

## ГИС ПРОГРАМИРАЊЕ – АУТОМАТИЗАЦИЈА ИЗРАЧУНА ХЕРФИНДАЛ-ХИРШМАНОВОГ ИНДЕКСА РЕГИОНАЛНЕ СПЕЦИЈАЛИЗАЦИЈЕ

Даворин Бајић<sup>1</sup>

<sup>1</sup>Природно-математички факултет, Универзитет у Бањој Луци, Република Српска

**Сажетак:** У раду је описан поступак аутоматизације израчуна Херфиндал–Хиршмановог индекса регионалне специјализације. ГИС програмирањем, коришћењем *Python* програмског језика, развијен је алат који омогућава аутоматизован начин израчуна и визуелизације резултата разматраног индекса, у *ArcGIS* софтверском окружењу. Поред дефинисања основних појмова везаних за регионалну специјализацију и описа математичког концепта Херфиндал-хиршмановог индекса, у раду је дат преглед комплетног програмерског поступка креирања алата за аутоматизацију кроз опис поступка писања кода. У резултатима рада објашњен је начин и практичан примјер примјене креираног алата на примјеру Републике Српске.

**Кључне ријечи:** ГИС, програмирање, Херфиндал–Хиршманов индекс, регионална специјализација, Python

Original scientific paper

## GIS PROGRAMMING – AUTOMATED CALCULATION OF THE HERFINDAHL-HIRSCHMAN INDEX OF REGIONAL SPECIALIZATION

Davorin Bajić<sup>1</sup>

<sup>1</sup>Faculty of Sciences, University of Banja Luka, Republika Srpska

**Abstract:** The paper describes the automated calculation of Herfindahl–Hirschman Index of regional specialization. Having used GIS programming and *Python* programming language, a tool has been designed to enable the automated calculation and visualization of the results of the observed index within *ArcGIS* software environment. Apart from the definition of basic regional specialization terms and the description of mathematical concept of Herfindahl–Hirschman Index, the paper outlines the programming procedure for designing a tool for the automated calculation by describing the code writing procedure. The results provide an explanation and a practical instance of the tool application in Republic of Srpska.

**Key words:** GIS programming, Herfindahl–Hirschman Index, regional specialization, Python.

---

### УВОД

Регионална специјализација одражава расподјелу секторских удјела (удјела дјелатности, подручја дјелатности, привредних грана) у укупној економији одређене територијалне јединице, обично у односу на остатак државе или регије. У статистичком смислу регионална специјализација се може

### INTRODUCTION

Regional specialization refers to the distribution of sectors (ratio of an activity, area of an activity or commercial branches) within the total economy of a specific territorial unit, usually in comparison with the rest of the state or region. Statistically, regional specialization may be defined as

дефинисати као дистрибуција "тежине" (број запослених, приходи и сл.) сектора дјелатности "i" у укупној привредној структури посматране територијалне јединице "j" (Aiginger, 1999). Појам регионална специјализација се раније употребљавао скоро искључиво за објашњавање регионалних односа у индустријском сектору, поготово у блоку социјалистичких држава које су фаворизовале овај сектор привреде. У новије вријеме појам регионалне специјализације се проширује и на остале секторе привреде. Појављује се у новим концептима, теоријама и правцима у регионалној економији, којима се покушавају објаснити регионални односи укупне економске расподеле, као што је концепт "Нове економске географије" нобеловца П. Кругмана (Krugman, 1991; Fujita et al., 1999). Велики број радова и студија је написан на тему регионалне специјализације. Због теоријског и методолошког доприноса у расвјетљавању наглашеног феномена нарочито треба истаћи радове сљедећих аутора: Corrado Gini (Gini, 1936), Masahisa Fujita, Paul Krugman i Anthony Venables (Fujita et al., 1999), Glen Ellison i Edward Glaeser (Ellison et al. 1997) итд.

Регионална специјализација је у великој мјери статистичка категорија коју је могуће квантификовати. Квантификацијом регионалне специјализације бавио се велики број истраживача, тако да је конструисан и велики број мјерних "алата" помоћу којих се врши поменута квантификација. Углавном се ради о различитим индексима и коефицијентима помоћу којих се покушавају разјаснити регионални односи у погледу специјализације или диверзификације економије, или појединих њених сектора. Индекси за мјерење регионалне специјализације су веома значајни у истраживањима у областима просторне економије и економске географије. Неки од значајнијих индекса помоћу којих је могуће вршити мјерење регионалне специјализације су: Krugman Index (Krugman, 1991), Hoover-Ballasa Index (Ballasa, 1965), Gini index (Gini, 1936), Hirschman-Herfindahl index (Hirschman, 1964), Theil index (Theil, 1967), Lilien index (Lilien, 1982), Индекс просторне концентрације

the distribution of 'weight' (number of employees, income, etc.) of an 'i' service sector within the total commercial structure of an observed territorial unit 'j' (Aiginger, 1999). The term regional specialization used to refer to regional relations within the industrial sector exclusively, especially in Socialist states which used to prefer this commercial sector. Lately, the term expanded to other commercial sectors. It appears within new concepts, theories, and movements of regional economy, which attempt to account for the regional relations of the total economic distribution such as in Nobel prize winner Krugman's concept of 'New economic geography' (Krugman, 1991; Fujita et al., 1999). There have been many papers and studies discussing regional specialization. In addition, the following authors largely contributed the field wither theoretically or methodologically: Corrado Gini (Gini, 1936), Masahisa Fujita, Paul Krugman and Anthony Venables (Fujita et al., 1999), and Glen Ellison and Edward Glaeser (Ellison et al. 1997).

Regional specialization is largely a statistical category that may be quantified. Many researchers have dealt with the quantification of regional specialization, which resulted in a large number of quantification 'tools'. The tools usually refer to different indices and coefficients that attempt to explain regional relations from the aspect of economy specialization or diversification (or some of the sectors). Indices for the calculation of regional specialization are highly pertinent for the spatial economy and economic geography studies. Some of the top indices are as follows: Krugman Index (Krugman, 1991), Hoover-Ballasa Index (Ballasa, 1965), Gini index (Gini, 1936), Hirschman-Herfindahl index (Hirschman, 1964), Theil index (Theil, 1967), Lilien index (Lilien, 1982), Index of spatial concentration of social-economic features (Бајић et al. 2013), etc.

Still, the calculation of some of these

социо-економских обилежја (Бајић et al. 2013) итд.

Поступак израчуна поменутих индекса често зна бити компликован у смислу израчуна, јер неки од њих захтијевају сложеније математичке и статистичке операције. Понављање поступка израчуна за индексе који се често употребљавају одузима пуно времена. Приказивање резултата израчуна индекса регионалне специјализације у класичном статистичком облику, у форми табела и дијаграма, није ефикасно, у смислу читљивости и анализе резултата. Обзиром да се приказују регионални/просторни односи резултате је најефикасније приказивати картографски у виду картограма и картодијаграма. Процес картографске интерпретације резултата у традиционалној картографији, такође, представља дуготрајан процес. Појавом ГИС технологија наведене процесе израчуна и интерпретације (визуелизације) статистичких мјера, као што су индекси регионалне специјализације, у великој мјери је могуће аутоматизовати. Аутоматизација израчуна и визуелизације се спроводи креирањем софтверских алата у ГИС окружењу. Једном креиран алат могуће је примијенити неограничено многу пута, на неограниченом броју рачунара. Креирањем алата за аутоматизацију омогућава се израчун индекса и приказивање резултата "на један клик".

У овом раду је приказан аутоматизован начин израчуна Хефиндал-Хиршмановог индекса (*HHI*). *HHI* се често употребљава као мјера регионалне специјализације. Поред тога што се користи као мјера регионалне специјализације, *HHI* се употребљава и као мјера величине предузећа у односу на укупну величину односног сектора привреде тј. као показатељ конкурентског односа међу предузећима у датом сектору привреде.

Аутоматизација израчуна и визуелизације резултата *HHI* реализована је креирањем алата унутар *ArcGIS* софтверског окружења. *ArcGIS* је једна од најзаступљенијих ГИС софтверских апликација у Свијету коју је развила фирма *ESRI*. *ArcGIS* омогућава креирање алата у свом

indices may be rather complicated as they may demand more complex mathematical and statistical operations. The repetition of some most common calculations takes time. The display of results of regional specialization indices in a classical statistical form via tables and diagrams complicates the interpretation and analysis of results. Since these are usually regional/spatial relations, it is most efficient to display the results by mapping them via cartograms and cart-diagrams. The process of map interpretation also takes too much time. Once GIS technology emerged, it became possible to automate the calculation and interpretation (visualization) of the statistical measures, such as the index of regional specialization. The automated calculation and visualization is conducted by creating software tools in GIS environment. Once the tool is created, it may be applied without limitations regarding the repetition or the number of computers. The creation of automated tools enables the index calculation and result display in only one click on the keyboard.

This paper presents the automated calculation of Herfindahl–Hirschman Index (*HHI*). *HHI* is widely used as measurement of regional regionalization. Apart from being used as measurement of regional specialization, it is also used as the measurement of enterprise size in comparison with the total size of referential commercial sector, i.e. as an indicator of competitive relation among enterprises within a specific commercial sector.

The automated calculation and visualization of *HHI* results was performed via a tool designed in the *ArcGIS* software environment. *ArcGIS* is one of most frequent GIS software applications worldwide developed by *ESRI* Company. *ArcGIS* enables the creation of tools via the following programming languages: *.NET*, *Python*, *C#* and *VBA*. GIS programming is becoming more and more important. The need for new tools and performance of complicated geo-

окружењу кориштењем програмских језика као што су: *.NET*, *Python*, *C#* и *VBA*. Примјена програмирања у ГИС-у добија све већи значај. Потреба за новим алатима и извршавањем компликованих геопросторних анализа у први план ставља програмерска знања. Обзиром на значај програмирања у ГИС-у у програмерској заједници развија се посебна област која се назива ГИС програмирање. ГИС програмирањем, кориштењем *Python* програмског језика, креиран је алат за израчун и визуелизацију резултата *HHI*.

*Python* (Пајтон) је интерпретерски, интерактивни, објектно-оријентисани програмски језик. *Python* код спрема се у текстуалне датотеке које завршавају на *.py* и које се могу преносити на било које платформе гдје се могу изводити. Осим стандардних типова података (бројеви, низови знакова и сл.) *Python* има уграђене типове података високог нивоа, као што су: листе, рјечници и сл. *Python* се може изводити у различитим окружењима. За развој програма најлакши је интерактивни начин рада у којем се програмски код пише наредбу за наредбом. Не постоји разлика у развојном и изведбеном начину: у првом се изводи наредба за наредбом, а у другом одједном читава скрипта. Синтакса језика је једноставна и чиста, тако да је програм прегледан и једноставан за читање. (Essert, 2007)

Поред наведених карактеристика *Python* програмског језика треба истаћи да његов објектно-оријентисани концепт омогућава читавање модула при рјешавању специфичних програмерских задатака. Пајтонова библиотека, при стандардној инсталацији, укључује преко 200 модула, што покрива све од функција оперативних система до WEB сервиса. Модули су својеврсни подпрограми који олакшавају и скраћују вријеме писања кода. Довољно је да се модули позову у скрипту преко одговарајућих команди чиме се аутоматски могу користити њихове функције у коду који развијамо. У програмирању алата за аутоматизацију израчуна *HHI* употребљена су два модула: *ArcPy* и *NumPy*.

*ArcPy* (<http://blogs.esri.com/esri/>

spatial analysis demand programming skills. Since programming has become a crucial part of GIS technology, a new GIS programming field emerged. GIS programming used *Python* programming language in order to create the tool for calculation and visualization of *HHI* results.

*Python* is an interpretive, interactive, object-oriented programming language. *Python* code is stored in textual computer files ending in *.py* and which can be transferred to any platform in which they can be executed. Beside standard data (numbers, sign sequences, etc.), *Python* has the integrated high level data such as lists, dictionaries, etc. It is possible to execute *Python* in different environment. For the development of the program, the interactive work mode in which program code writes one order after another is best suitable. There is no difference between the design and executive modes – the first performs one order after another and the latter deals with the whole script. The language syntax is simple and clear, which makes the program readable and synoptic (Essert, 2007).

Apart from the aforementioned *Python* features, we should say that its object-oriented concept enables module loading while solving specific program tasks. *Python* library, in standard installation, includes 200 modules, which covers both the operative system functions and WEB services. Modules are actually subprograms that make code writing easier and faster. All it takes is to withdraw the modules into the script by using specific commands, which automatically enables the usage of all the functions of the code. *ArcPy* and *NumPy* modules were used to design the tool for automated *HHI* calculation.

*ArcPy* (<http://blogs.esri.com/esri/arcgis/2010/02/12/what-is-arcpy/>) is a module for *Python* programming language intended for the application within *ArcGIS* software environment. Its basic task is to enable the efficient and productive geo-spatial data analysis, data conversion, data management, and automated data visualization. *NumPy*

arcgis/2010/02/12/what-is-arcpy/) је модул за *Python* програмски језик, намијењен за примјену унутар *ArcGIS* софтверског окружења. Основна улога овог модула је омогућавање ефикаснијих и продуктивнијих геопросторних анализа података, конверзије података, ефикасније управљање подацима и аутоматизована визуелизација података. *NumPy* (<http://www.numpy.org/>) је модул за програмски језик *Python* који омогућава подршку за велике вишедимензионалне низове и матрице, те сложене математичке функције и операције (Ascher et al., 1999).

(<http://www.numpy.org/>) is a module for *Python* programming language providing support for large multidimensional sequences and matrices, and complex mathematical functions and operations (Ascher et al., 1999).

## МАТЕРИЈАЛИ И МЕТОДЕ

Програмирање ГИС алата за израчун и визуелизацију засновано је на математичком концепту *HHI*. У математичком смислу *HHI* представља суму квадрата удјела дјелатности *j* у укупној привредној структури одређене територијалне јединице *i*. Математички израз разматраног индекса има сљедећу форму:

$$HHI = \sum_{j=1}^n (g_{ij})^2$$

$$g_{ij} = \frac{X_{ij}}{\sum_{j=1}^n X_{ij}} = \frac{X_{ij}}{X_i}$$

Гдје су:

- *HHI* - Херфиндал-Хиршман индекс просторне специјализације;
- *i* - Територијална јединица;
- *j* - Дјелатност/привредна грана/подручје дјелатности
- $X_{ij}$  - Број запослених (или други показатељ развијености) у дјелатности *j* и територијалној јединици *i*;
- $X_i$  - Број запослених (или други показатељ развијености) у територијалној јединици *i*;
- $g_{ij}$  - Удио дјелатности *j* у укупној привредној структури територијалне јединице *i*.

## MATERIALS AND METHODS

The GIS tool programming for the calculation and visualization is based on the mathematical *HHI* concept. Mathematically, *HHI* represents the sum of squared ratio of trade *j* within the total commercial structure of a specific territorial unit *i*. The mathematical expression of the observed index is as follows:

$$HHI = \sum_{j=1}^n (g_{ij})^2$$

$$g_{ij} = \frac{X_{ij}}{\sum_{j=1}^n X_{ij}} = \frac{X_{ij}}{X_i}$$

In which:

- *HHI* - Herfindahl–Hirschman Index of spatial specialization;
- *i*–territorial unit;
- *j* - trade/commercial branch/area of activity
- $X_{ij}$ –Number of employees (or another indicator of development) within the trade *j* and territorial unit *i*;
- $X_i$ –Number of employees (or another indicator of development) within the territorial unit *i*;
- $g_{ij}$ –Ratio of trade *j* within the total commercial structure of the territorial unit *i*.

Вриједности *HHI* могу се кретати у интервалу 0-1. Што су вриједности индекса веће од 0 то је степен специјализације већи, и обрнуто. Веће вриједности индекса указују на специјализацију, док мање вриједности указују на диверзификацију посматране привредне структуре у датој територијалној јединици. Дефинисани показатељ одражава степен диверзификационе уравнотежености у погледу броја и удјела структурних елемената у укупној структури појединих територијалних јединица. Степен диверзификационе уравнотежености је већи што је већи број структурних елемената, те што је њихов величински удио у структури посматране територијалне јединице уједначенији. У том случају његове вриједности су дистрибуисане око вриједности 0. Супротно, уколико је број структурних елемената мањи, те уколико је њихов величински удио у привредној структури одређене територијалне јединице неуједначенији (већа концентрација броја запослених у једној или мањем броју дјелатности) можемо говорити о специјализацији. У дефинисаном случају вриједности су веће од 0, што су вриједности ближе граничној вриједности 1 степен специјализације је већи. Показатељ не представља величинску категорију којом би се могао мјерити и упоређивати степен развијености структуре међу посматраним територијалним јединицама. Он је само мјера диверзификационе уравнотежености структурних елемената унутар једне посматране територијалне јединице, што показује степен регионалне специјализације унутар те територијалне јединице. На једној страни, индекс одражава степен структурне диверзификације/специјализације у свакој територијалној јединици засебно, а са друге стране, пропорционалне односе (не величинске) између територијалних јединица у погледу удјела појединих елемената у укупној структури. Рачунањем индекса за два или више временских периода могуће је утврдити правце и интензитет промјена регионалне специјализације.

Након дефинисања математичког концепта

*HHI* values may range from 0 to 1. The larger the index values from 0, the higher the specialization, and vice versa. Higher index values indicate the specialization, whereas lower values indicate the diversification of the observed commercial structure within a given territorial unit. The defined indicator reflects the level of diversified balance with reference to number and ratio of the structural elements within the total structure of specific territorial units. The level of diversified balance grows as the number of structural elements is bigger and the element ratio within the structure of the observed territorial unit is more equal. In this case, its values are distributed around the value of 0. Opposite to this, if the number of structural elements is smaller and the element ratio within the structure of the observed territorial unit is less equal (higher concentration of employees within one or a few trades), then we refer to the phenomenon as the specialization. Within this defined case, values are higher than 0 and the closer the values are to the limiting value of 1, the level of specialization is higher. The indicator does not represent a quantity category with which we may measure and compare the level of development of a structure among the observed territorial units. It is only a measure of diverse balance of structural elements within one observed territorial unit, which is indicated by the level of regional specialization within that territorial unit. On one hand, the index expresses the level of structural diversification/specialization within each separate territorial unit; on the other hand, it represents the proportional relation (not quantified ones) among the territorial units with reference to the ratio of individual elements within the total structure. By calculating the index for two or more time intervals, it is possible to estimate the direction and intensity of changes in regional specialization.

Once we defined the *HHI* mathematical concept, we designed the GIS tool for its automatization. The GIS tool was written as

*НИИ* приступило се креирању ГИС алата за његову аутоматизацију. ГИС алат развијен је као скриптни алгоритам, коришћењем објектно-оријентисаног програмског језика *Python*, уз коришћење *ArcPy* и *NumPy* модула у *ArcGIS* окружењу. Писање кода и развој скриптог алгоритма извршено је у *IDLE (Python)* развојном окружењу. У наставку је дат преглед изворног кода по сегментима и линију по линију са описом његових извршних функција.

На почетку скриптог алгоритма треба написати наредбу која "позива" модуле који ће бити коришћени у скрипти. Обзиром да су за развој дефинисаног алата коришћени модули *ArcPy* и *NumPy* у конкретном случају наредба за "позивање" модула има сљедећи облик:

```
1. import arcpy, numpy as np
```

Након наредбе за учитавање модула дефинисан је дио кода који дефинише параметре који се користе у алгоритму. Дефинисање параметара извршено је наредбом *arcpy.GetParameterAsText* која је садржана у модулу *ArcPy*. Ова наредба омогућава да се параметри накнадно позивају у извршном дијелу алгоритма, што значи да не морају бити дефинисани у самој скрипти. Дио кода којима се дефинишу параметри у нашем случају има сљедећи облик:

```
3. ulazna_tabela = arcpy.GetParameterAsText(0)
4. selektovanaTabela = arcpy.GetParameterAsText(1)
5. odabir_sqlUslov = arcpy.GetParameterAsText(2)
6. vektor_zaPridruzivanje = arcpy.GetParameterAsText(3)
7. input_join_field = arcpy.GetParameterAsText(4)
8. output_join_field = arcpy.GetParameterAsText(5)
9. fc = selektovanaTabela
```

Дефинисано је шест параметара: табела са улазним подацима (*ulazna\_tabela*), табела са селектованим подацима из улазне табеле (*selektovanaTabela*) - излазна табела, дефинисање упита за селектовање података из улазне табеле (*odabir\_sqlUslov*), векторски слој са границама територијалних јединица које ће бити укључене у анализу (*vektor\_zaPridruzivanje*), индексирана колона у табели

a script algorithm by using the object-oriented *Python* programming language and *ArcPy* and *NumPy* modules in *ArcGIS* environment. The code writing and development of script algorithm was performed in *IDLE (Python)* development environment. What follows is the overview of the source code per segments and line by line with the description of executive functions.

At the beginning of the script algorithm, a command should be written to 'sum up' all the modules that will be used within the script. As we used *ArcPy* and *NumPy* modules for the development of our tool, the 'sum up' command should be as follows:

After the module command, we defined a part of the code which determined the parameters used in the algorithm. The definition of the parameters was performed via *arcpy.GetParameterAsText* command, which *ArcPy* module already contained. This command enables the parameters to be later summed up during the algorithm execution, which means that they do not have to be defined within the script itself. The part of the code that defines the parameters in our specific case is as follows:

We defined six parameters: table with input data (*input\_table*), table with selected data from the input table (*selected Table*)—output table, definition the inquiry for the input table data selection (*odabir\_sqlUslov*), vector layer with the territorial unit borderlines that will be included into the analysis (*vector\_Mapping*), an indexed column in the attribute table with the input

атрибута векторског слоја (`input_join_field`), индексирана колона у табели са улазним подацима (`output_join_field`).

Сљедећа наредба (`arcpy.TableSelect_analysis`) која је употријебљена за писање кода такође је из модула *ArcPy*. Користи се за селектовање дијелова табела према задатом услову, путем креирања *SQL* упита, те креирање нове (излазне) табеле из селектованих података. Овом командом се омогућава да се из улазних табела селектују и користе само подаци који су нам потребни у анализи. Дакле, овом командом одабиремо из табеле улазних података податке које ћемо укључити у израчун *HHI*, нпр. ако у табели улазних података имамо податке о броју запослених по дјелатностима за већи низ година можемо одабрати само податке за једну годину коју желимо да укључимо у израчун. Дио кода који описује претходно дефинисано састоји се из сљедећег:

```
11. arcpy.TableSelect_analysis(ulazna_tabela, fc, odabir_sqlUslov)
```

У наредном дијелу кода уз употребу *Python for* петље (*for loops*), *if* изјава (*if statements*), те *ArcPy* наредби (`arcpy.ListFields`, `arcpy.AddField_management`), извршава се креирање нове колоне у новокреираној табели селектованих података - излазној табели. Овај дио кода креира колону под називом "HHI" у којој ће бити уписани резултати израчуна Херфиндал-Хиршмановог индекса. Код провјерава да ли у селектованој табели постоји колона са називом "HHI", ако не постоји креира се, ако постоји, програм даје обавјештење да колона већ постоји тако да треба провјерити селектовану табелу, са поновним селектовањем улазних података.

```
13. for field in arcpy.ListFields(fc):
14.     if field.name != 'HHI':
15.         arcpy.AddField_management(fc, 'HHI', 'FLOAT', "", "", "", "", "NULLABLE")
16.     else:
        arcpy.AddWarning("HHI already exists!")
```

Новокреирана колона по свим редовима у излазној табели сетована је на вриједности *NULL*. У програмском језику *Python* табеларне вриједности сетоване на вриједности *NULL* знају правити проблем код калкулација унутар табела и ажурирања истих. Стога је написан

`data` (`output_join_field`).

*ArcPy* module also contained the next command(`arcpy.TableSelect_analysis`) used for the code writing. It is used for the selection of table parts in accordance with the given order via a *SQL* inquiry and for the creation of a new (output) table on the basis of selected data. This command helps select and use only those table data that are necessary for the analysis. Therefore, by using this command, from the output table we select only the data necessary for the *HHI* calculation, e.g. if the input table contains data on the number of employees per a trade over many years, it is possible to select data for only one year and include them into the calculation. The part of the code referring to this is as follows:

What follows is that we use *Python for* loops, *if* statements, and *ArcPy* commands (`arcpy.ListFields`, `arcpy.AddField_management`), and create a new column within this new table of selected data—output table. This part of the code creates a column named "HHI" which will contain the results of Herfindahl–Hirschman Index. The code then searches the selected table for the "HHI" column; if it does not exist, it is then created; if it exists, the program sends a notification that the column is already there and that the selected table should be checked and the input data should be reelected.

The newly-created column is set to *NULL* values in all the lines within the output table. In *Python* programming language, if the table values are set to *NULL*, they may cause problems with the calculations and update of the tables. Hence, a part of the code is written that changes the



дио кода који мијења вриједности *NULL* у свим пољима новокреиране колоне "HHI". Поред претходно споменутих *Python* петљи и изјава, у овом дијелу кода кориштена је и наредба *ArcPy* модула (*arcpy.da.UpdateCursor*) помоћу које се врши ажурирање појединих колона у табелама. Овај дио кода написан је у сљедећем облику:

```
19. with arcpy.da.UpdateCursor(fc, ["HHI"]) as cursor:
20.     for row in cursor:
21.         if row[0] == None:
22.             row[0] = 0
23.             cursor.updateRow(row)
24.     del row
25.     del cursor
```

У даљем поступку програмирања дефинисаног ГИС алата дефинише се листа колона које ће бити укључене у израчун индекса. Овиме се избјегавају грешке које су могуће ако се у табели селектованих података налазе колоне са *STRING* и *OBJECTID* типом података. Листа се формира само на основу колона које имају типове података *DOUBLE*, *INTEGER* и *SINGLE*. Креирање листе извршено је помоћу *ArcPy* наредбе *arcpy.ListFields*. Овај дио кода приказан је као:

```
27. fieldNameList = []
28. fields = arcpy.ListFields(fc, "")
29. for field in fields:
30.     if field.type in ("Double", "Integer", "Single"):
31.         fieldNameList.append(field.name)
```

У наредном дијелу кода помоћу наредбе *np.array* из модула *NumPy*, те помоћу *Python* оператора, петљи и итерација извршава се израчун *HHI*. Помоћу *ArcPy* наредбе (*arcpy.da.UpdateCursor*) извршава се ажурирање претходно креиране колоне "HHI" у излазној табели са резултатима израчуна. Дио кода који описује израчун индекса има сљедећи изглед:

```
33. with arcpy.da.UpdateCursor(fc, (fieldNameList)) as cursor:
34.     for row in cursor:
35.         hi = sum(np.array([row[i]/(sum([row[i] for i in range(len(fieldNameList))
36.         )) for i in range(len(fieldNameList))])**2)
37.         row[-1] = hi
38.         cursor.updateRow(row)
39.     del row
40.     del cursor
```

*NULL* values in all the fields of the newly-created "HHI" column. Apart from the aforementioned *Python* loops and statements, this part of the code also used an *ArcPy* module command (*arcpy.da.UpdateCursor*) which updates specific columns in the table. This part of the code is written as follows:

The next step in the GIS tool definition defines the list of columns that will be included in the index calculation. It helps us avoid errors that may happen if the selected data tables contain *STRING* and *OBJECTID* columns. The list is created only on the basis of columns containing *DOUBLE*, *INTEGER* and *SINGLE* data types. The list design was performed by using *ArcPy* *arcpy.ListFields* command as follows:

In the next part of the code, *HHI* calculation is performed by using *NumPy* module *np.array* command, and *Python* operators, loops, and iterations. *ArcPy* command (*arcpy.da.UpdateCursor*) updates the previously created *HHI* column in the output table with the calculation results. The part of the code that describes the index calculation is as follows:

Последњи дио кода односи се на придруживање колоне "HHI" (са резултатима израчуна) из излазне табеле одабраном векторском слоју са садржаним територијалним јединицама које су биле укључене у израчун. Придруживање се врши ради визуализације резултата а обавља се преко наредбе `arcpy.JoinField_management` из *ArcPy* модула. Запис овог дијела кода је сљедећи:

```
41. arcpy.JoinField_management(vektor_zapridruzivanje, input_join_field, fc, output_join_field, "HHI")
```

На крају поступка писања кода формирана је *Python* скрипта, која извршава алгоритам израчуна *HHI*.

The last part of the code refers to the merge of "HHI" column (calculation results) from the output table and the selected vector layer containing territorial units that were included in the calculation. The merge is performed in order to visualize the results by using `arcpy.JoinField_management` command in *ArcPy* module, and it looks as follows:

In the end of code writing, a *Python* script is formed to execute the *HHI* calculation algorithm.

## РЕЗУЛТАТИ

Креирана скрипта са екстензијом *.py*, као крајњи резултат описаног програмерског поступка, преносива је на било који десктоп рачунар и могуће је њено покретање на било ком рачунару на коме је инсталирана *ArcGIS* десктоп апликација (прилагођено за *ArcGIS* 10.1 и веће верзије). Тиме је омогућен аутоматизован начин израчуна и визуелзације резултата *HHI*. Сам поступак инсталације и коришћења креираног алата је веома једноставан, те у том погледу нема ограничења за његово коришћење. Увођењем и параметризацијом *Python* скрипте у *ArcGIS* апликацију добија се ГИС алат чијом се употребом добијају два излазна резултата: излазна табела са приказаним селектованим/улазним подацима увезеним из улазне табеле и резултатима израчуна *HHI*, те векторски слој са придруженом колоном резултата *HHI*.

Поступак инсталације *Python* скрипте у *ArcGIS* окружење је једноставан и састоји се од неколико корака: *Catalog> MyToolboxes> New> Toolbox> Add> Script*. Параметризација скрипте приликом њеног увођења/инсталације у *ArcGIS* апликацију подразумијева дефинисање шест параметра (Слика 1): Први параметар (*Input Table*) - односи се на табелу са улазним подацима, вриједности на картици за подешавање треба поставити *Data Type - Table*,

## RESULTS

The final result of the programming is the script with an *.py* extension, and it is transportable to any desktop computer at which it may be launched if the desktop *ArcGIS* application is installed (adapted to *ArcGIS* 10.1 and larger versions). This enables the automated calculation and visualization of *HHI* results. The installation and usage of the created tool are very simple and there are no limitations for its practice. Once the *Python* script is introduced to *ArcGIS* application, a GIS tool is obtained and it provides us with two types of output results: the output table with selected/input data introduced from the input table and the results of *HHI* calculation, and the vector layer with the adjoined column of *HHI* results.

The *Python* script installation into the *ArcGIS* environment is a simple procedure comprising the following steps: *Catalog>MyToolboxes> New> Toolbox> Add> Script*. Parameterization of the script before the installation into the *ArcGIS* application took six parameters to define (Figure 1): First parameter (*Input Table*)—refers to the table with input data, values of the setting tab should be *Data Type - Table*, *Direction - Input*; Second parameter (*Output*

*Direction - Input*; Други параметар (*Output Table*) - односи се на табелу са излазним подацима, вриједности на картици за подешавање треба поставити *Data Type - Table, Direction - Output*; Трећи параметар (*Select Table*) - односи се на *SQL* упит помоћу кога вршимо селекцију улазних података, вриједности на картици за подешавање треба поставити *Data Type - SQL Expression, Direction - Input, Obtained Form - Input\_Table*; Четврти параметар (*Vector Layer*) - је улазни векторски слој са приказаним територијалним јединицама које ће бити укључене у анализу, вриједности на картици за подешавање треба поставити *Data Type - Feature Class, Direction - Input*; Пети параметар (*Input Join Field*) - односи се на индексирану колону у векторском слоју преко које ће бити извршено придруживање резултата *HHI* том слоју, вриједности на картици за подешавање треба поставити *Data Type - Field, Direction - Input, Obtained Form - Vector\_Layer*; Шести параметар (*Output Join Field*) - односи се на индексирану колону у излазној табели из које ће бити придружени резултати *HHI* векторском слојум, вриједности на картици за подешавање треба поставити *Data Type - Field, Direction - Input, Obtained Form - Inout\_Table*.

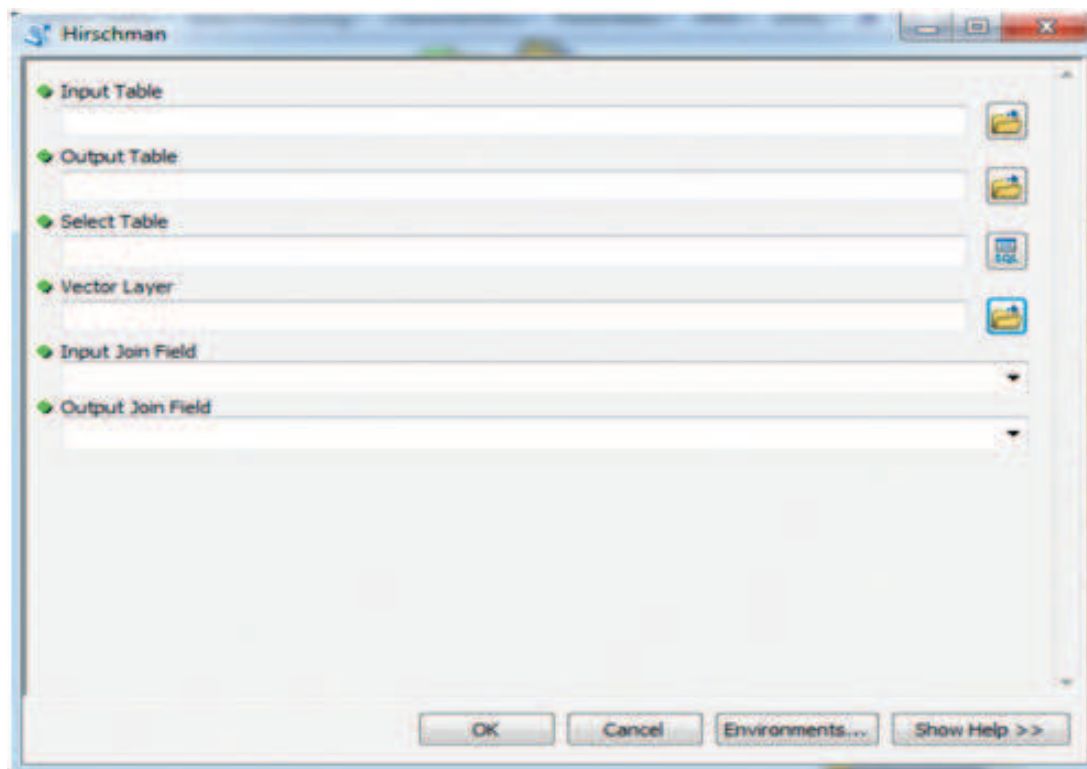
*Table*) –refers to the table with output data, values of the setting tab should be *Data Type - Table, Direction - Output*; Third parameter (*Select Table*) –refers to the *SQL* inquiry that helps us select the input data, values of the setting tab should be *Data Type - SQL Expression, Direction - Input, Obtained Form - Input\_Table*; Fourth parameter(*Vector Layer*) –is the input vector layer displaying the territorial units that will be included in the analysis, values of the setting tab should be *Data Type - Feature Class, Direction - Input*; Fifth parameter (*Input Join Field*) – refers to the indexed column in the vector layer through which the merge with *HHI* results will be performed, values of the setting tab should be *Data Type - Field, Direction - Input, Obtained Form - Vector\_Layer*; Sixth parameter (*Output Join Field*) –refers to the indexed column in the output table from which the *HHI* results will merge with the vector layer, values of the setting tab should be *Data Type - Field, Direction - Input, Obtained Form - Inout\_Table*.



Слика 1. Дефинисање параметара у ArcGIS апликацији  
Figure 1. Defining parameters in ArcGIS application

Покретањем увезене скрипте покреће се алат за израчунавање и визуелизацију *HHI*, којом приликом се отвара картица за унос назива параметара (Слика 2). Уношењем назива параметара (дефинисањем њихових путања) у понуђена поља у отвореној картици, те кликом на опцију „ОК“, покреће се извршна наредба којом се реализује дати алгоритам за израчунавање и визуелизацију *HHI*. Добијени резултати се читавају како су дефинисани код описа математичког концепта индекса.

The launch of the script sets off the tool for calculation and visualization of *HHI*, and the tab for the parameter input opens (Figure 2). Once the name of the parameter is entered (definition of its direction) into the appropriate field in the open tab, a click of „ОК“ starts the executive command for the given algorithm for calculation and visualization of *HHI*. The obtained results are then interpreted as was defined through the description of the mathematical concept of the index.



Слика 2. Изглед картице за покретање креираног алата у ArcGIS-у  
Figure 2. The layout of the tab for launching the created tool in ArcGIS

Из досадашњег излагања може се закључити да су за коришћење креираног алата неопходна два типа улазних података: улазна табела са подацима по територијалним јединицама који се планирају укључити у израчун и векторски слој на којима су приказане границе територијалних јединица. Концепт креираног алата је такав да је неопходно улазне податке похранити у *File Geodatabase* или *Personal Geodatabase*. Оба типа улазних података морају имати индексиране колоне, гдје сваки ред (територијална јединица) мора имати свој *ID*, који мора бити идентичан и у улазној табели и у векторском слоју. У смислу типа података *ID* треба да буде дефинисан као *SHORTINTEGER* или *OBJECTID*. Колоне у улазној табели које се односе на податке који ће бити укључени у израчун требају да буду дефинисане као *DOUBLE*, *INTEGER* или *SINGLE*. Евентуално остале колоне које се не односе на податке који ће бити укључени у израчун а које се налазе у улазној табели морају бити дефинисане као *STRING* или

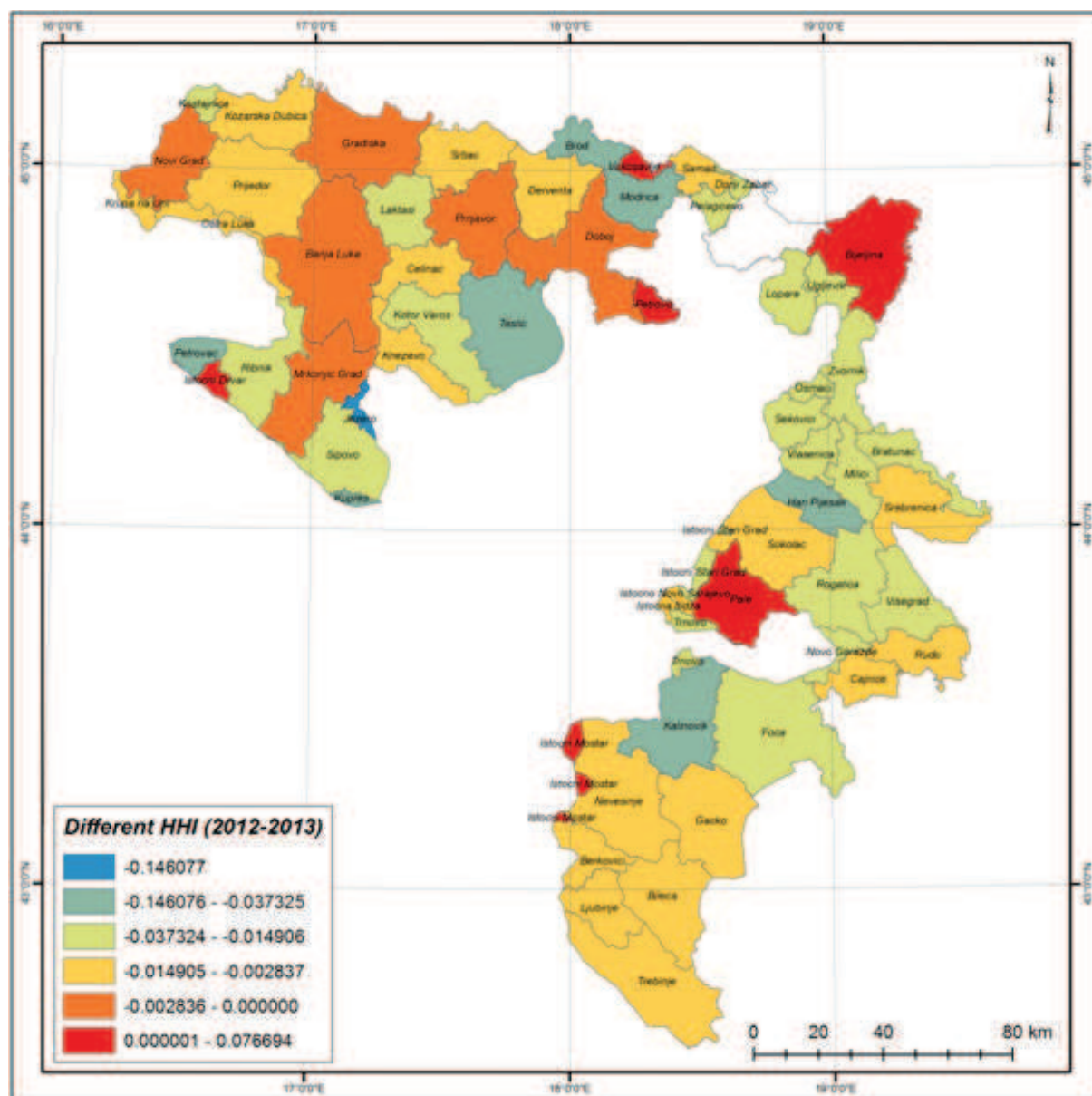
We may so far conclude that it takes two types of input data in order to use the created tool, i.e. the input table with the data for each territorial unit used in the calculation and the vector layer that displays the territorial unit borderlines. Due to the concept of the created tool, it is necessary to file the input data into *File Geodatabase* or *Personal Geodatabase*. Both types of input data must have indexed columns in which every line (territorial unit) must have its identical *ID* for both input table and vector layer. The *ID* should be defined either as *SHORTINTEGER* or *OBJECTID*. Columns in the input table referring to calculation data should be labeled as *DOUBLE*, *INTEGER* or *SINGLE*. In case there are columns that do not refer to calculation data but are still in the input table, these should be labeled as *STRING* or *SHORTINTEGER*.

To provide an instance of our tools usage, we performed an analysis of regional specialization per municipalities in Republic

*SHORTINTEGER.*

Као примјер примјене креираног алата извршена је анализа регионалне специјализације по општинама Републике Српске према броју запослених по подручјима дјелатности. Резултати израчуна приказани су на картографском прилогу (Слика 3).

of Srpska according to the number of employees in different trade areas. The results of one such calculation are displayed in the cartograph (*Figure 3*).



Слика 3. Разлике ННН по општинама у Републици Српској, 2012-2013. година  
Figure 3. Differences in HHI per municipalities in Republic of Srpska, 2012-2013

У наведеном примјеру извршен је израчун *HHI* за 2012. и 2013. годину а на основу добијених резултата израчуната је разлика *HHI* за двије посматране године (Слика 3). Према добијеним резултатима могу се утврдити правци развоја промјена привредне структуре по општинама у посматраном периоду. Општине са негативним вриједностима у посматраном периоду имале су тенденцију у равнотежавања привредне структуре, тенденцију која иде у правцу диверзификације привредне структуре. Док су обрнуто, општине са позитивним вриједностима разлике имале тенденцију концентрације броја запослених у мањем броју подручја дјелатности, тенденцију специјализације. Само је шест општина у Републици Српској у посматраном периоду имало позитивне резултате разлике, тј. имало тенденцију ка специјализацији. Из наведеног се може изцести закључак да је према карактеру регионалне специјализације структура привреде у Републици Српској, у посматраном периоду, имала тренд усмјерен ка диверзификацији (уједначавање броја запослених у подручјима дјелатности).

## ЗАКЉУЧАК

Истраживања просторних и регионалних обиљежја привреде су битан сегмент економских, економско-географских и просторно-планерских студија. Планирање развоја и просторне организације привреде у значајној мјери темељи се на идентификацији њених просторно-структурних карактеристика. Анализом просторно-структурних карактеристика привреде може се одговорити на питања везана за феномене просторне концентрације и регионалне специјализације. У методологији везаној за истраживања просторних обиљежја привреде постоји велики број квантитативних мјера помоћу којих се на егзактан начин могу детерминисати њена просторно-структурна обиљежја. Једна од квантитативних мјера помоћу које се идентификују просторно-структурна обиљежја

In our example, a *HHI* calculation for 2012 and 2013 was performed and the *HHI* difference between the two years was calculated (*Figure 3*). On the basis of results, it is possible to estimate the directions of changes in commercial structure per municipalities within an observed period. Municipalities with negative values within the observed period showed a tendency to harmonize the commercial structure, i.e. a tendency aiming at the diversified commercial structure. On the other hand, municipalities with positive values showed a tendency to concentrate the number of employees within a smaller number of trade areas, i.e. tendency of specialization. There were only six municipalities in Republic of Srpska that had the positive results of differences, i.e. they displayed tendency for specialization. Hence, we may infer that according to the type of regional specialization in Republic of Srpska within the observed period, the commercial structures focused on diversification (harmonization of number of employees within trade areas).

## CONCLUSION

Studies of spatial and regional features of commerce may be relevant to other economic, economic-geographic, and spatial planning studies. Planning the development and spatial organization of commerce is largely based upon the identification of its spatial-structural characteristics. The analysis of spatial-structural characteristics of commerce may provide an answer to many questions referring to phenomena of spatial concentration and regional specialization. Methodology of studies on spatial features of commerce comprises many quantitative measures that may accurately determine the spatial-structural characteristics. Herfindahl–Hirschman Index, which is described in this paper, is one such quantitative measure.

The application of quantitative measures,

привреде је и Херфиндал-Хиршманов индекс регионалне специјализације, који је описан у овом раду.

Развојем информационих технологија примјена квантитативних мјера којима се осликавају просторна обиљежја привредних структура, у великој мјери је олакшана. Раније је било потребно доста времена и математичког знања да се изврши израчуна појединих индекса или коефицијената из домена просторне концентрације или регионалне специјализације привреде. Поготово је био мукотрпан посао картографске интерпретације резултата наведених израчуна. Појавом ГИС технологија и картографска интерпретација је у великој мјери олакшана.

ГИС технологије омогућавају да се унутар постојећих ГИС софтверских апликација креирају алати за геопросторне анализе, потпуно прилагођени потребама корисника. Употребом програмских језика као што су *Python*, *C#*, и *.NET*, те посебно креираних модула за развој ГИС апликација унутар њих, програмерским поступцима могуће је развити ГИС алате посебне намјене који не постоје у актуелним ГИС софтверским апликацијама. Један од поступака развоја ГИС алата описан је у овом раду, односно приказан је поступак развоја ГИС алата за аутоматизацију израчуна Херфиндал-Хиршмановог индекса регионалне специјализације.

Поред значаја који овај рад може имати у смислу примјене развијеног алата у просторним и регионалним истраживањима, може се истаћи и значај рада у смислу примјене његових резултата као смјерница за развој нових ГИС алата, како у области просторне економије, тако и у осталим областима гдје ГИС има своју примјену.

which depict the spatial features of commercial structures, has been largely facilitated by the rise of information technology. It used to consume a lot of time and mathematical skills to calculate specific indices or coefficients within the domains of spatial concentration or regional specialization of commerce. The most difficult part was the mapping interpretation of results of earlier calculations. It was precisely GIS technology that largely facilitated the mapping interpretation.

GIS technology makes it possible to create tools for geospatial analyses within the existing GIS software application and adjust them to the users' needs. Using the programming languages such as *Python*, *C#*, and *.NET* and the specially designed modules for new GIS application development, it is possible to program GIS tools for special purposes, which do not exist within the current GIS software applications. One such procedure of developing a GIS tool is addressed to in this paper, i.e. we described the design of a GIS tool for an automated calculation of Herfindahl–Hirschman Index of regional specialization.

The results of our design may be relevant for the application in both spatial and regional studies but they may also benefit the development of new GIS tools within the field of spatial economy or in some other field where GIS is applicable.



ЛИТЕРАТУРА/REFERENCES:

- Aiginger, K. (1999). Do Industrial Structures Converge? A survey on the empirical literature on specialisation and concentration of industries. *WIFO working paper No. 116*
- Aiginger, K., Davies, D. (2004). Industrial specialization and geographical concentration: Two sides of the same coin? Not for the EU. *Journal of Applied Economics, Vol. VII, No. 2*
- Aiginger, K., Rossi-Hansberg, E. (2006). Specialization and concentration: a note on theory and evidence. *Empirica, vol. 33, Springer*
- Ascher, D., Dubois, P., Hinsen, K., Hugunin, J., Oliphant, T. (1999). *Numerical Python*. Berkley: The Regents of the University of California
- Bajić, D., Trbić, G., Stanojević, M., Gnjato, S. (2013). GIS Based Index as a Criterion of Spatial Concentration of the Social-economic Features. *Материјали међународној конференцији (Четвртая Ежегодная научная Ассамблея АРГО)*. Москва – Ростов-на-Дону: Ассоциация Российских географов обществоведов (АРГО), Институт географии РАН, Географический факультет МГУ имени М. В. Ломоносова, Северо-Кавказский НИИ экономических и социальных проблем Южного федерального университета
- Balassa, B. (1965). Trade Liberalization and Revealed Comparative Advantage. *Manchester School, 33, 99–123*
- Gini, C. (1936). On the Measure of Concentration with Special Reference to Income and Statistics. *Colorado College Publication, General Series No. 208*
- Goschin, Z., Constantin, D., Roman, M., Ileanu, B. (2009). Specialisation and Concentration Patterns in the Romanian Economy. *Journal of Applied Quantitative Methods, Vol. 4, No.1*
- Грчић, М. (1987). *Анализа просторне организације индустрије региона Београд*. Београд: Економски институт
- Ezcurra, R., Pascual, P., Rapun, M. (2006). Regional Specialization in the European Union. *Regional Studies, Vol. 40, No. 6*
- Ellison, G., Glaeser, E. (1997). Geographic concentration in u.s. manufacturing industries: A dartboard approach. *The Journal of Political Economy, Vol. 105, No.5*
- Essert, M. (2007). *Python – osnove. Digitalni udžbenik*. Osijek: Odjel za matematiku Sveučilišta Josipa Jurja Strossmayera
- Krugmana, P. (1991). *Geography and Trade*, Cambridge, London: The M.I.T. Press
- Lilien, D., M. (1982). Sectoral Shifts and Cyclical Unemployment. *Journal of Political Economy, no 90*
- Traistaru, I., Nijkamp, P., Longhi, S. (2002). Regional specialization and location of industrial activity in accession countries. *42nd congress of the European Regional Science Association*. Dortmund
- Theil, H. (1967). *Economics and Information Theory*. North-Holland Publishing Company
- Fujita, M., Krugman, P., Venables, A., J. (1999). *The Spatial Economy: Cities, Regions and International Trade*. Cambridge: MIT Press.
- Hallet, M. (2002). Regional Specialization and Concentration in the EU. *Regional Convergence in the European Union. Facts, Prospects and Policies*, Springer-Verlag
- Hirschman, O. (1964). The Paternity of an Index. *The American Economic Review* (American Economic Association) 54 (5).
- Campante, R., Quoc-Anh, Do (2009). A Centered Index of Spatial Concentration: Axiomatic Approach with an Application to Population and Capital Cities. *HKS Faculty Research Working Papers RWP09-005*