and Polish FUAs, which were not properly modeled, resulting in underestimation. However, the air quality change in Aguascalientes in Mexico, New York in USA or Modena in Italy, among many other FUAs, was successfully modeled. Second, the *change in the employment rate* was partially explained by means of spatio-temporal metrics (R^2 = 32%). The mean error was 3.3, accounting for 14% of the range in the change variable. When compared to the variability in the employment change, this is quite high (sd-NRMSE = 0.83), and the model slightly explained the inherent variation in the employment change within the FUAs. Figure 6.4 shows that the highest drops in employment rates (e.g., in Dublin in Ireland, and Benton and Washoe in the USA) were underestimated and much lower
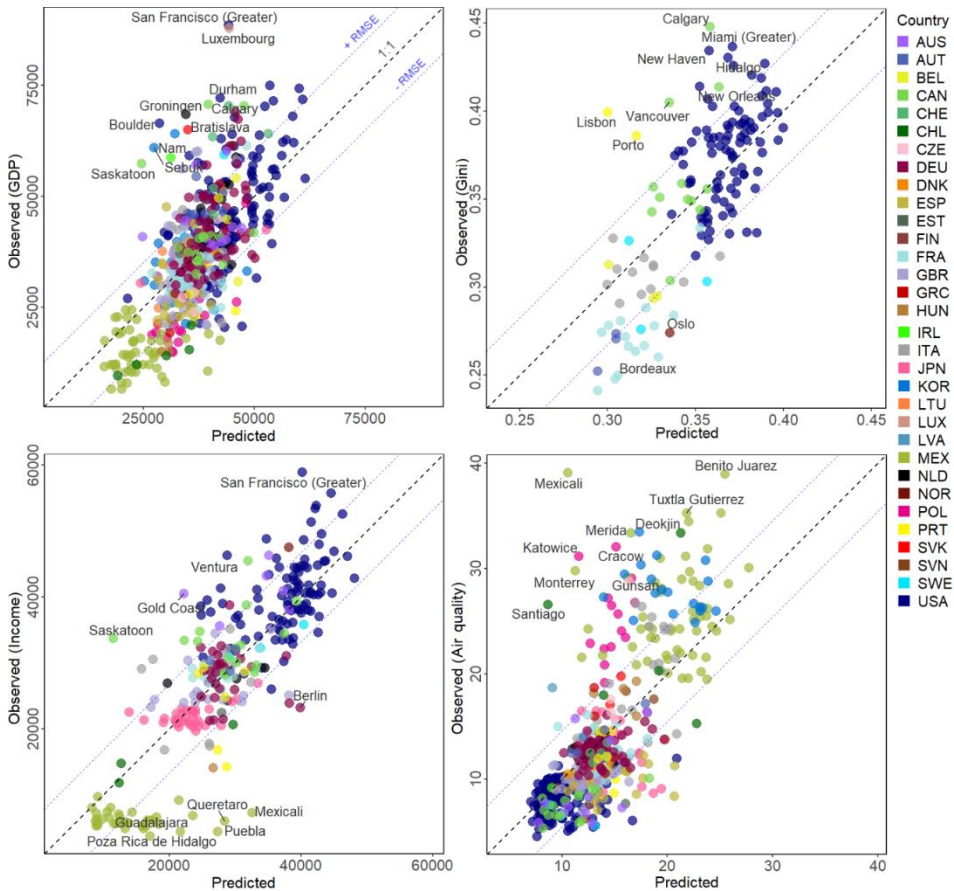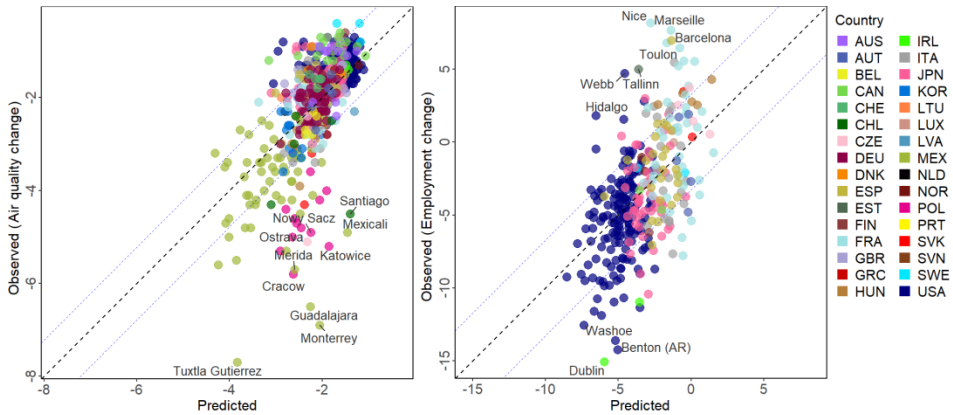
Figure 6.4. Observed versus predicted changes in variables according to the models. The more the FUAs, represented as points, that are closer to the 1:1 line, the better the estimation by the model. The black dashed lines show the 1:1 lines (lines of perfect fit), while the blue dashed lines show the root mean square error of the model (± RMSE). The labels for the ten FUAs with the highest errors are shown, to identify outliers. The units are air quality, PM2.5 in µg/m³, and employment, %.

rates were predicted. On the contrary, the greatest increases in employment in the study period were in Nice and Marseille in France, or Barcelona in Spain, which were also underestimated. In fact, the range of the predicted values (–8.5 to 1.5) was much lower than the range of actual employment change (–15 to 8.1), and the model was not able to properly capture this variation with the spatio-temporal metrics. Nevertheless, good estimates were made, for example, in Chicago, Washington and Dallas in the USA and Rouen and Seville in France and Spain, respectively.

6.3.3. Relevance of spatio-temporal metrics

Figure 6.5 portrays the importance and significance of the spatio-temporal metrics for the modeled variables. They are represented by the mean and standard deviation of the increase in the MSE when a metric is permuted, so that the higher the increase, the higher the importance. The two most important metrics that are key in all the models are the urban compactness (CU_{T2}) and the urban core compactness ($C_{UC\ T2}$); both measure the compact shape and aggregation level of the built-up and the core urban area. The changes in the built-up and urban core compactness (C_{CH} and $CU_{UC\ CH}$) are also important for the GDP and the change in air quality, and their effect on other socio-economic variables is lower or was not included. The dispersion index (DI_{T2}) is relevant for the estimation of the Gini and the change in air quality, but less so for the rest of variables, as well as its change over time (DI_{CH}), which
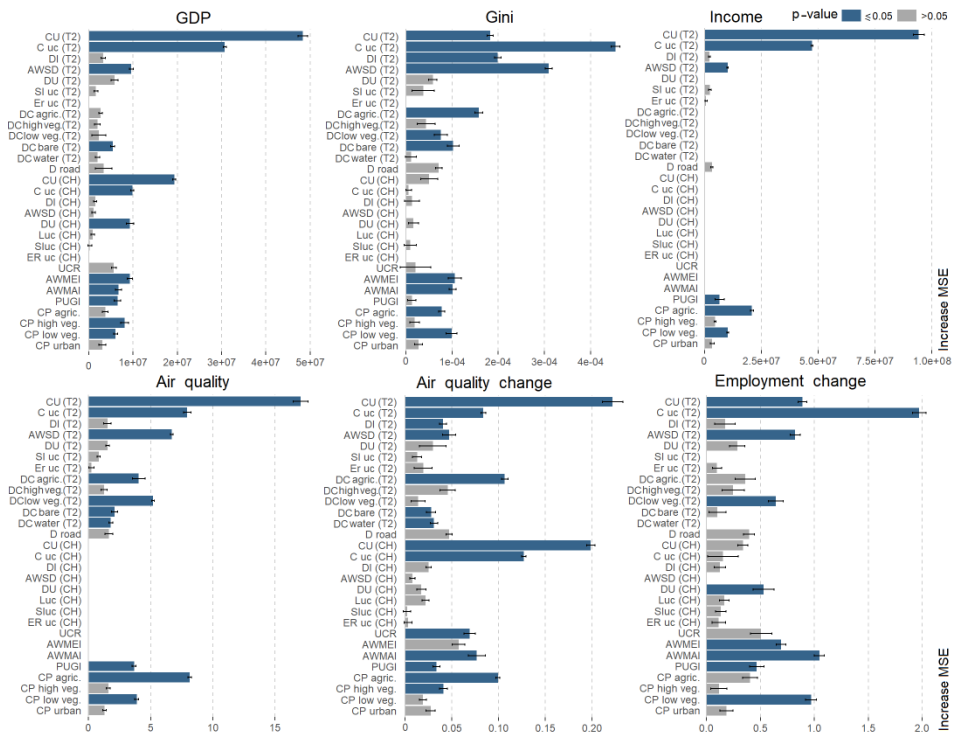
Figure 6.5. Spatio-temporal metrics' importance for the different regression models. The importance is represented by the mean and the standard deviation of the increase in the MSE (the units and final MSE of the model can be found in Table 6.2 and 6.3). Blue bars indicate statistically significant variables in the model. Where the bar is missing, the metric is not included in the model.

has a low influence. Another relevant metric is the centrality and concentration of the built-up elements relative to their centroid (AWSD$_{T2}$). This metric is very informative regarding the spatial configuration of the built-up areas in the FUAs, and its inclusion in the model improves the estimation of the Gini, GDP, air quality and change in employment rate; with regard to income and air quality change, its influence is lower but still significant (*p-values* < 0.05). On the contrary, the change in the centrality (AWSD$_{CH}$) presents very low importance; it is not significant and was even removed from the models for its negative effect. This could be due to the fact that the change is very low with the exception of in a few Japanese, Korean and Mexican cities that present significant changes in the concentrations of the built-up areas. The urban density (DU$_{T2}$) has a medium influence in all the models, but it is not significant enough. On the contrary, its change (DU$_{CH}$) is important for GDP and employment change. The elongation ratio was removed for its negative

influence in the GDP and Gini models, and it has a slight but non-significant importance for the rest of the models. The densities of the land covers are important for different indicators. The density of agricultural land ($DC_{agric.\ T2}$) influences the Gini, air quality and its change. Low vegetation land ($DC_{low\ veg.\ T2}$) contributes to the Gini, air quality and employment change. The density of the road network (D_{road}) improves the prediction of the Gini and changes in air quality and employment, but its contribution is not significant. Concerning the urban change rate (UCR), it only influences the change in air quality, as its impact on the rest of variables is not significant (*p-values* > 0.05). The densification of growth (AWMEI) also contributes in an intermediate manner to the GDP, Gini and change in employment rate. An important metric for the change in the variables is the accessibility of the new built-up elements to the road network (AWMAI), and it also contributes to the GDP and Gini. On the other hand, the imbalance between the built-up footprint and population growths (PUGI), which provides information not only about the inequality between newly developed land and demographic dynamics but also about urban sprawl, was significantly important for all the models except the Gini. Regarding the land cover change proportions, the agricultural land change ($CP_{agric.}$) is detected as important for estimating the Gini, income, air quality and its change in the FUAs, as well as low vegetation land change, which influences the GDP, Gini, income, air quality and employment change. The change in high vegetation land is important for air quality change and GDP per capita.

Analyzing the performance of relevant spatio-temporal metrics against the socio-economic variables complements the interpretation of the relationships found with the models. Therefore, the FUAs were split into five quantiles based on the socio-economic variable values, where quantile 1 groups low values, and 5, high values. Then, the standardized values of the selected spatio-temporal metrics were represented with boxplots (Figure 6.6). This figure shows a selection of spatio-temporal metrics whose relationships with socio-economic variables are described and analyzed in the discussion section. The full set of graphs representing the spatio-temporal metrics per socio-economic variable can be found in the Appendix B (Figure B.16 to Figure B.21).

6.4. Discussion

The combination of multi-source and multi-temporal datasets for almost six hundred functional urban areas across 32 countries led us to extract insights into the relationship between urban spatial patterns and socio-economic and environmental variables at a semi-global scale. By means of a machine learning algorithm, random forest regression, we were able to partially model some
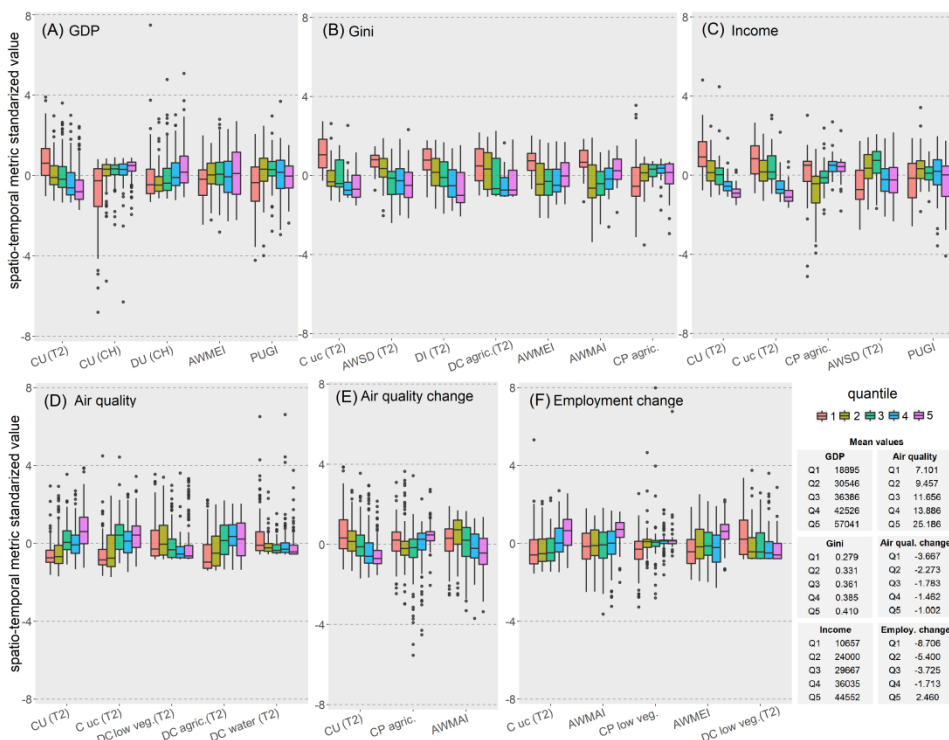
Figure 6.6. Boxplots of selected relevant spatio-temporal metrics sorted according to their importance. The socio-economic variables: (A) GDP; (B) Gini; (C) income; (D) air quality; (E) air quality change; and (F) employment rate change, are divided into five quantiles (Q1 to Q5, from low to high values), and the standardized values of the metrics are shown for each quantile. The table on the bottom-right shows the mean values of the socio-economic variables per quantile. Air quality measures the fine particulate matter (higher values mean more pollutants and lower air quality). The units are GDP and income, USD; Gini, ratio; air quality, μg/m³; and employment, %.

socio-economic variables and their change using spatio-temporal metrics extracted from geospatial databases. We explained between 68% and 44% of the variability of the income, Gini, GDP per capita and air quality variables with the sole use of spatial information. This central result proves that the spatial appearance of urban areas and their change are related to the socio-economic and environmental indicators for these areas.

We are aware that we have neither considered macro-economic or other overarching global developments nor considered intra-urban variabilities, but still, we can conclude that these relationships exist. With regard to their variations, we analyzed the relationships with the metrics for only two of them (i.e., air quality and employment rate), since many variables were not available

for two dates (such as income or Gini) or the change over time is not a good indicator of development, as is the case for GDP (OECD, 2020a). Nevertheless, we explained 41% and 32% of the variation in the air quality and employment rate, respectively, which suggests that the spatial component may relate partially to how these indicators change. Overall, however, we found that there are fundamental correlations between the spatial urban structure and socio-economic-ecological variables. Multi-temporal changes, however, cannot be estimated one-to-one from this correlation, since, for example, spatial urban structures are subject to certain inertia in contrast to economic developments.

The use of random forest regression has strengths and weaknesses. Its interpretability compared to that of parametric regression is reduced since the function is unknown. However, with the variable importance measure, it is possible to identify those independent variables that have strong influence in the model (Breiman, 2001b; Arribas-Bel et al., 2017), the ones with partial influence, and the ones adding noise or uncertainty.

We investigated the relationship between socio-economic, environmental and spatial variables and found evidence of their links. The compactness degree of built-up areas and their cores is highly associated with the average income in FUAs. In particular, more compact values are found in lower-income FUAs, while there are higher incomes in less compact, and thus more scattered, urban configurations (Figure 6.6, C). This assumption might be influenced by independent differences in compactness and income across countries. However, we found a similar negative correlation between income and compactness in the FUAs from the USA (Pearson's r = –42%), for instance, which shows that this trend is not only determined by geographical or cultural aspects. Salvati and Carlucci (2015) found that discontinuous settlements in Northern Italy (low compactness) had higher disposable incomes, and related the phenomenon to suburbanization processes typical in the developed and economically active regions of Europe. Besides, we measured nonlinearities, where a higher loss of agricultural land between 2000 and 2014, higher fragmentation of built-up areas and sprawl (more urban expansion than population growth) occurred in middle-income FUAs, while low- and high-income FUAs had built-up areas that were more spatially centralized and populations that outpaced built-up growth (Figure 6.6, C). Cities in countries with higher incomes have been previously related to higher levels of land consumption and urban fragmentation (Angel et al., 2011); however, this study disregarded income variation within cities from the same country. Income inequality, here measured with the Gini, was lower in the FUAs with compact urban cores that at the same time presented dispersed and more

spatially homogeneous built-up areas (Figure 6.6, B). These FUAs experienced higher densification and accessibility with urban growth between 2000 and 2014, which means more infilling and expansive urban growth closer to the road network. While the density of agricultural land was higher, they also lost higher proportions than more unequal FUAs in terms of income. In this sense, Boulant et al. (2016) claimed that the Gini was higher in larger cities, which usually provide more opportunities to dwellers but, in return, widen income inequalities. Meanwhile, Angel et al. (2011) related cities in countries with higher income inequalities to urban sprawl, in terms of lower population densities. Nevertheless, we did not find a significant relation between the Gini and PUGI index (which also accounts for sprawl). The GDP per capita was higher in less compact built-up shapes that experienced an increase in urban density between 2000 and 2014 (Figure 6.6, A). This trend was also found by Weilenmann et al, (2017), where wealth was positively related to higher urban densities and higher degrees of dispersion. We identified lower GDPs in compact FUAs that experienced dispersed growth with more population growth than built-up expansion between 2000 and 2014 (Figure 6.6, A). However, we found the positive correlation between GDP and the degree of urban centrality within Mexican FUAs not observed at the global level. Huang et al. (2007) also found a negative relationship between GDP per capita and compactness, stating that wealth brings more private motor vehicles and highways, which, in developed countries, contributes to the facilitation of life in outlying suburban areas; meanwhile, the lower motorization in developing countries results in more compact urban forms, as dwellers live close to their working places, usually in the inner city.

In the environmental dimension, air quality was better in FUAs with lower densities of agricultural land but higher densities of low semi-natural/natural vegetation land and water bodies (Figure 6.6, D). We also found a relationship between the pollution in the FUAs and compact shapes, both from the urban footprint and the urban core. The analysis of the compact shape of urban footprints has been proposed as a valuable indicator—besides population density, land-use mix, connectivity and accessibility—to be monitored in order to mitigate climate change. Angel et al. (2020) claimed that, other factors being equal, compact shapes reduce energy use and gas emissions. On the contrary, Bechle et al. (2011) did not find a significant correlation between compactness and NO2 concentration, but they did find such with leapfrog development and higher population densities. Regarding the change in air quality, more compact FUAs improved their air quality between 2000 and 2014, together with an increase in accessibility and a higher consumption of agricultural land as a consequence of urban growth (Figure 6.6, E). Last, concerning the employment rate change, positive rates were found in FUAs

with compact urban cores, a denser urban growth (i.e., infilling and expansive growth types) and an improvement in accessibility (Figure 6.6, F). This seems contradictory to the negative relationship between income and GDP, and built-up and urban core compactness; this may have a two-fold explanation: first, the subset of FUAs in the employment model does not represent the same geographical regions as in the GDP or income models; second, the OECD defines the employment rate as the ratio of the employed population over the working age population (OECD, 2020d), therefore, an increase in employment accompanied by a higher increase in the population of working age will result in a negative change. The employment rate model associated a higher drop in the employment rate with a higher density of low vegetation land together with greater consumption of low vegetation land due to urbanization between 2000 and 2014. Changes in employment have been previously related to LULC change in Portugal, where changes in land uses had a direct impact on labor (Meneses et al., 2017). In summary, we determined that built-up and urban core compactness are the most influential metrics for all the socio-economic variables analyzed, which has also been previously noticed by other authors (Angel et al., 2020; Ahlfeldt et al., 2018).

This analysis does not account for causality and should be interpreted cautiously; nonetheless, it helped to disentangle some relationships between the spatial patterns of functional urban areas and socio-economic indicators. Besides, the findings presented cannot be generalized to regions not covered in the analysis. The majority of the FUAs analyzed were chosen due to data availability in developed or high-income countries. Thus, we cannot assume the same relationships in developing or low-income countries until new models with more datasets are tested. In this sense, this study is a first step in exploring these global relationships and sub-models in certain regions.

In addition, some limitations should be considered when working with multi-temporal and global datasets. For example, the historical and cultural path dependencies of urban areas influence particular urban structures and land cover compositions. These influences should be considered when interpreting results at the global level. For instance, what might be considered a compact pattern in the USA versus Europe, and in high-income versus low-income countries or across continents, can be fundamentally different. Spatio-temporal metrics may have reflected those differences indirectly by means of the measured spatial patterns. Therefore, in future research, the inclusion of a categorical variable that groups FUAs with similar path dependencies or geographic-cultural contexts would be worth exploring.

On the other hand, the quality of the data is a crucial matter in this type of analysis. For instance, the GHSL used to describe the built-up areas had a

balanced accuracy of 86% (Corbane et al., 2019), which probably had an influence on the relationships found that remains unknown; however, with the interpretation of spatio-temporal metrics, we identified outliers that led to the detection of FUAs with classification errors, which were removed from the analysis, reducing the inclusion of potential errors in the models. In this direction, the use of spatio-temporal metrics linked to a boundary could be used to identify areas with anomalies and, therefore, potential errors in the GHSL database. In the realm of the OECD metropolitan areas dataset, it is still a challenge to model the variation over time, since multi-temporal data availability drastically decreases, and, when available, the data accumulate possible errors that variables might have for the two individual dates. Since different methods are applied to gather socio-economic data at the FUA level, such as aggregation or disaggregation from lower and higher levels, the reliability widely depends on the accuracy of these methods; thus, socio-economic variables are prone to uncertainties that we cannot quantify. It should be noted that the statistical data used in this study refer to data available in February 2020. After this date, OECD data are expected to be regularly updated and new cities, added to the database. However, this does not affect the proposed analysis, and the method still remains valid. Both statistical and geospatial open databases are dynamic, constantly being developed and improved; therefore, continuous changes over time are expected. Besides, statistics sometimes include estimates and assumptions; thus, data produced by different organizations for the same area are not hard facts and might differ, so they should be used with caution. However, since we compare data from the same database, we may assume that the data are consistent and the comparisons, solid. The analysis was restricted by the availability of statistical variables and geospatial data, but the inclusion of additional environmental variables, more suitable economic and social variables (e.g., employment and GDP) at the metropolitan level, and additional geo-information would be interesting to explore. Finally, the spatial boundaries used for extracting the urban spatial patterns of the EU-OECD FUAs rely on a consistent method for delineation; we recognized that due to various reasons such as the differing quality in datasets, the geometrical definition of the boundaries in some countries is not as fine as in others. For instance, Mexico, Chile and Japan showed coarser geometries than the USA or Europe, which might influence the spatio-temporal metrics, as the built-up areas were clipped using these boundaries.

The identification of socio-economic phenomena and their cross-comparison among regions, countries and continents by means of metrics derived from available geospatial databases for urban environments is increasingly feasible. These databases are continuously improving; their updates are becoming more

and more frequent since the processes are being automatized and an increasing number of satellites are providing freely available images with global coverage (e.g., the Landsat and Sentinel missions). In the foreseeable future, more comparable data with higher spatial and temporal resolutions will become available. Hence, the use of spatio-temporal metrics—describing urban spatial patterns and growth—linked to socio-economic and environmental indicators, and their change over the time, will help in improving the understanding of the drivers of the development in urban areas and their consequences at the global scale, which has been limited to date. Therefore, the proposed methodology, tested here with current semi-global data, could be extrapolated to a global scale as soon as more data become available. Furthermore, new spatial and socio-economic datasets at different levels should be explored soon, increasing the possibilities of new findings and analyses. Our preliminary outcomes show that there are common drivers and consequences of urban development within and across regions (e.g., the compactness of the built-up footprint influences or is related to household income, income inequality or GDP per capita in functional urban areas), indicating global trends. However, intra-urban variations should not be disregarded, since the high heterogeneity in terms of urban patterns and socio-economic factors existent within urban areas needs to be considered (Zhu et al., 2019; Lobo et al., 2020).

6.5. Conclusions

Monitoring the development of the built and natural elements in urban areas and the identification of their relationships with socio-economic-ecological processes allows for the comparison of these processes across regions. This will be beneficial for the elucidation of global development trends and will help in the design of more sustainable development policies. In this study, we quantified empirical and significant relationships between socio-economic-ecological indicators (income, inequality, GDP, employment rate and air quality) and spatio-temporal metrics describing the built and the natural environments. The latter were extracted from available geospatial databases in a multi-temporal manner. The spatial metrics represent the spatial organization of urban areas and LULC and their change over a period of time. They proved to be good descriptors of socio-economic and environmental processes in urban areas, tested in up to six hundred functional urban areas from 32 countries, reaching coefficients of determination varying from 32% to 68%.

Moreover, we identified the most important metrics for modeling socio-economic and environmental indicators: the compactness of built-up areas and their urban core are the spatial attributes that better relate to socio-economic

status. This could be used, for example, as a proxy of average household income in the analyzed FUAs. The concentration degree or built-up area relative to the center was important in all the models, especially for income inequality. Other relevant metrics were the dispersion index; the densification of growth and accessibility to roads, which quantify the urban growth spatial patterns in terms of their efficiency; and agricultural and low vegetation land cover densities and their change.

This first analysis aims to leverage the proliferation of long-term spatial and socio-economic databases in combination with machine learning methods, highlighting the high potential of open datasets for identifying general development growth trends in urban areas. The inclusion of more regions and higher resolution datasets will reinforce our observations. Since the availability of global datasets is an ongoing effort that many researchers and organizations are addressing, it will be feasible to identify more robust relationships in the near future at a global scale.

Chapter 7

Conclusions

7.1. Conclusions

Urban areas are environments of great interest since they are a critical part of global sustainability. Urban growth influences economic, social and environmental factors with a direct impact on inhabitants' health and well-being. Despite the growing relevance of urban environments, they are still widely unknown systems. The study of urban areas requires multi-disciplinary approaches, at multiple levels, across regions and over time. In this dissertation we aimed at contributing to the study of urban environments by developing tools and methods to efficiently characterize and monitor urban areas and their growth patterns, as well as relating their physical and socio-economic characteristics using geospatial and statistical databases. We tested these methods in different regions across the world, at multiple levels, with diverse databases and scales, spatially and temporally. This dissertation contributes to the analysis of urban environments and led us to achieve significant conclusions that are summarized in this chapter.

IndiFrag is a software tool specifically developed to analyze and monitor urban environments based on spatio-temporal metrics from multi-temporal LULC databases. This software provides researchers and planners with a tool to extract spatial information from LULC maps that can be used to monitor the development of urban areas, evaluate their growth and make informed decisions. This software tool is freely available which, together with the growing availability of EO-data and LULC databases, increases its potential use by the scientific community. In fact, IndiFrag has recently been used in studies by other authors (e.g., González-Yebra et al., 2018; Pili et al., 2019; Trinder and Liu, 2020), which shows its short-term impact on urban and natural environmental research.

Spatio-temporal metrics can be used in a wide range of applications and applied to multiple geospatial datasets (e.g., LULC, LCZ and built-up maps); however, their interpretation varies according to the MMU, thematic legend, unit and level of analysis. For this reason, it is important that comparative analyses based on spatio-temporal metrics be consistent with these factors. For instance, regarding the level of analysis, metrics at the intra-urban level provide specific and spatially explicit information on land use growth patterns, opposite to metrics at the inter-urban level that provide smoother or averaged values but allow for the cross-comparison of urban areas. Despite these differences, multi-level analyses are complementary. Analysis at local level is useful to identify discrepancies in growth patterns and distinctive features within an urban area, evidencing internal variabilities, which is important due to the intrinsic heterogeneity of urban areas. The use of broader levels provides an overview of the current state and development of urban areas and allows

for their comparison across space, which is important to find similarities between urban areas worldwide or to compare their growth patterns. In this context, the scope of analysis plays an important role. While local levels are closer to the administrative units used in local urban planning, regional or national decision-making may also require broader level approaches. Therefore, the level of analysis must be carefully selected according to the aim of the study, being restricted by the availability of data.

The population and urban growing imbalance index (PUGI) proposed in this thesis complements the multi-temporal LULC analysis based only on spatio-temporal metrics. This index provides information about demographic dynamics and urban growth, the inequality of their evolution, as well as patterns of land consumption per capita and degree of sprawl. This is a valuable information missed with the solely use of spatial metrics, and it has been shown to be relevant to evaluate the sustainability of growth patterns. Besides, this index can be used to quantify the inequality between any pair of variables, such as residential and green areas growth imbalance, and combined with spatio-temporal metrics supports the evaluation of urban planning and policy measures, along with monitoring the implementation of measures towards the fulfillment of the SDGs.

There is a relation between the urban spatial structure in cities and their quality of life and sustainability. This was demonstrated in North-Rhine Westphalia (Germany) by means of spatial metrics applied to Local Climate Zone (LCZ) maps. The metrics explained the variability of education, health, living conditions, labor, and transport indicators at inter-city level by means of statistical models. Grouping cities according to their socio-economic similarities allowed us to identify four groups with similar levels of quality of life, showing common internal spatial patterns. This confirms that the spatial structure of cities influences the quality of life of their inhabitants, and vice versa, although does not completely define it. The proposed method is transferable to other geographic areas, levels and datasets, and can be used to extract more empirical evidences of these relationships between the spatial structure and the socio-economic performance in different and larger urban environments. This knowledge is crucial for urban planning, and worth it to be explored deeper in other geographic areas.

Regarding growth patterns, a small subset of significant spatio-temporal metrics is sufficient to identify a set of well-defined urban growth spatial patterns from LULC maps. The spatial attributes that best discriminated between expansive, compact, dispersed, road-based and leapfrog patterns were the densification, compactness and concentration of growth. These attributes can be quantified based on two metrics, the Area-weighted mean

standard distance (AWSD) and the Area-weighted mean expansion index (AWMEI). These metrics can be used for the early identification of growth patterns based on LULC databases, which is useful for planners to anticipate to the consequences of unsustainable growth patterns and planning cities accordingly. However, it should be highlighted that spatio-temporal metrics were influenced by the urban form, which introduces noise in the classification of growth spatial patterns. Therefore, new approaches to diminish this influence when working with spatio-temporal metrics derived from multi-temporal LULC databases are needed.

Despite the current limited availability of open LULC, geospatial and statistical databases, we demonstrated the relationships between the spatial structure of urban environments and their socio-economic performance at a semi-global level, and this can be quantified by means of spatio-temporal metrics extracted from open databases available worldwide and statistical models. We found relationships between income, inequality, GDP and air pollution indicators and spatio-temporal metrics of built-up areas in 32 countries. By means of random forest regression models the most explanatory metrics for modeling these indicators were identified. These spatio-temporal metrics can be further used as proxies of socio-economic indicators at global scale. The compactness of built-up areas is one of the most relevant spatial attributes in all the models created, providing useful information to analyze and compare urban environments on a global scale when socio-economic data is not available.

Modeling variations in socio-economic indicators over time by means of spatio-temporal metrics of built-up areas had certain limitations at a semi-global level. Modeling socio-economic change still remains a challenge because consistent multi-temporal socio-economic databases at a global scale are scarce. Nevertheless, changes in air quality and employment rate indicators were partially explained using the metrics. Overall, a semi-global and cross-comparative analysis underlines the potential of long-term spatial and socio-economic databases for identifying and monitoring urban growth patterns on a broad scale. The proposed method can be extended to a more complete sample of geographical areas as soon as new data become available, and new databases could be included to extract spatial characteristics to be easily incorporated for modeling socio-economic indicators at a global scale.

This thesis contributes to a broader and better understanding of urban environments, their growth patterns and their relationships with social, economic and environmental processes. In general, this dissertation generates valuable knowledge and new methods for monitoring and evaluating the sustainability of growth in urban and peri-urban areas. In this context and on the basis of our findings and proposed methods, the expected availability of

long-term and better resolved geospatial and socio-economic data will provide new evidences in this field, which will bring a deeper understanding of how urban areas function and evolve, both physically and socio-economically, as well as their interrelations. This is relevant information to successfully monitor urban areas, evaluate planning policies and obtain empirical evidences across the world that will help managers and decision-makers to plan cities and urban areas towards their sustainable development.

7.2. Future research

Future research could span in several directions. The methods proposed in this thesis for analyzing relationships between the spatial structure in urban areas and socio-economic indicators are transferable. Thus, the models could be replicated or performed in larger regions within different context (e.g., including urban areas in low-income countries), using several levels of analysis, MMU, dates and databases, as soon as new data become available. This is very likely to happen thanks to all the efforts made in recent years regarding the development of global geospatial databases using EO-data, as well as the increasing availability of statistical databases. New analyses based on these methods would reinforce our findings and provide new insights and evidences in other geographical contexts and for not yet explored socio-economic indicators.

Another research line to explore further is the multi-temporal analysis. We encountered limitations in modeling socio-economic changes based on spatio-temporal metrics derived from geospatial databases due to the lack of long-term geospatial and statistical data. In this sense, in Europe, the current development of the LULC Urban Atlas 2018 database (besides the 2006 and 2012 versions) and its close link to the City Statistics database from Eurostat, brings the opportunity to start developing multi-temporal models for a period of 12 years for more than 300 urban areas. This will provide more solid conclusions and new insights regarding the relationships between LULC change patterns and quality of life than the preliminary explored relationships carried out in this thesis for a limited period of 6 years, based on the availability of data (Sapena et al., 2016).

The recent trend in the development of global maps of multi-temporal built-up areas, introduced in Chapter 1, could open a line of work. These datasets might be used for classifying urban areas into different urban forms and growth spatial patterns at multiple levels and across different regions by means of spatio-temporal metrics derived from the multi-temporal built-up maps. Then, two different directions might be taken. First, the combination of spatio-

temporal information with indices such as the PUGI (if population data is available), and use it to explore the consequences of different growth pattern on social, economic and environmental variables, applying appropriate statistical methods. Second, sensitivity analyses with varying spatial units and the influence of urban forms on spatio-temporal metrics and in the identification of growth patterns could be further explored (e.g., the inclusion of a categorical variable or some quantitative metric in the classification process including information of urban forms).

References

A

Abrantes, P., Rocha, J., Marques da Costa, E., Gomes, E., Morgado, P., Costa, N., 2019. Modelling urban form: A multidimensional typology of urban occupation for spatial analysis. Environment and Planning B: Urban Analytics and City Science, 46(1), 47-65. Doi:10.1177/2399808317700140.

Acuto, M., Parnell, S., Seto, K. C., Contestabile, M., Allen, A., Attia, S., Bai, X, et al., 2018. Science and the Future of Cities. Expert Panel Report on Science and the Future of Cities, 60. Doi:10.13140/rg.2.2.27706.64969

Adamczyk, J., Tiede, D., 2017. ZonalMetrics-a Python toolbox for zonal landscape structure analysis. Computers & Geosciences, 99, 91-99. Doi:10.1016/j.cageo.2016.11.005.

Ahlfeldt, G., Pietrostefani, E., Schumann, A., Matsumoto, T., 2018. Demystifying compact urban growth: Evidence from 300 studies from across the world, OECD Regional Development Working Papers, No. 2018/03, OECD Publishing, Paris, Doi:10.1787/bbea8b78-en.

Allen, L., Williams, J., Townsend, N., Mikkelsen, B., Roberts, N., Foster, C. Wickramasinghe, K., 2017. Socioeconomic status and non-communicable disease behavioural risk factors in low-income and lower-middle-income countries: a systematic review. The Lancet Global Health, 5(3), e277-e289. Doi:10.1016/S2214-109X(17)30058-X.

Altieri, L., Cocchi, D., Pezzi, G., Scott, E.M., Ventrucci, M., 2014. Urban sprawl scatterplots for Urban Morphological Zones data. Ecological Indicators. 36, 315-323. Doi:10.1016/j.ecolind.2013.07.011.

Anderson, W.P., Kanaroglou, P.S., Miller, E.J., 1996. Urban Form, Energy and the Environment: A Review of Issues, Evidence and Policy. Urban Studies, 33(1) 7-35.

Angel, S., Arango Franco, S., Liu, Y., Blei, A. M., 2020. The shape compactness of urban footprints. Progress in Planning, 139, 100429.

Doi:10.1016/j.progress.2018.12.001.

Angel, S., Blei, A.M., Parent, J., Lamson-Hall, P., Galarza-Sánchez,N., Civco, D.L., Qian Lei, R., Thom, K., 2016. Atlas of urban expansion – 2016 Edition. Ed: Lincoln Institute of Land Policy: Cambridge, USA, www.atlasofurbanexpansion.org.

Angel, S., Parent, J., Civco, D.L., Blei, A.M., 2011. Making room for a planet of cities. Ed: Lincoln Institute of Land Policy: Cambridge, USA,

Araya, Y.H., Cabral, P., 2010. Analysis and Modeling of Urban Land Cover Change in Setúbal and Sesimbra, Portugal. Remote Sensing, 2, 1549-1563. Doi:10.3390/rs2061549.

Arribas-Bel, D., Nijkamp, P., Scholten, H., 2011. Multidimensional urban sprawl in Europe: A self-organizing map approach. Computers, Environment and Urban Systems. 35, 263-275. Doi:10.1016/j.compenvurbsys.2010.10.002.

Arribas-Bel, D., Patino, J.E., Duque, J.C., 2017. Remote sensing-based measurement of Living Environment Deprivation: Improving classical approaches with machine learning. PLoS ONE, 12(5), e0176684. Doi:10.1371/journal.pone.0176684

Aubrecht, C., Steinnocher, K., Hollaus, M., Wagner, W., 2009. Integrating earth observation and GIScience for high resolution spatial and functional modeling of urban land use. Computers, Environment and Urban Systems, 33(1), 15-25. Doi:10.1016/j.compenvurbsys.2008.09.007.

B

Batchis, W., 2010. Urban Sprawl and the Constitution: Educational Inequality as an Impetus to Low Density Living. The Urban Lawyer, 42, 95-133. www.jstor.org/stable/27895768

Baumont, C., Ertur, C., and Le Gallo, J., 2004. Spatial analysis of employment and population density: the case of the agglomeration of Dijon 1999. Geographical analysis, 36(2), 146–176. Doi:10.1111/j.1538-4632.2004.tb01130.x.

Bechle, M.J., Millet, D.B., Marshall, J.D., 2011. Effects of Income and Urban Form on Urban NO2: Global Evidence from Satellites. Environmental Science & Technology, 45(11), 4914-4919. Doi:10.1021/es103866b.

Bechtel, B., Alexander, P.J., Böhner, J., Ching, J., Conrad, O., Feddema, J., Mills, G., See, L., Stewart, I., 2015. Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities. ISPRS International Journal of Geo-Information, 4(1), 199-219. Doi:10.3390/ijgi4010199.

Bechtel, B., Demuzere, M., Mills, G., Zhan, W., Sismanidis, P., Small, C., Voogt, J., 2019. SUHI analysis using Local Climate Zones - A comparison of 50 cities.

Urban Climate, 28, 100451. Doi:10.1016/j.uclim.2019.01.005.

Belsky, D. W., Caspi, A., Arseneault, L., Corcoran, D.L., Domingue, B.W., Harris, K.M., Houts, R.M., Mill, J.D., Moffitt, T.E., Prinz, J., Sugden, K., Wertz, J., Williams, B., Odgers, C., 2019. Genetics and the geography of health, behaviour and attainment. Nature Human Behaviour, 3(6), 576-586. Doi:10.1038/s41562-019-0562-1.

Bhatta, B., 2010. Causes and Consequences of Urban Growth and Sprawl, in: Analysis of Urban Growth and Sprawl from Remote Sensing Data. Springer: Heidelberg, Germany. 17-37. Doi:10.1007/978-3-642-05299-6.

Bosch, M., 2019. PyLandStats: An opensource Pythonic library to compute landscape metrics. PLoS ONE, 14(12), e0225734. Doi:10.1371/journal.pone.0225734.

Boulant, J., Brezzi, M., Veneri, P., 2016. Income Levels And Inequality in Metropolitan Areas: A Comparative Approach in OECD Countries. OECD Regional Development Working Papers, No. 2016/06, OECD Publishing, Paris, Doi:10.1787/5jlwj02zz4mr-en.

Breiman, L., 2001a. Random Forest. Machine Learning, 45, 5-32. Doi:10.1023/A:1010933404324.

Breiman, L., 2001b. Statistcal modeling: the two cultures. Statistical Science, 16(3), 199-231. Doi:10.1214/ss/1009213726.

Brown de Colstoun, E.C., Huang, C., Wang,P., Tilton, J.C., Tan, B., Phillips, J., Niemczura, S., Ling, P.Y., Wolfe, R.E., 2017. Global Man-made Impervious Surface (GMIS) Dataset From Landsat. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). Doi:10.7927/H4P55KKF. Accessed 05/06/2020.

Buchhorn, M., Bertels, L., Smets, B., De Roo, B., Lesiv, M., Tsendbazar, N.E., Masiliunas, D., Li, L., 2020a. Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Algorithm Theoretical Basis Document; 2020. Doi:10.5281/zenodo.3606361.

Buchhorn, M., Lesiv, M., Tsendbazar, N.E., Herold, M., Bertels, L., Smets, B., 2020b. Copernicus Global Land Cover Layers—Collection 2. Remote Sensing, 12, 1044. Doi:10.3390/rs12061044.

C

Camagni, R., Gibelli, M.C., and Rigamonti, P., 2002. Urban mobility and urban form: the social and environmental costs of different patterns of urban expansion. Ecological Economics, 40(2), 199–216. Doi:10.1016/S0921-

8009(01)00254-3.

Cao, W., Dong, L., Wu, L., Liu, Y., 2020. Quantifying urban areas with multi-source data based on percolation theory. Remote Sensing of Environment, 241, 111730. Doi:10.1016/j.rse.2020.111730.

Chen, J., Cao, X., Peng, S., Ren, H., 2017. Analysis and applications of GlobeLand30: a review. ISPRS International Journal of Geo-Information, 6(8), 230. Doi:10.3390/ijgi6080230.

Chen, M., Liu, W., Tao, X., 2013. Evolution and assessment on China's urbanization 1960-2010: Under-urbanization or over-urbanization? Habitat International, 38, 25-33. Doi:10.1016/j.habitatint.2012.09.007.

Chen, X., Nordhaus, W.D., 2011. Using luminosity data as a proxy for economic statistics. Proceedings of the National Academy of Sciences, 108(21), 8589-8594. Doi:10.1073/pnas.1017031108.

Chen, Y., Li, X., Zheng, Y., Guan, Y., Liu, X., 2011, Estimating the relationship between urban forms and energy consumption: A case study in the Pearl River Delta, 2005–2008. Landscape and Urban Planning, 102(1), 33–42. Doi:10.1016/j.landurbplan.2011.03.007.

Cheshire, P., Leunig, T., Nathan, M., Overman, H., 2012. Links between planning and economic performance: Evidence note for LSE Growth Commission. LSE Growth Commission and Institute for Government.

Chin, N., 2002. Unearthing the roots of urban sprawl: a critical analysis of form, function and methodology. Working Paper 47. Centre for Advanced Spatial Analysis (UCL): London, UK.

Clark, J.K., McChesney, R., Munroe, D.K., Irwin, E.G., 2009. Spatial characteristics of exurban settlement pattern in the United States. Landscape and Urban Planning, 90(3–4), 178–188. Doi:10.1016/j.landurbplan.2008.11.002.

Clarke, K.C., Gaydos, L., Hoppen, S., 1997. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. Environment and Planning B, 24, 247–261.

Commission for Environmental Cooperation (CEC), 2020. North America Land Change Monitoring System. Available from: http://www.cec.org/nalcms (Accessed 01/07/2020).

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote sensing of environment. 37, 35-46. Doi:10.1016/0034-4257(91)90048-B.

Copernicus, 2016. Mapping Guide for a European Urban Atlas.

http://land.copernicus.eu/user-corner/technical-library/.   (Accessed 16/09/2016).

Copernicus,   2020a.   Copernicus   services.   Available   from: https://www.copernicus.eu/ (Accessed 026/07/2020).

Copernicus, 2020b. Copernicus Land Monitoring Service. Available from: https://land.copernicus.eu/ (Accessed 01/07/2020).

Corbane, C., Florczyk, A., Pesaresi, M., Politis, P., Syrris, V., 2018a. GHS-BUILT R2018A - GHS built-up grid, derived from Landsat, multitemporal (1975-1990-2000-2014). European Commission, Joint Research Centre (JRC). [Dataset].   Doi: 10.2905/jrc-ghsl-10007.   Available   online: http://data.europa.eu/89h/jrc-ghsl-10007. (accessed 2/01/2020).

Corbane, C., Pesaresi, M., Kemper, T., Politis, P., Florczyk, A.J., Syrris, V., Melchiorri, M., Sabo, F., Soille, P., 2019, Automated global delineation of human settlements from 40 years of landsat satellite data archives. Big Earth Data, 3, 140-169, Doi:10.1080/20964471.2019.1625528.

Corbane, C., Politis, P., Syrris, V., Pesaresi, M., 2018b. GHS built-up grid, derived from Sentinel-1 (2016), R2018A. European Commission, Joint Research Centre   (JRC)   Doi: 10.2905/jrc-ghsl-10008.   PID: http://data.europa.eu/89h/jrc-ghsl-10008.

Corbane, C., Sabo, F., 2019. European Settlement Map from Copernicus Very High Resolution data for reference year 2015, Public Release 2019. European   Commission,   Joint   Research   Centre   (JRC)   Doi: 10.2905/8BD2B792-CC33-4C11-AFD1-B8DD60B44F3B.   PID: http://data.europa.eu/89h/8bd2b792-cc33-4c11-afd1-b8dd60b44f3b.

Cushman, S.A., McGarigal, K., Neel, M.C., 2008. Parsimony in landscape metrics: Strength, universality, and consistency. Ecological Indicators, 8(5). 691-703. Doi:10.1016/j.ecolind.2007.12.002.

D

De Leeuw, J., Georgiadou, Y., Kerle, N., De Gier, A., Inoue, Y., Ferwerda, J., Smies, M., Narantuya, D., 2010. The Function of Remote Sensing in Support of Environmental   Policy.   Remote   Sensing,   2,   1731-1750. Doi:10.3390/rs2071731.

Demuzere, M., Bechtel, B., Middel, A., Mills, G., 2019. Mapping Europe into local climate   zones.   PLOS   ONE,   14(4),   e0214474. Doi:10.1371/journal.pone.0214474

Deng, J., Desjardins, M.R., Delmelle, E.M., 2019. An interactive platform for the

analysis of landscape patterns: a cloud-based parallel approach. Annals of GIS, 25(2), 99-111. Doi:10.1080/19475683.2019.1615550.

Dewan, A.M., Yamaguchi, Y., Ziaur Rahman, M., 2012. Dynamics of land use/cover changes and the analysis of landscape fragmentation in Dhaka Metropolitan, Bangladesh. GeoJournal, 77, 315–330. Doi: 10.1007/s10708-010-9399-x.

DG REGIO, 2011. Cities of tomorrow-Challenges, visions, ways forward. Directorate-General for Regional Policy, European Commission, Luxembourg: Publications Office of the European Union. Doi:10.2776/41803.

Dong, T., Jiao, L., Xu, G., Yang, L., Liu, J., 2019. Towards sustainability? Analyzing changing urban form patterns in the United States, Europe, and China. Science of The Total Environment, 671, 632-643. Doi:10.1016/j.scitotenv.2019.03.269.

Donnay, J.P., Barnsley, M.J., Longley, P.A., 2000. Remote sensing and urban analysis: GISDATA 9. CRC Press. Doi:10.1201/9781482268119.

Dorning, M.A., Koch, J., Shoemaker, D.A., Meentemeyer, R.K., 2015. Simulating urbanization scenarios reveals tradeoffs between conservation planning strategies. Landscape and Urban Planning, 136, 28–39. Doi:10.1016/j.landurbplan.2014.11.011.

Dubovyk, O., Sliuzas, R., Flacke, J., 2011. Spatio-temporal modelling of informal settlement development in Sancaktepe district, Istanbul, Turkey.ISPRS Journal of Photogrammetry and Remote Sensing, 66, 235-246. Doi:10.1016/j.isprsjprs.2010.10.002.

Duque, J.C., Lozano-Gracia, N., Patino, J.E., Restrepo, P., Velasquez, W.A., 2019. Spatiotemporal dynamics of urban growth in Latin American cities: An analysis using nighttime light imagery. Landscape and Urban Planning, 191, 103640. Doi:10.1016/j.landurbplan.2019.103640

Duque, J.C., Patino, J.E., Ruiz, L.A., Pardo-Pascual, J.E., 2015. Measuring intra-urban poverty using land cover and texture metrics derived from remote sensing data. Landscape and Urban Planning, 135, 11-21. Doi:10.1016/j.landurbplan.2014.11.009.

E

Engelen, G., Lavalle, C., Barredo, J.I., Meulen, M., White, R., 2007. The MOLAND Modelling Framework for Urban and Regional Land-Use Dynamics. In: Modelling Land-Use Change. Springer. 297-319. Doi:10.1007/1-4020-

5648-6_17

ESA, 2017. Land Cover CCI Product User Guide Version 2. Tech. Rep., Available online: http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf. (Accessed 06/02/2020)

ESA, 2019. Applications - Mapping our global human footprint. https://bit.ly/35pqWzD (Accessed 26/07/2020)

Esch, T., Asamer, H., Bachofer, F., Balhar, J., Böttcher, M., Boissier, E., Hirner, A., Mathot, E., Marconcini, M., Metz-Marconcini, A., Permana, H., Soukup, T., Svaton, V., Üreyen S., Zeidler, J., 2018. New Prospects in Analysing Big Data from Space - The Urban Thematic Exploitation Platform. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, 2018, 8193-8196. Doi:10.1109/IGARSS.2018.8517493.

Esch, T., Taubenböck, H., Roth, A., Heldens, W., Felbier, A., Thiel, M., Schmidt, M., Müller, A. Dech, S., 2012. TanDEM-X mission: New perspectives for the inventory and monitoring of global settlement patterns. Journal of Applied Remote Sensing, 6(1), 061702. Doi:10.1117/1.JRS.6.061702.

ESPON, 2005. Potentials for polycentric development in Europe. ESPON 1.1.1 Project Report.

European and Global Forum for Geography and Statistics (EFGS), 2019. https://www.efgs.info/. (Accessed 10/03/2019).

European Commission (EC), 2012b. Soil sealing, Science for Environment Policy. In-depth Report.

European Commission (EC), 2013. Living well, within the limits of our planet: 7th EAP - The new general Union Environment Action Programme to 2020.

European Commission (EC), 2016. Supporting the Implementation of Green Infrastructure. Final Report, Rotterdam.

European Commission (EC), 2017. Urban Agenda for the EU - Orientation Paper - Sustainable Use of Land and Nature-based Solutions.

European Commission (EC), 2019. Growth, Internal market, Industry, Entrepreneurship and SMEs. https://ec.europa.eu/growth/tools-databases/regional-innovation-monitor/base-profile/north-rhine-westphalia. (Accessed 22/01/2019).

European Communities, 1999. European Spatial Development Perspective (ESDP). Towards balanced and sustainable development of the territory of the European Union. Office for Official Publications of the European Communities, Luxembourg.

European Environment Agency (EEA), 2011. Analysing and managing urban growth. European Environment Agency. http://www.eea.europa.eu/articles/analysing-and-managing-urban-growth. (Accessed 11/03/2015).

European Environment Agency (EEA), 2014. Urban morphological zones 2006. Available online: https://www.eea.europa.eu/data-and-maps/data/urban-morphological-zones-2006-1 (accessed on 3 June 2020)

European Environment Agency (EEA), 2016a. Urban Sprawl in Europe. Joint EEA-FOEN report. Publications Office of the European Union, Luxembourg. Doi:10.2800/143470.

European Environment Agency (EEA), 2016b. Copernicus Land Monitoring Service, Urban Atlas [dataset]. Available from: http://land.copernicus.eu/local/urban-atlas. (Accessed 01/06/2017).

European Union (EU), 2016. Urban Europe. Statistics on cities, towns and suburbs. Luxemburg: Publications office of the European Union. Doi:10.2785/91120.

Eurostat, 2016a. City statistics database. Available from: https://ec.europa.eu/eurostat/web/cities/data/database. (Accessed 15/10/2018).

Eurostat, 2016b. Cities (Urban Audit). https://ec.europa.eu/eurostat/web/cities/background. (Accessed 15/04/2020).

Eurostat, 2017. Final report of the expert group on quality of life indicators. Luxembourg: Publications Office of the European Union.

Eurostat, 2018. Statistics on European cities. http://ec.europa.eu/eurostat/statistics-explained/index.php/Statistics_on_European_cities. (Accessed 09/10/2018).

Eurostat, 2019. Regional statistics by NUTS classification. https://ec.europa.eu/eurostat/web/regions/data/database. (Accessed 15/01/2019).

F

Faisal, K., Shaker, A., Habbani, S., 2016. Modeling the relationship between the gross domestic product and built-up area using remote sensing and GIS data: A case study of seven major cities in Canada. ISPRS International Journal of Geo-Information, 5(3), 23. Doi:10.3390/ijgi5030023.

Feng, Y., Liu, Y., Batty, M., 2016. Modeling urban growth with GIS based cellular automata and least squares SVM rules: a case study in Qingpu-Songjiang area of Shanghai, China. Stochastic Environmental Research and Risk

Assessment, 30(5), 1387–1400. Doi:10.1007/s00477-015-1128-z.

G

Genuer, R., Poggi, J.M., Tuleau-Malot, C., 2015. VSURF: An R Package for Variable Selection Using Random Forests. The R Journal, 7(2) 19-33. Doi:10.32614/RJ-2015-018.

Georg, I., Blaschke, T., and Taubenböck, H., 2016. New spatial dimensions of global cityscapes: From reviewing existing concepts to a conceptual spatial approach. Journal of Geographical Sciences, 26(3), 355–380. Doi:10.1007/s11442-016-1273-4.

GEOSTAT, 2020. Eurostat, Geographical Information and maps. https://ec.europa.eu/eurostat/web/gisco/gisco-activities/integrating-statistics-geospatial-information/geostat-initiative. (Accessed 15/04/2020).

Ghafouri, B., Amiri, B.J., Shabani, A.A., Songer, M., 2016. Examining Relationships between Socioeconomic Factors and Landscape Metrics in the Southern Basin of the Caspian Sea. Environmental Modeling Assessment. Doi:10.1007/s10666-016-9503-9.

Gielen, E., Riutort-Mayol, G., Miralles i Garcia, J.L., Palencia Jiménez, J.S., 2019. Cost assessment of urban sprawl on municipal services using hierarchical regression. Environment and Planning B: Urban Analytics and City Science, Doi: 10.1177/2399808319869345.

Gielen, E., Riutort-Mayol, G., Palencia-Jiménez, J. S., Cantarino, I., 2018. An urban sprawl index based on multivariate and Bayesian factor analysis with application at the municipality level in Valencia. Environment and Planning B: Urban Analytics and City Science, 45(5), 888-914. Doi:10.1177/2399808317690148.

Gil-Yepes, J.L., Ruiz, L.A., Recio, J.A., Balaguer-Beser, Á., Hermosilla, T., 2016. Description and validation of a new set of object-based temporal geostatistical features for land-use/land-cover change detection. ISPRS Journal of Photogrammetry and Remote Sensing, 121, 77-91. Doi:10.1016/j.isprsjprs.2016.08.010.

Goerlich, F.J., Cantarino, I., Gielen, E., 2017. Clustering cities through urban metrics analysis. Journal of Urban Design, 22(5), 689-708. Doi:10.1080/13574809.2017.1305882.

Gong, C., Yu, S., Joesting, H., Chen, J., 2013. Determining socioeconomic drivers of urban forest fragmentation with historical remote sensing images. Landscape and Urban Planning, 117, 57-65.

Doi:10.1016/j.landurbplan.2013.04.009.

Gong, P., Li, X., Wang, J., Bai, Y., Chen, B., Hu, T., Liu, X., Xu, B., Yang, J., Zhang, W., Zhou, Y., 2020. Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. Remote Sensing of Environment, 236, 111510. Doi:10.1016/j.rse.2019.111510.

Gonzalez, J.J., Leboulluec, A., 2019. Crime Prediction and Socio-Demographic Factors: A Comparative Study of Machine Learning Regression-Based Algorithms. Journal of Applied Computer Science Mathematics, 13(1), 13-18. Doi: 10.4316/JACSM.201901002.

González-Yebra, Ó., Aguilar, M.A., Nemmaoui, A., Aguilar, F.J., 2018. Methodological proposal to assess plastic greenhouses land cover change from the combination of archival aerial orthoimages and Landsat data. Biosystems Engineering, 175, 36-51. Doi:10.1016/j.biosystemseng.2018.08.009

GRASS Development Team, 2017. Geographic Resources Analysis Support System (GRASS) Software, Version 7.2. Open Source Geospatial Foundation. Available from: http://grass.osgeo.org.

Grekousis, G., Mountrakis , G., Kavouras, M., 2015. An overview of 21 global and 43 regional land-cover mapping products. International Journal of Remote Sensing, 36(21), 5309-5335, Doi:10.1080/01431161.2015.1093195.

Grimm, N.B., Cook, E.M., Hale, R.L., Iwaniec, D.M., 2015. A broader framing of ecosystem services in cities. Benefits and challenges of built, natural or hybrid system function. Routledge Handbooks Online.

Güneralp, B., Reba, M., Hales, B.U., Wentz, E.A., Seto, K.C., 2020. Trends in urban land expansion, density, and land transitions from 1970 to 2010: a global synthesis. Environmental Research Letters, 15(4), 044015. Doi:10.1088/1748-9326/ab6669.

H

Haase, D., Kabisch, N., Haase, A., 2013. Endless urban growth? On the mismatch of population, household and urban land area growth and its effects on the urban debate. PLoS ONE, 8(6), 1-8. Doi:10.1371/journal.pone.0066531.

Hankey, S., Marshall, J.D., 2017. Urban Form, Air Pollution, and Health. Current Environmental Health Reports, 4(4) 491-503. Doi:10.1007/s40572-017-0167-7.

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-means clustering algorithm. Applied Statistics, 28, 100–108. Doi:10.2307/2346830.

He, C., Liu, Z., Gou, S., Zhang, Q., Zhang, J., Xu, L., 2019. Detecting global urban expansion over the last three decades using a fully convolutional network. Environmental Research Letters, 14(3), 034008. Doi:10.1088/1748-9326/aaf936.

Hennig, E. I., Schwick, C., Soukup, T., Orlitová, E., Kienast, F., Jaeger, J. A. G., 2015. Multi-scale analysis of urban sprawl in Europe: Towards a European de-sprawling strategy. Land Use Policy, 49, 483-498. Doi:10.1016/j.landusepol.2015.08.001.

Hermosilla, T., Ruiz, L. A., Recio, J. A., Cambra-López, M., 2012. Assessing contextual descriptive features for plot-based classification of urban areas. Landscape and Urban Planning, 106(1) 124-137. Doi:10.1016/j.landurbplan.2012.02.008.

Hermosilla, T., Ruiz, L. A., Recio, J. A., Estornell, J., 2011. Evaluation of automatic building detection approaches combining high resolution images and LiDAR data. Remote Sensing, 3(6), 1188-1210. Doi:10.3390/rs3061188.

Herold, M., Couclelis, H., Clarke, K.C., 2005. The role of spatial metrics in the analysis and modeling of urban land use change. Computers, Environment and Urban Systems, 29, 369-399. Doi:10.1016/j.compenvurbsys.2003.12.001.

Herold, M., Scepan, J., Clarke, K.C., 2002. The use of remote sensing and landscape metrics to describe structures and changes in urban land uses. Environment and Planning A: Economy and Space, 34, 1443–1458. Doi:10.1068/a3496

Hesselbarth, M.H., Sciaini, M., With, K.A., Wiegand, K., Nowosad, J., 2019. Landscapemetrics: an open-source R tool to calculate landscape metrics. Ecography, 42(10), 1648-1657. Doi:10.1111/ecog.04617.

Hoymann, J., Goetzke, R., 2016. Simulation and evaluation of urban growth for Germany including climate change mitigation and adaptation measures. ISPRS International Journal of Geo-Information, 5(7), 101. Doi:10.3390/ijgi5070101.

Huang, J., Lu, X.X., Sellers, J.M., 2007. A global comparative analysis of urban form: Applying spatial metrics and remote sensing. Landscape and Urban Planning, 82(4), 184-197. Doi:10.1016/j.landurbplan.2007.02.010.

I

Informal Ministerial Meeting on Urban Development Declaration, 2010, Declaration, Toledo, 22 June 2010,

http://www.mdrap.ro/userfiles/declaratie_Toledo_en.pdf. (Accessed 20/06/2017).

Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., Rodes, I., 2017. Operational high resolution land cover map production at the country scale using satellite image time series. Remote Sensing, 9(1), 95. Doi:10.3390/rs9010095.

Inostroza, L., Baur, R., Csaplovics, E., 2013. Urban sprawl and fragmentation in Latin America: A dynamic quantification and characterization of spatial patterns. Journal of environmental management, 115, 87-97. Doi:10.1016/j.jenvman.2012.11.007.

Irvine, J.M., Wood, R.J., McBee, P., 2017. Remote Sensing to Analyze Wealth, Poverty, and Crime. IEEE Applied Imagery Pattern Recognition Workshop. AIPR. Washington, DC, 1-9.

IVIE, 2013. Análisis de la situación económica, social y territorial de la Comunidad Valenciana. Documento base para definir la estrategia de la región en el período de programación 2014-2020. Resumen Ejecutivo. Conselleria de Hacienda y Modelo Económico, Generalitat Valenciana.

J

Jacobs, J., 1961. The Death and Life of Great American Cities. New York: Random House.

Jaeger, J.A.G., 2000. Landscape division, splitting index, and effective mesh size: New measures of landscape fragmentation. Landscape Ecology, 15, 115–130. Doi:10.1023/A:1008129329289.

Jaeger, J.A.G., Bertiller, R., Schwick, C., Cavens, D., Kienast, F., 2010b. Urban permeation of landscapes and sprawl per capita: New measures of urban sprawl. Ecological Indicators, 10 (2), 427-441. Doi:10.1016/j.ecolind.2009.07.010.

Jaeger, J.A.G., Bertiller, R., Schwick, C., Kienast, F., 2010a. Suitability criteria for measures of urban sprawl. Ecological Indicators, 10(2), 397-406. Doi:10.1016/j.ecolind.2009.07.007.

Jaeger, J.A.G., Schwick, C., 2014. Improving the measurement of urban sprawl: Weighted Urban Proliferation (WUP) and its application to Switzerland. Ecological Indicators, 38, 294-308. Doi:10.1016/j.ecolind.2013.11.022.

Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B., Ermon, S., 2016. Combining satellite imagery and machine learning to predict poverty. Science, 353. 6301. 790-794. Doi: 10.1126/science.aaf7894.

Jiao, L., 2015. Urban land density function: A new method to characterize urban expansion. Landscape and Urban Planning, 139, 26–39. Doi:10.1016/j.landurbplan.2015.02.017.

Jokar Arsanjani, J., Helbich, M., de Noronha Vaz, E., 2013. Spatiotemporal simulation of urban growth patterns using agent-based modeling: The case of Tehran. Cities, 32, 33–42. Doi:10.1016/j.cities.2013.01.005.

Jolliffe, I., 2002. Principal Component Analysis, Springer Ser. Statist, Springer-Verlag, New York. Doi:10.1007/b98835.

Ju, H., Zhang, Z., Zuo, L., Wang, J., Zhang, S., Wang, X., and Zhao, X., 2016. Driving forces and their interactions of built-up land expansion based on the geographical detector – a case study of Beijing, China. International Journal of Geographical Information Science, 30(11), 2188–2207. Doi:10.1080/13658816.2016.1165228.

Jung, M., 2016. LecoS - A python plugin for automated landscape ecology analysis. Ecological Informatics, 31, 18-21. Doi:10.1016/j.ecoinf.2015.11.006.

K

Kabisch, N., Haase, D., 2013. Green spaces of European cities revisited for 1990-2006. Landscape and Urban Planning, 110, 113-122. Doi:10.1016/j.landurbplan.2012.10.017.

Kasanko, M., Barredo, J.I., Lavalle, C., McCormick, N., Demicheli, L., Sagris, V., Brezger, A., 2006. Are European cities becoming dispersed? A comparative analysis of 15 European urban areas. Landscape and Urban Planning, 77, 111-130. Doi:10.1016/j.landurbplan.2005.02.003.

Ketchen, D.J., Shook, C.L., 1996. The application of cluster analysis in strategic management research: an analysis and critique. Strategic management journal, 17(6), 441-458.

Kibele, E. U., 2012. Regional mortality differences in Germany. Ed: Springer Science & Business Media. In: Springer, Netherlands. Doi:10.1007/978-94-007-4432-5.

Kompil, M., Aurambout, J.P., Ribeiro-Barranco, R., Barbosa, A., Jacobs-Crisioni, C., Pisoni, E., Zulian, G., Vandecasteele, I., Trombetti, M., Vizcaino, P., Vallecillo, S., Batista e Silva, F., Baranzelli, C., Rivero, I.M., Maes, J., Lavalle, C., 2015. European cities: territorial analysis of characteristics and trends - An application of the LUISA Modelling Platform. EU Reference Scenario 2013 - Updated Configuration 2014. EUR 27709 EN. Doi:10.2788/737963.

Krakover, S., Borsdorf, A., 2000. Spatial dynamics of urban expansion: The case of Innsbruck, Austria. Erde, 131(2), 125-141.

Krehl, A., Siedentop, S. 2019. Towards a typology of urban centers and subcenters–evidence from German city regions. Urban Geography, 40(1), 58-82. Doi:10.1080/02723638.2018.1500245.

Kuffer, M., Wang, J., Nagenborg, M., Pfeffer, K., Kohli, D., Sliuzas, R., Persello, C., 2018. The scope of earth-observation to improve the consistency of the SDG slum indicator. ISPRS international journal of geo-information, 7(11) 428. Doi: 10.3390/ijgi7110428.

L

Lang, S., Tiede, D., 2003. vLATE Extension für ArcGIS - vektorbasiertes Tool zur quantitativen Landschaftsstrukturanalyse. ESRI Conference 2003. Innsbruck, Austria., 8-10 October.

Lausch, A., Herzog, F., 2002. Applicability of landscape metrics for the monitoring of landscape change: issues of scale, resolution and interpretability. Ecological indicators, 2(1-2), 3-15. Doi:10.1016/S1470-160X(02)00053-5.

Lehrer, U., Wieditz, T., 2009. Condominium development and gentrification: The relationship between policies, building activities and socio-economic development in Toronto. Canadian Journal of Urban Research, 18(1) 140-161.

Leyk, S., Uhl, J.H., 2018. HISDAC-US, historical settlement data compilation for the conterminous United States over 200 years. Scientific data, 5, 180175. Doi: 10.1038/sdata.2018.175

Liang, H., Guo, Z., Wu, J., Chen, Z., 2020. GDP spatialization in Ningbo City based on NPP/VIIRS night-time light and auxiliary data using random forest regression. Advances in Space Research, 65(1), 481-493. Doi:10.1016/j.asr.2019.09.035.

Liang, X., Liu, X., Li, D., Zhao, H., Chen, G., 2018. Urban growth simulation by incorporating planning policies into a CA-based future land-use simulation model. International Journal of Geographical Information Science, 32, 2294–2316. Doi:10.1080/13658816.2018.1502441.

Liu, T., Yang, X., 2015. Monitoring land changes in an urban area using satellite imagery, GIS and landscape metrics. Applied Geography, 56, 42-54. Doi:10.1016/j.apgeog.2014.10.002.

Liu, X. P., Ma, L., Li, X., Ai, B., Li, S. Y., and He, Z. J., 2014. Simulating urban growth by integrating landscape expansion index (LEI) and cellular

automata. International Journal of Geographical Information Science, 28(1), 148–163. Doi:10.1080/13658816.2013.831097

Liu, X., Huang, Y., Xu, X., Li, X., Li, X., Ciais, Lin, P., Gong, k., Ziegler, A.D.,Chen, A., Gong, P., Chen, J., Hu, G., Chen, Y., Wang, S., Wu, Q., Huang, K., Estes, L., Zeng, Z., 2020. High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015. Nature Sustainability, 1-7. Doi:10.1038/s41893-020-0521-x.

Liu, X., Li, X., Chen, Y., Tan, Z., Li, S., Ai, B., 2010. A new landscape index for quantifying urban expansion using multi-temporal remotely sensed data. Landscape Ecology, 25, 671–682. Doi:10.1007/s10980-010-9454-5.

Liu, X., Liang, X., Li, X., Xu, X., Ou, J., Chen, Y., Li, S., Wang, S., Pei, F., 2017. A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. Landscape and Urban Planning, 168, 94-116. Doi:10.1016/j.landurbplan.2017.09.019.

Llausàs, A., Nogué, J., 2012. Indicators of landscape fragmentation: The case for combining ecological indices and the perceptive approach. Ecological Indicators, 15(1), 85-91. Doi:10.1016/j.ecolind.2011.08.016.

Lobo, J., Alberti, M.,Allen-Dumas, M., Arcaute, E., Barthelemy, M., Bojorquez-Tapia,L.A., Brail,S., Bettencourt,L.M., Beukes, A., Chen, W., Florida, R., Gonzalez, M., Grimm, N., Hamilton, M., Kempes, C., Kontokosta, C.E., Mellander, C., Neal, Z.P., Ortman, S., Pfeiffer, D., Price, M., Revi, A., Rozenblat, C., Rybski, D., Siemiatycki, M., Shutters, S.T., Smith,M.E., Stokes, E., Strumsky, D., West,G., White, D., Wu, J., Yang, V.C., York, A. Youn, H. 2020. Urban science: Integrated theory from the first cities to sustainable metropolises. Report submitted to the NSF on the Present State and Future of Urban Science. Doi:10.2139/ssrn.3526940.

Lowry, J.H., Lowry, M.B., 2014. Comparing spatial metrics that quantify urban form. Computers, Environment and Urban Systems, 44, 59-67. Doi:10.1016/j.compenvurbsys.2013.11.005.

M

MacLean, M.G., Congalton, R.G., 2013. PolyFrag : a vector-based program for computing landscape metrics. GIScience & Remote Sensing, 50, 591–603. Doi:10.1080/15481603.2013.856537.

MacLean, M.G., Congalton, R.G., 2015. A comparison of landscape fragmentation analysis programs for identifying possible invasive plant species locations in forest edge. Landscape Ecology, 30, 1241-1256. Doi:10.1007/s10980-015-0175-7.

Malaviya, S., Munsi, M., Oinam, G., Joshi, P. K., 2010. Landscape approach for quantifying land use land cover change (1972–2006) and habitat diversity in a mining area in Central India (Bokaro, Jharkhand). Environmental monitoring and assessment, 170(1-4), 215-229. Doi:10.1007/s10661-009-1227-8.

Marconcini, M., Metz-Marconcini, A., Üreyen, S., Palacios-Lopez, D., Hanke, W., Bachofer, F., Bachofer. F., Zeidler, J., Esch, T., Gorelick, N., Kakarla, A., Paganini, M., Strano, E., 2020. Outlining where humans live, the World Settlement Footprint 2015. Scientific Data, 7(1), 1-14. Doi:10.1038/s41597-020-00580-5

Marshall, S., 2005. Urban pattern specification. Institute of Community Studies, London.

Martinuzzi, S., Gould, W. A. Ramos-González, O. M., 2007. Land development, land use, and urban sprawl in Puerto Rico integrating remote sensing and population census data. Landscape and Urban Planning, 79, 288-297. Doi:10.1016/j.landurbplan.2006.02.014.

McCarty, J., Kaza, N., 2015. Urban form and air quality in the United States. Landscape and Urban Planning, 139, 168-179. Doi:10.1016/j.landurbplan.2015.03.008.

McGarigal, K., Cushman, S., Ene, E., 2012. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. http://www.umass.edu/landeco/research/fragstats/downloads/fragstats_downloads.html. (Accessed 05/04/2020).

Meentemeyer, R.K., Tang, W., Dorning, M.A., Vogler, J.B., Cunniffe, N.J., Shoemaker, D.A., 2013. FUTURES: Multilevel Simulations of Emerging Urban–Rural Landscape Structure Using a Stochastic Patch-Growing Algorithm. Annals of the American Association of Geographers, 103(4), 785–807. Doi:10.1080/00045608.2012.707591.

Meijer, J.R., Huijbregts, M.A., Schotten, K.C., Schipper, A.M., 2018. Global patterns of current and future road infrastructure. Environmental Research Letters, 13(6), 064006. Doi: 10.1088/1748-9326/aabd42. Available online: www.globio.info. (Accessed 01/11/2019).

Meneses, B.M., Reis, E., Pereira, S., Vale, M.J., Reis, R., 2017. Understanding Driving Forces and Implications Associated with the Land Use and Land Cover Changes in Portugal. Sustainability, 9, 351. Doi:10.3390/su9030351.

Montanges, A.P., Moser, G., Taubenböck, H., Wurm, M, Tuia, D., 2015. Classification of urban structural types with multisource data and structured

models. Joint Urban Remote Sensing Event. JURSE. Lausanne. Doi:10.1109/JURSE.2015.7120489.

Moulds, S., Buytaert, W., Mijic, A., 2015. An open and extensible framework for spatially explicit land use change modelling: The lulcc R package. Geoscientific Model Development, 8(10), 3215–3229. Doi:10.5194/gmd-8-3215-2015.

Mouratidis, K., 2018. Built environment and social well-being: How does urban form affect social life and personal relationships? Cities, 74, 7-20. Doi:10.1016/j.cities.2017.10.020.

Musa, S.I., Hashim, M., Reba, M.N.M., 2016. A review of geospatial-based urban growth models and modelling initiatives. Geocarto International, 32(8), 1–21. Doi:10.1080/10106049.2016.1213891.

Mveyange, A., 2015. Night lights and regional income inequality in Africa. The United Nations University World Institute for Development Economics Research. UNU-WIDER. Working Paper, No. 2015/085, Helsinki. Doi:10.35188/UNU-WIDER/2015/974-9.

N

Nabielek, K., Hamers, D., Evers, D., 2016. Cities in Europe. PBL Netherlands Environmental Assessment Agency, The Hague. Report. Available from: http://www.pbl.nl/en/publications/cities-in-europe (Accessed 15/10/2018).

National Geomatics Center of China (NGCC). 2020. GlobeLand30 dataset. Available from: www.globeland30.org (Accessed 01/07/2020).

Nazarnia, N., Schwick, C., Kopecky, M., Soukupd, T., Orlitova, E., Kienast, F., Jaeger, J.A.G., 2016. Urban Sprawl Metrics (UMS) Toolset - User manual, First edition. Url: www.wsl.ch/zersiedelung

O

OECD, 2012. Redefining Urban: A New Way to Measure Metropolitan Areas, OECD Publishing. Doi: 10.1787/9789264174108-en. (Accessed 01/05/2020).

OECD, 2017a. The Governance of Land Use in OECD Countries: Policy Analysis and Recommendations, OECD Publishing, Paris, Doi:10.1787/9789264268609-en.

OECD, 2018, Rethinking Urban Sprawl:Moving Towards Sustainable cities, OECD Publishing, Paris. Doi:10.1787/9789264189881-en

OECD, 2019a. Metropolitan areas, OECD Regional Statistics [database]. Doi: 10.1787/data-00531-en. (Accessed 22/11/2019).

OECD, 2019b. The metropolitan Database. Metadata and release notes. Available online: http://stats.oecd.org/wbos/fileview2.aspx?IDFile=4aed3009-6020-48f3-8eeb-e01a8e5f61c4. (Accessed 07/02/2020).

OECD, 2020a. Gross domestic product (GDP) (indicator). Doi: 10.1787/dc2f7aec-en. (Accessed 01/05/2020).

OECD, 2020b. Income inequality (indicator). Doi: 10.1787/459aa7f1-en. (Accessed 01/05/2020).

OECD, 2020c. Air pollution exposure (indicator). Doi: 10.1787/8d9dcc33-en. (Accessed 01/05/2020).

OECD, 2020d. Employment rate (indicator). Doi: 10.1787/1de68a9b-en. (Accessed 01/05/2020).

OECD. 2017b. How's Life? 2017: Measuring Well-being. OECD Publishing, Paris.

Oldekop, J.A., Sims, K.R., Karna, B.K., Whittingham, M.J., Agrawal, A., 2019. Reductions in deforestation and poverty from decentralized forest management in Nepal. Nature Sustainability, 2(5), 421-428. Doi:10.1038/s41893-019-0277-3.

Oliveira, V., 2016. Urban Morphology. An introduction to the Study of the Physical Form of cities. Switzerland: Springer.

Olsen, J.R., Nicholls, N., Mitchell, R., 2019. Are urban landscapes associated with reported life satisfaction and inequalities in life satisfaction at the city level? A cross-sectional study of 66 European cities. Social Science & Medicine, 226, 263-274. Doi:10.1016/j.socscimed.2019.03.009.

Olsen, L. M., Dale, V. H., Foster, T., 2007. Landscape patterns as indicators of ecological change at Fort Benning, Georgia, USA. Landscape and urban planning, 79(2), 137-149. Doi:10.1016/j.landurbplan.2006.02.007.

Otto, S.A., 2019. How to normalize the RMSE. Available online: https://www.marinedatascience.co/blog/2019/01/07/normalizing-the-rmse/. (Accessed 12/01/2020).

P

Patino, J. E., Duque, J. C., 2013. A review of regional science applications of satellite remote sensing in urban settings. Computers, Environment and Urban Systems, 37(1), 1-17. Doi:10.1016/j.compenvurbsys.2012.06.003.

Patino, J.E., Duque, J.C., Pardo-Pascual, J.E., Ruiz, L.A., 2014. Using remote sensing to assess the relationship between crime and the urban layout. Applied Geography, 55, 48-60. Doi: 10.1016/j.apgeog.2014.08.016.

Patz, J.A., Daszak, P., Tabor, G.M., Aguirre, A.A., Pearl, M., Epstein, J., Wolfe, N.D., Kilpatrick, A.M., Foufopoulos, J., Molyneux, D., Bradley, D. J., and Working Group on Land Use Change and Disease Emergence, 2004. Unhealthy landscapes: Policy recommendations on land use change and

infectious disease emergence. Environmental Health Perspectives, 112(10), 1092–1098. Doi:10.1289/ehp.6877.

Paul, S.S., Coops, N.C., Johnson, M.S., Krzic, M., Chandna, A., Smukler, S.M., 2020. Mapping soil organic carbon and clay using remote sensing to predict soil workability for enhanced climate change adaptation. Geoderma, 363, 114177. Doi:10.1016/j.geoderma.2020.114177.

Pazúr, R., Feranec, J., Štych, P., Kopecká, M., Holman, L., 2017. Changes of urbanised landscape identified and assessed by the Urban Atlas data: Case study of Prague and Bratislava. Land Use Policy, 61, 135-146. Doi:10.1016/j.landusepol.2016.11.022.

Petrasova, A., Petras, V., Van Berkel, D., Harmon, B.A., Mitasova, H., Meentemeyer, R.K., 2016. Open source approachto urban growth simulation. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B7, 953–959. Doi:10.5194/isprsarchives-XLI-B7-953-2016.

Pickard, B. R., Van Berkel, D., Petrasova, A., Meentemeyer, R. K., 2017. Forecasts of urbanization scenarios reveal trade-offs between landscape change and ecosystem services. Landscape Ecology, 32, 617–634. Doi:10.1007/s10980-016-0465-8.

Pili, S., Serra, P., Salvati, L., 2019. Landscape and the city: Agro-forest systems, land fragmentation and the ecological network in Rome, Italy. Urban Forestry & Urban Greening, 41, 230-237. Doi:10.1016/j.ufug.2019.04.016

Plexida, S.G., Sfougaris, A.I., Ispikoudis, I.P., Papanastasis, V.P., 2014. Selecting landscape metrics as indicators of spatial heterogeneity—A comparison among Greek landscapes. International Journal of Applied Earth Observation and Geoinformation, 26, 26-35. Doi:10.1016/j.jag.2013.05.001.

Poelman, L., Dijkstra, H., 2015. European cities - the EU-OECD functional urban area definition. http://ec.europa.eu/eurostat/statistics-explained/index.php/Archive:European_cities_%E2%80%93_the_EU-OECD_functional_urban_area_definition. (Accessed 15/06/2017).

Probst, P., Wright, M. N., Boulesteix, A. L., 2019. Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(3), e1301. Doi:10.1002/widm.1301.

Q

Qiu, C., Schmitt, M., Geiß, C., Chen, T.H.K., Zhu, X.X. , 2020. A framework for

large-scale mapping of human settlement extent from Sentinel-2 images via fully convolutional neural networks. ISPRS Journal of Photogrammetry and Remote Sensing, 163, 152-170. Doi:10.1016/j.isprsjprs.2020.01.028.

R

R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.R-project.org/.

Reis, J.P., Silva, E. a., Pinho, P., 2015. Spatial metrics to study urban patterns in growing and shrinking cities. Urban Geography, 3638, 1-26. Doi:10.1080/02723638.2015.1096118.

Reis, J.P., Silva, E.A., Pinho, P., 2014. Measuring space: A review of spatial metrics for urban growth and shrinkage. In: The Routledge Handbook of Planning Research Methods, 279-292. New york, Routledge.

Rempel, R.S., Kaukinen, D., Carr, A.P., 2012. Patch Analyst and Patch Grid. Ontario Ministry of Natural Resources, Centre for Northern Forest Ecosystem Research, Thunder Bay. http://www.cnfer.on.ca/SEP/patchanalyst/

Ribeiro-Barranco, R., Batista e Silva, F., Marin-Herrera, M., Lavalle, C., 2014. Integrating the MOLAND and the Urban Atlas geo-databases to analyze urban growth in European cities. Journal of Map & Geography Libraries, 10, 305-328. Doi:10.1080/15420353.2014.952485.

Rimal, B., Zhang, L., Keshtkar, H., Wang, N., Lin, Y., 2017. Monitoring and Modeling of Spatiotemporal Urban Expansion and Land-Use/Land-Cover Change Using Integrated Markov Chain Cellular Automata Model. ISPRS International Journal of Geo-Information, 6, 288. Doi:10.3390/ijgi6090288.

Rosa, I.M.D., Ahmed, S.E., Ewers, R.M., 2014. The transparency, reliability and utility of tropical rainforest land-use and land-cover change models. Global Change Biology, 20(6), 1707–1722. Doi:10.1111/gcb.12523.

S

Salat, S. 2011. Cities and Forms: On Sustainable Urbanism. Ed: Hermann.

Salvati, L., Carlucci, M., 2015. Patterns of Sprawl: The Socioeconomic and Territorial Profile of Dispersed Urban Areas in Italy, Regional Studies, 50(8), 1346-1359. Doi: 10.1080/00343404.2015.1009435.

Salvati, L., De Rosa, S., 2014. Hidden Polycentrism or Subtle Dispersion? Urban growth and long-term sub-centre dynamics in three Mediterranean cities. Land Use Policy, 39, 233–243. Doi:10.1016/j.landusepol.2014.02.012

Salvati, L., Venanzoni, G., Serra, P., and Carlucci, M., 2016. Scattered or polycentric? Untangling urban growth in three southern European metropolitan regions through exploratory spatial data analysis. The Annals of Regional Science, 57(1), 1–29. Doi:10.1007/s00168-016-0758-5.

Sandborn, A., Engstrom, R.N., 2016. Determining the Relationship Between Census Data and Spatial Features Derived From High-Resolution Imagery in Accra, Ghana. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(5), 1970-1977. Doi:10.1109/JSTARS.2016.2519843.

Sapena, M., Ruiz, L.A., 2015a. Descripción y cálculo de índices de fragmentación urbana: Herramienta IndiFrag. Revista de Teledección 43, 77-89. Doi: 10.4995/raet.2015.3476.

Sapena, M., Ruiz, L.A., 2015b. Analysis of urban development by means of multi-temporal fragmentation metrics from LULC data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-7/W3. 1411-1418. Doi:10.5194/isprsarchives-XL-7-W3-1411-2015.

Sapena, M., Ruiz, L.A., 2018. Caracterización de los patrones espaciales del crecimiento urbano aplicando índices espacio-temporales de los usos del suelo. Perspectivas multidisciplinares en la sociedad del conocimiento. XVIII Congreso de Tecnologías de la Información Geográfica, 580-590, 20-22 June, València.

Sapena, M., Ruiz, L.A., 2019. Analysis of land use/land cover spatio-temporal metrics and population dynamics for urban growth characterization. Computers, Environment and Urban Systems, 73, 27-39. Doi:10.1016/j.compenvurbsys.2018.08.001.

Sapena, M., Ruiz, L.A., Goerlich, F.J., 2016. Analysing relationships between urban land use fragmentation metrics and socio-economic variables. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B8, 1029-1036. Doi:10.5194/isprs-archives-XLI-B8-1029-2016

Sapena, M., Ruiz, L.A., Joó, K., 2017. Modelos de simulación de expansión urbana a partir de imágenes de satélite: Adecuación al análisis temporal de la fragmentación de los usos del suelo. Nuevas plataformas y sensores de teledetección. XVII Congreso de la Asociación Española de Teledetección., 311-314, 4-7 October, Murcia.

Schindler, S., Poirazidis, K., Wrbka, T., 2008. Towards a core set of landscape metrics for biodiversity assessments: a case study from Dadia National Park, Greece. Ecological indicators, 8(5), 502-514.

Doi:doi:10.1016/j.ecolind.2007.06.001.

Schneider, A., Woodcock, C.E., 2008. Compact, Dispersed, Fragmented, Extensive? A Comparison of Urban Growth in Twenty-five Global Cities using Remotely Sensed Data, Pattern Metrics and Census Information. Urban Studies, 45, 3, 659–692. Doi:10.1177/0042098007087340.

Schumm, S.A., 1956. Evolution of Drainage Systems and Slopes in Badlands at Perth Amboy, New Jersey. Geol. Soc. Am. Bull., 67, 597-646. Doi:10.1130/0016-7606(1956)67[597:EODSAS]2.0.CO;2.

Schwarz, N., 2010. Urban form revisited-Selecting indicators for characterising European cities. Landscape and Urban Planning, 96, 29–47. Doi:10.1016/j.landurbplan.2010.01.007.

SEDAC. NASA Socioeconomic Data and Applications Center. U.S. Census Grids, 2020, https://sedac.ciesin.columbia.edu/. (Accessed 15/04/2020).

Seto, K.C., Golden, J.S., Alberti, M., Turner, B.L., 2017. Sustainability in an urbanizing planet. Proceedings of the National Academy of Sciences, 114(34), 8935-8938. Doi:10.1073/pnas.1606037114.

Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality. complete samples. Biometrika, 52(3-4) 591-611.

Sharaf, S.A., Serra, P., Saurí, D., 2018. A district and sector land-use and landscape analysis of urban sprawl in Al Ain municipality (United Arab Emirates): Just a quick conversion from sand to a built-up environment?. Applied Geography, 95, 88-100. Doi:10.1016/j.apgeog.2018.04.008.

Shi, L., Taubenböck, H., Zhang, Z., Liu, F., Wurm, M., 2017. Urbanization in China from the end of 1980s until 2010 – spatial dynamics and patterns of growth using EO-data. International Journal of Digital Earth, 12(1), 78–94. Doi:10.1080/17538947.2017.1400599.

Shi, Y., Sun, X., Zhu, X., Li, Y., Mei, L., 2012. Characterizing growth types and analyzing growth density distribution in response to urban growth patterns in peri-urban areas of Lianyungang City. Landscape and Urban Planning, 105(4), 425–433. Doi:10.1016/j.landurbplan.2012.01.017.

Siedentop, S., Fina, S., 2010. Monitoring urban sprawl in Germany: towards a GIS-based measurement and assessment approach. Journal of Land Use Science, 5(2), 73–104. Doi:10.1080/1747423X.2010.481075.

Šímová, P., Gdulová, K., 2012. Landscape indices behavior: A review of scale effects. Applied Geography, 34, 385-394. Doi:10.1016/j.apgeog.2012.01.003.

Sims, K.R., Thompson, J.R., Meyer, S.R., Nolte, C., Plisinski, J. S., 2019. Assessing the local economic impacts of land protection. Conservation Biology, 33(5), 1035-1044. Doi:10.1111/cobi.13318.

Soares-Filho, B.S., Coutinho Cerqueira, G., Lopes Pennachin, C., 2002. DINAMICA - A stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. Ecological Modelling, 154(3), 217–235. Doi:10.1016/S0304-3800(02)00059-5.

Sohl, T.L., Sayler, K.L., Drummond, M.A., Loveland, T.R., 2007. The FORE-SCE model: a practical approach for projecting land cover change using scenario-based modeling. Journal of Land Use Science, 2(2), 103–126. Doi:17474230701218202.

Song, Y., Shao, G., Song, X., Liu, Y., Pan, L., Ye, H., 2017. The relationships between urban form and urban commuting: An empirical study in China. Sustainability, 9(7), 1150. Doi:10.3390/su9071150.

Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. American Meteorological Society, 93, 1879-1900. Doi:10.1175/BAMS-D-11-00019.1.

Stiller, D., Stark, T., Wurm, M., Dech, S., Taubenböck, H., 2019. Large-scale building extraction in very high-resolution aerial imagery using Mask R-CNN. In 2019 Joint Urban Remote Sensing Event. JURSE. 1-4. IEEE. Doi:10.1109/JURSE.2019.8808977.

Stokes, E.C., Seto, K.C., 2019. Characterizing and measuring urban landscapes for sustainability. Environ. Res. Lett., 14(4),045002. Doi:10.1088/1748-9326/aafab8.

Sun, C., Wu, Z.F., Lv, Z.Q., Yao, N., Wei, J.B., 2013. Quantifying different types of urban growth and the change dynamic in Guangzhou using multi-temporal remote sensing data. International Journal of Applied Earth Observation and Geoinformation, 21, 409-417. Doi:10.1016/j.jag.2011.12.012.

Sun, X., Crittenden, J. C., Li, F., Lu, Z., Dou, X., 2018. Urban expansion simulation and the spatio-temporal changes of ecosystem services, a case study in Atlanta Metropolitan area, USA. Science of the Total Environment, 622, 974–987. Doi:10.1016/j.scitotenv.2017.12.062.

T

Tapiador, F.J., Avelar, S., Tavares-Corrêa, C., Zah, R., 2011. Deriving fine-scale socioeconomic information of urban areas using very high-resolution satellite imagery. International Journal of Remote Sensing. 32, 6437-6456.

Doi:10.1080/01431161.2010.512928.

Taubenböck, H., 2019. Remote Sensing for the Analysis of Global Urbanization. DLR-Research Report 2019-10. Habilitation Thesis, University Würzburg. 10, 600.

Taubenböck, H., Debray, H., Qiu, C., Schmitt, M., Wang, Y., Zhu, X.X. 2020. Seven city types representing morphologic configurations of cities across the globe. Cities, 105, 102814. Doi:10.1016/j.cities.2020.102814

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., Dech, S. 2012. Monitoring of mega cities from space. Remote Sensing of Environment, 117, 162-176. Doi:10.1016/j.rse.2011.09.015.

Taubenböck, H., Ferstl, J., Dech, S.,2017. Regions set in stone—Delimiting and categorizing regions in Europe by settlement patterns derived from EO-data. ISPRS International Journal of Geo-Information, 6(2) 55. Doi:10.3390/ijgi6020055.

Taubenböck, H., Wiesner, M., Felbier, A., Marconcini, M., Esch, T., Dech, S., 2014. New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. Applied Geography, 47, 137–153. Doi:10.1016/j.apgeog.2013.12.002.

Taubenböck, H., Wurm, M., Setiadi, N., Gebert, N., Roth, A., Strunz, G., Birkmann, J., Dech, S., 2009. Integrating remote sensing and social science. Joint Urban Remote Sensing Event. Doi:10.1109/URS.2009.5137506.

Terando, A. J., Costanza, J., Belyea, C., Dunn, R. R., McKerrow, A., Collazo, J. A., 2014. The southern megalopolis: Using the past to predict the future of urban sprawl in the Southeast U.S. PLoS One, 9(7). Doi:10.1371/journal.pone.0102261.

Tian, G., Jiang, J., Yang, Z., Zhang, Y., 2011. The urban growth, size distribution and spatio-temporal dynamic pattern of the Yangtze River Delta megalopolitan region, China. Ecological Modelling, 222(3), 865–878. Doi:10.1016/j.ecolmodel.2010.09.036.

Tian, G., Ma, B., Xu, X., Liu, X., Xu, L., Liu, X., Xiao, L., Kong, L., 2016. Simulation of urban expansion and encroachment using cellular automata and multi-agent system model—A case study of Tianjin metropolitan region, China. Ecological Indicators, 70, 439–450. Doi:10.1016/j.ecolind.2016.06.021.

Tong, X., Feng, Y., 2019. A review of assessment methods for cellular automata models of land-use change and urban growth. International Journal of Geographical Information Science, 1–33. Doi:10.1080/13658816.2019.1684499.

Tonkiss, F., 2013. Cities by design: the social life of urban form. Ed: Polity Press, Cambridge, UK.

Torrens, P.M., 2009. Cellular Automata. In: Kitchin, Roband Thrift, N. (Eds.) International Encyclopedia of Human Geography. London, 1-4.

Travisi, C.M., Camagni, R., Nijkamp, P. 2010. Impacts of urban sprawl and commuting: a modelling study for Italy. Journal of Transport Geography, 18(3) 382-392. Doi:10.1016/j.jtrangeo.2009.08.008.

Trinder, J., Liu,Q. 2020. Assessing environmental impacts of urban growth using remote sensing, Geo-spatial Information Science, Doi:10.1080/10095020.2019.1710438

Tuia, D., Moser, G., Wurm, M., Taubenböck, H., 2017. Land use modeling in North Rhine-Westphalia with interaction and scaling laws. Joint Urban Remote Sensing Event, Dubai, United Arab Emirates. Doi:10.1109/JURSE.2017.7924542.

U

Uhl, J.H., Zoraghein, H., Leyk, S., Balk, D., Corbane, C., Syrris, V., Florczyk, A.J. , 2020. Exposing the urban continuum: implications and cross-comparison from an interdisciplinary perspective. International Journal of Digital Earth, 13(1), 22-44, Doi:10.1080/17538947.2018.1550120

UN-Habitat, 2013. Urban Planning for City Leaders. ISBN 978-92-1-132505-8. In: Nairobi. https://unhabitat.org/91566. (Accessed 05/02/2018)

UN-Habitat, 2015. The Economics of Urban Form: A Literature Review. In: Nairobi.

UN-Habitat, 2016. World Cities Report 2016 - Urbanization and Development: Emerging Futures.

United Nations (UN), 2015. General Assembly, Transforming our world : the 2030 Agenda for Sustainable Development, 21 October 2015, A/RES/70/1, available at: https://www.refworld.org/docid/57b6e3e44.html. (Accessed 23/04/2020).

United Nations (UN), 2020. World Social Report 2020: Inequality in a Rapidly Changing World, UN, New York, Doi:10.18356/7f5d0efc-en.

United Nations (UN), Department of Economic and Social Affairs, Population Division, 2019. World Urbanization Prospects: The 2018 Revision. ST/ESA/SER.A/420. New York: United Nations

United Nations (UN), Department of Economic and Social Affairs, Population Division. 2018. World Urbanization Prospects: The 2018 Revision, Key Facts.

Ustaoglu, E., Williams, B., Petrov, L. O., Shahumyan, H., Van Delden, H., 2018. Developing and assessing alternative land-use scenarios from the MOLAND Model: A scenario-based impact analysis approach for the evaluation of rapid rail provisions and urban development in the Greater Dublin Region. Sustainability, 10(1), 61. Doi:10.3390/su10010061.

Uuemaa, E., Antrop, M., Roosaare, J., Marja, R., Mander, Ü., 2009. Landscape Metrics and Indices: An Overview of Their Use in Landscape Research. Living Reviews in Landscape Research, 3, 1–28. Doi:10.12942/lrlr-2009-1.

Uuemaa, E., Mander, Ü., Marja, R., 2013. Trends in the use of landscape spatial metrics as landscape indicators: a review. Ecological Indicators, 28, 100-106. Doi:10.1016/j.ecolind.2012.07.018.

V

Van de Voorde, T., Jacquet, W., Canters, F., 2011. Mapping form and function in urban areas: An approach based on urban metrics and continuous impervious surface data. Landscape and Urban Planning, 102(3), 143-155. Doi:10.1016/j.landurbplan.2011.03.017

Van de Voorde, T., Van der Kwast, J., Poelmans, L., Canters, F., Binard, M., Cornet, Y., Engelen, G., Uljee, I., Shahumyan, H., Williams, B., Convery, S. Lavalle, C., 2016. Projecting alternative urban growth patterns: The development and application of a remote sensing assisted calibration framework for the Greater Dublin Area. Ecological Indicators 60:1056–1069. DOI:10.1016/j.ecolind.2015.08.035.

Venerandi, A., Quattrone, G., Capra, L. 2018. A scalable method to quantify the relationship between urban form and socio-economic indexes. EPJ Data Science. Doi:10.1140/epjds/s13688-018-0132-1.

Verburg, P.H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., Mastura, S.S.A., 2002. Modeling the spatial dynamics of regional land use: The CLUE-S model. Environmental Management, 30(3), 391–405. Doi:10.1007/s00267-002-2630-x.

Villeneuve, P.J., Jerrett, M., Su, J.G., Burnett, R.T., Chen, H., Wheeler, A.J., Goldberg, M.S., 2012. A cohort study relating urban green space with mortality in Ontario, Canada. Environmental Research, 115, 51–58. Doi:10.1016/j.envres.2012.03.003.

Vogt, P., Riitters, K., 2017. GuidosToolbox: Universal digital image object analysis. European Journal of Remote Sensing, 50(1), 352-361. Doi:10.1080/22797254.2017.1330650

W

Wang, P., Huang, C., Brown de Colstoun, E.C., Tilton, J.C., Tan, B., 2017. Global Human Built-up And Settlement Extent (HBASE) Dataset From Landsat. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). Doi:10.7927/H4DN434S. (Accessed 05/02/2020).

Watmough, G.R., Atkinson, P.M., Saikia, A., Hutton, C.W., 2016. Understanding the evidence base for poverty–environment relationships using remotely sensed satellite data: An example from Assam, India. World Development, 78, 188-203. Doi:10.1016/j.worlddev.2015.10.031.

Wegmann, M., Leutner, B.F., Metz, M., Neteler, M., Dech, S., Rocchini, D., 2018. r. pi: A grass gis package for semi-automatic spatial pattern analysis of remotely sensed land cover data. Methods in Ecology and Evolution, 9(1), 191-199. Doi:10.1111/2041-210X.12827.

Wei, Y.D., Ewing, R., 2018. Urban expansion, sprawl and inequality. Landscape and Urban Planning, 177, 259-265. Doi:10.1016/j.landurbplan.2018.05.021.

Weigand, M., Staab, J., Wurm, M., Taubenböck, H., 2020. Spatial and semantic effects of LUCAS samples on fully automated land use/land cover classification in high-resolution Sentinel-2 data. International Journal of Applied Earth Observation and Geoinformation, 88, 102065. Doi:10.1016/j.jag.2020.102065.

Weigand, M., Wurm, M., Dech, S., Taubenböck, H., 2019. Remote sensing in environmental justice research—a review. ISPRS International Journal of Geo-Information, 8(1), 20. Doi:10.3390/ijgi8010020.

Weilenmann, B., Seidl, I., Schulz, T., 2017. The socio-economic determinants of urban sprawl between 1980 and 2010 in Switzerland. Landscape and Urban Planning, 157, 468-482. Doi:10.1016/j.landurbplan.2016.08.002.

Wentz, E.A., Anderson, S., Fragkias, M., Netzband, M., Mesev, V., Myint, S.W., Quattrochi, D., Rahman, A., Seto, K.C., 2014. Supporting Global Environmental Change Research: A Review of Trends and Knowledge Gaps in Urban Remote Sensing. Remote Sensing, 6, 3879-3905. Doi:10.3390/rs6053879.

Wentz, E.A., York, A.M., Alberti, M., Conrow, L., Fischer, H., Inostroza, L, Jantz, C., Pickett, S.T.A, Seto, K.C., Taubenböck, H., 2018. Six fundamental aspects for conceptualizing multidimensional urban form: A spatial mapping perspective. Landscape and Urban Planning, 179, 55-62. Doi:10.1016/j.landurbplan.2018.07.007.

Whitmee, S., Haines, A., Beyrer, C., Boltz, F., Capon, A.G., de Souza Dias, B.F., et al., 2015. Safeguarding human health in the Anthropocene epoch: report of The Rockefeller Foundation–Lancet Commission on planetary health. The Lancet, 386(10007), 1973-2028. Doi: 10.1016/S0140-6736(15)60901-1.

Wilkinson, D.A., Marshall, J.C., French, N.P., Hayman, D.T.S., 2018. Habitat fragmentation, biodiversity loss and the risk of novel infectious disease emergence. Journal of the Royal Society Interface, 15, 20180403. Doi:10.1098/rsif.2018.0403.

Williams, K., 2014. Urban form and infrastructure: a morphological review. London: Government Office for Science. Report No.:URN GS/14/808.

Williams, K., Burton, E., and Jenks, M., 2000. Achieving Sustainable Urban Form. London and New York: Spon Press.

Wilson, E.H., Hurd, J.D., Civco, D.L., Prisloe, M.P., Arnold, C., 2003. Development of a geospatial model to quantify, describe and map urban growth. Remote Sensing of Environment, 86, 275–285. Doi:10.1016/S0034-4257(03)00074-9.

Wissen Hayek, U., Efthymiou, D., Farooq, B., von Wirth, T., Teich, M., Neuenschwander, N., Grêt-Regamey, A., 2015. Quality of urban patterns: Spatially explicit evidence for multiple scales. Landscape and Urban Planning, 142, 47–62. Doi:10.1016/j.landurbplan.2015.05.010.

Wu, Y., Li, S., Yu, S., 2016. Monitoring urban expansion and its effects on land use and land cover changes in Guangzhou city, China. Environmental Monitoring and Assessment, 188, 54. Doi:10.1007/s10661-015-5069-2.

Wurm, M., d'Angelo, P., Reinartz, P. Taubenböck, H. 2014. Investigating the Applicability of Cartosat-1 DEMs and Topographic Maps to Localize Large-Area Urban Mass Concentrations. in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(10), 4138-4152. Doi:10.1109/JSTARS.2014.2346655.

Wurm, M., Stark, T., Zhu, X. X., Weigand, M., Taubenböck, H., 2019a. Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. ISPRS journal of photogrammetry and remote sensing, 150, 59-69. Doi:10.1016/j.isprsjprs.2019.02.006

Wurm, M., Taubenböck, H., 2018. Detecting social groups from space – Assessment of remote sensing-based mapped morphological slums using income data, Remote Sensing Letters, 9(1), 41-50. Doi:10.1080/2150704X.2017.1384586.

Wurm, M., Weigand, M., Stark, T., Goebel, J., Wagner, G. G., Taubenböck, H.

2019b. Modelling the impact of the urban spatial structure on the choice of residential location using 'big earth data' and machine learning. Joint Urban Remote Sensing Event, Vannes, France. Doi:10.1109/JURSE.2019.8808942.

Y

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S. M., Case, A., Costello, c., Dewitz, J., Fry, J., Funk, M., Granneman, B., Liknes, G.C., Rigge, M., Xian, G., 2018. A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. ISPRS Journal of Photogrammetry and Remote Sensing, 146, 108-123. Doi:10.1016/j.isprsjprs.2018.09.006.

Yu, M., Huang, Y., Cheng, X., Tian, J., 2019. An ArcMap plug-in for calculating landscape metrics of vector data. Ecological informatics, 50, 207-219. Doi:/10.1016/j.ecoinf.2019.02.004

Z

Zaragozí, B., Belda, A., Linares, J., Martínez-Pérez, J. E., Navarro, J. T., Esparza, J., 2012. A free and open source programming library for landscape metrics calculations. Environmental Modelling & Software, 31, 131-140. Doi:10.1016/j.envsoft.2011.10.009.

Zhao, M., Cai, H., Qiao, Z., Xu, X., 2016. Influence of urban expansion on the urban heat island effect in Shanghai. International Journal of Geographical Information Science, 30(12), 2421–2441. Doi:10.1080/13658816.2016.1178389.

Zhou, Y., Li, X., Asrar, G. R., Smith, S. J., Imhoff, M., 2018. A global record of annual urban dynamics (1992–2013) from nighttime lights. Remote Sensing of Environment, 219, 206-220. Doi:10.1016/j.rse.2018.10.015.

Zhu, Z., Zhou, Y., Seto, K.C., Stokes, E.C., Deng, C., Pickett, S.T. A., Taubenböck, H., 2019. Understanding an urbanizing planet: Strategic directions for remote sensing. Remote Sensing of Environment, 228, 164–182. Doi:10.1016/j.rse.2019.04.020.

Zohdy, S., Schwartz, T.S., Oaks, J.R., 2019. The coevolution effect as a driver of spillover. Trends Parasitology, 35(6), 399-408. Doi:10.1016/j.pt.2019.03.010.

Appendix A

The IndiFrag tool: spatio-temporal metrics

This appendix describes the software tool IndiFrag that has been developed as part of this thesis. IndiFrag collects spatio-temporal metrics from the literature and automatizes their computation in urban environments based on categorical maps (e.g., LULC maps). These metrics can be used for several applications, such as quantifying urban form, spatio-temporal patterns, urban growth, LULC change, and landscape ecology. The tool has been created to cover the main shortcomings from other existing tools in urban applications, which were presented in the introduction and summarized in Table 1.2. As a result of the first objective of this thesis, the tool has been used in all the analyses conducted, from chapter 3 to 6, and has been made freely available to the scientific community to extend its usage. The tool was presented and described in detail in Sapena and Ruiz (2015).

The tool is registered in the Patents and Software Catalog of the Universitat Politècnica de València since 2016, under the reference code R-18020-2016, and it can be downloaded from http://cgat.webs.upv.es/software/ with an open license for non-commercial purposes. The downloaded file is a zip with: (1) the ArcGIS ToolBox; (2) the extensive user guide in Spanish; (3) a short version in English; (4) tutorials; (5) the corresponding data necessary to perform the tutorials, and (6) a table describing the metrics. The tool was designed to be easily applied and as simple as possible, including a complete *Help* from ArcGIS describing all the options and metrics available in the ToolBox.

The tool is composed of three modules. The first module consists of the spatial analysis, where spatial metrics are computed for one-date categorical map; then, a sub-module calculates the differences between the spatial metrics computed independently from two categorical maps. The second module conducts the multi-temporal analysis, where metrics are directly extracted from two categorical maps from different dates. The third module creates geometric super-objects for analyses in several regions. There are four options: grids, hexagons, pies, and rings, based on Adamczyk and Tiede (2017). The appearance of the Toolbox and the spatial analysis window are shown in Figure A.1.
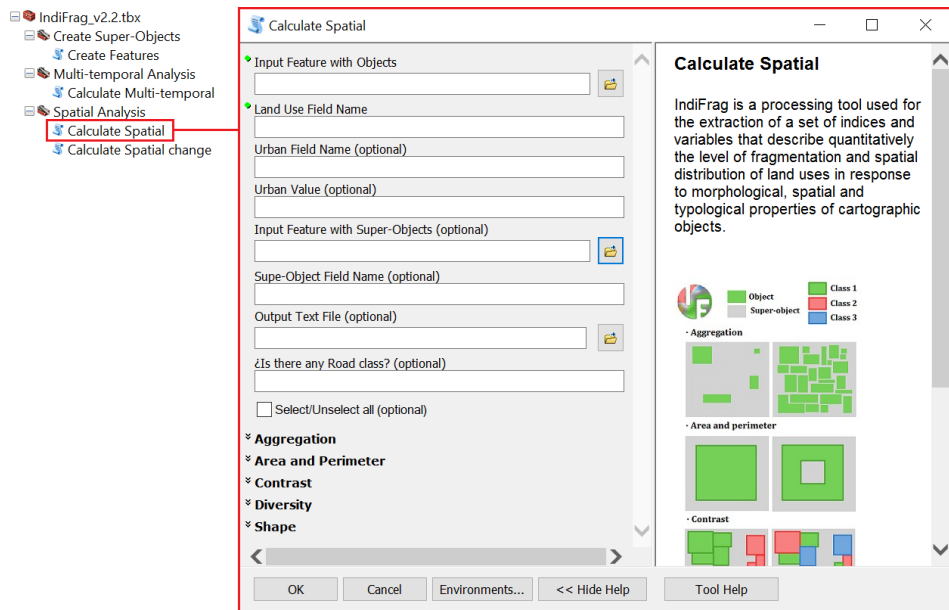

Figure A.1. Example of IndiFrag ToolBox modules and the window of the spatial analysis.

Regarding the development of the tool, it encompasses all the functionalities necessary to suit the analyses of this thesis, but also aiming to be general and useful for additional applications beyond this thesis. Since its first publication, there were several updates, including new metrics and improving geo-processing functions. The main characteristics of IndiFrag are listed as follows:

> It works with any categorical map in vector format, which allows for using topological relations between cartographic objects. For example, it considers the spatial properties from two adjacent objects with the same class, which is not possible using raster data.

> Metrics are computed at three hierarchy levels: (i) the super-object level describes spatio-temporal relationships between objects from various classes within the individual boundaries of multiple regions. (e.g.: territorial units, grids, hexagons, etc.); (ii) the class level describes spatio-temporal relationships between a set of objects from the same class (e.g.: land use, land cover, habitat, etc.); and (iii) the object level describes spatio-temporal characteristics of each cartographic object, despite their class or super-object (e.g.: cadastral plot, building, etc.). Figure A.2 illustrates an example of the three levels.

> When a shapefile with super-objects is available, the analysis is conducted for several regions in the study area (Figure A.2, B). If not available, the study area is considered a super-object.

> The results are quantitative values stored in the input shapefiles and in a text file. Object and class level metrics are stored in the shapefile with the cartographic objects and class field. Super-object levels are stored in the super-object shapefile.

> IndiFrag includes more than fifty spatio-temporal metrics. First, spatial metrics are calculated for one-date categorical map, and metrics are divided into five semantic groups according to the spatial attributes they measure. The groups are: Area and perimeter, shape, aggregation, diversity, contrast, based on a previous classification from McGarigal et al (2012). Second, multi-temporal metrics are computed based on two categorical maps from different dates.

> A sub-module computes spatio-temporal metrics as the differences of spatial metrics for each class and super-object. The differences are stored in the attribute table of the super-objects, easing the mapping of spatial changes within a GIS software. If the sub-module is applied for one-date categorical map, the class level metrics are transferred to the super-object shapefile, which also facilitates the mapping of results.

> Specifically for large-scale categorical maps, when a class represents the road network, there is an option to remove the influence of roads
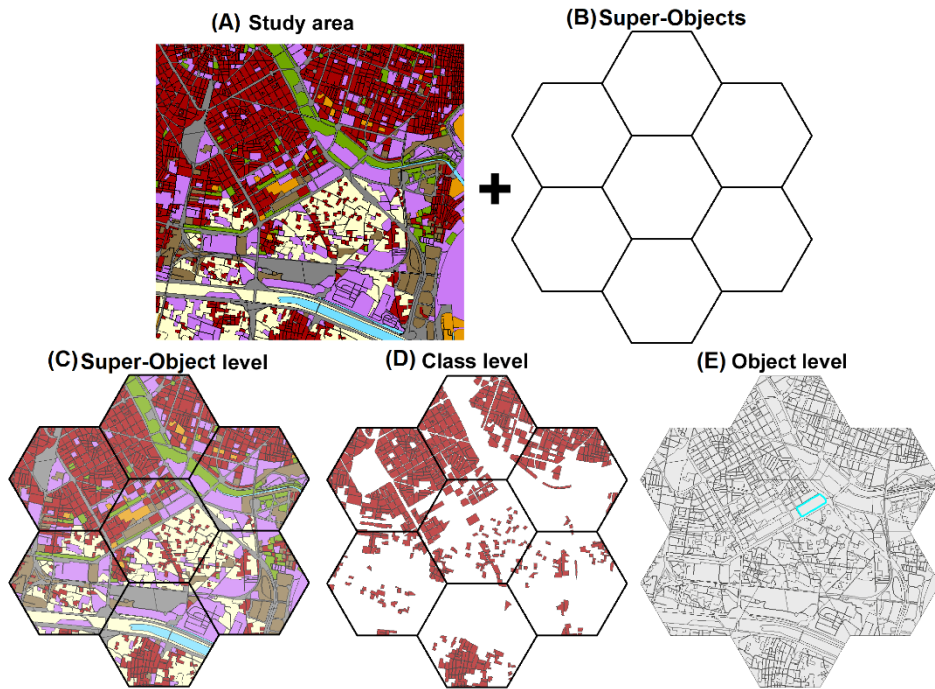
Figure A.2. Example of the input data in IndiFrag: (A) a LULC map in vector format for the study area; and optionally, (B) a shapefile with super-objects. Three hierarchy levels used for computing spatio-temporal metrics: (C) super-object level, where metrics are computed for each super-object independently, (D) class level, where metrics are computed for each class and super-object independently, and (E) object level, where metrics are calculated for each cartographic object (e.g., in light blue) without considering the super-object or their class.

> in the contrast analyses. An example is shown in the contrast section (A.1.5).
> IndiFrag creates graphical outputs for some metrics to complement and enrich the interpretation of the results by means of maps and graphs.
> The multi-temporal module includes a LULC change analysis, creating LULC change matrices, mapping conversions between classes and identifying changed objects.
> A module creates geometric super-objects that can be used instead of territorial or administrative units. The options are: hexagons, grids, pie or rings.

It should be noted that not all of the functionalities and metrics described in this appendix are included in IndiFrag v.2.1, the version available at http://cgat.webs.upv.es/software/. All functionalities are included in a more

recent version, IndiFrag v.3.1, which is not yet available on the website, but will be published soon.

The following sections describe in detail the metrics included in IndiFrag. First, spatial metrics (mono-temporal) are organized by semantic groups and described. Second, multi-temporal metrics are described. Finally, Table A.1 summarizes all the metrics described.

A.1. Spatial metrics

Spatial metrics are mono-temporal indicators that convert spatial patterns into quantitative terms; quantify the spatial structure, configuration, and heterogeneity of landscapes by means of categorical maps at specific scales (Herold et al., 2005; Llausàs and Nogué, 2012).

A.1.1. Area and perimeter

This group collects metrics related to the size and perimeter of the objects, as well as their density. Area and perimeter are among the simplest spatial attributes, but they are very useful when assessing and comparing super-objects.

A.1.1.1. Area (Area)

It represents the extent of a surface. It is computed at three levels: for every single object (m^2), for all the objects within a class (km^2) and for the super-object itself (km^2). This is a basic metric used as input in several spatio-temporal metrics.

A.1.1.2. Perimeter (Perim)

It represents the line that shapes the boundary of an object. It is computed at three levels: for every single object (m), for all the objects within a class (km) and for the entire super-object itself (km). It does not consider boundary duplicity, when two objects are adjacent only the common edge is considered. This is a basic metric used as input in several spatio-temporal metrics.

A.1.1.3. Total perimeter (PerimT)

It considers the perimeter of the super-object plus the sum of the perimeter of the objects, without duplication of boundaries (McGarigal et al., 2012). This metric is computed at the super-object level (km^2).

A.1.1.4. Urban density (DU)

Ratio between the total urban area (combining urban classes) and the super-object area (%); it measures the percentage of urbanization (Romano et al.,

2010; Colaninno et al., 2011b). It is computed at the super-object level.

$$DU=\frac{A_u}{A_T} \tag{A.1}$$

A_U= total area of urban objects (m²)

A_T= total area of the super-object (m²)

A.1.1.5. Class density (DC)

Ratio between a class area and the super-object area (%), it quantifies the percentage of the super-object occupied by each class. It is computed at the class level.

$$DC=\frac{\sum_{i=1}^{n}(A_i)}{A_T} \tag{A.2}$$

A_i = area of the object i from a class (m²)

n = number of objects in the class.

A.1.1.6. Road density (D_{road})

Ratio between the total length of roads in meters and the super-object area in km² (m/km²). It is computed at the super-object level.

$$D_{road}=\frac{\sum_{i=1}^{nr}(L_i)}{A_T} \tag{A.3}$$

L_i = length of the road segment i (m)

nr = number of roads segments in the super-object.

A.1.1.7. Mean object size (MS)

It equals the average of the size of the objects from a class or super-object (m²) (Irwin and Bockstael, 2007). It is computed at the class and the super-object levels.

$$MS=\frac{\sum_{i=1}^{n}(A_i)}{n} \tag{A.4}$$

A_i = area of the object i from a class/super-object (m²)

n = number of objects in the class/super-object.

A.1.1.8. Edge density (ED)

It equals the sum of the perimeter of the objects from a class or super-object (without duplication of boundaries) divided by the area of the super-object (m/m²) (Herold et al., 2002; McGarigal et al., 2012). It is computed at the class and super-object levels.

$$ED = \frac{\sum_{i=1}^{n}(P_i)}{A_T}$$ (A.5)

P_i = perimeter of the object i from a class/super-object (m)

A.1.1.9. Area-weigthed mean urban fragmentation index (AWUFI)

It evaluates the fragmentation due to the presence of urban areas in natural landscapes. It is weighted by the area of super-objects in the study area in order to perform comparisons between the fragmentation values (Astiaso et al., 2013; Romano 2002; Romano and Tamburini, 2006). It is computed at the super-object level. When there is only one super-object the weighted value will be the same as the non-weighted one. The obstruction coefficient has to be stablished for each class. Some examples are:

- High obstruction coefficient (O_c=1): industrial, commercial, road, rail networks, airports; construction sites classes.
- Medium-high obstruction coefficient (O_c=0.8): continuous urban fabric classes, high-dense residential classes.
- Medium obstruction coefficient (O_c=0.6): discontinuous urban fabric, low-dense residential classes.
- Low obstruction coefficient (O_c=0.4): green urban areas, sport and leisure facilities classes.

$$UFI_i = \frac{\sum_{i=1}^{m}(L_{max} \cdot \sum_{i=1}^{n}(A_i) \cdot O_c)}{A_{Ti}}$$ (A.6)

$$AWUFI_i = \frac{\sum_{i=1}^{ns}(UFI_i \cdot A_{Ti})}{\sum_{i=1}^{ns}(A_{Ti})}$$ (A.7)

A_{Ti} = area of the super-object i in the study area (m²).

L_{max} = maximum length of the object from the class analyzed in the super-object i (m).

O_c = obstruction coefficient of the class analyzed.

m = number of classes in the super-object i.

ns = number of super-object in the study area.

A.1.1.10. Porosity (P)

It measures the total ratio of open spaces in a class compared to the class area (%). It is the percentage of holes in a class and measures discontinuity (Reis et al., 2015). It is computed at the class level.

$$P = \frac{\sum_{h=1}^{nh}(A_h)}{\sum_{i=1}^{n}(A_i)} \cdot 100 \qquad (A.8)$$

A_h = area of the hole within a class (m^2).

nh = number of holes in the class.

A.1.1.11. Urban porosity (PU)

It has the same formula as (A.8) for urban classes. It is computed at the super-object level (%).

A.1.1.12. Largest object index (LOI)

It represents the percentage of a class that is represented by the largest object. It ranges from 0 to 100 (%) (McGarigal et al., 2012). It is computed at the class level. This metric can also be computed for the largest urban core (L_{UC}). The urban core is defined based on the urban morphological zone definition of the EEA (2014) as a set of urban elements within a distance of 200 m.

$$LOI = \frac{A_{max}}{\sum_{i=1}^{n}(A_i)} \cdot 100 \qquad (A.9)$$

A_{max} = area of the largest object in a class or the largest urban core (m^2).

A.1.1.13. Second largest object index (SLOI)

It represents the percentage of the class that is represented by the second largest object following the same equation as (A.9) (%). It is computed at the class level. It can be also computed for the second largest urban core (SL_{UC}).

$$SLOI = \frac{A_{2nd\ max}}{\sum_{i=1}^{n}(A_i)} \cdot 100 \qquad (A.10)$$

$A_{2nd\ max}$ = area of the second largest object in a class or the second largest urban core (m^2).

A.1.2. Shape

This group represents a collection of metrics that quantify the shape of objects. The interaction between the shape of an object and its size has great influence

in the internal processes of the class and the super-object. The shape is an attribute that provides information about the morphology of the objects.

A.1.2.1. Boundary dimension (Dim_B)

It represents the relationship between the object area and the perimeter and measures complexity and homogeneity of the shape of the objects at the class level (Wu et al., 2013). It computes a minimum quadratic adjustment of the distribution of the objects, using the area and perimeter as variables, and being C a constant (A.11). Dim_B gives the boundary dimension by means of the slope of the linear function (Figure A.3). The value varies between one and two, the larger the number the more complex and heterogeneous the shapes of the objects within a class. This metric, generates graphical outputs with one plot per class with the linear function, where the slope of the line (in red) is the value of Dim_B.

$$\ln(A_i) = \frac{2}{Dim_B} \cdot \ln(P_i) + \ln(C) \tag{A.11}$$
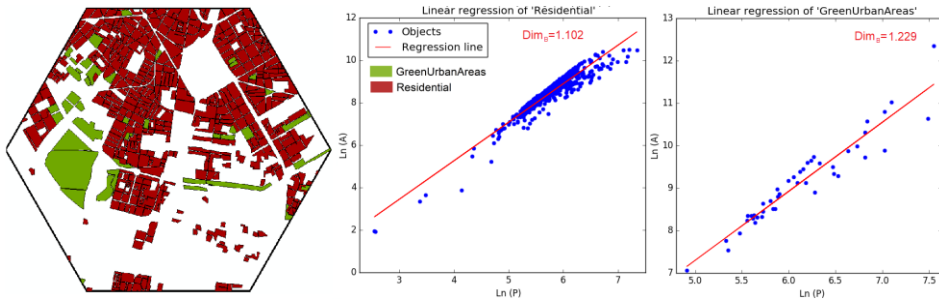

Figure A.3. Example of the distribution of each object (blue points) from a class based on the logarithm of the area (Y-axis) and perimeter (X-axis). It shows the Dim_B for the Residential and Green urban areas classes in a super-object (left). The plots on the right are graphical outputs from IndiFrag when Dim_B is computed.

A.1.2.2. Fractal dimension (FD)

Normalized shape index based on perimeter-area relationships, where the perimeter and area are transformed with logarithm. It can be computed at the three levels. At class or super-object levels, it consists of the sum of all objects within a class or a super-object (Frenkel and Ashkenazi, 2008; Gong et al., 2013; Herold et al., 2002; McGarigal et al., 2012).

$$FD_O = \frac{2 \cdot \ln(0.25 \cdot P_i)}{\ln(A_i)}; \quad FD_{C,SO} = \sum_{i=1}^{n} \frac{2 \cdot \ln(0.25 \cdot P_i)}{\ln(A_i)} \tag{A.12}$$

where object level (O), class level (C), and super-object level (SO).

A.1.2.3. Area-weighted mean fractal dimension (AWFD)

It equals the area-weighted-mean fractal dimension for the objects in a class. It is computed at the class level (Gong et al., 2013; Herold et al., 2002; McGarigal et al., 2012).

$$AWFD= \sum_{i=1}^{n} \left[\left(\frac{2 \cdot \ln(0.25 \cdot P_i)}{\ln(A_i)} \right) \cdot \left(\frac{A_i}{\sum_{i=1}^{n} A_i} \right) \right] \qquad (A.13)$$

A.1.2.4. Shape index (SI)

Perimeter-area normalized ratio of the shape of objects. The shape is compared to a square with the same size (Frenkel and Ashkenazi, 2008; Jiang et al., 2007; McGarigal et al., 2012). It can be computed at the three levels.

$$SI_O= \frac{0.25 \cdot P_i}{\sqrt{A_i}}; \; SI_{C,SO}= \sum_{i=1}^{n} \frac{0.25 \cdot P_i}{\sqrt{A_i}} \qquad (A.14)$$

A.1.2.5. Perimeter-area mean ratio (PAR)

It is a perimeter-area ratio of objects. It measures how the perimeter of an object increases per unit area (Irwin and Bockstael, 2007; McGarigal et al., 2012). It can be computed at the three levels.

$$PAR_O= \frac{P_i}{A_i}; \; PAR_{C,SO}= \frac{\sum_{i=1}^{n} \left(\frac{P_i}{A_i} \right)}{n} \qquad (A.15)$$

A.1.2.6. Elongation ratio (ER)

It quantifies the elongation of the largest object in a class. This metric is commonly used in hydrology (Schumm, 1956), it measures the elongation of the largest object by dividing the diameter of the circumference with the same area as the object by the largest side of the object. It ranges from 0-1. Closer values to zero show elongated shapes. It is computed at the class level.

$$ER= \frac{2 \cdot \sqrt{A_{max}/\pi}}{L_{max}} \qquad (A.16)$$

L_{max}= length of the larges side from the boundary box covering the largest object from a class (m).

A.1.3. Aggregation

This group represents a collection of metrics that show the tendency of objects to aggregate spatially, showing the spatial structure of a region. The term aggregation encompasses other concepts such as dispersion, subdivision and fragmentation.

A.1.3.1. Number of objects (NOb)

Number of objects in a class or super-object. If available, road class objects are not counted.

A.1.3.2. Object density (DO)

Number of objects divided by the area of the super-object at the class and super-object levels. It measures the number of objects by unit area (Gong et al., 2013; Herold et al., 2002; Irwin and Bockstael, 2007; McGarigal et al., 2012). The higher the value, the denser.

$$DO = \frac{n}{A_T} \qquad (A.17)$$

A.1.3.3. Leapfrog (LPF)

Ratio between the total area of isolated objects from a class (i.e., objects located separately at a distance from the rest of objects of a class) and the area of the class (%) (Frenkel and Ashkenazi, 2008). It quantifies the percentage of isolated objects in class. It is computed at the class level.

$$LPF = \frac{A_{out}}{\sum_{i=1}^{n}(A_i)} \qquad (A.18)$$

A_{out} = area of isolated objects of a class (m²)

A_i = area of the object i from a class (m²)

A.1.3.4. Area-weighted mean standard distance (AWSD)

It measures the degree to which objects of a class are concentrated around their centroid. It equals the average of the distances from objects to the centroid in a super-object (km) (Colaninno et al., 2011a). It is computed at the class and super-object levels. Low values show a concentrated distribution of objects and thus centrality, while high values show disaggregation.

It can be normalized by the shape and size of the super-object by means of the maximum distance of a regular grid with the same area as the super-object to the centroid, the normalized index ranges from 0 to 100 (%).

$$AWSD= \sqrt{\left(\frac{\sum_{i=1}^{n}(A_i \cdot (x_i - \bar{x})^2)}{A_T}\right) + \left(\frac{\sum_{i=1}^{n}\left(A_i \cdot \left(y_i - \bar{y}\right)^2\right)}{A_T}\right)} \qquad (A.19)$$

$$AWSD_N = \frac{AWSD}{SD_{max}} \cdot 100 \qquad (A.20)$$

x_i , y_i= coordinates of the centroid of an object i (m)

$\bar{x}$, $\bar{y}$= coordinates of the centroid of objects from a class/super-object (m)

SD_{max}= maximum Standard distance, computed based on a grid covering the same area as the super-object.

A.1.3.5. *Euclidean nearest neighbor mean distance (ENND)*

It quantifies object isolation by measuring the average distance between nearest objects from a class in a super-object (m), it is computed at the class level (Gong et al., 2013; McGarigal et al., 2012). Higher distances show more isolation.

$$ENND = \frac{\sum_{i=1}^{n}(D_{ij})}{n} \qquad (A.21)$$

D_{ij}= distance between an object i and its nearest object j (from boundary to boundary) of the same class (m).

A.1.3.6. *Effective mesh size (EMS)*

It represents the size of the objects when the class or super-object are divided into n areas with the same degree of the class or super-object division (km^2). It measures the probability that two random points in an area are connected to each other (EEA, 2011; Jaeger, 2000; McGarigal et al., 2012).

$$EMS = \frac{\sum_{i=1}^{n}\left(A_i^2\right)}{A_T} \qquad (A.22)$$

A.1.3.7. *Cohesion (COHE)*

It measures the physical connectivity of the objects at the class or super-object levels (Congalton, 2013; McGarigal et al., 2012). It value is higher when objects are more aggregated.

$$COHE=\frac{1-\left(\frac{\sum_{i=1}^{n}(P_i)}{\sum_{i=1}^{n}(P_i\cdot\sqrt{A_i})}\right)}{1-\left(\frac{1}{\sqrt{A_T}}\right)}$$

(A.23)

A.1.3.8. Splitting index (SPL)

It represents the number of objects when dividing the class or super-object into equal parts, with the same degree of class or super-object division (Jaeger, 2000; McGarigal et al., 2012). It is computed at the class and super-object levels. It equals one when there is only one object, its value increases when smaller objects are subdividing the class or super-object into smaller and fragmented parts.

$$SPL=\frac{A_T^2}{\sum_{i=1}^{n}(A_i^2)}$$

(A.24)

A.1.3.9. Coherence degree (CD)

It represents the probability for two points to be in the same object in a class or super-object. It is computed at the class and super-object levels (Jaeger, 2000; McGarigal et al., 2012).

$$CD=\sum_{i=1}^{n}\left(\frac{A_i}{A_T}\right)^2$$

(A.25)

A.1.3.10. Radius dimension (Dim$_R$)

It measures the centrality of a class. It describes the radial spatial distribution and variation of objects around a designated center point, reflecting the centrality of objects. It shows a cumulative or diffuse distribution of the objects with respect to the given point. It is defined by the total area of the objects within a class by equidistant radius, and represents the density change of a particular class based on the distance to a center. The radius is a set of equidistant circumferences centered at the given center. It is computed by performing a minimum quadratic adjustment of the objects (using the class area within a radius and the radius as variables) where the Dim$_R$ is obtained from the slope of the regression line (Wu et al., 2013). This metric generates a graphical output with the values of the class area in ring (Y-axis) and the radius of the equidistant ring (X-axis). The Dim$_R$ is the inverse of the slope of the line in red (Figure A.4).

$$\ln A_c(r)=Dim_R\cdot\ln(r)+\ln(C)$$

(A.26)

r=radius from the center point.

$A_c(r)$ = area of the class within the circumference with radius r (m²).
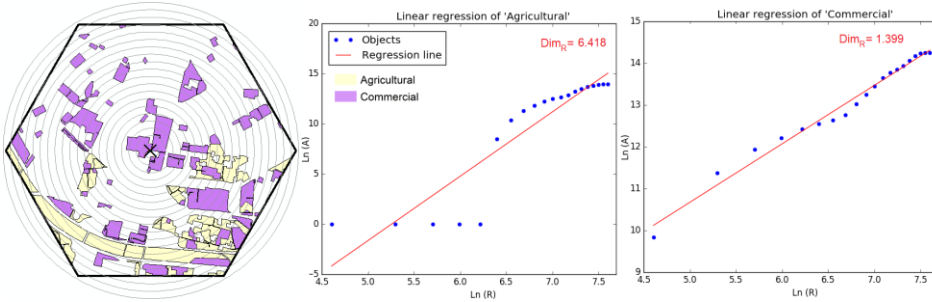

Figure A.4. Example of the distribution of each class based on the logarithm of the area and radius to the centroid, using as distance 100 m, it shows the Dim_R for the Agricultural and Commercial classes on a super-object (left). The plots on the right are outputs of IndiFrag when Dim_R is computed.

A.1.3.11. Class compactness (C)

This metric measures the compactness of the objects in a class (Zhang et al., 2016). High values show more compact class. The value ranges from 0 to 100, being 100 a circumference (%). It is computed at the class level.

$$C = \frac{2 \cdot \sqrt{\pi \cdot \sum_{i=1}^{n}(A_i)}}{\sum_{i=1}^{n}(P_i)} \cdot 100 \tag{A.27}$$

A.1.3.12. Urban compactness (CU)

It is the compact degree of the urban cover, similar to (A.27) but accounting only for urban objects. It is computed at the super-object level. High values indicate a regular shape of the urban cover. The value ranges from 0 to 100, being 100 a circumference (%).

$$CU = \frac{2 \cdot \sqrt{\pi \cdot A_U}}{P_U} \cdot 100 \tag{A.28}$$

A_U= Area of the urban cover (m²).

P_U= perimeter of the urban cover (m).

A.1.3.13. Dispersion index (DI)

Taubenböck et al. (2014) proposed the dispersion index as a function of the number of objects (Nob) and the largest object (A.9). If a class in the super-object is represented by one single coalescent object, it represents a compact

class. If a class has the maximum possible number of individual objects, the class would be the most dispersed. It is computed at the class level.

$$Nob_N = \frac{NOb-1}{NOb_{max}-1} \cdot 100 \qquad (A.29)$$

$$DI = \frac{Nob_N+(100-LOI)}{2} \qquad (A.30)$$

NOb_{max}= maximum number of possible objects in the class. When data come from vector format, it equals the area of the class divided by the minimum mapping unit. If data come from raster, it is the total number of pixels of the class.

A.1.4. Diversity

This group of metrics quantifies the diversity of classes within a super-object. It describes the internal composition of the super-object without the influence of the spatial configuration of the objects. Diversity has two components, richness, which refers to the number of classes present in the super-object, and uniformity, which refers to the distribution of classes in a super-object.

A.1.4.1. Number of classes (Ncl)

It counts the total number of classes in the super-object.

A.1.4.2. Shannon diversity (DSHAN)

It measures the diversity of classes in a super-object computed by the total number of classes and the percentages of the super-object covered by each class. Higher values show more diversity. It equals minus the sum of the proportional abundance of each class multiplied by that proportion (Colaninno et al., 2011; McGarigal et al., 2012). It is computed at the super-object level.

$$DSHAN = -\sum_{i=1}^{m} [P \cdot (\ln(P))] \qquad (A.31)$$

P= percentage of the super-object occupied by a class i (A_i/A_T)(%)

A.1.4.3. Shannon evenness (USHAN)

It measures the number of classes in a super-object and their relative percentages. It is calculated by dividing the DSHAN by its maximum value (McGarigal et al., 2012, Romano et al., 2010). It is computed at the super-object level.

$$USHAN=\frac{-\sum_{i=1}^{m}(P\cdot(\ln(P))}{\ln(m)} \tag{A.32}$$

m = number of classes in the super-object.

A.1.4.4. Simpson diversity (SIMP)

It represents the probability that two objects selected randomly belong to different classes. It is close to zero when the distribution of areas among classes is uneven or there is only one class. It is one when the distribution of the areas of the classes is more even (McGarigal et al., 2012), computed at the super-object level.

$$SIMP=1-\sum_{i=1}^{m}(P^2) \tag{A.33}$$

A.1.4.5. Density-diversity (DD)

It quantifies the diversity and uniformity of classes. Unlike DSHAN, it increases as the distribution of classes becomes more uniform, instead of with the total number of classes, i.e.: it prioritizes uniformity. It can be calculated at class level or super-object level. At the class level it measures the division of a class area with respect to the class area in a super-object. At the super-object level, it equals the sum of the areas of each class divided by the area of the most representative class (i.e., the class with largest area) (Batty et al., 2004; Escolano, 2009)

$$DD=\sum_{i=1}^{m}\left(\frac{\sum_{i=1}^{n}(A_i)}{\max(\sum_{i=1}^{n}(A_i))}\right) \tag{A.34}$$

A.1.4.6. Relative functional fragmentation index (RFFI)

It measures the level of functional fragmentation in a super-object by means of the ratio between the total number of classes in the super-object and the total number of classes in the study area (Marinescu and Avram, 2012). It is computed at the super-object level.

$$RFFI=\frac{(R_v-m)}{(R_v-1)} \tag{A.35}$$

R_v= number of classes in the total study area beyond the study area (having into account all the classes from all the super-objects).

A.1.4.7. Absolute functional fragmentation index (AFFI)

It represents the ratio between the super-object perimeter and the sum of the perimeters for each class (Marinescu and Avram, 2012). It measures the level of functional and structural integration within the perimeter of the super-object.

$$AFFI = \frac{P_T}{\sum_{i=1}^{m} \left(\sum_{i=1}^{n} (P_i) \right)}$$

(A.36)

P_T = perimeter of the super-object (m).

A.1.5. Contrast

This group quantifies the differences between adjacent or continuous objects from different classes within a super-object. The contrast between the objects and their neighbor objects can influence internal processes within the super-object.

A.1.5.1. Boundary contrast ratio (BCR)

It equals the sum of the segment lengths of an object adjacent to an object from a different class, divided by the perimeter of the object (Irwin and Bockstael, 2007; McGarigal et al., 2012). It ranges between zero and one, being zero when the object is contiguous to objects from the same class, and one when the entire perimeter of the object is shared with objects from other classes.

$$BCR = \frac{l_{i,j}}{P_i}$$

(A.37)

$l_{i,j}$ = length of the edges shared between two objects from different classes *i j* (m). At class or super-object level equals the sum of the total lengths (m).

For those categorical maps with a class assigned to the roads, an optional functionality has been included that "removes" the influence of the roads in the contiguity analysis between objects. The algorithm detects the axes of the roads and increases the size of the surrounding objects to these axis, and then analyzes the $l_{i,j}$ based on the extended objects for each class (Figure A.5).
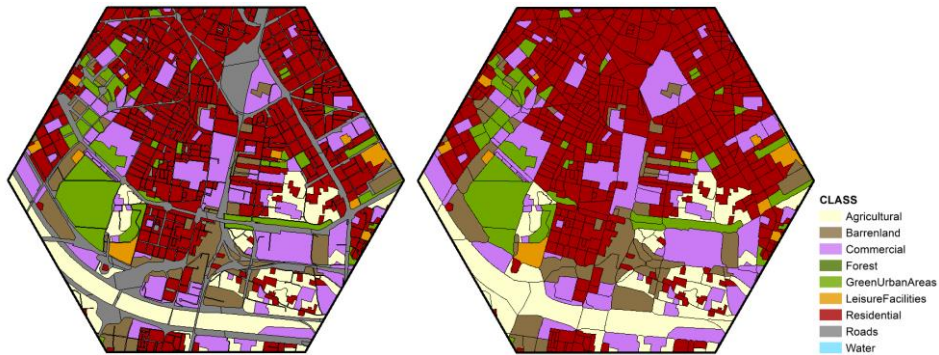
Figure A.5. Example of the optional removal of the road class for computing the contrast ratio between classes. The example on the left is an original LULC map, on the right the extended objects are shown after removing the roads.

A.2. Spatio-temporal metrics

The extraction of spatial metrics at different dates allows for the analysis of changes in terms of the variation of spatial distribution and urban morphology. In order to complement the temporal analysis, IndiFrag includes a set of multi-temporal metrics from the literature that directly measure changes from two categorical maps (Sapena and Ruiz, 2015b). Therefore, there are two sets of metrics in the spatio-temporal analysis, those obtained as the difference of the spatial metrics extracted independently from two maps (equations (3.1) to (A.37)), and metrics extracted simultaneously from two maps (equations (A.39) to (A.49)).

A.2.1. Difference between spatial metrics

IndiFrag v.3.1 includes a sub-module where the difference between the spatial metrics, previously computed for two categorical maps, is calculated straightforward. The differences for each metrics by class are stored in the super-object attribute table, facilitating the creation of maps in GIS software.

A.2.2. Multi-temporal metrics

A.2.2.1. Land use change development (LUC)

Ratio of change of the percentages of class areas weighted by their exploitation degree between two dates (Pan et al., 2011). It is computed at the super-object level. When LUC is greater than zero the super-object is in a state of development, when this is negative it is in a process of readjustment or underdevelopment.

$$LUD= \sum_{i=1}^{n}(E \cdot P) \cdot 100 \qquad (A.38)$$

$$LUC= \frac{LUD_{t2}-LUD_{t1}}{LUD_{t1}} \qquad (A.39)$$

E = exploitation degree of the class, where: (E=1) barren or unused land; (E=2) forest, water, grass; (E=3) agricultural land; and (E=4) urban, mine, roads, etc.

P= proportion of the super-object occupied by class i (A_i/A_T) (%)

A.2.2.2. Change proportion (CP)

The ratio represents the change of a class with respect to the size of the super-object. It is useful to assess the relative area of change for the different classes, showing their expansion intensity (Yin et al., 2011). It is computed at the class level.

$$CP= \frac{A_{t2}-A_{t1}}{A_T} \cdot 100 \qquad (A.40)$$

A_{t1}= area of a class in the first date (m²)

A_{t2}= area of a class in the second date (m²)

A.2.2.3. Landscape expansion index (LEI)

Liu et al. (2010) proposed several indices to quantify urban dynamics, and they were slightly modified and implemented in *IndiFrag*. First, the objects are classified in two types: new developed objects and old objects. Objects that existed in both times are old, while those that did not exist in the first time are new. Subsequently, new objects are classified in three growth types: infilling, edge-expansion, and outlying (Liu et al., 2010; Sun et al., 2013; Wilson et al., 2003). For the classification, we used the ratio between the length of the common edge between the new and old neighboring objects and the perimeter of the new object (Figure A.1).

$$LEI= \frac{l_w}{P_w} \cdot 100 \qquad (A.41)$$

l_w= length of the edge between a new object and old neighboring object (m).
P_w= perimeter of a new object (m).

It ranges from 0 to 100, where:

- Infilling growth: $100 \geq LEI > 50$

- Edge-expansion growth: 50 ≥ LEI > 0
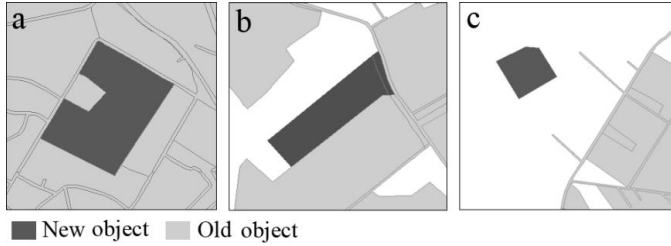- Outlying growth: LEI = 0.

Figure A.1. Three growth types: (a) infilling, (b) edge-expansion, and (c) outlying.

The index is computed at the object level, and statistics and graphical outputs are created at the class level (Figure A. 6).
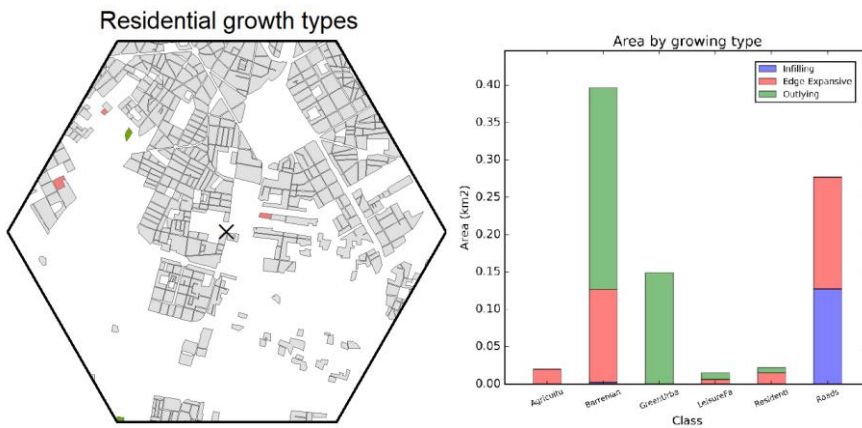

Figure A. 6. Example of the residential growth types (left) and plot created in IndiFrag (right), quantifying the total amount of each growth type per class for the study area.

A.2.2.4. Mean expansion index (MEI)

It equals to the average of the LEI of all new objects in a class (Liu et al., 2010; Sun et al., 2013; Wilson et al., 2003). It is computed at the class level.

$$MEI = \frac{\sum_{w=1}^{nw} LEI_w}{nw} \tag{A.42}$$

nw = number of new objects

A.2.2.5. Area-weighted mean expansion index (AWMEI)

It equals to the area-weighted sum across all new objects of a class. It

quantifies the aggregation properties of the objects within a class (Liu et al., 2010; Sun et al., 2013; Wilson et al., 2003). It is computed at the class level. A high value of AWMEI indicates a densification and adjacency of growth, and therefore a compacted pattern of the urban growth, while a low value represents a scattered trend.

$$\text{AWMEI} = \sum_{w=1}^{nw} \left(\text{LEI}_w \cdot \frac{A_w}{A_{TW}} \right) \tag{A.43}$$

A_w = area of a new object (m^2).

A_{TW} = total area of the new objects from a class (m^2).

A.2.2.6. Area new (A_{new})

Total area of new objects from a class (km^2), computed at the class level.

A.2.2.7. Area Infilling (A_{inf})

Total area of new objects from a class with infilling growth type (km^2), computed at the class level.

A.2.2.8. Area edge-expansion (A_{exp})

Total area of new objects from a class with edge-expansive growth type (km^2), computed at the class level.

A.2.2.9. Area outlying (A_{out})

Total area of new objects from a class with outlying growth type (km^2), computed at the class level.

A.2.2.10. Change rate (CR)

This index measures the dynamic characteristics of urban expansion, is calculated for the class level and is used to study the change over a period of time. It is calculated considering that the change is not constant in time (%) (Malaviya et al., 2010).

$$\text{RC} = \frac{1}{(t_2 - t_1)} \cdot \ln\left(\frac{A_{t2}}{A_{t1}}\right) \cdot 100 \tag{A.44}$$

$(t_2 - t_1)$ = number of years between the two datasets

A.2.2.11. Change area (A)

The area of change is the total area of a class changed in a period of time (gain or loss in km^2) (Tian et al., 2014), computed at the class level.

$$A_c = A_{t2} - A_{t1} \tag{A.45}$$

A.2.2.12. Change area ratio (Ar)

It represents the area of change per year. This gives a value of the rate of change in km^2 per year for each class (Tian et al., 2014), computed at the class level.

$$A_r = \frac{A_c}{t_2 - t_1} \tag{A.46}$$

A.2.2.13. Urban change ratio (UCR)

Proportion of growth by class relative to the area at the first date (%), computed at the class level.

$$UCR = \frac{A_{t2} - A_{t1}}{A_{t1}} \cdot 100 \tag{A.47}$$

A.2.2.14. Area-weighted mean accessibility index (AWMAI)

It quantifies the accessibility of new objects of a class when data from the road network is available. It measures the inversed area-weighted-mean distance from new objects of a class to the closest road (%). Higher values show more accessibility with closer objects to the roads, while lower values show more isolation. It is computed at the class level.

$$AWMAI = \sum_{w=1}^{nw} \left(\frac{A_w}{(D_w^R + 1) \cdot A_{TW}} \right) \cdot 100 \tag{A.48}$$

$(D_w^R + 1)$ = distance between a new object w from a class to the closest road. The unit is summed to avoid that objects adjacent to a road with a distance of zero create an undefined function.

A.2.2.15. Disaggregation index (DIS)

It measures the extent to which growth occurs at a significant distance from existing objects from the same class (m). The higher the value, the more spread and isolated distribution the new objects. A value close to zero will show aggregated growth (Reis et al., 2015). It is computed at the class level.

$$DIS = \frac{\sum_{w=1}^{nw} (D_w^{old})}{nw} \tag{A.49}$$

D_w^{old} = distance between a new object and the closest old object from the same class (m)

A.2.2.16. Centroid displacement (CNT)

In an attempt to characterize the evolution of urban expansion, Jing and Jianzhong (2011) proposed to measure the city center coordinate displacement (CNT). This metric represents the tendency of aggregation of a class with respect to the geometric center and its displacement. Thus, the distance between the two geometric centroids in two dates is computed (m). It is computed at the class level.

$$CNT=\sqrt{(x_{t2}-x_{t1})^2+(y_{t2}-y_{t1})^2} \tag{A.50}$$

x_{t1}, y_{t1} = coordinates of the class center in the first dates

x_{t2}, y_{t2} = coordinates of the class center in the second dates

A.2.2.17. Concentric circle

The concentric circle analysis creates a graphical output used to quantify the spatial distribution of growth by class in the study area, without considering the super-objects (Figure A.7). It measures the area of a class in two dates based on several distances to a given point considered the center of the data (Yin et al., 2011). The equidistance of concentric circles can be stablished depending on the scale and size of the study area, and the center is introduced as an input parameter in IndiFrag.
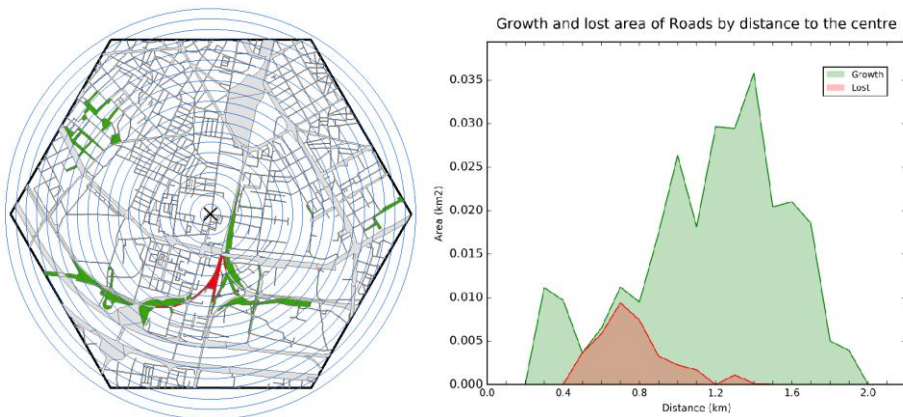
Figure A.7. Global concentric ring growth analysis. Example of the growth of a class according to the distance to the center. This plot is created by the tool for each class.

A.2.2.18. Sector analysis

The sector analysis creates a graphical output used to quantify the spatial

distribution of growth by class in the study area, without considering the super-objects (Figure A.8). It measures the area of a class in two dates based on 16 sectors with angles of 22.5 degrees, measuring the change at different orientations with respect to the center point (Yin et al., 2011).
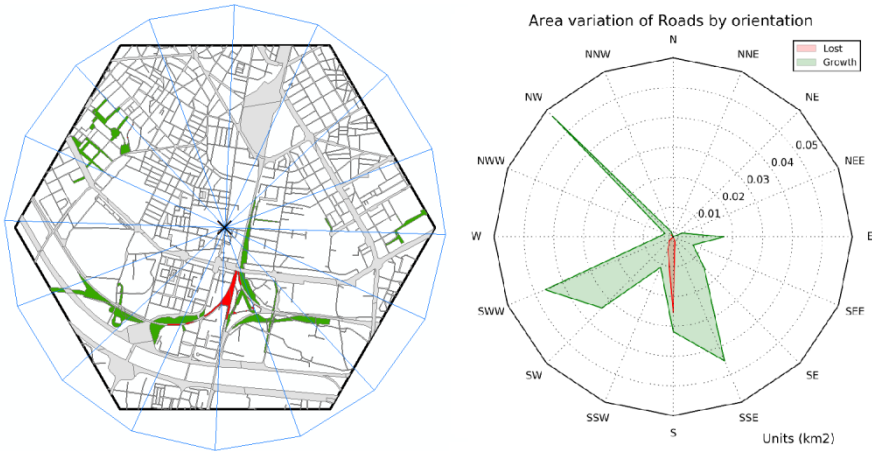

Figure A.8. Sector growth analysis. Example of the growth of a class according to the orientation of change. This plot is created by the tool for each class.

A.2.2.19. LULC change analysis

A land-use/land-cover change analysis is conducted using two categorical maps. This analysis is done at different levels. At the object level, the class at each date is stored and objects are classified as change/no-change, which allows for the creation of great detail change maps (Figure A.9). At the class and super-object levels, the total area of change and no-change per class and super-object are calculated. The output is a shapefile with the objects and their changes, as well as two change matrices with the summary of the changes per class and super-object (Figure A.10).
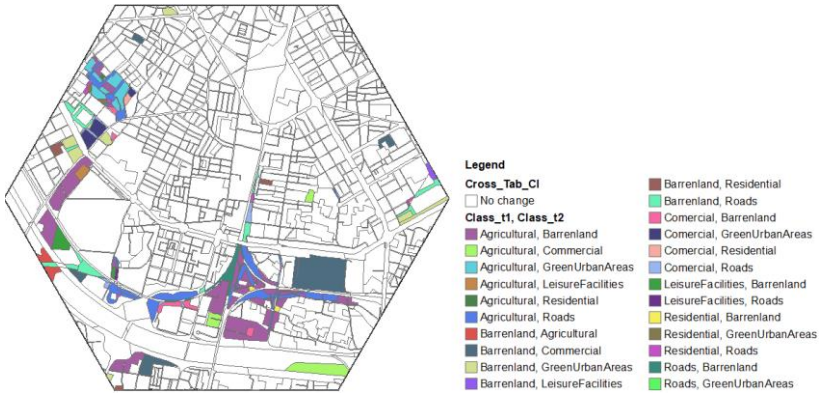
Figure A.9. Example of LULC change map at the object level.

	FID *	Class_t1	Class_t2	Descrip	NAME	FREQUENCY	SUM_Area
	1	Agricultural	Agricultural	NoChange	Hex_24	47	1132647.04351
	2	Agricultural	Barrenland	Change	Hex_24	30	311615.294434
	3	Agricultural	Commercial	Change	Hex_24	3	66657.739258
	4	Agricultural	GreenUrbanAreas	Change	Hex_24	7	52909.14624
	5	Agricultural	LeisureFacilities	Change	Hex_24	1	9219.672852
	6	Agricultural	Residential	Change	Hex_24	1	1843.84729
	7	Agricultural	Roads	Change	Hex_24	37	168184.918348
	8	Barrenland	Agricultural	Change	Hex_24	2	19923.65918
	9	Barrenland	Barrenland	NoChange	Hex_24	30	350444.808838

Cross_Tab_Cl_SO

Figure A.10. Example of a change matrix table at the class level (Class_t1 and Class_t2) and super-object level (NAME).

Table A.1. IndiFrag v3.1 Description of spatio-temporal metrics included in IndiFrag software tool. Categorised in groups: Area and perimeter, shape, aggregation, diversity, contrast and multi-temporal. The group, name, formula, definition, units, reference and level are reported, where O means object level, Cl class level and SO super-object level. Acronyms of formulas are below the table.

Eq.	Name	Formula	Definition	Unit	Reference	Level
Area and perimeter						
-	Area	Area	Area of the object, class or super-object	$m^2 \mid km^2$	-	O,Cl,SO
-	Perimeter	Perim	Perimeter of the object, class or super-object (without boundary duplicity)	$m \mid km$	-	O,Cl,SO
-	Total perimeter	PerimT	Perimeter of the super-object plus the total perimeter of the objects (without boundary duplicity)	km	-	SO
(A.1)	Urban density	$DU=\dfrac{A_u}{A_T}$	Ratio between urban area and super-object area	%	Romano et al., 2006	SO
(A.2)	Class density	$DC=\dfrac{\sum_{i=1}^{n}(A_i)}{A_T}$	Ratio between a class area and super-object area	%	-	Cl
(A.3)	Road density	$D_{road}=\dfrac{\sum_{i=1}^{nr}(L_i)}{A_T}$	Ratio between the total length of roads in meters and super-object area	m/km^2	-	SO
(A.4)	Object mean size	$MS=\dfrac{\sum_{i=1}^{n}(A_i)}{n}$	Equals the average of the size of the objects from a class or super-object	m^2	Irwin y Bockstael, 2007	Cl,SO
(A.5)	Edge density	$ED=\dfrac{\sum_{i=1}^{n}(P_i)}{A_T}$	Equals the sum of the perimeter of the objects from a class or super-object divided by the area of the super-object	m/m^2	Herold et al., 2002; McGarigal et al., 2012	Cl,SO
(A.6) (A.7)	Area-weighted mean urban fragmentation index	$AWUFI_i=\dfrac{\sum_{i=1}^{ns}(UFI_i \cdot A_{Ti})}{\sum_{i=1}^{ns}(A_{Ti})}$ $UFI_i=\dfrac{\sum_{i=1}^{m}(L_{max} \cdot \sum_{i=1}^{n}(A_i) \cdot O_c)}{A_{Ti}}$	Evaluates the habitat fragmentation due to the presence of urban areas, it is weighted by SO area in order to perform comparisons between the fragmentation values	None	Astiaso et al., 2013; Romano 2002; Romano and Tamburini, 2006	SO
(A.8)	Porosity	$P=\dfrac{\sum_{h=1}^{nh}(A_h)}{\sum_{i=1}^{n}(A_i)} \cdot 100$	Ratio between the open space area (area of holes within the class) and the area of the class.	%	(Reis et al., 2015)	Cl
(A.8)	Urban porosity	PU	Ratio between the urban open space area (area of holes within the urban cover) and the area of the urban area.	%	(Reis et al., 2015)	SO
(A.9)	Largest object index	$LOI=\dfrac{A_{max}}{\sum_{i=1}^{n}(A_i)} \cdot 100$	Percentage of a class that represents the largest object.	%	(McGarigal et al., 2012)	Cl

(A.10)	Second largest object index	$SLOI=\dfrac{A_{2nd\ max}}{\sum_{i=1}^{n}(A_i)}\cdot 100$	Percentage of the class that represents the second largest object.	%	-	CI

Shape

(A.11)	Boundary dimension	$\ln(A_i)=\dfrac{2}{Dim_B}\cdot \ln(P_i)+\ln(C)$ graph	Represents the relationship between the object area and the perimeter, it measures the complexity and randomicity of classes	None	Wu et al, 2013	CI
(A.12)	Fractal dimension	$FD=\sum_{i=1}^{n}\dfrac{2\cdot \ln(0.25\cdot P_i)}{\ln(A_i)}$	A normalized shape index based on perimeter-area relationships in which the perimeter and area are transformed with logarithm	None	Frenkel and Ashkenazi, 2008; Gong et al., 2013; Herold et al., 2002	O, CI, SO
(A.13)	Area-weighted mean fractal dimension	$AWFD=\sum_{i=1}^{n}\left[\left(\dfrac{2\cdot \ln(0.25\cdot P_i)}{\ln(A_i)}\right)\cdot \left(\dfrac{A_i}{\sum_{i=1}^{n}A_i}\right)\right]$	Equals the average fractal dimension of objects in the class or super-object, weighted by the area of the object	None	Gong et al., 2013; Herold et al., 2002; McGarigal et al., 2012	CI
(A.14)	Shape index	$SI=\sum_{i=1}^{n}\dfrac{0.25\cdot P_i}{\sqrt{A_i}}$	A normalized ratio of object perimeter to area in which the complexity of object shape is compared to a square of the same size	None	Frenkel and Ashkenazi, 2008; Jiang et al., 2007	O, CI, SO
(A.15)	Perimeter-area mean ratio	$PAR=\dfrac{\sum_{i=1}^{n}\left(\dfrac{P_i}{A_i}\right)}{n}$	Describes the relationship between the object area and perimeter, and thus describes how object perimeter increases per unit increase in object area	None	Irwin and Bockstael, 2007, ; McGarigal et al., 2012	O,CI,SO
(A.16)	Elongation ratio	$ER=\dfrac{2\cdot \sqrt{A_{max}/\pi}}{L_{max}}$	Measures the elongation of the largest patch dividing the diameter of the circumference with the same area than the object by the largest side of the object	%	Schumm, 1956	CI

Aggregation

-	Number of Objects	Nob	Number of objects in a class or super-object (except road objects)	n°	-	CI,SO
(A.17)	Object density	$DO=\dfrac{n}{A_T}$	Number of objects divided by the area of the super-object	n°/km²	Gong et al., 2013; Herold et al., 2002; Irwin and Bockstael, 2007; McGarigal et al., 2012	CI,SO

(A.18)	Leapfrog	$LPF=\dfrac{A_{out}}{\sum_{i=1}^{n}(A_i)}$	Ratio between the area of leapfrog or isolated objects from a class located separately at a distance from the rest of the class and the area of the whole class	%	Frenkel and Ashkenazi, 2008	CI
(A.19)	Area-weighted mean standard distance	$AWSD=\sqrt{\left(\dfrac{\sum_{i=1}^{n}(A_i\cdot(x_i-\bar{x})^2)}{A_T}\right)\cdot\left(\dfrac{\sum_{i=1}^{n}\left(A_i\cdot(y_i-\bar{y})^2\right)}{A_T}\right)}$	Degree to which objects from are concentrated around their centroid. Equals the average of the distance from objects to the centroid in a super-object	km	Colaninno et al., 2011	CI,SO
(A.20)	Normalized Area-weighted mean standard distance	$AWSD_N=\dfrac{AWSD}{SD_{max}}\cdot100$	Normalized by means of the maximum distance of a regular grid with the same area as the super-object to the centroid	%	-	CI
(A.21)	Euclidean nearest neighbor mean distance	$ENND=\dfrac{\sum_{i=1}^{n}(D_{ij})}{n}$	Quantifies object isolation. Equals the average distance between nearest objects from the same class in a super-object	m	Gong et al., 2013; McGarigal et al., 2012	CI
(A.22)	Effective mesh size	$EMS=\dfrac{\sum_{i=1}^{n}(A_i^2)}{A_T}$	Size of the objects when the super-object is divided into n areas with the same degree of super-object division	km²	EEA, 2011; Jaeger, 2000; McGarigal et al., 2012	CI, SO
(A.23)	Cohesion	$COHE=\dfrac{1-\left(\sum_{i=1}^{n}(P_i)\Big/\sum_{i=1}^{n}(P_i\cdot\sqrt{A_i})\right)}{1-\left(1\Big/\sqrt{A_T}\right)}$	Measures the physical connection of the objects	None	Maclean and Congalton, 2013; McGarigal et al., 2012	CI,SO
(A.24)	Splitting index	$SPL=\dfrac{A_T^2}{\sum_{i=1}^{n}(A_i^2)}$	Number of objects when dividing the super-object into equal parts, with the same degree of super-object division	None	Jaeger, 2000; McGarigal et al., 2012	CI,SO
(A.25)	Coherence degree	$CD=\sum_{i=1}^{n}\left(\dfrac{A_i}{A_T}\right)^2$	Represents the probability that two points are in the same object in a super-object	%	Jaeger, 2000; McGarigal et al., 2012	CI,SO
(A.26)	Radius dimension	$\ln A_c(r)=Dim_R\cdot\ln(r)+\ln(C)$ graph	Defined by the total object class area and its radius to depict the density change radiating outward from	None	Wu et al., 2013	SO

			a center point. An effective evaluation criterion to assess the centrality of land use patterns			
(A.27)	Class compactness	$C=\dfrac{2\cdot\sqrt{\pi\cdot\sum_{i=1}^{n}(A_i)}}{\sum_{i=1}^{n}(P_i)}\cdot100$	Quantifies the degree of spatial aggregation of the objects of a class	%	Zhang et al., 2016	CI
(A.28)	Urban compactness	$CU=\dfrac{2\cdot\sqrt{\pi\cdot A_U}}{P_U}\cdot100$	Quantifies the degree of spatial aggregation of the urban type objects	%	Zhang et al., 2016	SO
(A.29) (A.30)	Dispersion Index	$DI=\dfrac{Nob_N+(100-LOI)}{2}$ $Nob_N=\dfrac{NOb-1}{NOb_{max}-1}\cdot100$	As a function of the number of objects (Nob) and the largest object (A.10)	-	Taubenböck et al., 2014	CI

Diversity

-	Number of classes	NCl	Number of classes in the super-object	No.	-	SO
(A.31)	Shannon diversity	$DSHAN=-\sum_{i=1}^{m}[P\cdot(\ln(P))]$	Equals minus the sum of the proportional abundance of each object class multiplied by that proportion	None	Colaninno et al., 2011, McGarigal et al., 2012	SO
(A.32)	Shannon's evenness	$USHAN=\dfrac{-\sum_{i=1}^{m}(P\cdot(\ln(P)))}{\ln(m)}$	Covers the number of classes in a super-object and their relative abundances. It is calculated by dividing DSHAN by its maximum	None	McGarigal et al., 2012, Romano et al., 2010	SO
(A.33)	Simpson diversity	$SIMP=1-\sum_{i=1}^{m}(P^2)$	Probability that two random objects are from different classes	None	McGarigal et al., 2012	SO
(A.34)	Density-diversity	$DD=\sum_{i=1}^{m}\left(\dfrac{\sum_{i=1}^{n}(A_i)}{\max(\sum_{i=1}^{n}(A_i))}\right)$	Equals the sum of the amount of a class as proportion of its maximum	None	Batty et al., 2003, Escolano, 2009	CI,SO
(A.35)	Relative functional fragmentation	$RFFI=\dfrac{(R_v-m)}{(R_v-1)}$	Points out the level of functional fragmentation in a super-object by the ratio of the number of classes in the super-object to the number of classes of the whole study area	None	Marinescu and Avram, 2012	SO
(A.36)	Absolute functional fragmentation	$AFFI=\dfrac{P_T}{\sum_{i=1}^{m}(\sum_{i=1}^{n}(P_i))}$	Level of functional and structural integration within the perimeter. Ratio between the super-object and the sum of class perimeters	None	Marinescu and Avram, 2012	SO

Contrast

(A.37)	Boundary contrast ratio	$BCR=\dfrac{l_{i,j}}{P_i}$	Equals the sum of the segment lengths of an object (class or super-object) adjacent to different classes, divided by the perimeter of the object (or the class or super-object depending on the level)	%	Irwin and Bockstael, 2007, McGarigal et al., 2012,	O,Cl,SO
Multi-temporal						
(A.38) (A.39)	Land use change development	$LUD=\displaystyle\sum_{i=1}^{n}(E\cdot P)\cdot100$ $LUC=\dfrac{LUD_{t2}-LUD_{t1}}{LUD_{t1}}$	It is based on the exploitation degree of classes that are classified into four levels. It is calculated on the basis of the ratio of class area to the super-object area with the consideration of weighted values of each class type level between the two dates	None	Pan et al., 2011	SO
(A.40)	Change proportion	$CP=\dfrac{A_{t2}-A_{t1}}{A_T}\cdot100$	Expansion intensity. Ratio between the change area of a class and the area of the super-object	%	Yin et al., 2011	Cl
(A.41)	Landscape expansion index	$LEI=\dfrac{l_w}{P_w}\cdot100$ graph	Categorizes new objects in: infilling, edge-expansion, and outlying types by comparing perimeters between new and old objects	%	Liu et al., 2010; Sun et al., 2013; Wilson et al., 2003	O
(A.42)	Mean expansion index	$MEI=\dfrac{\sum_{w=1}^{nw}LEI_w}{nw}$	Equals to the mean of the LEI of all new objects	None	Liu et al., 2010; Sun et al., 2013; Wilson et al., 2003	Cl
(A.43)	Area-weighted mean expansion index	$AWMEI=\displaystyle\sum_{w=1}^{nw}\left(LEI_w\cdot\dfrac{A_w}{A_{TW}}\right)$	Equals to the sum, across all new objects, previously classified according to LEI, multiplied by the proportional area of the new object	None	Liu et al., 2010; Sun et al., 2013; Wilson et al., 2003	Cl
-	Area new	A_{new}	Total area of new objects from a class. It does not take into account lost objects.	km²	-	Cl
-	Area infilling	A_{inf}	Total area of new objects from a class with infilling growth type	km²	-	Cl
-	Area edge-expansion	A_{edg}	Total area of new objects from a class with edge-expansive growth type	km²	-	Cl
-	Area outlying	A_{out}	Total area of new objects from a class with outlying growth type	km²	-	Cl
(A.44)	Change rate	$RC=\dfrac{1}{(t_2-t_1)}\cdot\ln\left(\dfrac{A_{t2}}{A_{t1}}\right)\cdot100$	Annual rate of change calculated as it was not linear	%	Malaviya et al., 2010	Cl
(A.45)	Change area	$A_c=A_{t2}-A_{t1}$	Equals the difference of the areas of a class in the studied period	km²	Tian et al., 2014	Cl

(A.46)	Change area ratio	$A_r = \dfrac{A_c}{t_2 - t_1}$	Equals the difference of the areas of a class in the studied period divided by the number of years	km²/year	Tian et al., 2014	CI
(A.47)	Urban change ratio	$UCR = \dfrac{A_{t2} - A_{t1}}{A_{t1}} \cdot 100$	Proportion of growth by class relative to the area in the first time	%	-	CI
(A.48)	Area-weighted mean accessibility index	$AWMAI = \sum\limits_{w=1}^{nw} \left(\dfrac{A_w}{(D_w^R + 1) \cdot A_{TW}} \right) \cdot 100$	Measures the inversed area-weighted-mean distance from new objects of a class to the closest road.	%	-	CI
(A.49)	Disaggregation index	$DIS = \dfrac{\sum_{w=1}^{nw} (D_w^{old})}{nw}$	Measures the extent to which growth occurs at a significant distance from the existing class. It is the mean distance from every new object to the class.	m	(Reis et al., 2015)	CI
(A.50)	Centroid displacement	$CNT = \sqrt{\dfrac{(x_{t2} - x_{t1})^2 +}{(y_{t2} - y_{t1})^2}}$	Distance and orientation between the geometric centers of a class (t_1 and t_2)	m	Jing and Jianzhong, 2011	CI
	Concentric circle	graph	Quantity and spatial distribution of a class change, it measures areas at different distances (given) with respect to a center point	km²	Yin et al., 2011	CI
	Sector analysis	graph	Quantity and spatial distribution of a class change, it measures areas at different orientation (22.5 degrees) with respect to a center point	km²	Yin et al., 2011	CI
	LULC change	Change tables and maps	A LULC change analysis is conducted using two LULC maps.	-	-	O, CI, SO

i = object
A_i = area of the object i (m²)
P_i = perimeter of the object i (m)
A_T = total area of the super-object (m² | km²)
P_T = perimeter of the super-object (m | km)
L_i = length of the road i (m)
nr = number of roads (segments) in the super-object.
n = number of objects in a class or super-object

m = number of classes in the super-object
ns = number of super-objects in the study area
A_{out} = area of isolated objects in a class (m²)
A_h = area of the hole within the class (m²).
nh = number of holes in the class.
L_{max} = length of the larges side from the boundary box covering the largest object from a class (m)
A_{max} = area of the largest object in the class or largest urban core

(m²).

$A_{2nd\ max}$ = area of the second largest object in the class or second largest urban core (m²).

x_i, y_i = coordinates of the centroid of an object (m)

$\bar{x}$, $\bar{y}$ = mean of the coordinates of the centroid of the objects of a class (m)

SD_{max} = maximum Standard distance, computed through a grid covering the same area as the super-object.

O_c = obstruction coefficient, where: (1) industrial, commercial, road, airports; barren land, (0.8) continuous urban; (0.6) discontinuous urban and; (0.4) green areas; sport and leisure facilities.

D_{ij} = distance between an object i and its nearest object j (from boundary to boundary) of the same class (m)

A_U = total area of urban objects (m²)

P_U = total perimeter of urban objects (m)

NOb_{max} = maximum number of possible objects in the class. When data come from vector format, it equals the area of the class divided by the minimum mapping unit, if data come from raster, it is the total number of pixels of the class.

r = radius from the center point to the circle (m)

$A_c(r)$ = area of a class inside the circle with radius r (m²)

C = constant

P = proportion of the super-object occupied by the class i, (A_i/A_T) (%)

R_v = number of classes in the total study area (having into account all the SO)

$l_{i,j}$ = length of the shared edges between two objects from different classes i,j (m). At class or super-object level equals the sum of the total lengths (m)

A_{t1} = area of a class in the first date (m²)

A_{t2} = area of a class in the second date (m²)

w = new object

l_w = length of the edge between a new object and an old object (m)

P_w = perimeter of a new object (m)

A_w = area of a new object (m²)

A_{Tw} = total area of the new objects from a class (m²)

n_w = number of new objects

D_{wc} = distance between a new object and the nearest old object from the same class (m).

E = exploitation degree of the class, where: (1) barren/unused land; (2) forest, water, grass; (3) agricultural land; and (4) urban, mine, roads, etc.

(D_w^R+1) = distance between a new object w from a class to the closest road. The unit is summed to avoid that objects adjacent to a road with a distance of zero create an undefined function.

D_w^{old} = distance between a new object and the closest old object from the same class (m)

x_{t1}, y_{t1} = coordinates of the city center in the first date

x_{t2}, y_{t2} = coordinates of the city center in the second date

Appendix B

Ancillary maps, figures and graphs

This appendix presents additional maps, figures and graphs supporting and complementing the findings of chapters 3 to 6.

B.1. Spatio-temporal maps for the analysis at intra-urban level
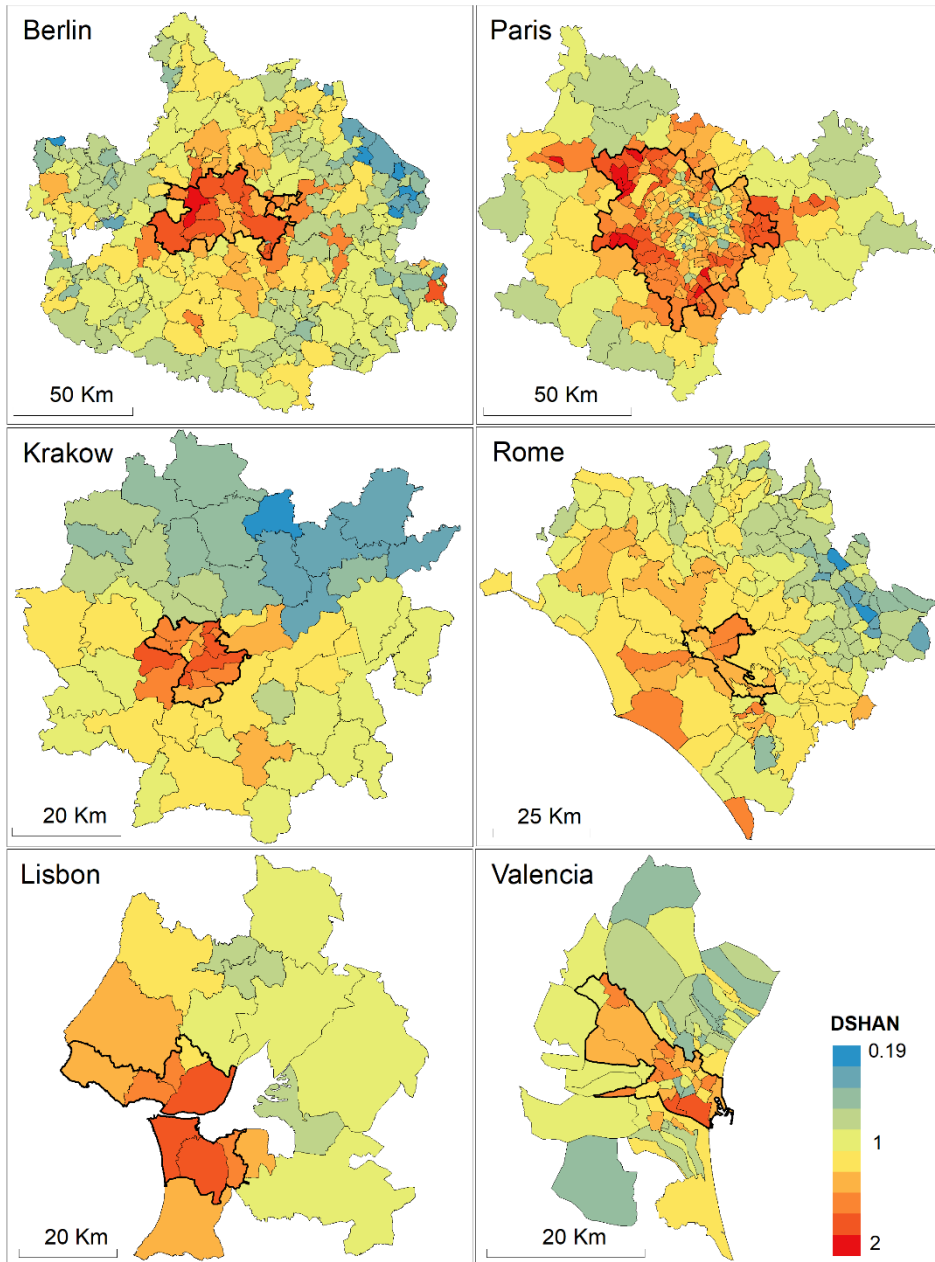
Figure B.1. Shannon diversity (DSHAN) in 2012 for Berlin, Paris, Krakow, Rome, Lisbon and Valencia at the administrative level.
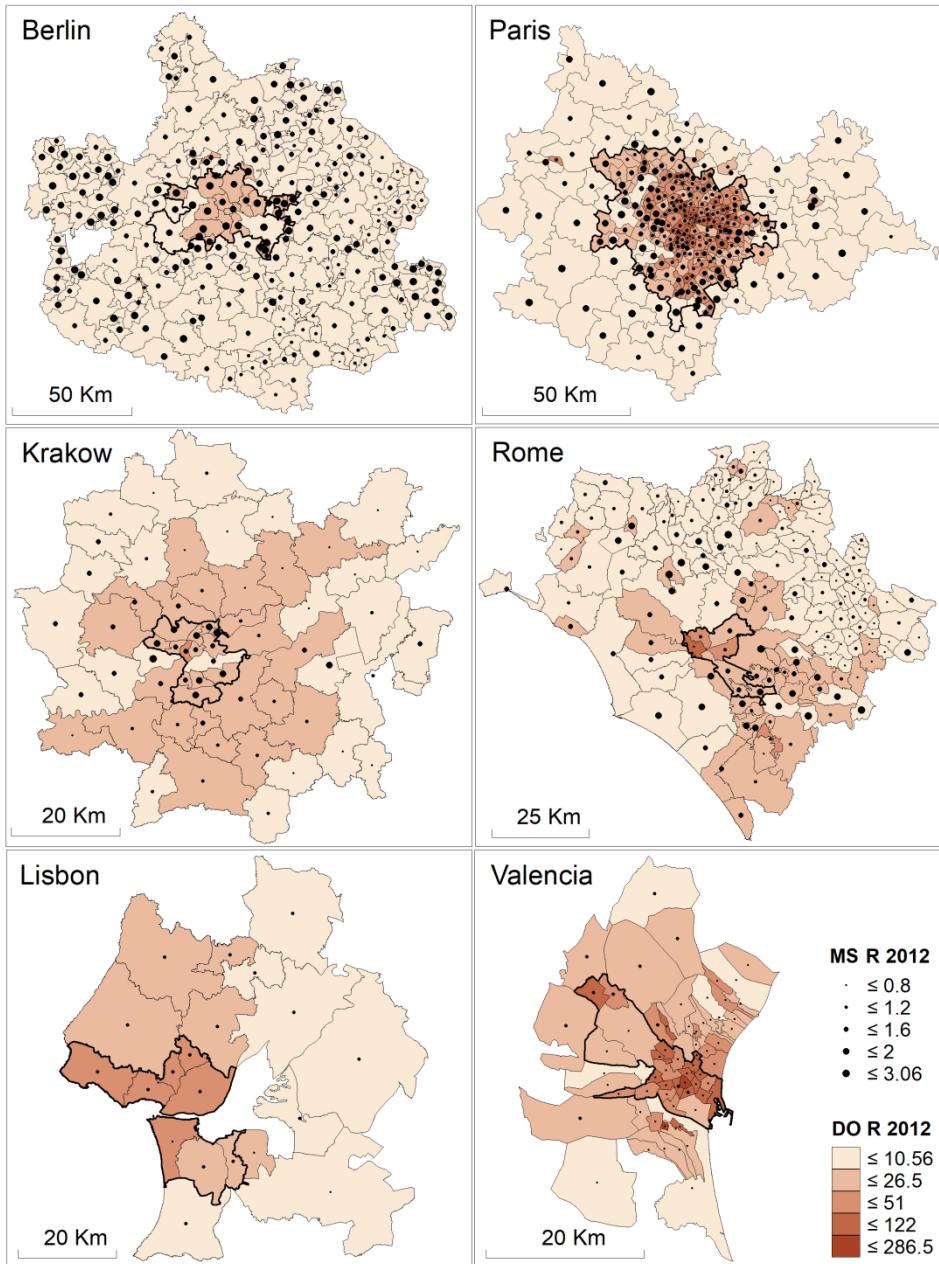
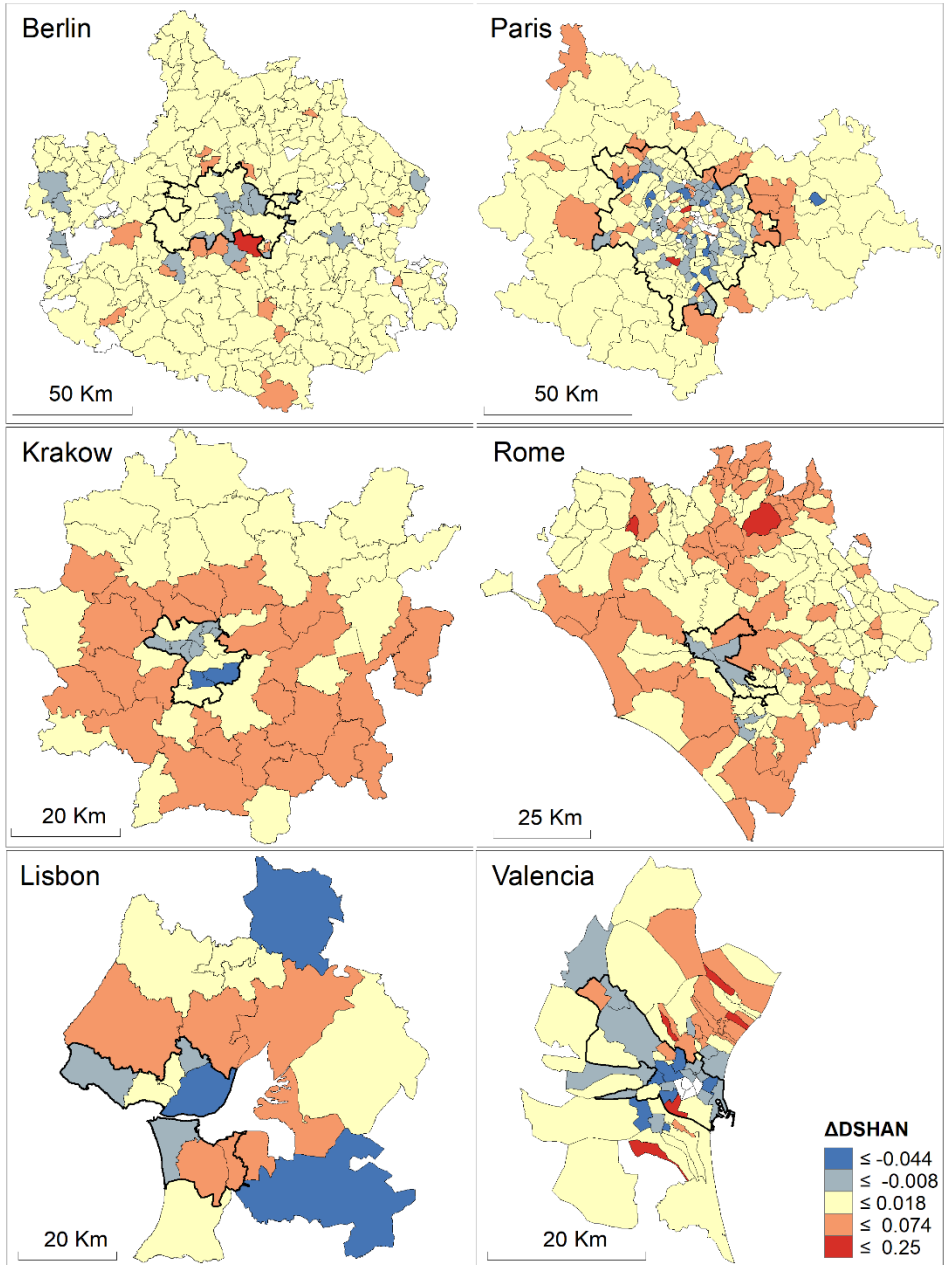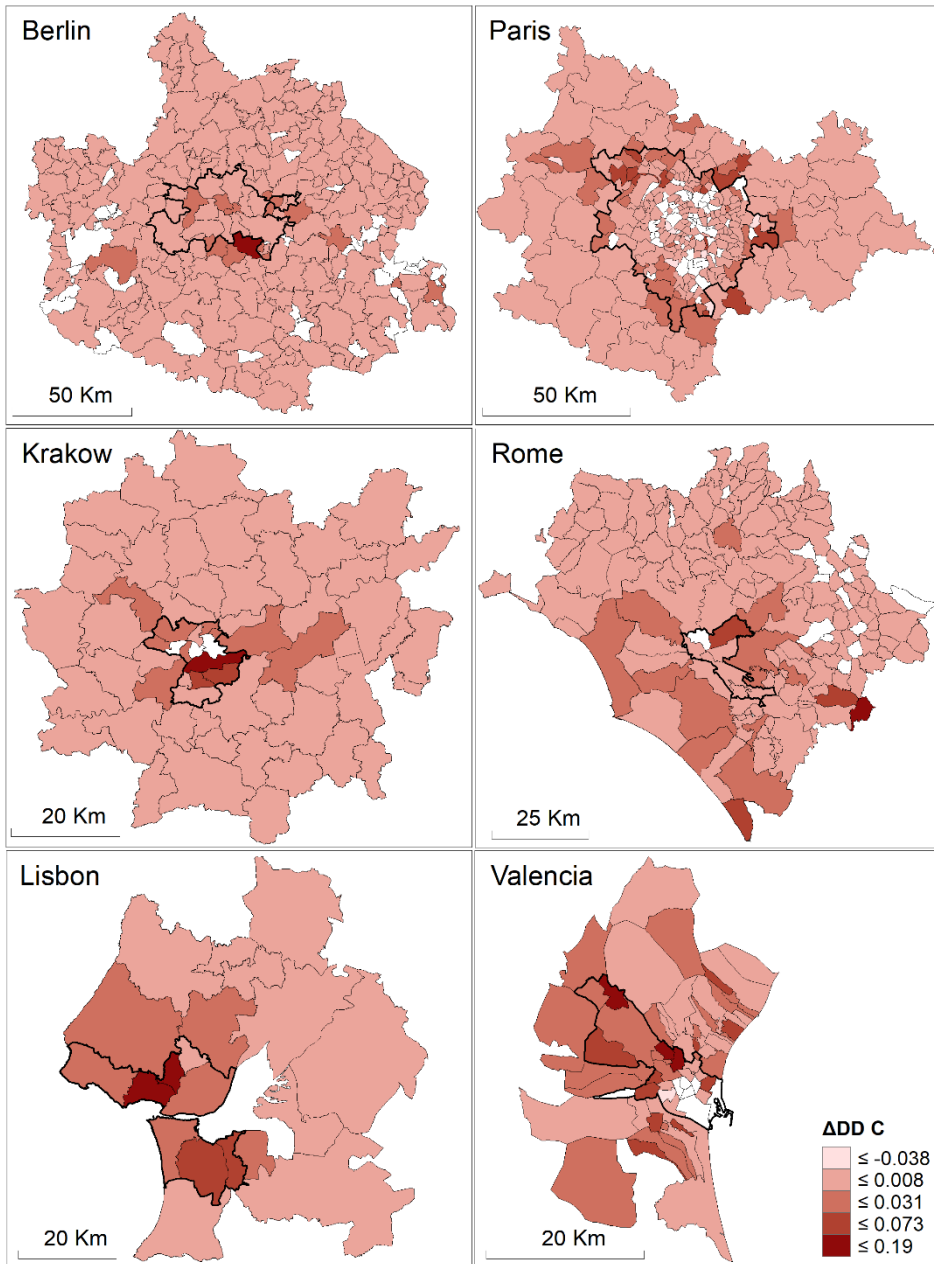Figure B.2. Mean object size (MS) and object density (DO) of residential class in 2012 for Berlin, Paris, Krakow, Rome, Lisbon and Valencia at the administrative level.
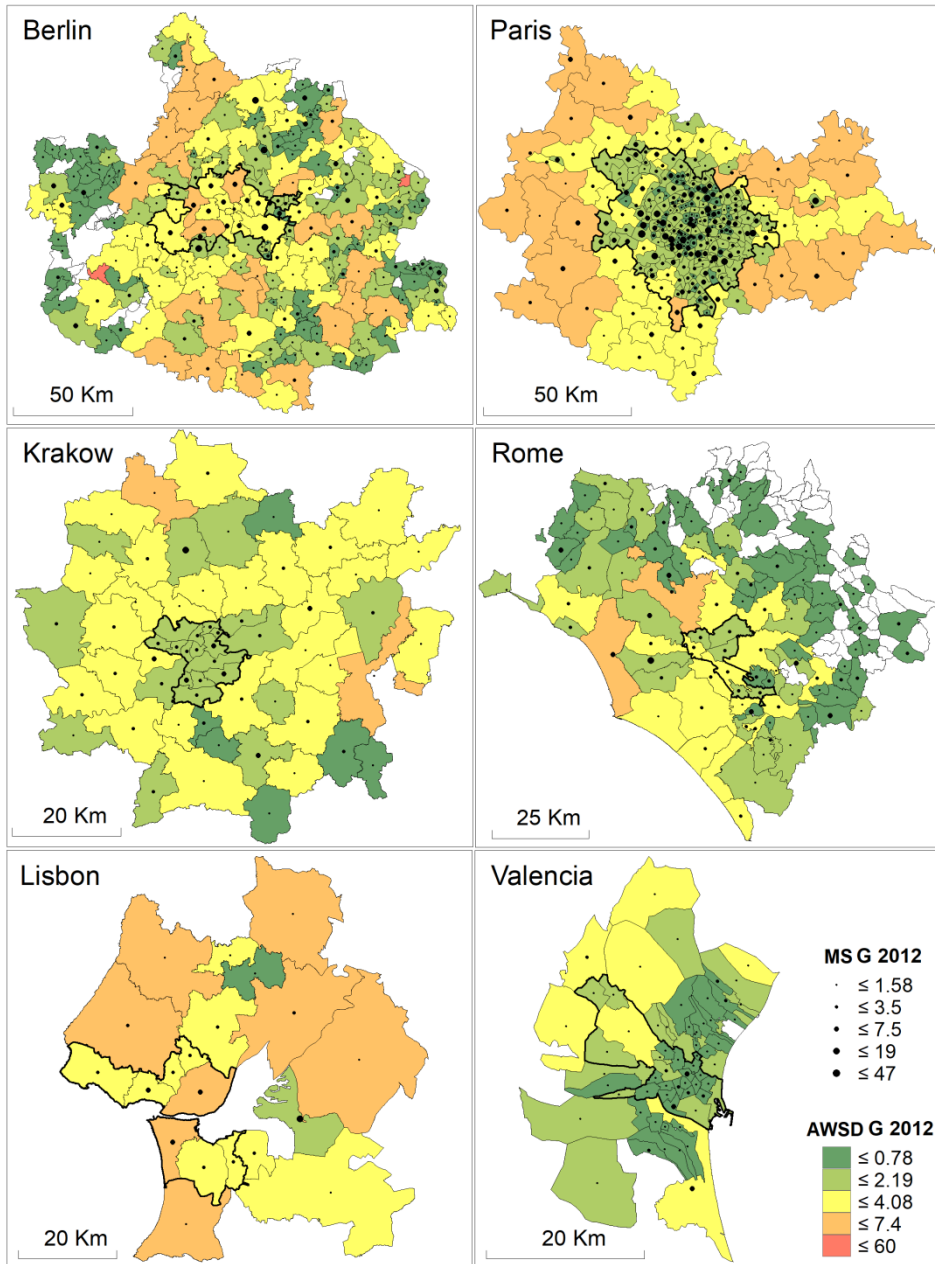
Figure B.3. Mean object size (MS) and area-weighted standard distance (AWSD) of green urban areas (G) class in 2012 for Berlin, Paris, Krakow, Rome, Lisbon and Valencia at the administrative level.

Figure B.4. Change of Shannon diversity (ΔDSHAN) between 2006 and 2012 for Berlin, Paris, Krakow, Rome, Lisbon and Valencia at the administrative level.

Figure B.5. Change of density-diversity (ΔDD) of commercial and industrial (C) class between 2006 and 2012 for Berlin, Paris, Krakow, Rome, Lisbon and Valencia at the administrative level.
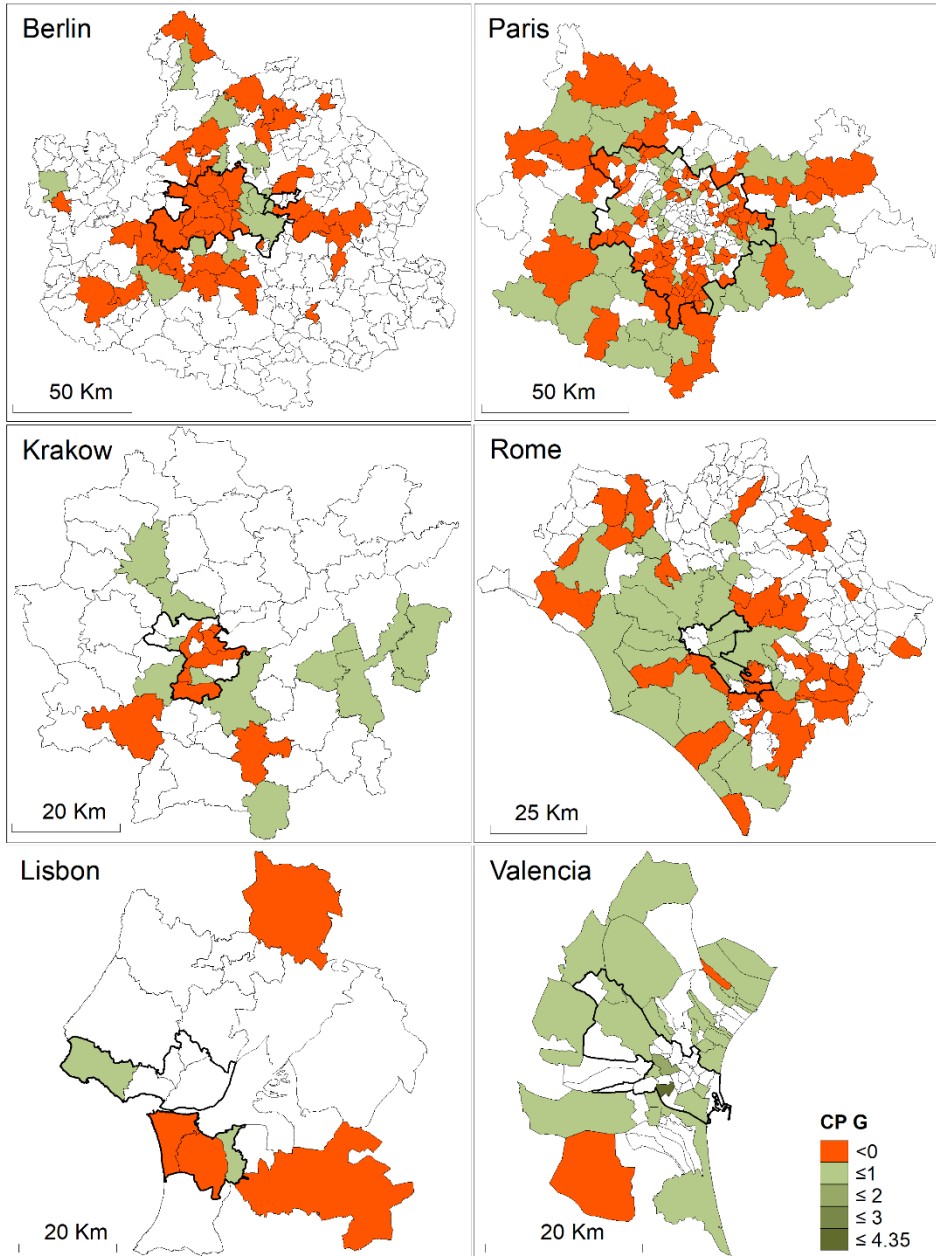
Figure B.6. Change ratio (CP) of green urban areas (G) class between 2006 and 2012 for Berlin, Paris, Krakow, Rome, Lisbon and Valencia at the administrative level.
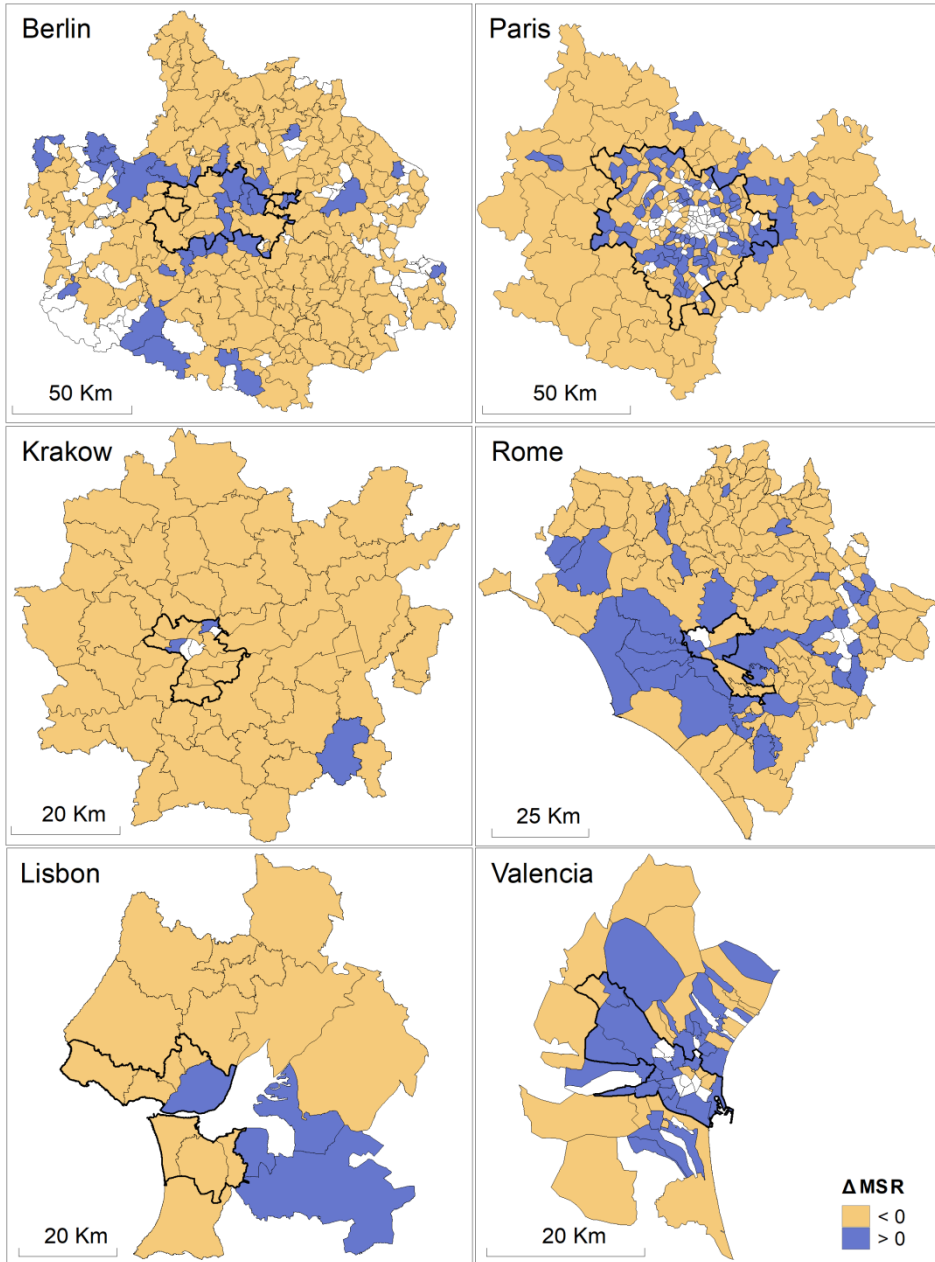
Figure B.7. Change of mean patch size (ΔMS) of residential (R) class between 2006 and 2012 for Berlin, Paris, Krakow, Rome, Lisbon and Valencia at the administrative level.
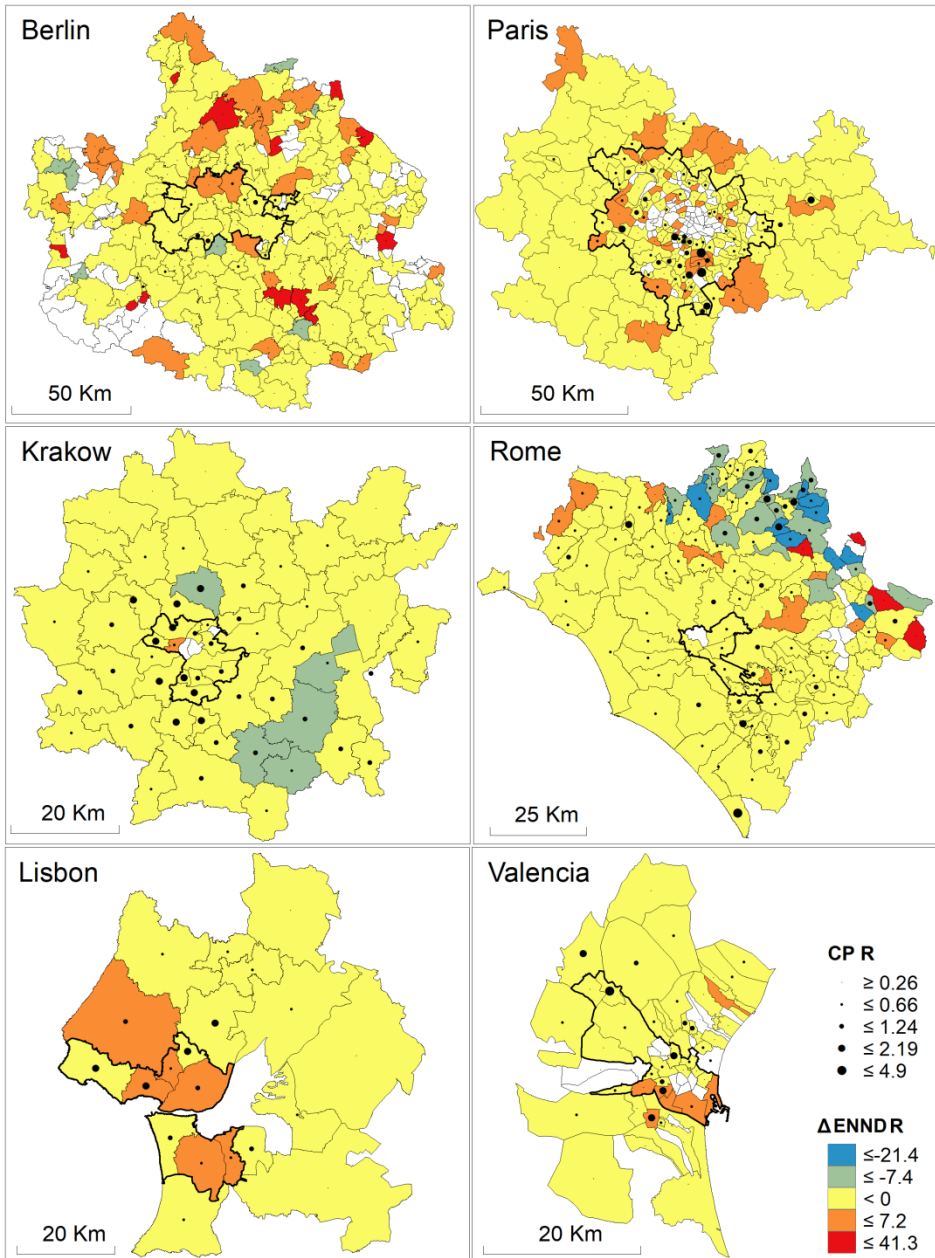
Figure B.8. Change ratio (CP) and change of mean Euclidean distance of the nearest neighborhood (ΔENND) of residential (R) class between 2006 and 2012 for Berlin, Paris, Krakow, Rome, Lisbon and Valencia at the administrative level.
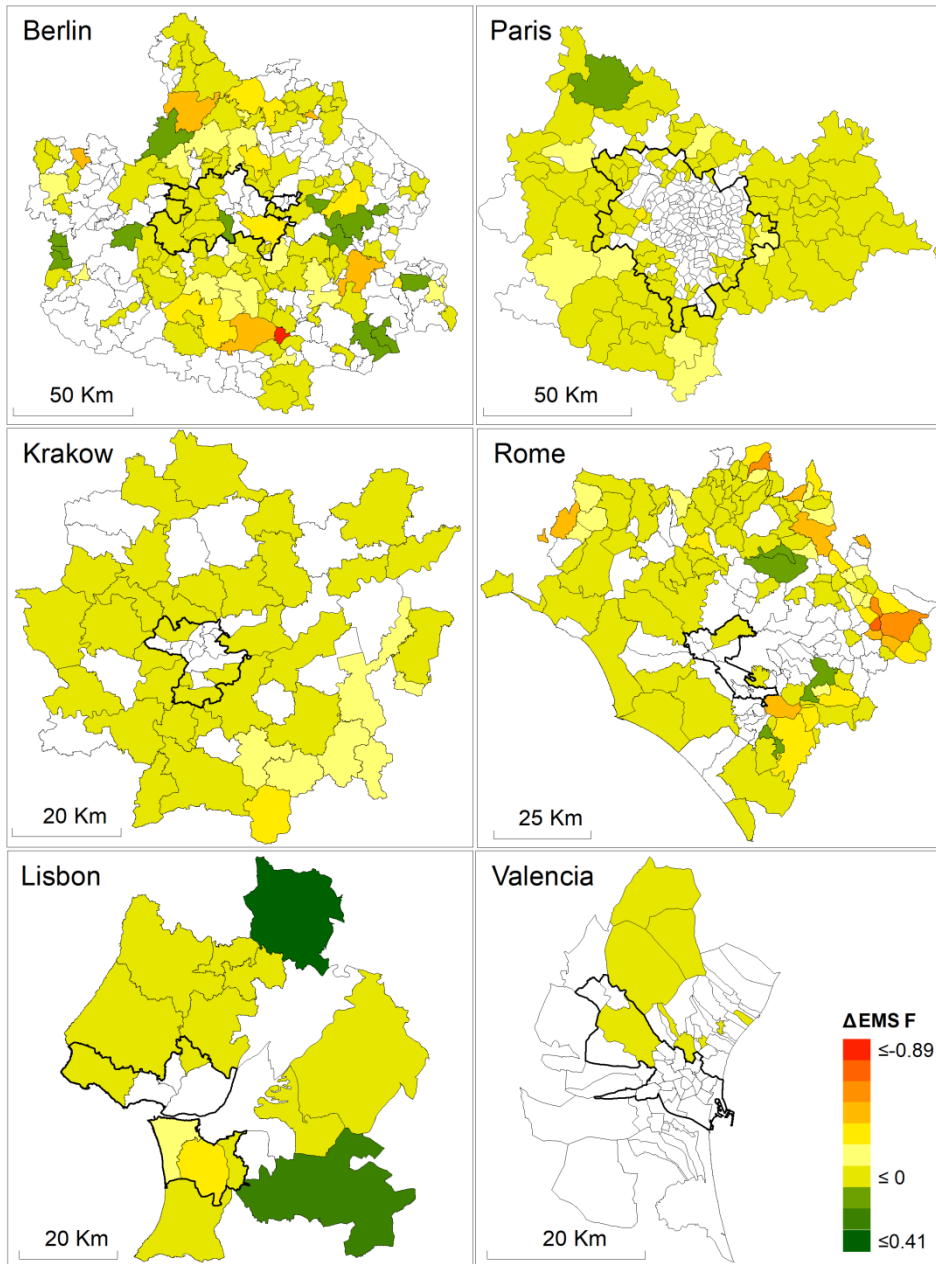
Figure B.9. Change of effective mesh size (EMS) of forest (F) class between 2006 and 2012 for Berlin, Paris, Krakow, Rome, Lisbon and Valencia at the administrative level.

B.2. Population and urban growth imbalance (PUGI) scatterplots
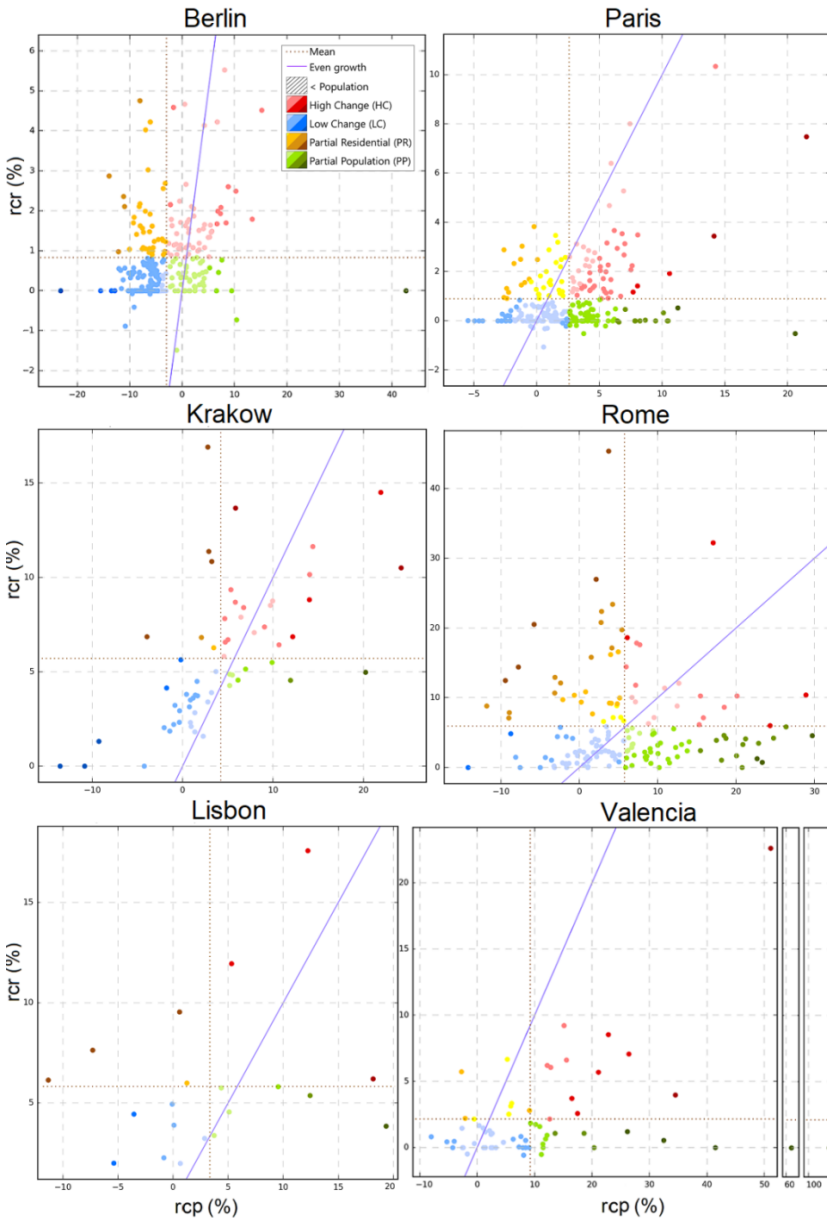
Figure B.10. Population and urban growth imbalance (PUGI) scatterplots between 2006 and 2012. The administrative units are classified in change types by color, the hue depends on the distance to the even growth line (PUGI), where darker color mean more inequality between residential and population growths.

B.3.   Simulated scenarios of urban growth spatial patterns
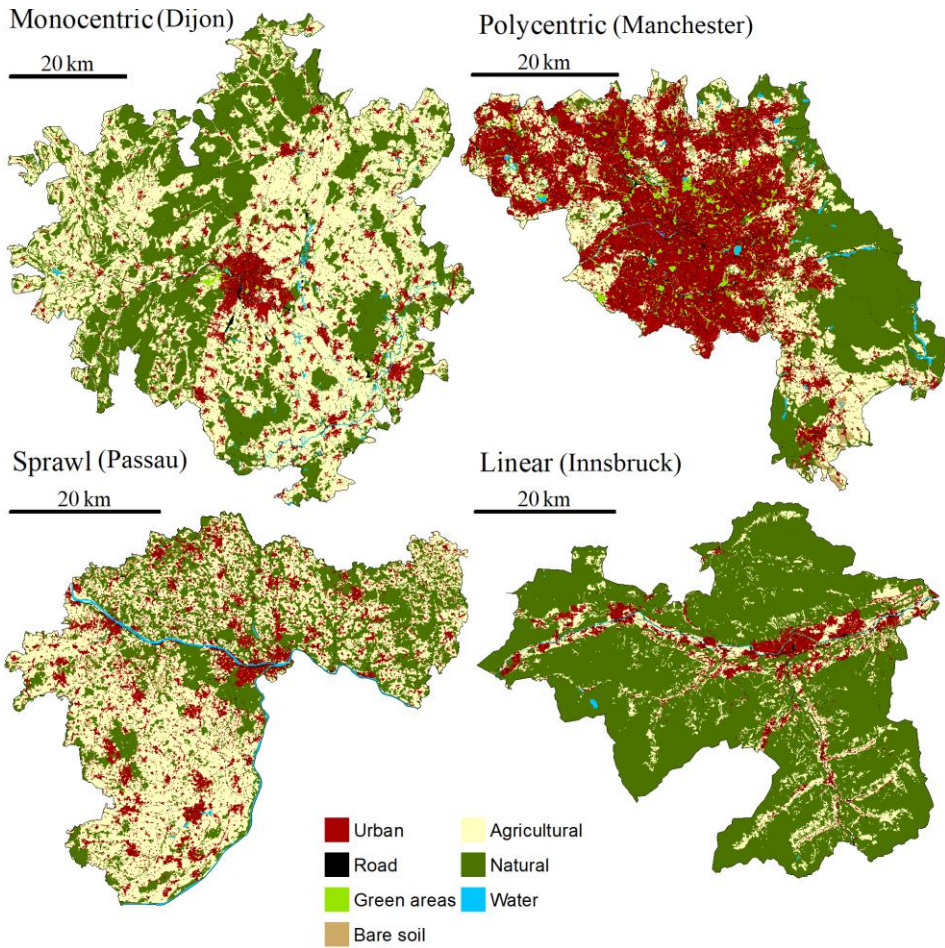
Figure B.11. Scenarios of expansive urban growth from monocentric, polycentric, sprawl and lineal urban forms.
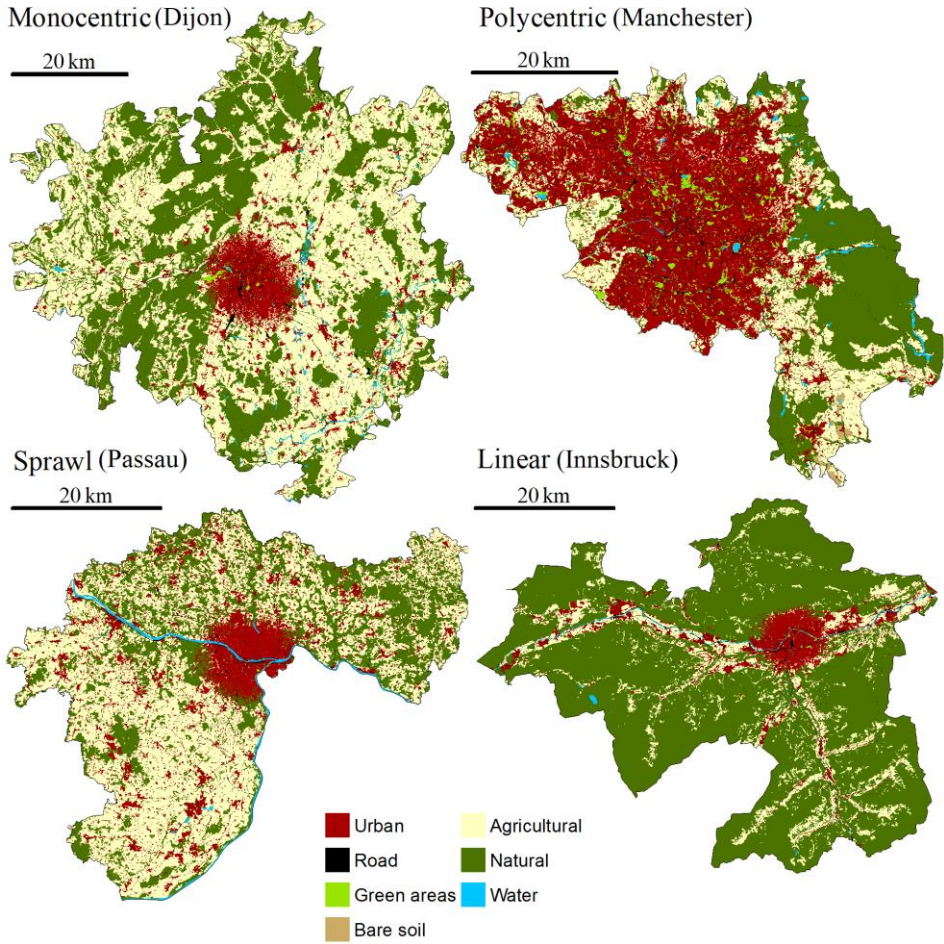
Figure B.12. Scenarios of compact urban growth from monocentric, polycentric, sprawl and lineal urban forms.
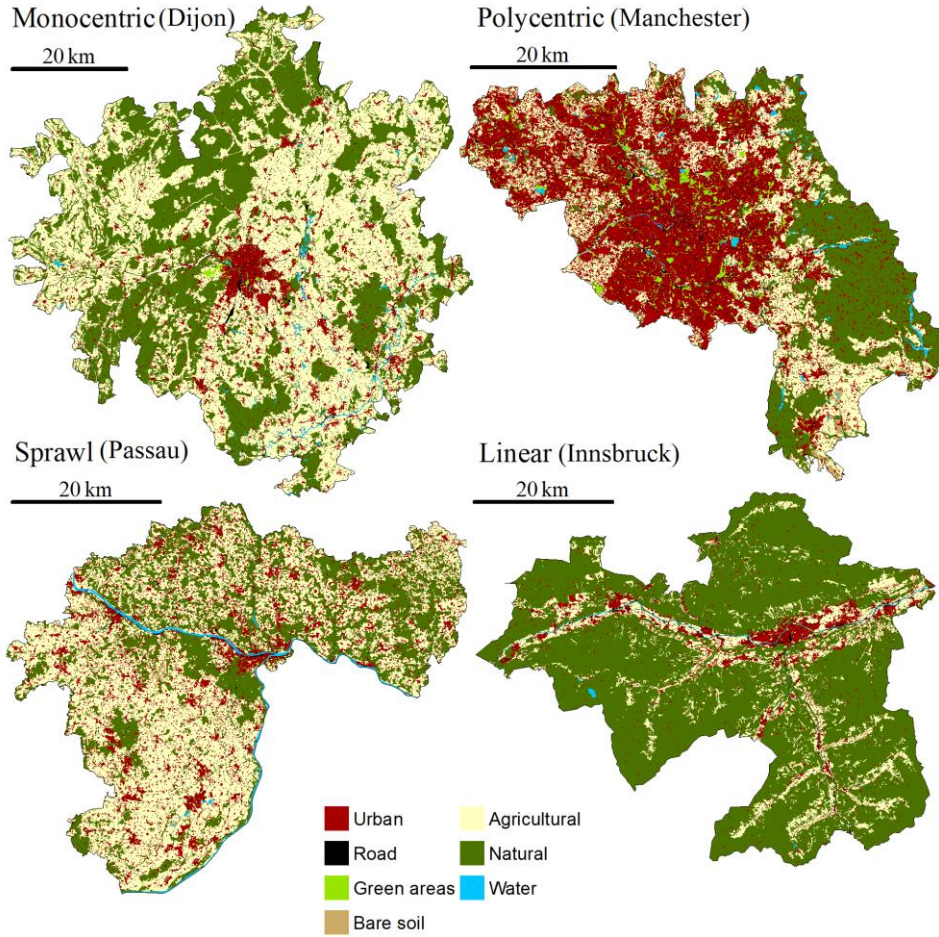
Figure B.13. Scenarios of dispersed urban growth from monocentric, polycentric, sprawl and lineal urban forms.
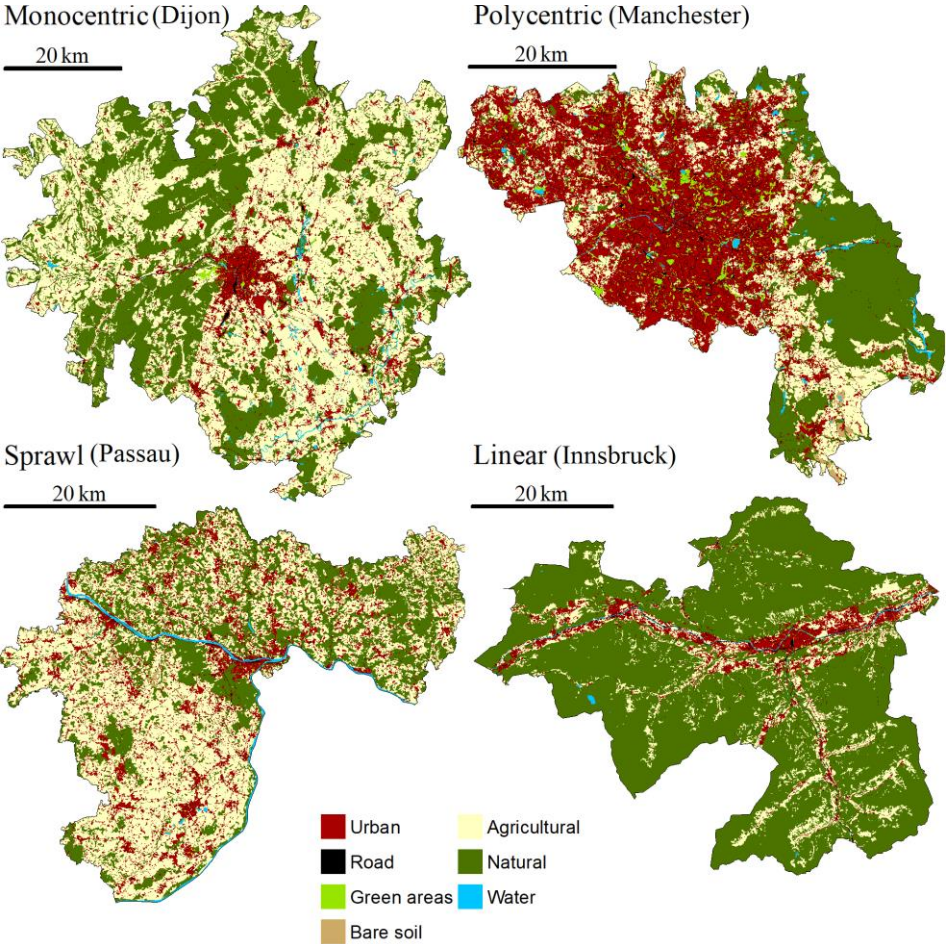
Figure B.14. Scenarios of road-based urban growth from monocentric, polycentric, sprawl and lineal urban forms.
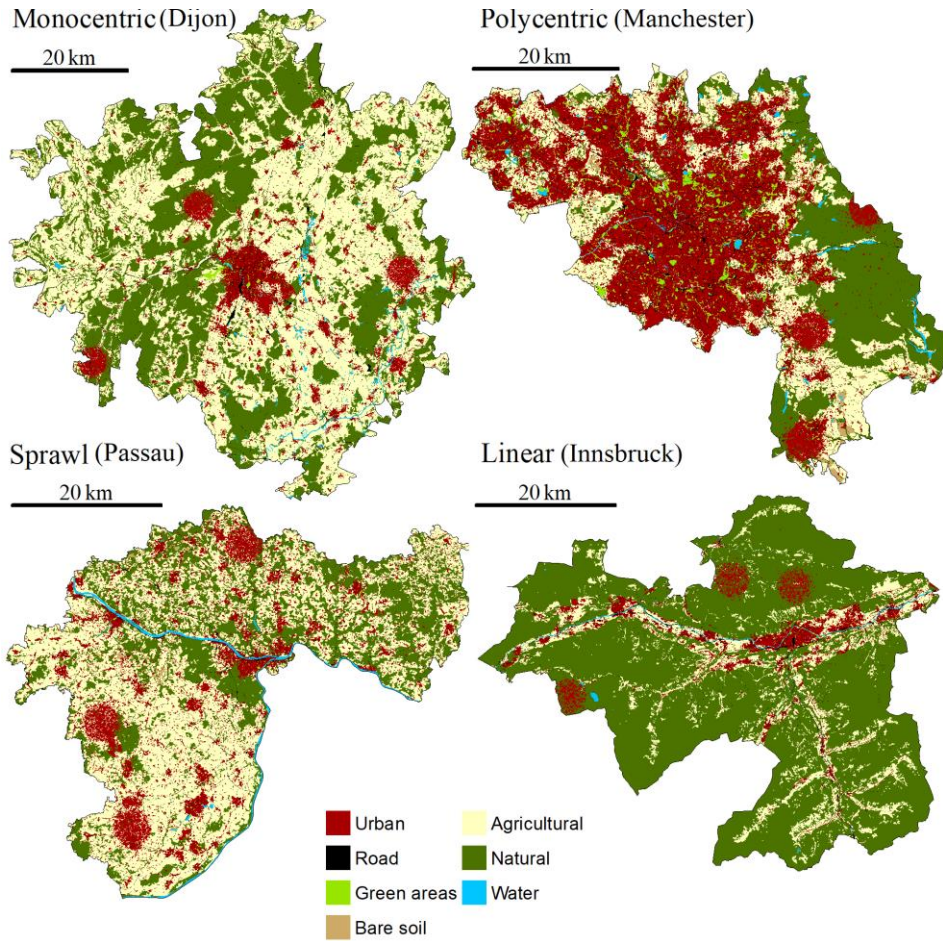
Figure B.15. Scenarios of leapfrog urban growth from monocentric, polycentric, sprawl and lineal urban forms.

B.4. Boxplots of relevant spatio-temporal metrics sorted according to their importance per socio-economic variable.
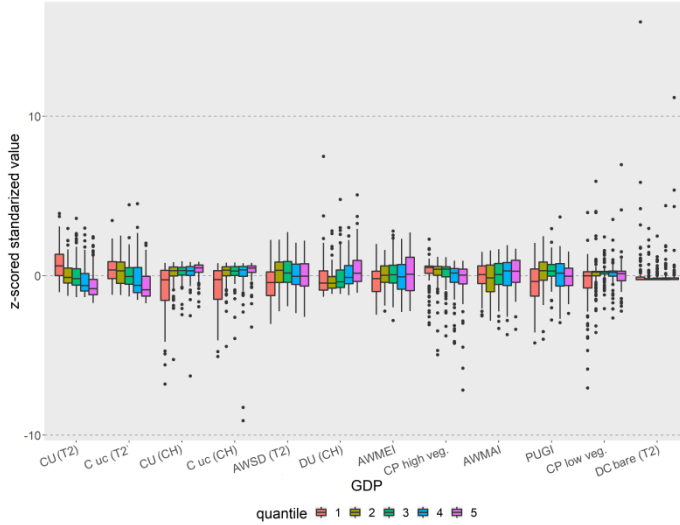
Figure B.16. GDP boxplots of relevant spatio-temporal metrics sorted according to their importance. GDP is divided into five quantiles (Q1 to Q5, from low to high values) and the standardized values of metrics are shown for each quantile.
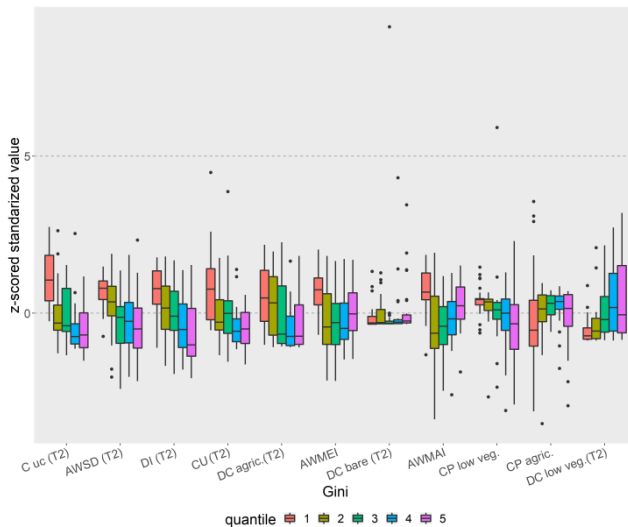
Figure B.17. Gini boxplots of relevant spatio-temporal metrics sorted according to their importance. Gini is divided into five quantiles (Q1 to Q5, from low to high values) and the standardized values of metrics are shown for each quantile.
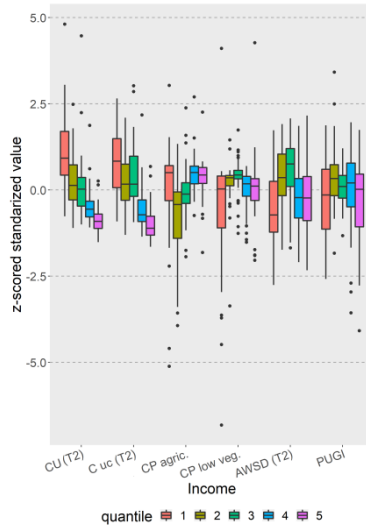
Figure B.18. Income boxplots of relevant spatio-temporal metrics sorted according to their importance. Income is divided into five quantiles (Q1 to Q5, from low to high values) and the standardized values of metrics are shown for each quantile.
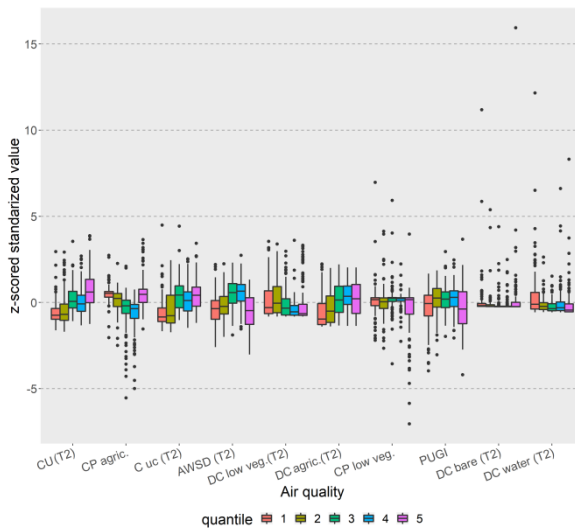

Figure B.19. Air quality boxplots of relevant spatio-temporal metrics sorted according to their importance. Air quality is divided into five quantiles (Q1 to Q5, from low to high values) and the standardized values of metrics are shown for each quantile. Air quality measures the fine particulate matter (higher values mean more pollutants and lower air quality)
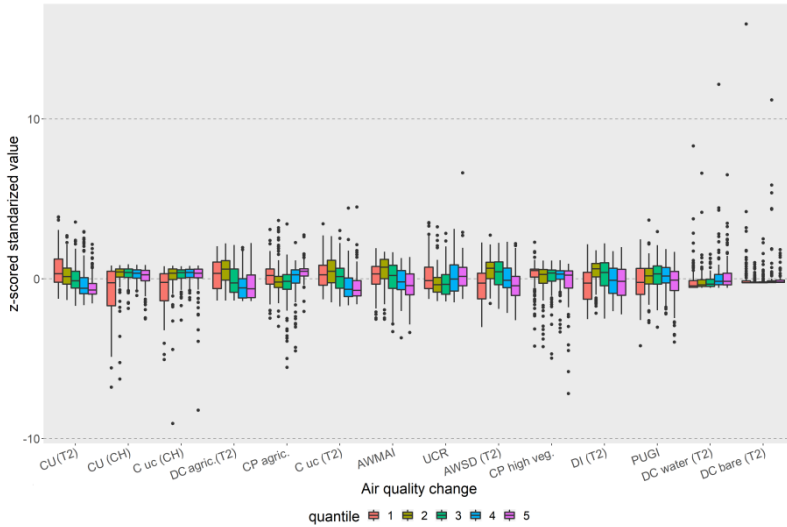
Figure B.20. Air quality change boxplots of relevant spatio-temporal metrics sorted according to their importance. Air quality change is divided into five quantiles (Q1 to Q5, from low to high values) and the standardized values of metrics are shown for each quantile. Air quality measures the fine particulate matter (higher values mean more pollutants and lower air quality).
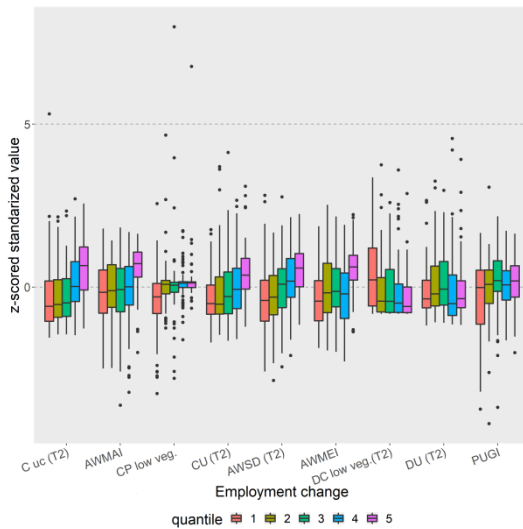

Figure B.21. Employment change boxplots of relevant spatio-temporal metrics sorted according to their importance. Employment change is divided into five quantiles (Q1 to Q5, from low to high values) and the standardized values of metrics are shown for each quantile.