Author's Original Manuscript – preprint

The Version of Record of this manuscript has been published and is available in **International Journal of Geographical Information Science**. http://www.tandfonline.com/.
Doi: 10.1080/13658816.2020.1817463

Identifying urban growth patterns through land-use/land-cover spatio-temporal metrics: Simulation and analysis

Marta Sapena [a,b],*, Luis A. Ruiz [a]

[a] Geo-Environmental Cartography and Remote Sensing Group, Department of Cartographic Engineering, Geodesy and Photogrammetry, Universitat Politècnica de València, Camí de Vera s/n, 46022, Valencia, Spain
[b] German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, 82234 Wessling, Germany

*corresponding author
Sapena M: marta.sapena-moll@dlr.de ORCID: https://orcid.org/0000-0003-3283-319X
Ruiz LA: laruiz@cgf.upv.es ORCID: https://orcid.org/0000-0003-0073-7259

ABSTRACT

The spatial pattern of urban growth determines how the physical, socio-economic and environmental characteristics of urban areas change over the time. Monitoring urban areas for the early identification of spatial patterns facilitates assuring their proper development and counteracting unsustainable trends. In this paper, we assess the use of spatio-temporal metrics from land-use/land-cover maps to identify growth patterns by means of GIS techniques. We applied land use change models to simulate 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). Then, we computed the spatio-temporal metrics for the simulated scenarios, selected the most relevant by applying discriminant analysis and classified the growth patterns using clustering methods. Two metrics, Weighted mean expansion and Weighted Euclidean distance, 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 for monitoring and evaluating the management of developing urban areas.

Keywords: spatio-temporal metrics; urban form; urban simulation; land use change model; growth pattern;

Introduction

The sustainability of developing and developed urban areas is an ongoing concern worldwide. The United Nations defined seventeen goals for ensuring a sustainable future for people and the planet, and the goal eleven focuses on the environmental, social and economic sustainability of cities. The knowledge of how urban areas are spatially configured and their variations is essential to successfully monitor the urbanization impacts on the environment and their socio-economic effects (Schneider and Woodcock 2008, Siedentop and Fina 2010, Reis *et al.* 2016, Salvati *et al.* 2016). In this regard, remote sensing and geographic information system (GIS) techniques are valuable assets to conduct such studies (Liu *et al.* 2010, Ju *et al.* 2016, Zhao *et al.* 2016, Abrantes *et al.* 2019). Remote sensing provides an important source of geographic information for urban studies, while GIS allows for its spatial analysis. Their combined use has proven to be efficient in analysing urban form (Song *et al.* 2017), monitoring urban dynamics (Wu *et al.* 2016), and modelling land use change (Liu *et al.* 2014), among other urban applications.

On the one hand, scholars have relied on land use change (LUC) models to propose and validate methodologies that aim to reverse unsustainable trends in cities (Musa *et al.* 2017). These LUC 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). Thus, Van de Voorde *et al*. (2016) and Ustaoglu *et al.* (2018) simulated alternative scenarios under different planning strategies to foresee their implications and serve as a tool for planning cities accordingly; Dorning *et al.* (2015) and Sun *et al.* (2018) simulated different development scenarios under various planning strategies to assess the effectiveness of regional natural resources conservation plans and to explore optimal strategies for improving ecosystem services; and Hoymann and Geotzke (2016) evaluated the effect of policy measures to mitigate climate change and developed new strategies based on simulated urban development scenarios. The use of simulation strategies is mainly due to the scarce availability of long time-series and high-resolution land use databases, which are the basis to monitor urban development and to evaluate growth patterns. LUC 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).

The form of the urban environment affects the population in many aspects. The influence it has in transport systems, commuting choices (Song *et al*. 2017), energy consumption (Chen *et al*. 2011), air quality, and health (Hankey and Marshall 2017), among many others, has been demonstrated so far. Moreover, not only there is a wide diversity of urban shapes and sizes, but also their spatial development is manifold, conditioned by the history of the territorial development, shape, topography, geography, economic and social development, land use policies, etc. (Schneider and Woodcock 2008, European Union 2016). Different scholars have evidenced relationships between urban form and urban development with their sustainability, as collected in Williams *et al.* (2000).

There is a growing interest in developing methods and indicators able to detect growth trends, which will be a source of information for planners and policy makers. Urban growth has been characterized using a diversity of GIS methods. Reis *et al.* (2016) 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 the 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, thereby using these values to describe growth patterns as a diffusion or coalescence growth process. Other studies used a straightforward index to quantify the adjacencies between urban and newly urban patches, classifying them into infill, edge-expansion and outlying growth types (Liu *et al*. 2010, Shi *et al*. 2012), which may serve as a basis for more complex pattern classification. Jiao (2015) proposed different indicators to characterize urban growth in nearly thirty Chinese cities, measuring the urban land density decline, the urban compactness, expansion rate and degree of sprawl. Two recent studies proposed new methods for characterizing urban expansion. 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). Spatial and spatio-temporal metrics were used in nearly all of the cited studies. Even if it is common to find redundant information when working with large set of metrics (Chen *et al*. 2011, Sapena and Ruiz, 2019), they seem to be successful in quantifying the growth and determining its type. However, these studies focused on various degrees of compact-sprawl growth (e.g.: Tian *et al*. (2011) and Jiao (2015) classified urban development processes as compact, sprawl or intermediate phase), instead of a more detailed classification of growth types. Since the consequences of the urban growth differ according to their pathways (Williams *et al*. 2000, Bhatta 2010, European Union 2016), the identification of different types of growth patterns will allow for more complex analyses of growth trends and the assessment of their consequences, which will eventually improve the understanding of their interrelationships.

Monitoring and characterizing urban growth spatial patterns will contribute to a better understanding of past and present trends, allowing planners to make informed and better decisions for the future in order to minimize social, economic and environmental impacts of urban development. The purpose of this study is to identify an efficient subset of spatio-temporal metrics for the identification of different urban growth spatial patterns, to evaluate them in a diversity of baseline urban forms, and to assess the influence of the initial urban form in the identification of such patterns.

Materials and methods

Figure 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 1.1), and apply a land use change model for simulating five urban growth patterns from the baseline forms (Figure 1.2 and 1.3), this provides a wider range of possible scenarios to evaluate the metrics. Afterwards, the extraction and selection of spatio-temporal metrics for every simulated scenario are described (Figure 1.4 and 1.5A). Following, growth patterns are classified using the spatio-temporal metrics (Figure 1.5B), and the results are interpreted and described, including the influence of the initial urban forms in identifying growth classes (Figure 1.5C).

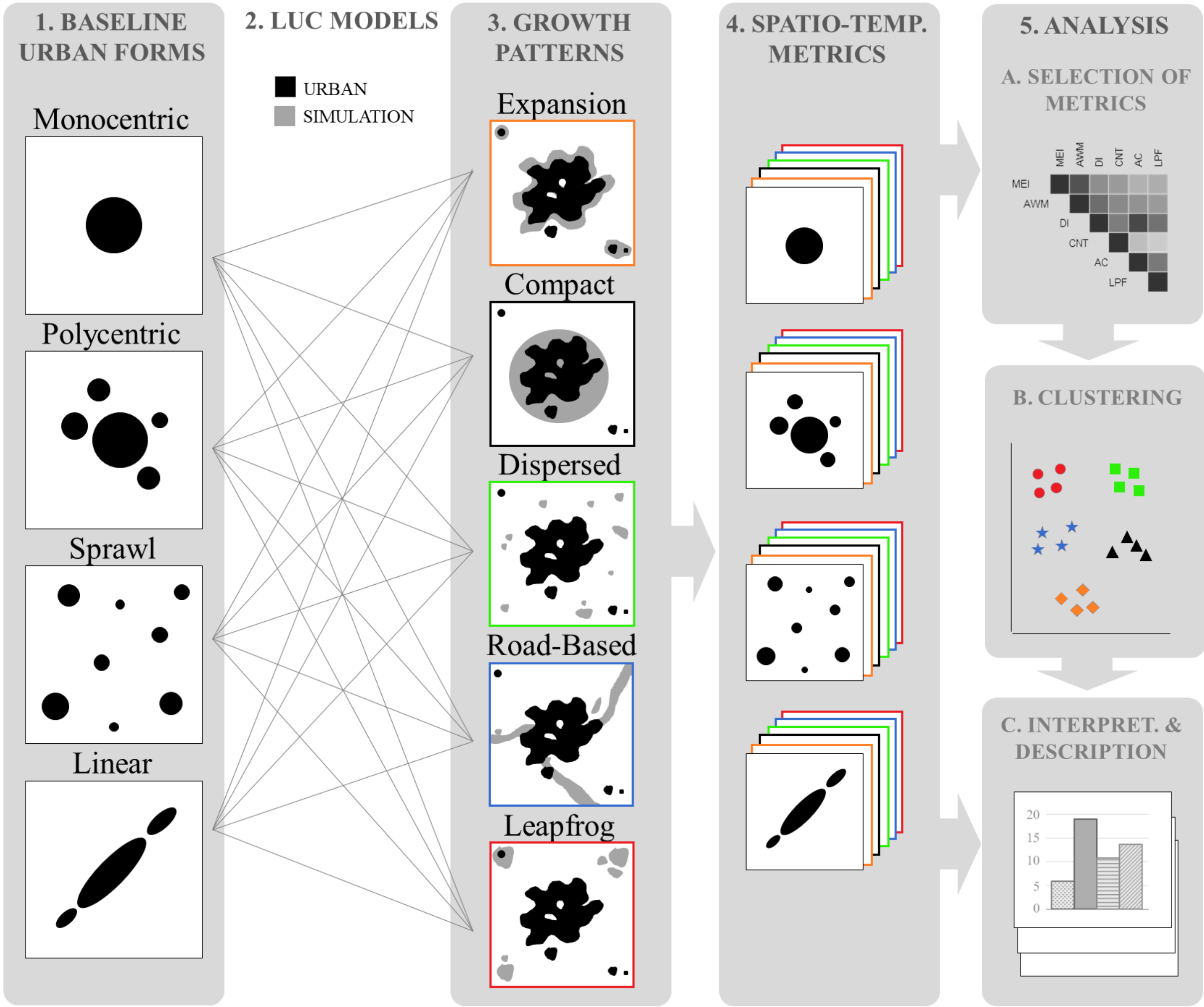

Figure 1. Workflow of the methodology: (1) Definition and selection of four initial urban areas having four different urban forms. (2) Application of the land use change (LUC) model for the simulation of (3) five urban growth spatial patterns. (4) Computation of spatio-temporal metrics for the twenty pairs of baseline-growth simulated scenarios. (5A) Selection of meaningful subset of metrics, (5B) classification of growth patterns using the metrics, and (5C) interpretation of results.

Definition of urban forms and growth spatial patterns

Urban 97form refers to the spatial configuration of the physical built environment and human activities (Georg *et al*. 2016, Abrantes *et al*. 2019). In this paper, we consider the urban form as the static physical configuration of the urban cover. We define four theoretical spatial types of urban forms extracted from 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 1). The urban growth spatial pattern is a dynamic process of urban development that, in some cases, modifies the initial urban form. The spatial patterns of urban growth are manifold and have been described using different nomenclature. We summarized 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) in five types: compact, dispersed, expansion, leapfrog and road-based (Table 1). It must be considered that both, form and growth pattern defined, are pure theoretical types and they are often combined in real urban areas.

Table 1. Name and description of urban forms and growth spatial patterns that are combined by means of a LUC 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 centre increases. Consists of a dominant city and several dependant 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 subcentres 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 monofunctional 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 centre.	(Marshall 2005, Georg *et al*. 2016, Nabielek *et al.* 2016)
Urban growth spatial pattern	**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 development.	(Camagni *et al.* 2002, Wilson *et al*. 2003, Marshall 2005, Schneider and Woodcock 2008)
	Dispersed	When low-density urban development 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)
	Expansion	It increases the built-up area from the boundaries of the urbanized area, fostering a greater extension of the urban layout. 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)
	Leapfrog	When secondary new centres 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,	(Camagni *et al*. 2002, Chin 2002, Wilson *et al*. 2003, Marshall 2005, Salvati *et a*l. 2016)

	compact and dense development.	
Road-based	The urban development 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)

Data

Four functional urban areas were selected as working data, comprising the cities and their commuting zones. 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 2016). 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 2): (a) Dijon, France, as an example of monocentric agglomeration, according to Baumont *et al.* (2014). (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 exurban 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 (Krajiver 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 development scenarios.

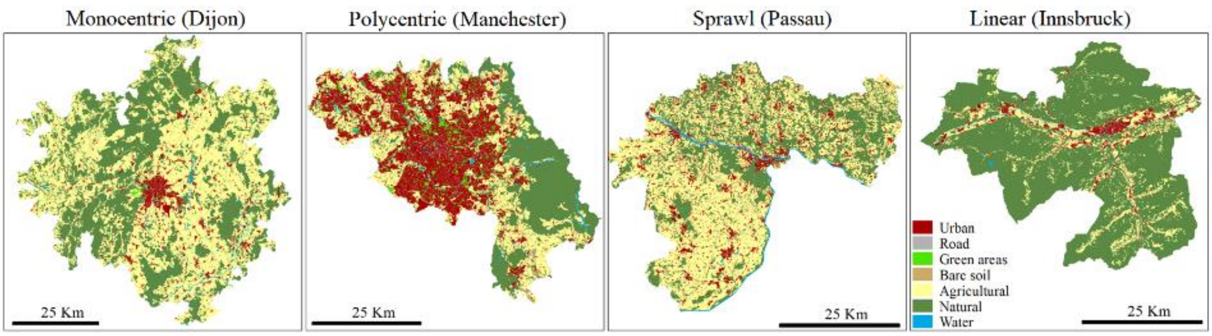

Figure 2. The four urban areas representing different baseline urban forms. Source: Urban Atlas 2012 (EEA, 2016), with an aggregated legend.

For simulating development scenarios we used LULC data from the Urban Atlas dataset for the year 2012. The Urban Atlas is a two-date, detailed and harmonised LULC dataset in vector format (scale 1:10,000) for large European Functional Urban Areas, built in the context of the Copernicus European Earth Observation programme (http://land.copernicus.eu/local/urban-atlas). 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 centres from Eurostat (https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data) were also used.

Land use change model

Even though image classification techniques in remote sensing are continuously improving, we still lack high-resolution long-time series of LULC datasets. A promising initiative in Europe is the Urban Atlas dataset, which provides high-resolution LULC data covering more than 300 urban areas for 2006 and 2012. However, this period is still insufficient for detecting reliable growth trends. Therefore, we created longer LULC time-series using the LUC model FUTURES (FUTure Urban-Regional Environment Simulation model). FUTURES was suitable to simulate long-term urban growth spatial patterns from different baseline forms, creating alternative synthetic growth scenarios by altering a few factors (Dorning *et al*. 2015).

LUC model and factors of urban growth

The model FUTURES is a cellular automata, stochastic and patch-based LUC model based on the logistic regression method, and was implemented in GRASS GIS (Meentemeyer *et al.* 2013, GRASS development team, 2017). It requires an urban mask and geographic, economic, and social factors that determine where the development is likely to occur. We used FUTURES because it allows for 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.

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 development (Figure 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 centre, 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 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 (Meentemeyer *et al*. 2013): 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 meters, based on the concept of Urban Morphological Zones defined by Goerlich and Cantarino (2013). Since topographic conditions may limit or ease urban development, 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 development of 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 development 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 centres (leapfrog growth). All the aforementioned factors were scaled to a range from zero to one to avoid the influence of the measurement units (Figure 3).

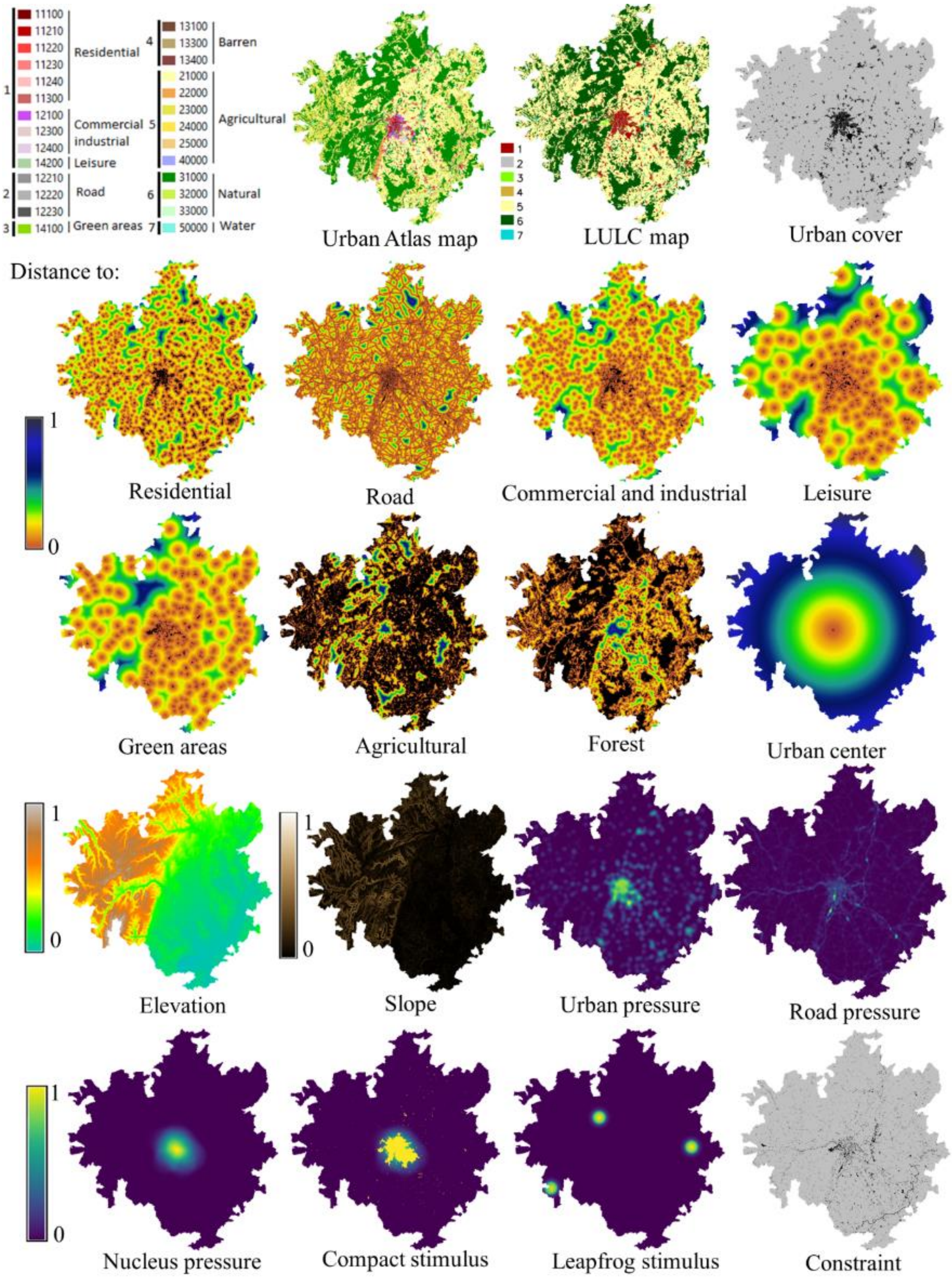

Figure 3. Example of factors computed for the monocentric form. On the top left, reclassification of Urban Atlas legend (five digits, see the interpretation in https://land.copernicus.eu/user-corner/technical-library/urban-atlas-2012-mapping-guide-

new/) into nine classes for computing factors (distances to, pressures, stimulus and constraint), and into seven classes for creating the reference LULC map.

Simulated urban growth patterns

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

- *Expansion growth* represents an expansion of the existing urban cover from the urban fringe.
- *Compact growth* encourages infill growth and land use recycling, prioritizing the development of open land inside urban areas, nearby the urban nucleus, and bare soil.
- *Dispersed growth* follows a scattered, isolated and uncontrolled urban development beyond developed areas.
- *Road-based growth* occurs when the urban development takes place nearby the road network.
- *Leapfrog growth* creates new urban centres at a considerable distance from the developed area.

In order to have different scenarios simulating urban development in different pathways, we computed twenty different models, one for each combination of baseline form (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 dependent variable and the factors as independent variables. The factors vary according to the simulation pattern (see Figure 4).

(2) Retraining the model discarding those factors not statistically significant according to their *p-values* (see Figure 4).

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

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

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

(6) Defining the demand based on the spatial area of development 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 a twenty-five percent was established, due to the fact that this urban area was initially highly developed (the demand of development in number of pixels is shown in Figure 4).

(7) Iterative allocation of development using the Monte Carlo method until the demand is achieved. First a potential seed is located. Second, based on suitability of contiguous pixels and a random size and shape from the calibration step list, the patch is finally allocated (Figure 4, right).

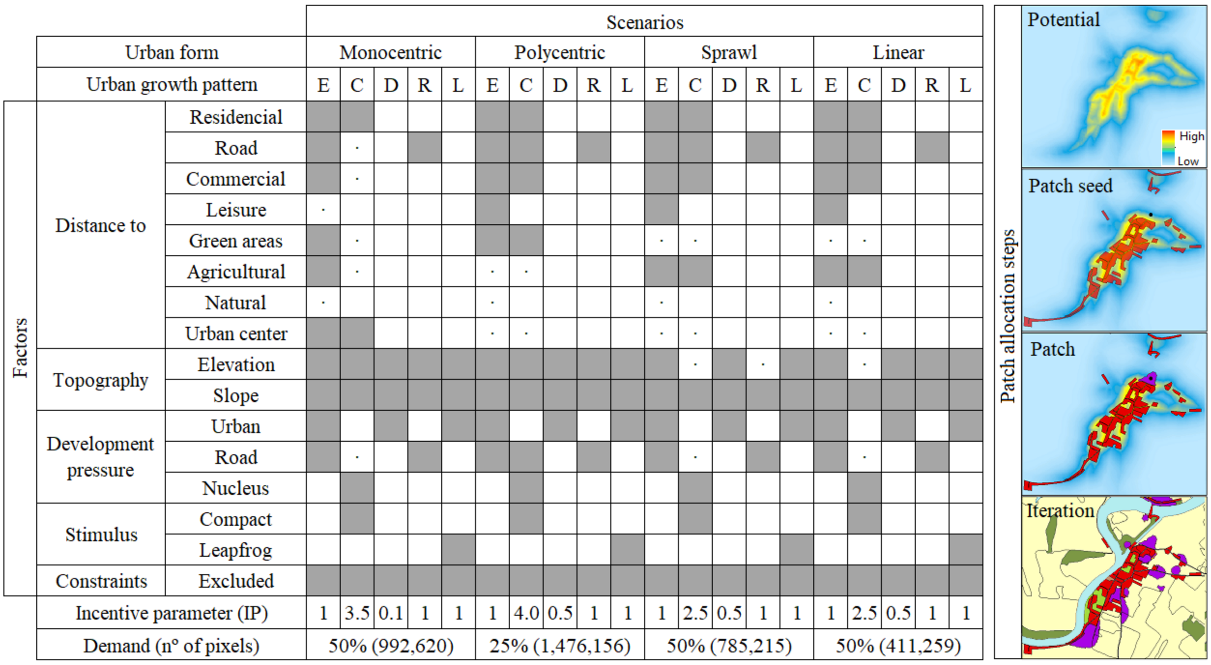

Figure 4. (Left) Factors used in the LUC model are in grey (rows) for the twenty simulated scenarios (columns). White grids were not included in the model, while the dot (·) means that the factor was not statistically significant. Growth patterns: (E) expansion, (C) compact, (D) dispersed, (R) road-based, and (L) leapfrog. The incentive parameter and the demand are shown on the last two rows. (Right) Example of the patch allocation steps.

Computing spatio-temporal metrics

We computed twenty-four spatio-temporal metrics related to the aggregation and spatial distribution of land use change for each simulated growth scenario, using the IndiFrag software (Sapena and Ruiz 2015). This software, available at http://cgat.webs.upv.es/software/, is a tool that extracts an exhaustive set of fragmentation, spatial and temporal indices from LULC data.

When working with many spatial metrics or variables, as in this case, it is expected to find high correlations, making difficult the interpretation and introducing noise in the classification process. Therefore, it is advisable to remove correlated metrics and keep only the most informative (Uuemaa *et al.* 2009, Schwarz 2010). Hence, we conducted a correlation analysis of the spatio-temporal metrics to discard those metrics with strong correlation (ρ≥0.8) and avoid redundancies in the spatial information. As a result, only the eleven metrics described in Table 2 were used. Some spatio-temporal metrics were computed as the difference between those obtained from the initial LULC maps and the final simulated scenarios (spatial metrics in Table 2). Others were computed as direct spatio-temporal metrics from both maps. Thus, a set of eleven metrics was obtained from the twenty urban development scenarios (Table 2).

Table 2. Description of the selected spatio-temporal metrics. Formulas can be found in doi:10.4995/raet.2015.3476 and Reis *et al*. (2016).

	Metric	**Description**
Spatial metrics	**Leapfrog (LPF)**	Proportion of isolated urban patches. It is considered isolated when the distance to the closest patch is higher than 20 m (avoiding two patches separated by roads to be isolated).
	Porosity (P)	The ratio of open space (area of holes within the land cover) compared to the total land cover area (Reis *et al*. 2016).
	Weighted Euclidean Distance (DEP)	Concentration degree, or area-weighted mean distance of patches to the centroid of the land cover (kilometres).
	Mean nearest neighbourhood distance (DEM)	Mean distance between nearest patches (meters). It is considered adjacent when the distance to the closest patch is lower than 20 m.
	Compactness (C)	The ratio between area and perimeter. Measures the shape complexity of the urban cover.
	Radius dimension (DimR)	Measures the centrality of the urban cover with respect to the urban center.
	Effective mesh size (TEM)	Measures landscape connectivity. Lower values mean more fragmentation (hectares).
	Splitting index (IS)	The number of patches when dividing the cover into equal size parts with the same division.
Spatio-temporal metrics	**Weighted mean expansion (AWM)**	Weighted growth compactness. It is the area-weighted mean of the proportion adjacencies between new urban patches and the urban cover.
	Disaggregation (DI)	Mean distance from new urban patches to the closest patch of the urban cover (metres) (Reis *et al*. 2016).
	Centroid displacement (CNT)	The distance between the geometrical centroid of the urban cover at two different times (metres).

The metrics with area and length units might be affected by the scale, size, and boundary effect. However, as input data have the same spatial resolution, the scale will not affect DEM, CNT and DI metrics in this comparative analysis, since they measure relative distances. The influence of size and boundary were tested normalizing DEP and TEM by dividing their values by the radius and area of the circumference with the same area than the boundary. Since these values did not change significantly (correlation coefficients $\rho_{(\mathrm{DEP-nDEP})}$ =0.98 and $\rho_{(\mathrm{TEM-nTEM})}$ =0.99), we used the non-normalized metrics to ease further interpretation.

Urban growth spatial pattern classification

In order to harmonise the differences in units, the values of the metrics were standardised to mean zero and standard deviation one. From the pre-selected metrics (Table 2), 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 relevant metrics selected were: the weighted mean expansion index (AWM_{urban}), the variation of the weighted Euclidean distance (DEP_{urban}), the disaggregation index (DI_{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. These analyses were applied using the R statistical software (R Core Team 2019).

Finally, in order to assess how the baseline urban form influences the classification of growth patterns, two outputs were compared: the classification error rates per urban form and the behaviour of metrics per urban form using graphs.

Results

Categorization of urban growth spatial patterns

As a result of the LUC modelling we created twenty growth scenarios, whose distinctive features are shown in Figure 5. These scenarios recreate the behaviour of five growth patterns that may happen in a developing area, and how these patterns will progress on different urban areas with specific baseline urban forms. The spatio-temporal metrics were extracted from these scenarios.

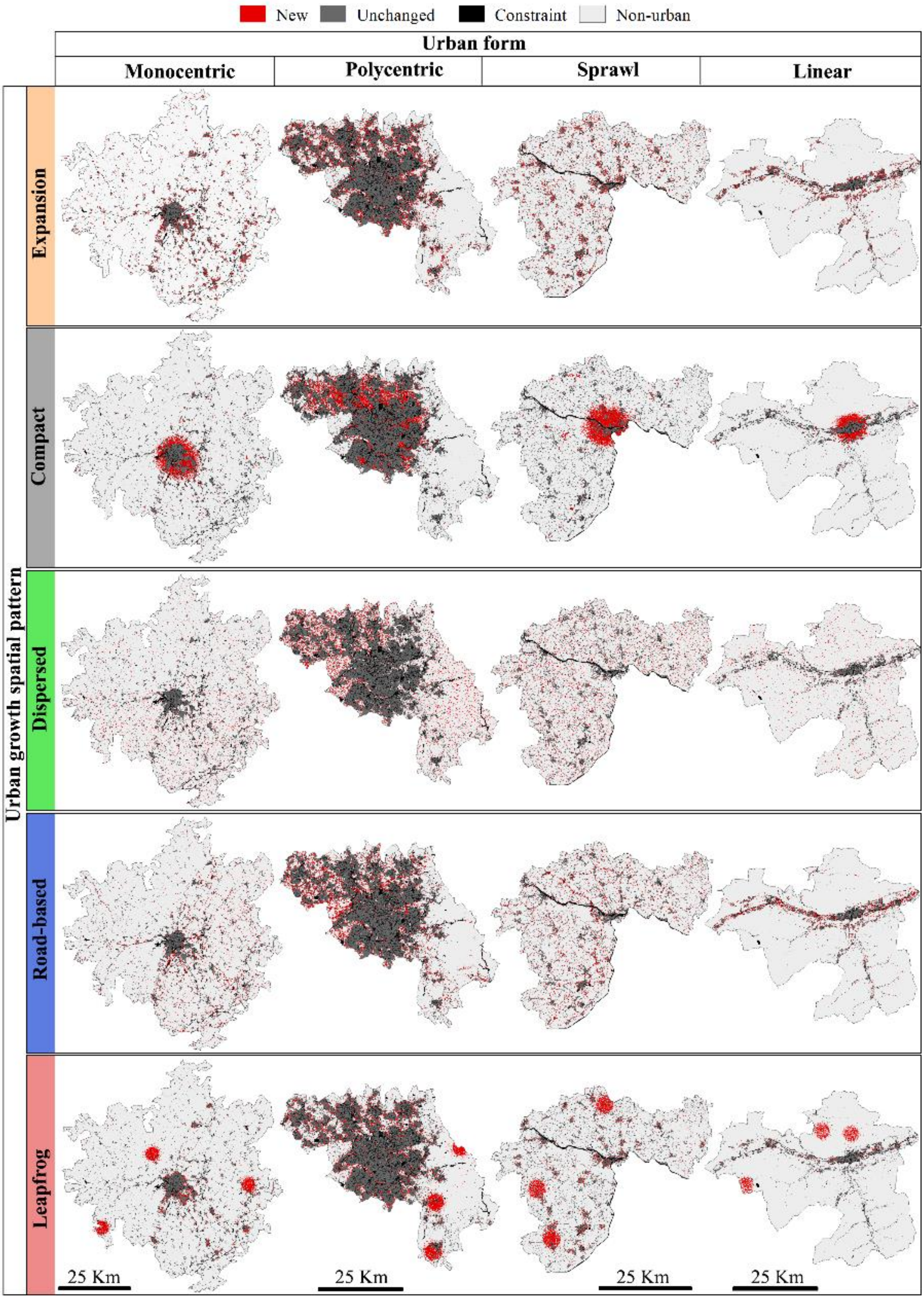

Figure 5. Scenarios with the simulated urban growth in red, following five spatial patterns (rows) from the four baseline urban forms (columns). The baseline urban covers are shown in dark grey and constraints in black, showing the areas restricted for development.

Figure 6 shows the distribution of scenarios by means of the standardized values of the selected spatio-temporal metrics (AWM_{urban}, DI_{urban}, C_{urban}, and DEP_{urban}), where the baseline urban forms and growth patterns are represented with different shape and colour, respectively. The distances between growth pattern scenarios on the space represented by each pair of metrics are inversely related to their similarities. Observing the combination by pairs of metrics, Figure 6 suggests the contribution of metrics for the identification of growth patterns. DEP_{urban} discriminates well the compact pattern and AWM_{urban} the expansion, with some exception. The DI_{urban} splits the dispersed pattern from the rest, even if sometimes it is 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.

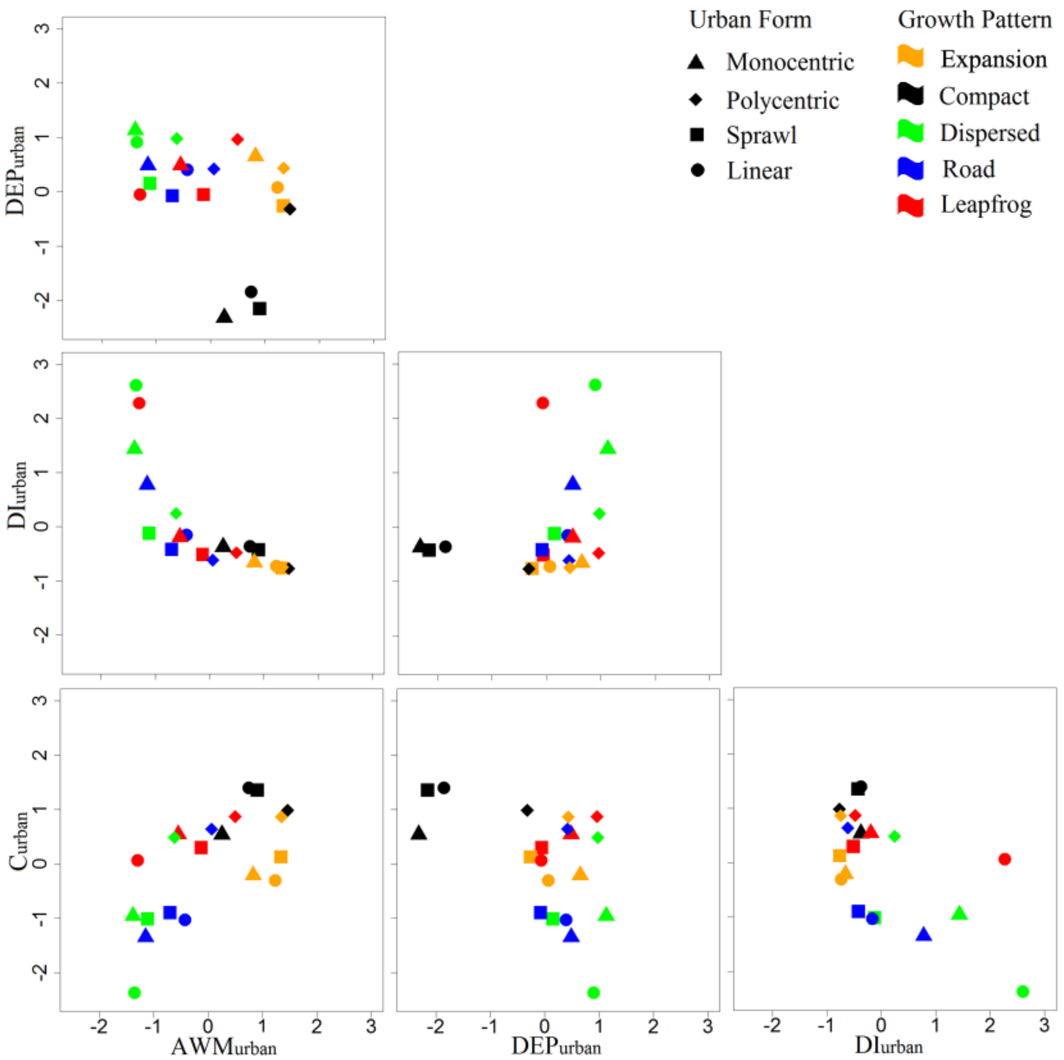

Figure 6. The distribution of simulated growth scenarios according to the combination of the standardised values of AWM_{urban}, DI_{urban}, DEP_{urban}, and C_{urban} metrics. The colour represents the simulated growth pattern, while the symbol is the initial urban form.

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

Table 3 shows the classification errors of the clustering method for each scenario, using AWM_{urban} and DEP_{urban}. The omission error (OE) gives the proportion of underclassification of a pattern, while the commission error (CE) informs about the overclassification of a pattern. Accordingly, the expansion pattern is the one with higher 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 6 (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.

The comparison of the centroids of the classified clusters against the actual patterns shows the highest difference in the growth adjacency (AWM_{urban}) of the road-based pattern. This is because even if the road-based growth patterns are quite clustered by means of AWM_{urban} and DEP_{urban} they are overlapped by the leapfrog growth (Figure 6). Consequently, only two scenarios were identified correctly, which displaces the centroid of the cluster to the left, however, as DEP_{urban} centroids are quite similar for road-based and leapfrog patterns, the differences are least in this metric (Table 3). With regards to the rest, the centroids are quite similar (**Error! Reference source not found.**).

Table 3. Classification of scenarios into five clusters (colour) using AWM_{urban} and DEP_{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 AWM and DEP_{urban}.

	MONO	POLY	SPRA	LINE	OE	CE	Cluster centroid AWM_{urb}	Cluster centroid DEP_{urb}	Pattern centroid AWM_{urb}	Pattern centroid DEP_{urb}	Euclidean distance
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							

Influence of the urban form in growth spatial pattern categorization

According to error rate per urban form (UFE) from Table 3, 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 DEP_{urban} has a different behaviour against the other forms and was classified as road-based with lower mean value of DEP_{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 3).

When interpreted individually, the adjacency and concentration degrees of urban growth (AWM_{urban} and DEP_{urban}) present different responses depending on baseline forms (Figure 7). The expansion and compact growths have similar values of AWM_{urban}, but different values of the changes of DEP_{urban}. On the other hand, the leapfrog growth from the linear form had an unexpected value of AWM_{urban} compared to the rest of the simulations, as also seen in Figure 6. This scenario has the particularity that the hilly areas are not urbanised along with the fact that the simulated leapfrog pattern projected randomly the new urban clusters – in the hilly areas – and, consequently, the adjacency of new urban elements to previous urban areas are much lower compared to the rest of scenarios. The polycentric form is characterized for being highly urbanized and compact. Therefore, it not only has higher values of AWM_{urban} in all patterns, due to higher probabilities of growth adjacent to the urban elements, but also DEP_{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 said above the already spread urban cover together with new isolated urban patches slightly increase the distance to the centroid. In fact, the changes in DEP_{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 are highlighting the notable influence that urban form has on the identification of spatial patterns.

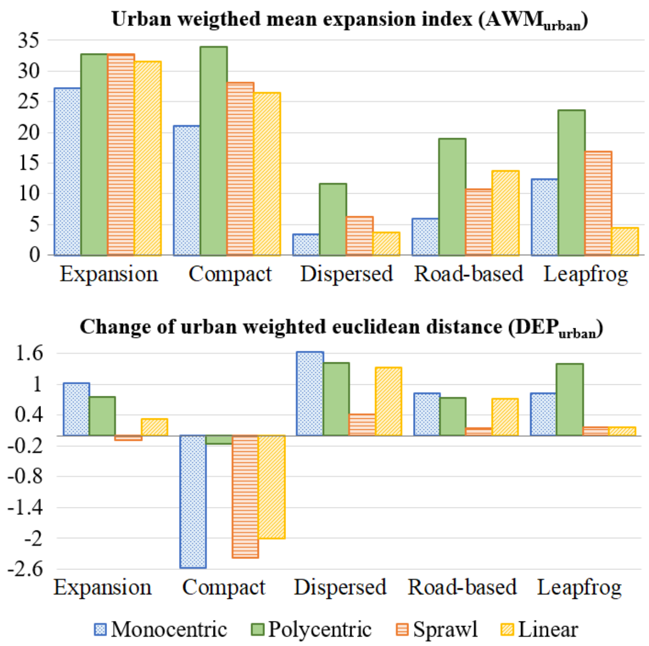

Figure 7. Values of the spatio-temporal metrics used in the cluster analysis. Metrics are grouped by growth pattern, and colours represent the baseline forms.

Discussion

The more well-informed and efficient decisions are made in urban planning and management practices, the better urbanization challenges will be addressed. It has been widely discussed that urban growth has diverse impacts on environmental, social, and economical aspects according to their spatial characteristics (e.g.: Williams *et al*. 2000, Oliveira 2016, Zhao *et al.* 2016, Wei and Ewing 2018). Remote sensing and Geographic Information Science can provide data and methods that facilitate monitoring and evaluating the development of urban areas. In this direction, this paper proposes a methodology for the early identification of five different growth patterns in urban areas based on a meaningful subset of spatio-temporal metrics derived from LULC data. To the authors’ knowledge, it differs from other studies as it attempts to identify growth types/classes rather than degrees between compactness and dispersion (e.g.: Tian *et al*. 2011, Liu *et al*. 2010, Jiao 2015).

There is a vast amount of spatial metrics for the quantitative analysis of urban landscapes (Reis *et al*. 2016). Since our aim was to identity growth spatial patterns, we focused on metrics that quantify aggregation, spatial relations, and their variations. In this context, extracting an exhaustive set of spatio-temporal metrics followed by a selection of the most relevant attending to these characteristics was revealed as efficient for this purpose.

Some authors previously mentioned that spatio-temporal metrics are complementary when conscientiously selected, and their combined use enriches the study of urban areas and their dynamics (Arribas-Bel *et al*. 2011, Abrantes *et al*. 2019, Sapena and Ruiz 2019). From the four final metrics selected, AWM_{urban} is particularly helpful to discriminate the expansion growth as it quantifies adjacencies, DI_{urban} separates disperse growth as it measures distance to old urban patches, C_{urban} detects road-base and disperse growths as they tend to be less compact in shape, and DEP_{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 and discriminate the five growth spatial patterns analysed, which has a practical relevance for their use in monitoring urban growth. The change in the concentration degree of the urban cover (DEP_{urban}) is the metric that individually better identifies patterns. It measures the area-weighted distances of the urban elements with respect to the urban centroid in two data and then quantifies its variation. Negative values mean more concentration, while positive values mean the fragmentation of the urban cover; higher values suggest fragmentation in the peri-urban area. When combined with the degree of adjacency of the urban growth (AWM_{urban}), which quantifies urban densification and growth compactness (higher values mean denser and more compact areas), the identification of patterns improves. The complementarity of these two metrics allows 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 centre, the second quantifies the level of aggregation of the new development.

The use of graphs to represent the spatio-temporal metrics enhances the differences between monocentric, polycentric, sprawl and linear forms when analysing the same urban growth patterns. According to our analysis, the polycentric and linear initial urban forms are the ones adding more uncertainty into the categorization of growth patterns. Therefore, when applying spatio-temporal metrics for growth pattern classification, the influence of the urban form should be considered together with the widely known sensitiveness of spatio-temporal metrics to the size, scale and boundary effect (Uuemaa *et al.* 2009; Reis *et al*. 2016). In our study, the scale did not affect metrics, as data had the same resolution and the sizes of the urban areas were similar. Regarding the urban form, approaches to overcome its influence in the classification of growth patterns are still required for the correct identification of development trends. In this sense, the inclusion of the baseline urban form as a qualitative variable in the classification procedure would be worth to investigate in future research to improve the discrimination of growth patterns.

Regarding growth patterns analysed, we made a synthesis of those described in the literature in order to assess if they can be identified by means of spatio-temporal metrics derived from LULC data and using clustering methods. However, categorization 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 development and on the urban baseline form. To avoid this growth pattern mix usually present in the reality, and to overcome the problem of lack of long-term and high-resolution LULC databases, we created different scenarios by simulating urban growth using FUTURES model. The use of simulated scenarios also provide transferability to other geographical areas, despite their differences in morphology or growth types.

The increasing availability of frequent and updated urban data, in particular those related to LULC, will open new opportunities in this field, requiring tools and methods, as well as interpretable indicators to efficiently characterize urban growth. Eventually, when databases and LULC data increase, new studies based on real development cases, instead of simulations, can be conducted.

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 this pair of spatio-temporal metrics, will allow its relationship with their environmental, social and economic impacts; consequently, an empirical cause-effect relationship will be determined by means of statistical models, which will provide a better understanding of complex development processes in urban environments and their consequences.

Conclusion

The development of methodologies for the description and quantification of urban growth is useful to monitor urban areas, to diminish the consequences of fast developing and to improve planning and sustainability of urban systems. In the absence of long-term LULC data at high-resolution, we simulated urban growth of different cities and scenarios to answer the question of whether spatio-temporal metrics derived from LULC maps are able to identify urban growth patterns, and to analyse the influence of different initial or baseline urban forms in this classification. As a result, 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 its usefulness for the categorization of urban growth spatial patterns in diverse urban environments 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 development trends and thus, the ability of foresee their consequences, will be valuable for land use planning in urban and peri-urban areas.

Data and codes availability statement

The data and codes that support the findings of this study are available with a DOI at https://doi.org/10.6084/m9.figshare.c.4853124.

Supplementary material

The following are available online at https://doi.org/10.6084/m9.figshare.12871481. Figure S1 to S5. Simulated growth scenarios from the four baseline urban forms (monocentric, polycentric, sprawl and linear). The baseline urban covers are shown in black, while growth is shown in different colors.

Acknowledgments

This study was partially funded by the Spanish Ministerio de Economía y Competitividad and FEDER, in the framework of the project CGL2016-80705-R.

References

Abrantes, P., *et al.*, 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.

Arribas-Bel, D., Nijkamp, P., and Scholten, H., 2011. Multidimensional urban sprawl in Europe: A self-organizing map approach. *Computers, Environment and Urban System*, 35 (4), 263–275. doi:10.1016/j.compenvurbsys.2010.10.002.

Bhatta, B., 2010. Causes and Consequences of Urban Growth and Sprawl. In: Springer. *Analysis of Urban Growth and Sprawl from Remote Sensing Data*. Berlin, Heidelberg, 17–36.

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.

Chen, Y., *et al.*, 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.

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., *et al.*, 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.

Dorning, M.A., *et al.*, 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.

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

European Environment Agency (EEA), 2016. Data from: Copernicus Land Monitoring Service, Urban Atlas [dataset]. Available from: http://land.copernicus.eu/local/urban-atlas [Accessed 1 June 2017].

European Union, 2016. *Urban Europe. Statistics on cities, towns and suburbs*. Luxemburg: Publications office of the European Union. doi:10.2785/91120.

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.

Goerlich, F.J. and Cantarino, I., 2013. Zonas de morfología urbana: coberturas del suelo y demografía. Madrid: Fundación BBVA.

GRASS Development Team, 2017. Geographic Resources Analysis Support System (GRASS) Software, Version 7.2. Open Source Geospatial Foundation. Available from: http://grass.osgeo.org.Hankey, S. and 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. and Wong, M.A. 1979. Algorithm AS 136: A K-means clustering algorithm. *Applied Statistics*, 28, 100–108. doi:10.2307/2346830.

Hermosilla, T., *et al*., 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.

Hoymann, J. and 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.

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.

Ju, H., *et al.*, 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.

Liang, X., *et al.*, 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, X., *et al.*, 2010. A new landscape index for quantifying urban expansion using multi-temporal remotely sensed data. *Landscape Ecology*, 25 (5), 671–682. doi:10.1007/s10980-010-9454-5.

Liu, X. P., *et al.*, 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.

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

Meentemeyer, R.K., *et al.*, 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.

Musa, S.I., Hashim, M., and 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.

Nabielek, K., Hamers, D., and 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 Oct 2018].

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

Reis, J.P., Silva, E. A., and Pinho, P., 2015. Spatial metrics to study urban patterns in growing and shrinking cities. *Urban Geography*, 37 (2), 246-271. doi:10.1080/02723638.2015.1096118.

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/.

Salvati, L., *et al.*, 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.

Sapena, M. and Ruiz, L.A., 2015. Description and extraction of urban fragmentation indices: The Indifrag tool. *Revista de Teledetección*, 43, 77–89. doi:10.4995/raet.2015.3476.

Sapena, M. and 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.

Schneider, A. and 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.

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.

Shi, L., *et al.*, 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., *et al.*, 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. and 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.

Song, Y., *et al.*, 2017. The relationships between urban form and urban commuting: An empirical study in China. *Sustainability*, 9 (7), 1150. doi:10.3390/su9071150.

Sun, X., *et al.*, 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.

Taubenböck, H., *et al.*, 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.

Terando, A.J., *et al.*, 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., *et al.*, 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.

Tong, X. and 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.

Ustaoglu, E., *et al.*, 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., *et al.*, 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.

Van de Voorde, T., *et al.*, 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.

Wei, Y.D. and Ewing, R. 2018. Urban expansion, sprawl and inequality. *Landscape and Urban Planning*, 177, 259-265.

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

Wilson, E.H., *et al.*, 2003. Development of a geospatial model to quantify, describe and map urban growth. *Remote Sensing of Environment*, 86(3), 275–285. doi:10.1016/S0034-4257(03)00074-9.

Wu, Y., Li, S., and 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.

Zhao, M., *et al.*, 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.