Journal of
Social and Economic Statistics

*Vol. 12, No. 2, Winter 2023*
**DOI**: 10.2478/jses-2023-0007

# *GREEN AND SMART URBAN DEVELOPMENT: A COMPARATIVE STUDIES BETWEEN CITIES OF ROMANIA, CANADA AND DENMARK*

**Laura-Elena ILINU[a], Maria HOROIU[b], Alin Cristian MARICUŢ[c], Giani Ionel GRĂDINARU[d]**

[a)] *Bucharest University of Economic Studies, Bucharest, Romania.*
[b)] *Bucharest University of Economic Studies, Bucharest, Romania*
[c)] *Bucharest University of Economic Studies, Bucharest, Romania.*
[d)] *Bucharest University of Economic Studies, Bucharest, Romania.*

**Abstract**

*Due to the fact that the planet's resources are limited and human exploitation has led to unprecedented environmental pollution, sustainability has become a concept of great importance in recent years, especially in the context of very rapid and large-scale urban development. The green city is a form of sustainable city focused mainly on the creation of green spaces, which helps, among other things, to reduce pollution, to combat climate change and to create a more favorable environment for people. Green infrastructure is the main element that characterizes this type of sustainable city, the dynamics of the use of the term in specialized studies showing an upward trend. Interest in the notion of green city has seen a major increase in the last 8 years, highlighting the need to create a more nature-friendly way of urban development. The country that stands out regarding its contribution in terms of studies carried out on the theme of green city is China, while Romania is one of the countries where this subject is very little researched. A cluster analysis of cities in Romania, Denmark and Canada provides a valuable perspective, namely that Romanian cities are the most polluted and have very few green spaces per capita, suggesting the existence of problems with government policies to transform the cities into ones that respect the environment more.*

**Keywords**: sustainability, green city, bibliometric analysis, cluster analysis, random forest
**JEL Classification: R00**

[*]Corresponding author, **Alin-Cristian Maricuț** – alin.maricut@csie.ase.ro

## 1. Introduction

The concept of green and sustainable cities has gained significant attention in recent years as a promising solution to the challenge of urban sustainability. In order to achieve sustainable urban development, an interconnected triad including economy, society and nature is needed to establish a socio-economic system that does not negatively affect the natural environment (Dur and Bunker, 2014).

Urban green spaces provide a wide range of ecosystem services that can help address many urban challenges and improve the quality of life of urban dwellers, especially their health, as these green spaces are diverse in terms of size, plant cover or variety of species (Dahmann et al., 2010). Adequate access to these urban green spaces is a key factor in ensuring adequate living conditions and promoting a healthy environment in urban areas (Wüstemann et al., 2017).

Starting from the hypothesis of the existence of economic, social, environmental and development gaps the paper aims to analyze Romania in terms of green infrastructure, the main objectives considered in this article being to conduct a bibliometric analysis and a comparative analysis of the cities from Romania, Canada and Denmark. The bibliometric analysis aims at verifying the dynamics of the subject over a period of time, its evolution and exhaustion, and the comparative analysis has the purpose of identifying the main differences between cities from Romania, Denmark and Canada.

The structure of the paper is first described through a review of the specialized literature, in order to lay the foundations for a comprehensive understanding of the field of sustainable cities and green cities. The paper is outlined in five chapters, the state of the article being followed by the description of the methods and the case study developed to fulfill the two objectives. Chapter 5 contains the practical part, which will address the research hypotheses, including the question whether Romania can be considered a sustainable country, which encompasses green and sustainable cities. The first part of this chapter focuses on determining an optimal number of clusters and their characteristics, while the other one classifies the cities included in the analysis into one of the previously developed clusters.

## 2. State of the article

Green cities are cities with a significant amount of green space and are designed to be more sustainable and environmentally friendly. These cities often have parks, gardens, recreation areas and other green spaces that enhance the quality of life and improve the well-being of residents.

The functions provided by ecological systems are beneficial for people and the urban environment in numerous ways. To address the challenges related to climate change, natural disasters, food and water security, health and well-being of the people, and economic and social development, ecosystem-based approaches have been grouped under the generic term of "nature-based solutions". (Cohen-Shacham et al., 2016).

Breuste (2023) highlights how achieving green cities is now a global goal and promoting green spaces and nature in cities is an important part of this approach. Many cities recognize the value of urban nature and encourage the reintegration and protection of urban green spaces, such as parks, forests, community gardens and blue infrastructure. He also defines the green city as being in harmony with nature, where all its forms of nature are part of the green infrastructure and are maintained and expanded for the benefit of the city's inhabitants. Strategic and integrated planning for the protection, development and management of urban nature at the city level is of particular importance as the green city must be a realistic program, not just a vision.

The benefits of green cities include reducing the effects of urban warming, improving air quality, increasing biodiversity, enhancing the well-being of residents and promoting an active and healthy lifestyle. For example, Nutsford et al. (2013) investigated whether proximity to urban green spaces is associated with people's mental health. They examined the relationship between access to green spaces and the number of treatments for anxiety or mood disorders among residents of the city of Auckland in New Zealand. A total of 3149 small area units were analyzed and six measures of access to green spaces were used. Results showed that increased access to green spaces was associated with fewer treatments for anxiety and mood disorders, and benefits could be linked to both active participation in nearby green spaces and observable green spaces in the surrounding environment.

Air quality is crucial for the survival of city dwellers because it is a fundamental measure of environmental quality. All living beings have the right to breathe clean and fresh air, but unfortunately, the air is often polluted, especially in big cities, posing a serious threat to human life.

At the same time, urban ecosystems can benefit from high air quality due to the presence of green areas, which contribute to increased humidity levels, air filtration cooling, oxygen production and reduction of greenhouse gas emissions (Barris, 2005). Zupancic et al. (2015), also supports this idea, highlighting that urban green spaces are widely recognized for their ability to naturally remove pollutants from the air and provide the city with a cooler and cleaner atmosphere, leading to the idea that one of the main advantages of green cities is by far the improvement of air quality.

Badiu et al. (2019) aimed to examine the effectiveness of the use of urban green space per capita as a quantitative indicator for assessing urban green infrastructure, which is considered a crucial element in improving the quality of life and creating a sustainable framework for cities. The case study is carried out for Romania, a country that has a target of 26 $m^2$ of urban green space per capita in all cities. A sample of 38 cities in Romania was drawn and compared with other databases. Multiple linear regression was used to determine the contribution of geographical and socio-economic factors to urban green space per capita, and multiple correspondence analysis was performed to identify differences between cities in terms of urban green space proportions. The results showed that the goal of 26 $m^2$ of green area per capita in all Romanian cities is not feasible and the characteristics of each city should be taken into account in order to achieve sustainability objectives. Hence, urban ecological planning should focus more on the development of urban green infrastructure models that are tailored to each type of urban area.

A study on the level of knowledge of urban society in Romania regarding the concept of urban green infrastructure showed that Romanians are not sufficiently informed about the meaning, implementation and management of this concept, which can partly explain the unsuccessful policies. This analysis explained the lack of successful policies and measures related to urban green areas in Romania and their decreasing coverage after the fall of communism (Gavrilidis, 2020).

This study was based on a survey that involved sending 970 questionnaires to target groups (citizens, public authorities, and economic agents), and 300 valid responses were collected, representing a validation rate of 30.9% (citizens 37.2%, public authorities 42.2%, and economic agents 20.6%). The analysis of the data collected from the three groups showed that 41% of the respondents were familiar with the concept of urban green infrastructure. Of those who are aware of this concept, the majority are public authorities (33% of the total responses and 78% within the group) and economic agents (13% of the total responses and 61% within the group). In contrast, respondents in the citizens' group are at the opposite end, with only 12% of the participants in this group claiming to have heard of this concept, which represents 5% of the total responses.

This represents a small percentage considering the importance of the subject and the fact that, in the context of climate change and the threat of urban heat waves, urban green spaces are essential to provide clean air, reduce noise and counteract high air temperatures through their cooling capacity. These green spaces can bring significant benefits in terms of urban quality of life, as well as in protecting the health and well-being of the population (Bolund and Hunhammar, 1999).

Green areas are often included in the environmental chapter of development strategies for Romanian cities, but there is a unbalanced approach. The emphasis is primarily on the types and sizes of green spaces, and local administrations frequently fail to provide solid strategies for their development and management. Ecological development strategies in Romania often contain unrelated targets and poorly defined indicators for monitoring progress. The urban green model is not traditionally integrated into the planning process of Romanian cities, focusing exclusively on the total area of green spaces, without considering their functionality and benefits (Raymond et al., 2017).

One of the reasons for these failures can be represented by Romania's history from a political regime perspective. Although there are factors that promote the development of green and equitable cities, such as EU policies, debates or funding, the availability, accessibility and attractiveness of urban green and blue spaces remain a low priority topic in former socialist countries. In these post-communist countries, including Romania, the concept of environmental justice is still at an early stage of development, hindered by the prevalence of market-oriented and neoliberal tendencies in society, privatization and the prioritization of private interests (Kronenberg, 2020).

## 3. Methodology

Bibliometric analysis was used to study the evolution of the concept of green city. A subfield of scientometrics is bibliometrics, which uses statistical and mathematical techniques to analyze the activity of academic publications (Popescu et al., 2022). With the help of bibliometrics, one can determine the interest in various themes, as well as the progress of certain concepts.

The bibliometric analysis carried out in this paper aims to identify the elements that are closely related to the concept of the green city, to evaluate the development of research on this topic and to identify the countries that have made a significant contribution to the advancement of this concept. An adequate and relevant database is crucial for the success of such an analysis. Therefore, Web of Science (WoS) is a highly comprehensive bibliographic database, suitable for many fields of study (Pranckutė, 2021). As it was the first international bibliographic database, over time, it has become the most important source of bibliographic data used for article selection, research appraisal and bibliometric analysis (Li et al., 2018).

In the current study, the top 2000 articles were selected, both those focusing on sustainable cities and those with green cities as their theme. The WoS database was used and the articles chosen fall within the Q1 and Q2 quartiles. A quartile is a category of scientific papers defined by bibliometric indicators that reflect the level of citation (García et al., 2012). As a result of this ranking, each article falls into one of four quartiles: Q1 reflects the top 25% most cited articles, and Q4 denotes the last 25% of articles in this ranking.

VOSviewer and R were used to analyse the 2000 articles. VOSviewer is a program designed for the creation, visualization, and analysis of bibliometric maps of science (van Eck and Waltman, 2010). With its help, co-occurrence analysis of keywords was carried out in order to determine the concepts that underlie the sustainable city or are interrelated with the term "green city". Thus, the colour of the node indicates which cluster the word corresponds to, the size of the node represents the frequency of the term's occurrence and the thickness of the line indicates the strength of the relationship between two words.

Using the Biblioshiny application from the R Bibliometrix package, the data needed to create graphs highlighting the evolution of the concepts of sustainable and green cities were downloaded. Therefore, the visual analysis platform, R, facilitated the creation of graphs that present the situation of the most cited countries, the progress of the use of certain keywords, the number of articles written over time, as well as the visualization of the most used keywords, all of which contributed to achieving the first objective of this paper.

To address the second objective of the paper, data for the year 2020 were collected using the OECD Statistics website for 28 cities, including 8 from Romania, 4 from Denmark and 16 from Canada. In the sample, these cities were selected respecting the quality assurance criteria of official statistics, such as: availability, comparability, accuracy and timeliness of statistical data. In addition, all cities from the 3 countries with available data based on OECD Statistics were selected. The cities from these countries were selected to achieve the main objective of the research, the identification of structural differences regarding the concept of a sustainable city. Therefore, the sample includes cities from two countries of the European Union (Romania and Denmark), but which present different models of economic development, Romania having a development model specific to ex-communist countries, while Denmark presents a development model economic center also on its environmental and sustainability issues. In addition, cities from Canada were also analyzed in this research, in order to have a clear overall picture of the differences between two countries of the European Union, which are located at the extremes as a model of economic development and an economically developed country from North America, such as Canada. For the notion of green city, data regarding the area of green spaces per capita were gathered, while for air pollution, the average population exposed to PM2.5 particles, which are particles smaller than 2.5 microns in diameter, were collected. Another variable for which data were collected is an economic one, represented by GDP per capita, and the fourth variable included in the analysis is related to digitalization, represented by the average downloading speed (Table no. 1).

**Table no. 1: Description of the variables included in the analysis**

| Variable name | Unit of measurement | Reference period | Source |
|---|---|---|---|
| Area of green spaces per capita | m2 per capita | 2020 | OECD Statistics |
| Average population exposed to PM2.5 | % | 2020 | OECD Statistics |
| GDP per capita | USD, constant prices, constant PPP, base year 2015 | 2020 | OECD Statistics |
| Average downloading speed | Deviation (%) from national average | 2020 | OECD Statistics |

*Source: Own processing using OECD Statistics website,* https://stats.oecd.org

Thus, the grouping of the cities taken into consideration was carried out according to the 4 variables chosen, using cluster analysis. Aimed at classifying the observations into groups that are different from each other but homogeneous within, this analysis allows for the investigation of similarities and differences between units (Căutișanu, 2016). Initially, a hierarchical clustering method was applied, represented by the Ward method, the objective for this being to identify the optimal number of clusters to be used in the next stage of the analysis. The best clustering method is the Ward method, because at a given iteration two classes are brought together that have the minimum Ward distance of all possible combinations (Cumatrenco, 2007). Another clustering method used is K-means, a non-hierarchical method, which aims to minimize the variation within clusters and maximize variation between clusters (Celebi et al., 2013). The Euclidean distance method is used for the K-means algorithm. Thus, there is initially a fixed number of clusters, and then iteratively the observations closest to the cluster centroid are added. The algorithm ends when the clusters do not change their configuration from one iteration to another.

As the variables do not have the same unit of measure and one of them has outliers, i.e. GDP per capita, the data were standardized with the range method in order to be used for the clustering algorithm. Additionally, using the SAS Studio website, both clustering methods were performed on the units included in the sample.

The random forest algorithm was proposed in 2001 by Leo Breiman and is a widely used machine learning method for classifying observations (Speiser et al., 2019). The method uses a number of randomly created decision trees, thus being uncorrelated, and makes predictions by averaging their decisions (Biau and Scornet, 2016). This algorithm has also helped achieve the second objective by providing new insights into the characteristics of the cities analyzed.

## 4. The evolution of the concepts of sustainable city and green city

For the purpose of visualizing the evolution of the two concepts, a bibliometric analysis was conducted, in which the first 2000 articles on sustainable cities were downloaded and uploaded into the VOSviewer program, enabling the clustering of keywords. The first cluster, the one in red, includes items referring to some of the sustainable development objectives, such as health, energy, transport, economy, climate change and land use. The second cluster, represented by the colour green, highlights that technology and innovation are at the heart of sustainable, eco and smart cities. The blue cluster indicates the need for digitization for a city to become sustainable (Figure 1).

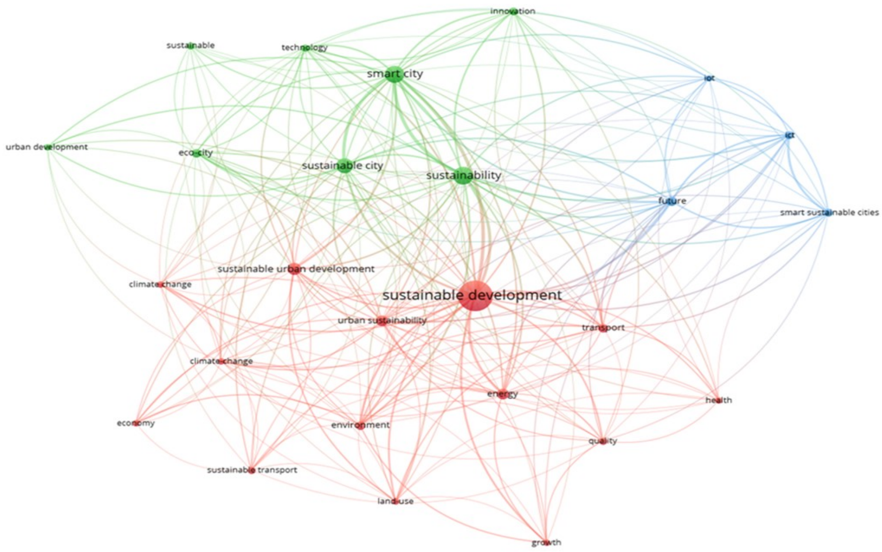


**Figure no.1: Sustainable city - clustering of keywords**
*Source: Own processing using VOSviewer based on Web of Science data*

Furthermore, the first 2000 articles on the subject of green cities were also downloaded and keyword clustering was performed using VOSviewer. The first cluster, with the nodes represented in red, contains items that have an impact on the environment, with some keywords being energy, land use, green infrastructure, climate change and sustainable development. The second cluster consists of terms related to nature, such as vegetation, parks, green spaces, biodiversity and conservation (Figure 2).


**Figure no.2: Green city - clustering of keywords**
*Source: Own processing using VOSviewer based on Web of Science data*

Out of the 2000 articles analyzed, approximately 300 of them were published in 2022. The publication of such articles gradually increased from 1960 to 2013, after which the phenomenon gained momentum and the importance devoted to this topic increased drastically. Thus, until 2005 very few articles were published, but from 2015 onwards, the number of publications focusing on the subject of green cities has increased exponentially (Figure 3).

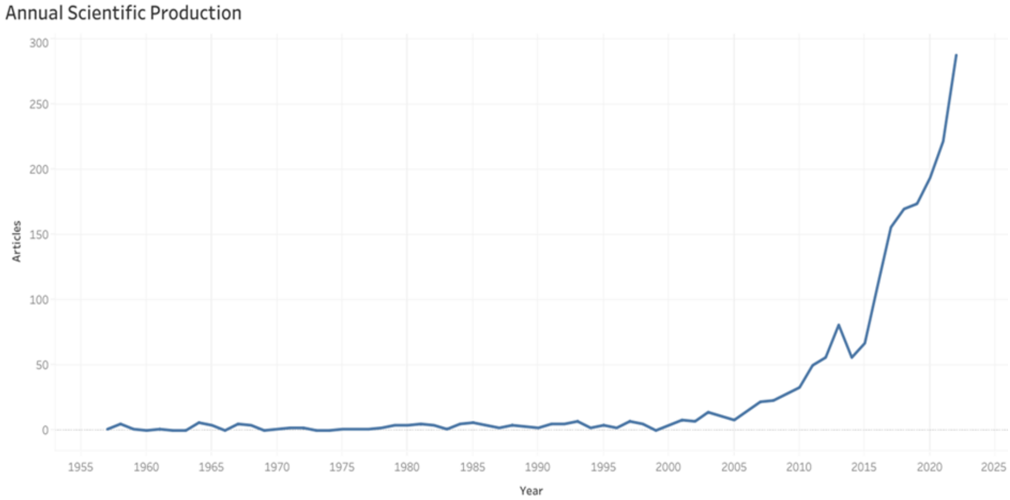


**Figure no.3: Evolution of articles on the green city from 1956 to 2022**
*Source: Own processing using RStudio based on Web of Science data*

Most articles were written in China, 1092, as this country has the largest population, but also a high interest in the subject due to numerous pollution-related issues. The second country is the United States, with 449 articles written, followed by Germany with 189 papers, the UK with 179, and Spain with 137 (Figure 4).

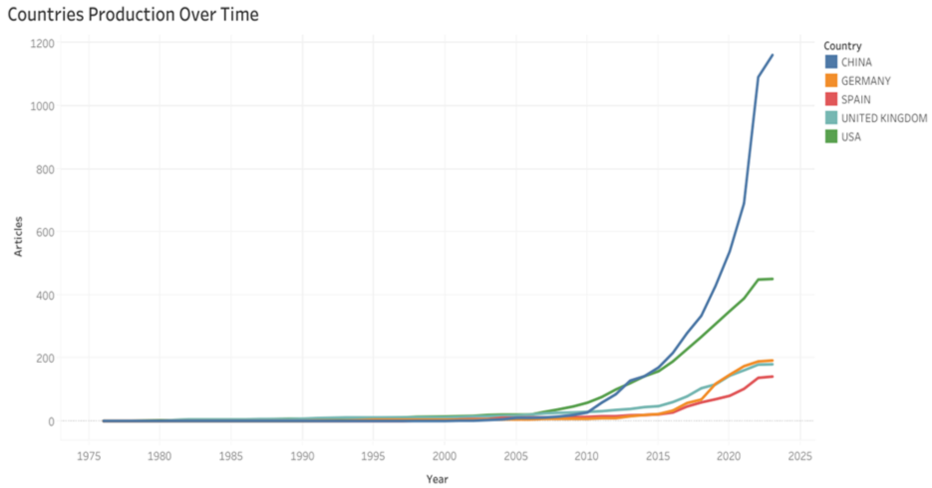


**Figure no.4: Evolution of green city articles by country, 1956-2022**
*Source: Own processing using RStudio based on Web of Science data*

The most cited countries in studies regarding green cities are China, with approximately 7500 citations, followed by the United States, Germany, Spain and Australia. Romania ranks 15$^{th}$ in this ranking with about 400 citations (Figure 5).

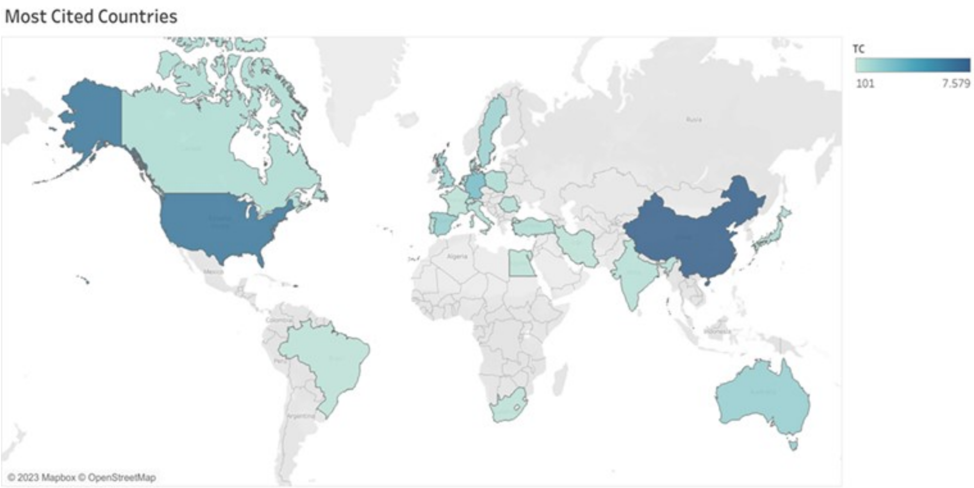

**Figure no.5: Most cited countries**
*Source: Own processing using RStudio based on Web of Science data*

An analysis of the 2000 articles downloaded shows that most authors are from China. Out of the 581 scientific papers, 85 85 of them were written in collaboration with authors from various countries. The United States ranks second, with 33 articles co-authored with foreign authors. In this ranking, Romania ranks 18$^{th}$, with only 21 articles written, out of which only 7 were written in partnership (Figure 6).

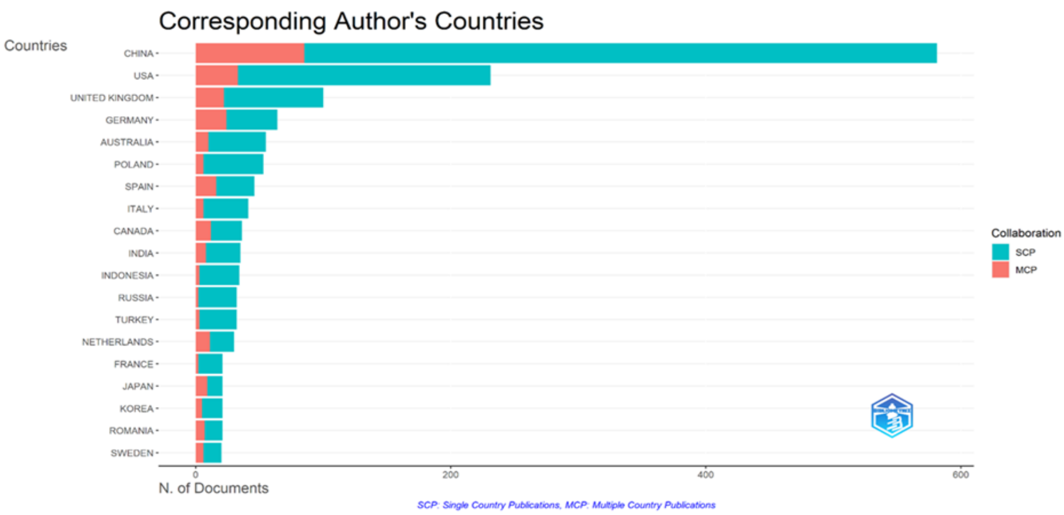

**Figure no.6: Authors' countries of origin**
*Source: Own processing using RStudio based on Web of Science data*

To describe the concept of a green city, a word cloud graphic was created. This graphic illustrates the frequency of certain words, suggesting what best describes a green city. The larger the word in the visual image, the more frequent it was in the analyzed articles.


**Figure no.7: Word Cloud**

*Source: Own processing using RStudio based on Web of Science data*

The concept of a green city is based on ideas such as green infrastructure (Figure 7), defined as an interconnected network of green spaces that preserves the roles and functions of natural ecosystems (Benedict and McMahon, 2002), green spaces, represented by trees, parks, etc., or sustainable development, defined as development that meets the needs of the present without compromising the ability of future generations to meet their own needs (WCED, 1987). These words best describe the concept of green cities addressed in the selected articles and highlight important features of a sustainable environment.

The statements made above are supported by the graph depicting the evolution of the concepts over time. Therefore, it can be observed that until 2016, the keyword "green space" was the most commonly found, and then the term green infrastructure gained momentum, becoming the most frequently used word (Figure 8).

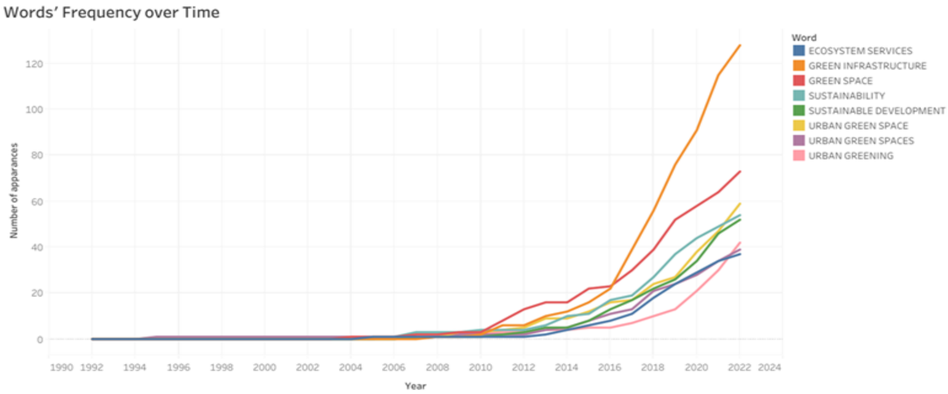


**Figure no.8: Word Cloud**

*Source: Own processing using Tableau based on Web of Science data*

## 5. Case study

A smart city is described as an intellectual capacity that encompasses various socio-technical and socio-economic aspects. These aspects lead to different types of smart cities, such as the green city, which focuses on urban infrastructure for environmental protection and reducing $CO_2$ emissions; the interconnected city, associated with digitization and broadband internet; the smart city, capable of generating value-added information through real-time data processing; the innovative city, and the knowledge city, which refers to the ability to generate innovation based on competent and creative human capital (Zygiaris, 2013).

The case study presents an analysis of the green city concept in Romania compared to two other countries, namely Denmark and Canada. These countries were chosen according to a ranking based on the Environmental Performance Index (EPI). In this ranking, Denmark occupies the first position, with the highest EPI score of 77.90. Romania ranks $30^{th}$ with a score of 56, while Canada is ranked 49th with a score of 50.

Environmental Performance Index (EPI) 2022 provides a data-driven summary of the state of sustainability worldwide. It utilizes 40 performance indicators across 11 issue categories and ranks 180 countries based on their performance in addressing climate change, environmental health, and ecosystem vitality. EPI provides scores that highlight countries' environmental performance and offers practical guidance for those aspiring towards a sustainable future (Wolf et al., 2022).

The overall ranking of the Environmental Performance Index (EPI) indicates the capacity of countries to adapt to environmental issues and brings to attention those that demonstrate optimal proficiency in tackling these challenges.

In the comparative analysis of these cities, the Ward method, a hierarchical clustering method based on the four grouping variables, was used. To determine the number of clusters in which the cities from Romania, Denmark, and Canada are grouped into, several criteria can be utilized. One such criterion is the Cubing Clustering Criterion (CCC), which suggests an optimal number of four clusters, reaching a maximum value of approximately 1.25 at point 4. Another criterion is represented by Pseudo F, which describes the ratio between the variation between clusters to the variation within a cluster. A low value of the F criterion indicates high within-group variation, while a high value suggests low within-group variation and high between-group variation. According to this criterion, the optimal number of clusters is three. The same conclusion is supported by the Pseudo T Squared criterion. Analyzing the graph from right to left, the most abrupt increase in the T Squared value is observed at point 3 (Figure 9).

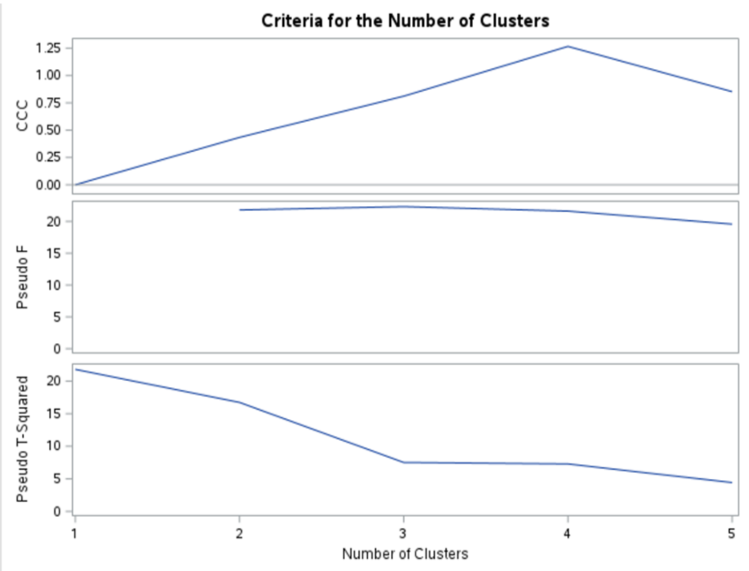


**Figure no.9: City classification tree**

*Source: Own processing using SAS based on OECD data*

Another tool used to characterize the clustering of cities is the classification tree, also known as a dendrogram. It serves as a criterion for selecting an optimal number of clusters and provides valuable insights about their composition. According to the dendrogram, 3 or 4 clusters can be outlined so that the variation between clusters is high and the variation within clusters is low.

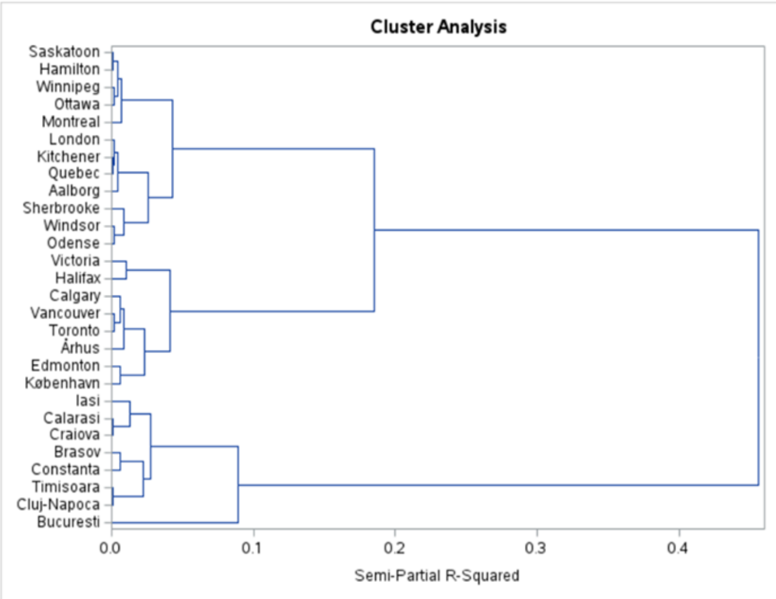


**Figure no.10: City classification tree**

*Source: Own processing using SAS based on OECD data*

Considering an optimal number of clusters equal to 4, from this graphical representation, it can be seen that Bucharest, the capital of Romania, exhibits distinctive characteristics as it forms a single cluster. Additionally, it becomes evident that cities from Romania group together in a separate cluster, showcasing unique qualities compared to the other cities in the two countries. Cluster 3 predominantly consists of Canadian cities, including the capital of Canada, Ottawa, Hamilton, Sherbrooke, Winnipeg, Montreal, London, Saskatoon, Quebec, Windsor, as well as two cities from Denmark, Odense and Aalborg. In cluster 4, the capital of Denmark, Copenhagen, is grouped together with Aarhus, another city from Denmark, along with seven Canadian cities: Vancouver, Toronto, Victoria, Halifax, Kitchener, Edmonton, and Calgary (Figure 10).

Another method that facilitates comparisons between the 28 cities is K-Means, a non-hierarchical clustering method. Since it has been determined that the optimal number of clusters using the Ward method is 4, the cities have been divided into 4 clusters. Therefore, the first cluster contains a single city, the second cluster contains 11 cities, the third cluster contains 7 cities, and the fourth cluster contains 9 cities. The cities in cluster 2 are similar to those in cluster 4, with the distance between the cluster centroids being 0.57 (Table no.2).

**Table no. 2: Description of Clusters**

| Cluster | Frequency | RMS Standard Offset | Nearest Cluster | Distance between clusters |
|---|---|---|---|---|
| 1 | 1 | - | 3 | 0.9360 |
| 2 | 11 | 0.1414 | 4 | 0.5664 |
| 3 | 7 | 0.1597 | 4 | 0.8389 |
| 4 | 9 | 0.4716 | 2 | 0.5664 |

*Source: Own processing based on SAS results*

The grouping into 4 clusters is determined by all 4 variables used in the analysis, as the R-Square values exceed the threshold of 0.5. The variable that contributes the most to the clustering of cities is the average population exposed to PM 2.5 particles, as it has the highest R-Square coefficient, 0.84. Another variable that has a considerable influence on the clustering is the average downloading speed, with an R-Square of 0.71 (Table no. 3).

**Table no. 3: Statistics for variables**

| Variable | R Square |
|---|---|
| GDP per capita | 0.686864 |
| Area of green spaces per capita | 0.666366 |
| Average downloading speed | 0.707869 |
| Average population exposed to PM 2.5 | 0.841903 |
| Total | 0.729335 |

*Source: Own processing based on SAS results*

The average values of the cluster variables help in making comparisons between the 4 groupings, and it is observed that Bucharest is the most polluted city, with reasonable green spaces per capita, the highest average downloading speed, and the highest GDP per capita. The cities in Romania, all belonging to cluster 3, have the lowest GDP per capita, few green spaces per capita, a good average downloading speed, and are moderately polluted. Additionally, the cities in cluster 2 have the greenest spaces per capita, the lowest average downloading speed, and are less polluted. The cities in cluster 4 are the least polluted, have a high GDP per capita, numerous green spaces and a very good average downloading speed (Table no. 4).

**Table no. 4: Cluster Average**

| Cluster | GDP per capita | Area of green spaces per capita | Average data download speed | Average population exposed to PM 2.5 |
|---|---|---|---|---|
| 1 | 1 | 0,74 | 0,95238 | 1 |
| 2 | 0,57295 | 0,6669 | 0,16234 | 0,16695 |
| 3 | 0,2326 | 0,14722 | 0,54762 | 0,68104 |
| 4 | 0,70061 | 0,55167 | 0,70106 | 0,12557 |

*Source: Own processing based on SAS results*

In order to classify the 28 cities, the Random Forest algorithm was also used. The classification error of the model is 25%, with 7 out of the 28 cities being incorrectly assigned to other clusters. Furthermore, it can be observed that the classification error is minimized when 500 decision trees are used (Figure 11).

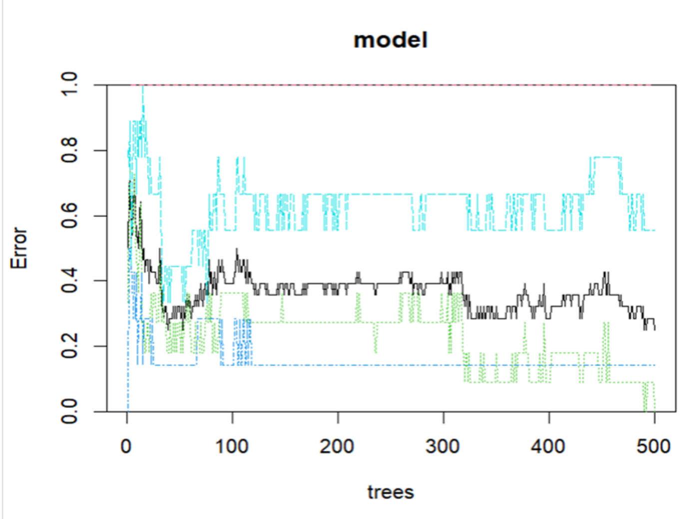


**Figure no.11: Evolution of model classification error as a function of the number of additional trees used**

*Source: Own processing using R based on clustering results*

It is observed that the city in the first cluster, Bucharest, has been classified into cluster 3, which consists of Romanian cities. This highlights the capital city's affiliation with a similar pattern as other cities in the country, indicating a potential issue with the development policies. Out of the 7 Romanian cities that constitute cluster 3, only one city has been assigned into a different cluster, namely cluster 4. This city is Constanța, and the misclassification could be attributed to the presence of many port cities in cluster 4, with the exception of Kitchener, Edmonton, and Calgary, indicating that the city located on the Black Sea coast shares similarities with cities like the capital cities of Denmark, København, or Århus, Cluster 2 has a classification error rate of 0%, meaning that all units belonging to that cluster have been correctly assigned. Over half of the observations in cluster 4 have been erroneously allocated to cluster 2, highlighting the fact that cities like Toronto, Vancouver, Calgary, Edmonton, and Victoria share common elements with cities such as Ottawa, Odense, Montreal, etc (Table no.5).

**Table no. 5: Cluster Average**

| | **1** | **2** | **3** | **4** | **Classification error** |
|---|---|---|---|---|---|
| **1** | 0 | 0 | 1 | 0 | 1 |
| **2** | 0 | 11 | 0 | 1 | 0 |
| **3** | 0 | 0 | 6 | 1 | 0,14 |
| **4** | 0 | 5 | 0 | 4 | 0,56 |

*Source: Own processing using R based on clustering results*

The table below shows that the average download speed makes the highest contribution to this classification, with an average decrease in the Gini coefficient of 5.1. This variable is followed by the average population exposed to PM 2.5 particles (4.92), while the green space area per capita contributes the least to the clustering of cities compared to the other three variables (Table no. 6).

**Table no. 6: Average decrease in Gini coefficient**

| Variable | Average decrease in Gini coefficient |
|---|---|
| GDP per capita | 4,18 |
| Area of green spaces per capita | 4,1 |
| Average downloading speed | 5,1 |
| Average population exposed to PM 2.5 | 4,92 |

*Source: Own processing using R based on clustering results*

As the classification model developed using the Random Forest algorithm has a rather large classification error of 25%, further considerations should be given to additional variables related to the concept of a green city, such as energy consumption per capita, water consumption per capita, proportion of potable water, proportion of CO2 emissions, etc. Additionally, a principal component analysis could be performed to reduce the dimensionality of the feature space. Furthermore, a larger sample size consisting of a greater number of cities would significantly contribute to improving the results of the method.

## 6. Conclusions

The current study aimed to track the evolution of interest in the theme of the green city and identify the characteristics of cities in Romania, Denmark, and Canada.

The bibliometric analysis facilitated a perspective on the interconnected elements of the green city concept as well as its popularity. Thus, it is observed that the assertions made by Breuste (2020) are supported by the co-occurrence analysis of keywords, where vegetation, green spaces, and parks are identified as the main elements underlying green cities.

This analysis also highlights the positive impact of these cities, as they contribute to sustainable development and improve the quality of life for residents through the creation of green infrastructure, a point emphasized by Badiu et al. (2019). Furthermore, while fewer than 25 articles were published on this topic between 1960 and 2005, the importance given to this subject has exponentially increased, with approximately 300 articles published in 2022 alone. Prior to 2013, the United States held the highest number of research papers on this topic, but China has since surpassed it, with 1,092 papers originating from the country with the largest population, indicating both the research capacity and interest in this subject. Regarding citation counts and collaborative works with authors from different countries, China comes on top first in these rankings, followed by the United States. The word cloud graph highlights the concepts underlying the green city, which include green infrastructure with significant roles in ecosystem conservation, green spaces contributing to residents' health, and sustainable development. Until 2016, the term "green infrastructure" was not widely used, with "green spaces" being the predominant term, indicating increased attention given to this concept in recent years.

The clustering of the 28 cities based on the four variables helped identify the characteristics of cities in Romania, Denmark, and Canada. Bucharest proved to be an exceptional case, forming a single cluster and being the most polluted city with the fewest green spaces per capita. However, it is also the city with the highest GDP per capita and the highest average downloading speed. The other seven cities in Romania, which form a distinct cluster, are both highly polluted and have few green spaces per capita. The capital of Romania shares similar characteristics with these cities considering these two variables. The difference lies in the fact that at the national level, Bucharest represents an extreme value in terms of quantitative analysis.

The cluster that includes the capital of Canada (Ottawa) and two of the four cities in Denmark included in the analysis (Aalborg and Odense), namely Cluster 2, have the highest number of green spaces per capita, are less polluted, and have the lowest average downloading speed. The cities that form Cluster 4, including the capital of Denmark (Copenhagen), have the lowest levels of air pollution, ample green spaces per capita, and a very good average downloading speed.

In addition to these findings, the random forest algorithm highlighted that Bucharest is similar to the other cities in Romania, suggesting a potential issue with government policies regarding the transformation of current cities into more sustainable and greener ones. One city in Romania that stands out is Constanța, which is classified in Cluster 4 along with numerous port cities such as Copenhagen, Aarhus, Vancouver, Toronto, Victoria, and Halifax. This highlights the similarity between the Romanian city and the other ports in the two countries concerning the four variables. However, since the classification error of the algorithm is quite high, this analysis can undoubtedly be improved, not only by adding more cities but also by increasing the number of relevant variables.

The originality of the work is represented by the simultaneous use of two methods of classification of statistical units. Through this iterative classification, homogeneous clusters and implicitly the development directions of the cities in the same cluster can be determined, especially in the case of cities that are classified both by the first method and by the second method in the same cluster. However, the research results present some limitations, such as the small units of the sample, but also of the variables intended to synthesize the relationship between economic development and sustainable development. Therefore, the research will be extended on a larger sample, which will include cities from several member states of the European Union, but also cities from the rest of the continents, in order to have a more complex analysis. Also, new variables will be added that will illustrate with greater accuracy the relationship between economic development and the sustainable development of urban settlements. In addition, in order to identify the influence of the international and national legislative framework, an exploratory factorial analysis will be carried out.

## References

Badiu, D.L., Onose, D.A., Niță, M.R. and Lafortezza, R., 2019. From "Red" to Green? A Look into the Evolution of Green Spaces in a Post-Socialist City. *Landscape and Urban Planning*, 187, 156–164. https://doi.org/10.1016/j.landurbplan.2018.07.015

Barıs, M., 2005. Urban planning, urban ecosystem and trees, planning, J.Chamber of City Plann

Benedict, M.A. and McMahon, E.T., 2002. Green infrastructure: smart conservation for the 21st century. *Renewable resources journal*, 20(3), pp.12-17.

Biau, G. and Scornet, E., 2016. Rejoinder on: A random forest guided tour. *Test*, 25, doi: https://doi.org/10.1007/s11749-016-0488-0

Bolund, P. and Hunhammar, S., 1999. Ecosystem services in urban areas. *Ecological economics*, 29(2), pp.293-301. doi: https://doi.org/10.1016/S0921-8009(99)00013-0

Breuste, J., 2023. The Green City: General Concept, Making Green Cities, 2023, doi: https://doi.org/10.1007/978-3-030-73089-5_1

Căutișanu, C., 2016. Analiza cluster a procesului de îmbătrânire demografică în euroregiunea „Siret-Prut-Nistru", dezvoltarea economico-socială durabilă

Celebi, M.E., Kingravi, H.A. and Vela, P.A., 2013. A comparative study of efficient initialization methods for the k-means clustering algorithm, *Expert Systems with Applications*, doi: https://doi.org/10.1016/j.eswa.2012.07.021

Cetin, C. and Karafakı, F.C., 2019. The influence of green areas on city-dwellers'perceptions of air pollution: The case of Nigde city center, *Journal of Environmental Biology*, doi: http://doi.org/10.22438/jeb/41/2(SI)/JEB-26

Cohen-Shacham, E., Walters, G., Janzen, C. and Maginnis, S. (eds.), 2016. Nature-based solutions to address global societal challenges. *IUCN: Gland, Switzerland*, xiii + 97pp https://doi.org/10.2305/IUCN.CH.2016.13.en

Cumatrenco, A., 2007. Algoritmi de analiză cluster de tip ierarhic, Economica

Dahmann, N., Wolch, J., Joassart-Marcelli, P., Reynolds, K. and Jerrett, M., 2010. The active city? Disparities in provision of urban public recreation resources. *Health & place*, 16(3), pp.431-445. doi. https://doi.org/10.1016/j.healthplace.2009.11.005

Dur, F., Yigitcanlar, T. and Bunker, J., 2014. A spatial-indexing model for measuring neighbourhood-level land-use and transport integration. *Environment and Planning B: Planning and Design*, 41(5), pp.792-812., doi: https://doi.org/10.1068/b39028

García, J. A., Rodriguez-Sánchez, R., Fdez-Valdivia, J. and Martinez-Baena, J., 2012. On first quartile journals which are not of highest impact, *Scientometrics* volume, 90, doi: https://doi.org/10.1007/s11192-011-0534-3

Gavrilidis, A.-A., 2020. Planning the "unknown": Perception of urban green infrastructure concept in Romania. Urban Forestry & Urban Greening, Volume 51

Kronenberg, J. (2020), Environmental justice in the context of urban green space availability, accessibility, and attractiveness in postsocialist cities, *Cities*, Volume 106

Li, K., Rollins, J. and Yan, E., 2018. Web of Science use in published research and review papers 1997–2017: a selective, dynamic, cross-domain, content-based analysis, *Scientometrics*, 115, doi:10.1007/s11192-017-2622-5

Van Eck, N. and Waltman, L., 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), pp.523-538. DOI: 10.1007/s11192-009-0146-3

Nutsford, D., Pearson, A.L. and Kingham, S., 2013. An ecological study investigating the association between access to urban green space and mental health. Public health, 127(11), pp.1005-1011. doi: https://doi.org/10.1016/j.puhe.2013.08.016

Popescu, D.V., Dima, A., Radu, E., Dobrotă, E.M., and Dumitrache, V.M., 2022. Bibliometric Analysis of the Green Deal Policies in the Food Chain. *Amfiteatru Economic*, 24(60), pp. 410-428. doi: 10.24818/EA/2022/60/410

Pranckutė, R., 2021. Web of Science (WoS) and Scopus: The Titans of Bibliographic Information in Today's Academic World. *Publications*, 9(1), doi: https://doi.org/10.3390/publications9010012

Raymond, C.M., Frantzeskaki, N., Kabisch, N., Berry, P., Breil, M., Nita, M.R., Geneletti, D. and Calfapietra, C., 2017. A framework for assessing and implementing the co-benefits of nature-based solutions in urban areas. *Environmental Science & Policy*, 77, pp.15-24. doi: https://doi.org/10.1016/j.envsci.2017.07.008

Speiser, J.L., Miller, M.E., Tooze, J. and Ip, E., 2019. A comparison of random forest variable selection methods for classification prediction modeling. *Expert systems with applications,* 134, pp.93-101. doi: https://doi.org/10.1016/j.eswa.2019.05.028

WCED, U. 1987. *Our common future—The Brundtland report, Report of the World Commission on Environment and Development,*

Wolf, M. J., Emerso, J. W., Esty, D. C., de Sherbinin A., Wendling, Z. A. and al., 2022. *2022 Environmental Performance Index*, New Haven, CT: Yale Center for Environmental Law & Policy. epi.yale.edu

Wüstemann, H., Kalisch, D. and Kolbe, J., 2017. Access to urban green space and environmental inequalities in Germany. *Landscape and Urban Planning*, Volume 164, pp.124-131. doi: https://doi.org/10.1016/j.landurbplan.2017.04.002

Zupancic, T., Westmacott, C. and Bulthuis, M., 2018. *The Impact of Green Space on Heat and Air Pollution in Urban Communities: A Meta-Narrative Systematic Review*. [online]. David Suzuki Foundation. Available at: https://davidsuzuki.org/science-learning-centre-article/impact-green-space-heat-air-pollution-urban-communities/ [Accessed June 2023].

Zygiaris, S., 2013. Smart city reference model: Assisting planners to conceptualize the building of smart city innovation ecosystems. Journal of the Knowledge Economy, 4, pp.217-231., doi: https://doi.org/10.1007/s13132-012-0089-4