Changes in the buildings' density and its connection with population

Case study of Central Bohemia Region between 2010 and 2018

Greta Kukeli

Introduction

The political changes of the 1990s that took place in the CEE countries resulted in changes concerning the territory. Central Bohemia was one of the regions that were mostly affected by suburbanization and migration processes. Due to the proximity of the capital city of Prague and the good accessibility through different transport modes, this region has attracted investors because of the cheap land and taxes and land available, but also people who saw suburban places as a promise to have their dream house in the greenery. Therefore, Central Bohemia became the host of the highest waves of suburbanization in Czechia beginning in the mid-1990s and ongoing, especially two districts Prague-East and Prague-West (Ouředníček, 2007), which are considered as Prague's hinterland.

All these transformations have led towards changes in the physical aspects of the settlements that directly or indirectly affect the lives of residents and users and the social ties among them. For this reason, the establishment, measurement and continual monitoring of indicators that deal with the physical environment is essential. Density is one of these indicators, which have high influence and can be found in almost all territorial development plans and policies of all scales (Longley, Mesev, 2002). On the other hand, density is one of the most discussed indicators too (Galster et al., 2001). This because there are different ways how to conceptualize and measure it. Focusing only on urban density (this study does not focus on population density), the possibility of measuring it varies from: number of buildings per total area (for example city area), number of dwellings per area of study, number of buildings per built-up area and other varieties used in drafting of master plans when it gets elaborated like FAR (floor area ratio), etc. We can say that the adaptable way to measure such an indicator depends totally on your data, scale, and purpose of analyses. Still, all their results show interesting information concerning the development of the settlements that condition our daily and future life.

The aim of this map is to measure the density of buildings in the territory of Central Bohemia and analyze its change during the period 2010–2018 by arguing its possible connections with the population concentration and the location of transport infrastructure. This will contribute to understanding where the construction and development that has been happening during this period. To realize this map of density study, we used the Kernel Density (KD) tool available on ArcMap 10.6 Spatial Analyst toolbox. This method can be considered as a nonconventional way of measuring of the density. However due to the increase of popularity of spatial software and availability of data in this format, this tool and similar ones have started to slowly integrate and even change the conceptualization frame of such research (Longley, Mesev, 2002). One of the main issues of calculating density has been the determination of the area that will be taken under study, and this causes different results and interpretations. These areas in most cases are administrative borders that equalize the values of density for the whole area giving general values and avoiding the possibility to identify diversity within them (Hudeček et al., 2019). ArcMap offers other tools to measure density, however, among them KD is the best at reducing the effects of boundaries. This attribute has made it important in solving similar problems that have to do with events or features' concentration (Leigh et al., 2019). However, to demonstrate how this technique works we have included a conceptualization scheme below (Figure 1). The KD tool divides the area in a grid whose cell size depends on the research interest and the density values within them will be presented in a raster format. Nonetheless, the implementation of the bandwidth as said by Leigh et al. (2019), has reduced the effect of the grid and any other borders. The kernel function is applied on each point within the grid cells and a kernel is lifted upon them. The function has its highest value on the points' location and goes down until it reaches zero at the end of the bandwidth distance. Kernel heights of all the points within one grid cell are gathered and represented by colors based on their values. Due to the bandwidth, we consider the contribution of all the features within its distance as direct influences on the cell results. Basically, the result is not created only on the individual cell value but also on its neighborhood. The selection of this element of the KD analysis is crucial for the result and there are different ways to select one. Based on the work done on the Kernel Density Estimation function which directly contributes to this tool (Gramacki, 2017) the method which is used for this map is considered as data-driven selection and is explained in the methodology section.

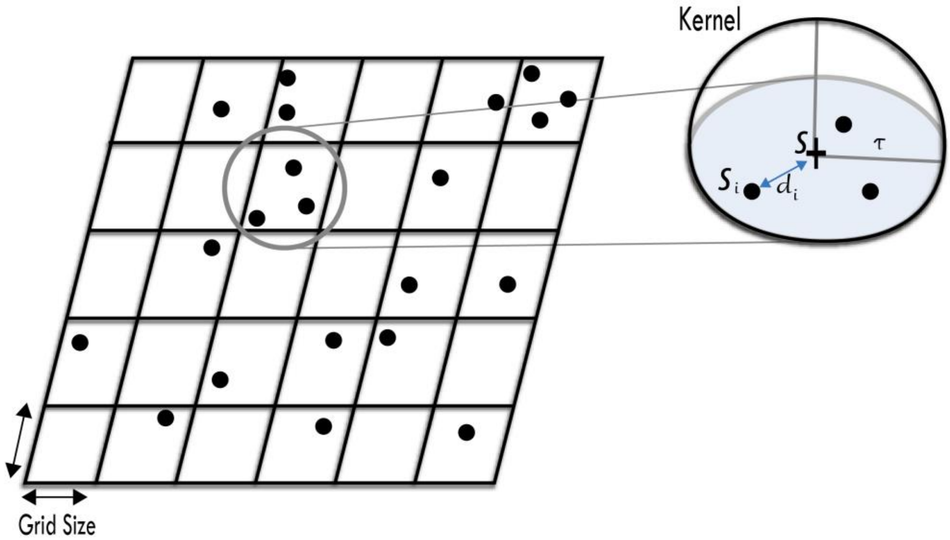

Figure 1: Kernel Density conceptualization scheme.
Source: Leigh et al., 2019.

Data and methods

The map employs data from the Czech Statistical Office (CZSO) regarding population numbers for the municipalities and from Register of Census Tracts and Buildings (RCTB), which is also official information provided by the CZSO through years. Also, for the infrastructure, we used the shapes taken from ArcČR500 database version 3.3. The data taken from RCTB consist of point data representative of each building containing diverse information such as location, type of construction, correct address, number of dwellings, etc. In this map, we work with the point data of two different periods, 2010 (447,473 data points) and 2018 (484,635 data points). The structure of the dataset for 2018 shows that 78.8% of the buildings are of residential character and only 21.2% nonresidential. Due to this clear indicator of the region's development character, we can consider this as a study of residential density in Central Bohemia.

In order to realize Kernel Density analysis, we calculated the Bandwidth (radius) for both years under the evaluation. The proper distance to apply as a radius should be the one that demonstrates some level of spatial clustering patterns. This distance should be adequate for the research scale that we are interested in too. The tool that helps in this is Incremental Spatial Autocorrelation (Spatial Statistic toolbox) which runs Moran's I for different distances testing the level of clustering. However, this tool also requires a beginning distance and incremental value to start. We applied the Calculate Distance Band from Neighbor Count tool (Spatial Statistic toolbox) which gives the minimum, average and maximum distance of one point from the specified number of neighbors. We wanted to know these values for each point from one neighbor only. The results that came for both years are similar to each other, with a slight decrease in the values for 2018. The average distance is 25m, the maximum one is 2440m and the minimum one is zero.

Because our dataset is very large and by implementing Incremental Spatial Autocorrelation (ISA) in such conditions we risk running out of memory, we had to integrate the data by using the Integrate tool (Data Management toolbox). This means that points of similar coordinates within a specific distance from each other are joined and represented as one. We integrated points within the 10m distance and then used the Collect Events tool (Spatial Statistic toolbox) to count them and create a new shape with a weighted field in the attribution table showing the number of integrated points per location. This weighted field was used later in the application of KD. The ISA is run on the integrated data using the maximum distance (produced by Calculate Distance band tool) as beginning distance and the average one as the incremental distance. The beginning distance is the value from where the Moran's I calculation will start while the incremental one shows how much the beginning distance will be incremented. We ran the tool for 10 distance bands.

The results of ISA results come in a graph where we have z-scores and distances. The z-scores show the level of the clustering in the data for each distance. Depending on the dataset the ISA graph can have different appearances with ups and downs or be a continuously growing line. The results for both years showed that the line graphs of z-scores even though they have high positive values never reach a peak. The peak represents the 'perfect distance' to be used in calculations, but this doesn't devalue the other distances and their z-scores. This is caused by large distances among the buildings, considering those who are farer as outliers. The possible way to solve this issue would be to identify and remove these buildings, run the Calculate Distance Band tool again to get a new maximum distance. ISA will be rerun with the new coefficient by using a bigger incremental distance and running it for many distances until we find a new distance that would cover all the outliers. This is the method that would help to get a peak in the z-score line and not run out of memory but in our case, approximately 1/3 of the points could be considered as outliers. Trying to figure out a perfect incremental distance that would lead us to a better distance than those given by the ISA already would perhaps lead as to values that are beyond our research interest scale. In such situations, based on the recommendations of ESRI and GIS users community, it is important to understand that there is not only one perfect distance to run the ISA and the fact that the z-score graph never peaks shows that there are a high level of processes happening simultaneously on different scales in the area. This means that all the distances received by the ISA tool can be used and give exact results. The distance used for both years KD is 2440 m (first values of ISA result for both years) and as a population field, the weighted field created by the integration of data.

To specify the Kernel Density grid cell, we selected the data-oriented approach which consists of using the tools default version. This version selects the shortest value among the width or the height of the output extent, in our case the border of Central Bohemia, and divides it by 250. For both years this value was 471 m however the final results of the KD are in km^2. So, all the results of KD can be understood as the density of buildings per km^2. The individual raster of both years is on the same units, grid cell size, and bandwidth. To find how much the density values have changed between the two years, we deducted the values of 2010 from 2018 by using the Raster Calculator tool (Spatial Analyst toolbox) and the results of it are represented in a new raster file.

The next step included work with the population data where the difference was calculated between 2018 and 2010 based on the Crude's rate of population formula, and the results are in percent per mile. In order to make a classification of these values we divided those into six categories (Table 1) based on Sturges' rule for determining bars. By joining these results with the original shape on ArcMap (Excel to Table tool of Conversion toolbox) we gave different symbologies to each category to facilitate recognition and for purpose of analyses. As one of the map's aim is to see if there is any connection between the population and density change in the region, we measured their correlation (Figure 1). Through the use of the Zonal Statistic tool (Spatial Analyst toolbox), we could connect the KD raster values with the municipalities' shape and subtract statistics from it. The statistic that was used in the calculation of the correlation was the Mean value which consists of the average values of the raster within each municipality. To understand this relationship better and be sure of the correlation result, we reanalyzed the relationship through the Geographic Weighted Regression tool (Spatial Statistic toolbox). This allows us to see whether one variable is able to predict the other, in our case if population change can affect density.

The infrastructure layers of roads, railroads, and train stations were also added to the population change and Kernel density raster. However, we selected to show only those roads whose medium width is above 10,5 m based on the official data, since we would like to focus on major features of infrastructure that are simultaneously representative of the major state policies of development through years. Therefore, all the roads with a medium width smaller than 10,5 m were not taken into consideration in the result discussion or shown on the map.

Results

The results show that the density of buildings has generally increased during the period 2000–2018 but there is a relatively wide spectrum of municipalities with growth and decline of density values. Municipalities with an increase of density are mainly situated close to Prague's border which seems to attract all types of increase (population and the construction of buildings) and around train stations and areas trespassed by main roads. As can be seen clearly in the map, positive values of density are positioned in a way that creates continuous corridors along railways and roads. This is an indicator of how the state's policies and investments of different periods in infrastructure shape the present and future development in the territory influencing even on the enhancement of process like suburbanization but not only (Stanilov, 2007). The increase of buildings´ density in more peripheral parts of the region can be connected with the location of specific natural features or job places such is the case of the area surrounding Mladá Boleslav. The increase of density seems more prominent close to the northern border of Prague. The density's values are high on the south of this border also, but the development does not appear to concentrate only around it but extend southern into the Central Bohemia following the infrastructure and favored also by the soft territory.

Category	1	2	3	4	5	6
Percentage group	<0%	0–20.21%	20.21-40.1%	40.1-60%	60-78.89%	78.89%<
Percentage of change 2010–2018	17.9%	31.5%	24.6%	12.8%	6.1%	6.8%
Number of municipalities	209	367	286	149	71	80

Table 1: Categories of municipalities based on the population percentage change

The legend of the map uses six categories of population change within the municipalities of the Central Bohemia Region (Table 1) and the most frequent is the second category with population change between 0–20.21% manifesting in approximately 31.5% of the municipalities. The location of these municipalities is homogenously spread in the region. However, in the region, we have municipalities with negative growth too (either to migration or administrative divisions). They are the third category with around 17.9% and as it can be seen from the map their location is peripheral and far from the main transportation corridors. The three groups where the population has grown above 40.1% and overpassing 79%, are mainly located close to Prague's border or in important locations. From the position of the municipalities' categories, we see that infrastructure is an important feature in the territory whose impact in development should be always taken into consideration (Haider, Miller, 2000).

The map integrates data concerning density numbers and the population in the region of Central Bohemia, and the results show that there is a correlation between them (Figure 1). The correlation between them is approximately 60% (57%). The number is not bigger due to the fact that 1/5 of the buildings are of nonresidential character so they are not connected directly with the population increase. This shows that the change in density is not entirely affected by the increase of population but also by the economic activity. Nonetheless, density and population have a positive correlation showing that there is a strong connection between the two indicators.

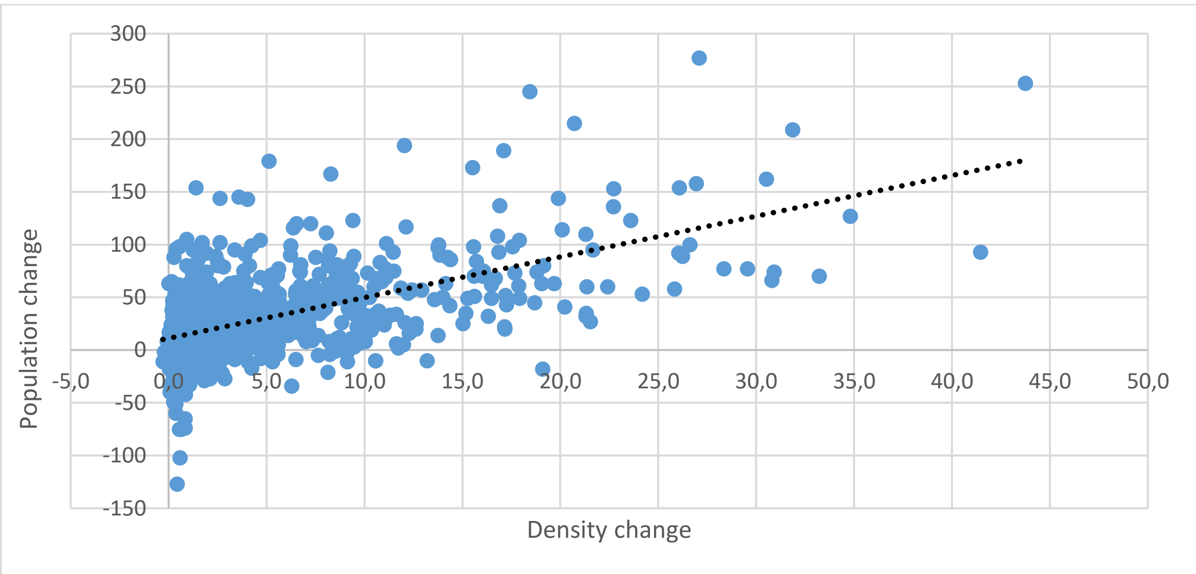

Figure 1: Correlation between density change and population change 2010–2018

The result of the Geographic Weighted Regression is similar to the correlation values (50%). However, this value shows the existence of other variables that influence the changes in the density and we should work on their identification. Due to these results, we can assume that in Central Bohemia features of interest such are infrastructure and points of interest (natural, economic, etc.) affect population numbers and construction interest by increasing/decreasing the density values. Future investigation should be made into identifying all the variables and measure their effect in the density change. This region in eight years has had an increase of 37 162 buildings which means 387 new constructions each month. The densification can be seen even by the reduced values of average and maximum distances among building points when comparing 2010 and 2018. The highest change in density was with 62 buildings/km^2, however, these changes are not present everywhere in the territory. Nonetheless, this map shows areas of development interest for the public and investors serving as an orientation also for administrations of different levels, policymakers, and investors.

References:

GALSTER, G., HANSON, R., RATCLIFFE, M. R., WOLMAN, H., COLEMAN, S., FREIHAGE, J. (2001): Wrestling Sprawl to the Ground: Defining and measuring an elusive concept. *Housing Policy Debate,* 12(4), 681–717.

GRAMACKI, A. (2017): Bandwidth Selectors for Kernel Density Estimation. In: Gramacki, A. (ed.): *Nonparametric Kernel Density Estimation and its Computational Aspects. Studies in Big Data.* Springer, Cham, pp. 63-83.

HAIDER, M., MILLER, E. J. (2000): Effects of Transportation Infrastructure and Location on Residential Real Estate Values: Application of Spatial Autoregressive Techniques. *Transportation Research Record: Journal of the Transportation Research Board,* 1 January, 1722(1): 1–8.

HUDEČEK, T., HNILIČKA, P., DLOUHÝ, M., LEŇO CUTÁKOVÁ, L., LEŇO, M. (2019): Urban structures, population density and municipal expenditures: An empirical study in the Czech Republic. *Urban Studies,* 56(16), 3450–3465.

LEIGH, J., DUNNETT, S., JACKSON, L. (2019): Predictive police patrolling to target hotspots and cover response demand. *Annals of Operations Research,* 283(1-2): 395–410.

LONGLEY, P., MESEV, V. (2002): Measurement of density gradients and space-filling in urban systems. *Papers in Regional Science,* 81(1): 1–28.

OUŘEDNÍČEK, M. (2007): Differential Suburban Development in the Prague Urban Region. *Geografiska Annaler: Human Geography,* 89B(2): 111−125.

STANILOV, K. (2007): Housing trends in Central and Eastern European cities during and after the period of transition. In: Stanilov, K. (ed.): *The Post-Socialist City. Urban Form and Space Transformations in Central and Eastern Europe after Socialism*. Springer, Dordrecht, pp. 173–190.