Information Society and its Development in the Czech Republic

JITKA KOMARKOVA, PAVEL SEDLAK, KATERINA LANGROVA
Faculty of Economics and Administration, Institute of System Engineering and Informatics
University of Pardubice
Studentska 95, Pardubice, 532 10
CZECH REPUBLIC
Jitka.Komarkova@upce.cz, Pavel.Sedlak@upce.cz

Abstract: - Development of information society has been strongly supported by all governments including the government of the Czech Republic. Recently, ways how to evaluate efficiency of the investments and development of information society have been searched for. Paper deals with utilization of geographic information systems, spatial analyses and exploratory spatial data analyses (both global and local indicators) for evaluation of development of information society and participation of the citizens. Within the case study development in all 14 regions of the Czech Republic during last few years is evaluated. Evaluation is based on available data which were provided by Czech Statistical Office. One of the results is that Prague (the capital) was a leader in the first years but during last few years the other regions have made the gap on the capital smaller so many differences across the country have been decreased.

Key-Words: - eGovernment, Information Society, eParticipation, Spatial Analyses, ESDA, GIS, DEMO-net.

1 Introduction

Information society has introduced a new idea - an easy and remote access to information for all citizens and openness of public administration to the citizens. One of the aims is to increase participation of citizens in public life [14]. Participation of citizens in public life is the main idea of democracy. Nevertheless, in prevailing systems of governments, citizens are mainly allowed to significantly participate only by voting once in several years. In periods between infrequent elections, citizens have only limited power to change things they do not like. On the other side, the idea that all members of a community participate in reaching common decisions is unfeasible and not realizable. It has been proved that in social groups which are larger than a village or small town it is impractical let each particular person to co-decide on issues of public concern. Before the emergence of information and communication technologies (ICT) these arguments seemed more convincing, especially for very large units such as cities or states. However, with modern technology it is theoretically possible to allow all citizens to inform themselves about public issues and to participate on public affairs [12].

Today, eGovernment services can be provided at the three following levels [16]:

- Publish
- Interact
- Transact.

European Union (EU) has provided a strong support to implementation of ICT into public administration at all levels to strengthen development of information society and eGovernment. The last initiative in this field is initiative "i2010 – A European Information Society for growth and employment" [8]. Many methodologies and frameworks for ICT implementation into government have been developed as well [17, 18].

On the other hand, even a user-friendly environment is not a sufficient condition to ensure participation of inhabitants in public deals. Danish project Public Participation GIS can be used as an example of a project targeted to the all citizens but resulting in a participation of only limited group of citizens (middle-age and well-educated males with income above average in this case) [7]. So, there is a question how to get the citizens involved in public deals regardless of the way of their participation. Aim of the paper is to provide some appropriate methods for evaluation of development of information society and participation of the citizens in the Czech Republic during last few years and by this way contribute to evaluation of efficiency of investments into information society development.

DEMO-net Network of Excellence project (FP6-2004-IST-4-027219, see www.demonet.org) is one of projects funded by European Commission. This project is aimed at eParticipation issues [11] and it provided a framework for this research.

2 Information Society Development in the Czech Republic and its Evaluation

Some evaluation criteria, indicators and measures have been proposed to allow evaluation of information society

and comparison of countries, e.g. [1], but because of short time period of information society development and spatial dimension which should be taken into account as well, development of information society must be evaluated by means of many different methods and approaches. Utilization of text-mining methods and artificial intelligence [9] is an example of another suitable method for getting information from many text documents which can be used in some cases as well. A special set of indicators has been proposed to evaluate mobile Government [6].

In this paper, utilization of spatial analyses, spatial statistics and visualisation methods provided by geographic information systems (GIS) is proposed, namely standard deviational ellipse, centroids, various cartograms and exploratory spatial data analysis (ESDA).

3 Evaluation of Information Society Development in the Czech Republic

As it was mentioned previously, the aim is to evaluate development of information society and interest of the citizens in eGovernment services at regional level. Evaluation will be based on available data.

3.1 Used Methods and Tools

Various spatial analyses and statistical methods were used for evaluation of information society development in the Czech Republic. Data were treated by means of the following software packages: ArcGIS 9.2 (ArcInfo) and GeoDA 0.9.5-i5. Specifically, standard deviational ellipse, centroids, various cartograms and exploratory spatial data analysis (ESDA) were used to analyse available data. All calculations were based on Euclidean distance.

Standard deviational ellipse calculation is based on the standard distance calculation. The standard distances are calculated separately in the x and y directions. Standard deviation of both x and y coordinates to the mean centre are calculated to define both axes of an ellipse. Ellipses allow users to measure the orientation of a spatial distribution of a phenomenon. Identification of directional trends can be useful for example for modelling of phenomena spreading and for prediction of their future behaviour. The one and three standard deviational ellipses were used in this study. Utilization of both one and three standard deviation ellipses allows to take into account approximately 99 percent of all features [3].

Centroids represent weighted mean centre or centre of gravity. They are quite useful to observe movement of centre of gravity during a certain time period, i.e. movement of phenomena around the studied area within the given time period [3].

Exploratory spatial data analysis (ESDA) takes explicitly into account spatial dependence of the data. It is the main difference between ESDA and exploratory data analysis (EDA). In more detailed view ESDA is a collection of techniques to describe and visualize spatial distributions, identify not typical locations or spatial outliers, discover patterns of spatial association, clusters or hot spots, and suggest spatial regimes or other forms of spatial heterogeneity [2, 13]. Namely, global Moran's I and Getis Ord General G statistics were used to find out whether data are clustered. Then, two local indicators of spatial association were used: Anselin Local Moran's I and Getis Ord Gi* statistics were used to find outliers and hot/cold spots. Getis Ord Gi* statistics is useful to identify significant patterns of local spatial association (e.g. hot/cold spots – their attribute value significantly differs from the attribute values of neighbours). Anselin Local Moran's statistics can be more useful because it can identify significant patterns of local spatial association (e.g. outliers) too. Further, it can help to evaluate the stability of the global Moran's I when a global association is found [2].

Besides, some basic methods coming from exploratory data analysis were used too, e.g. box plots. Results were again visualised by means of a map.

Unfortunately, there was one limitation – there are only 14 regions in the Czech Republic so only 14 records were available for analytical purposes. It of course limited analytical possibilities.

Null hypothesis is that there is no spatial association, i.e. values of attributes of a feature do not depend on attribute values of neighbouring features.

3.2 Czech Republic – Basic Information

Czech Republic is a small country in the central Europe. It shares borders with Germany (810 km), Poland (762 km), Austria (466 km) and Slovakia (265 km). It has approximately 10 millions of inhabitants and its area is 78 866 sq. km. Czech Republic is a landlocked country approximately 300 km from the sea.

Democratic government has been reestablished in the country in the beginning of the nineties of the 20th century after 40 years of communistic governance. However, some parts of the previous government system stayed in use to do not interrupt life in the country. But it was impossible to keep the mixture of the new and old system forever. This was one of the reasons for a reform of public administration system which has been in progress since 1998. District level of public administration was cancelled by January 1, 2003 and was replaced by regional level (regions are larger then former districts; see Fig. 1).

Fig. 1 – Czech Republic and its 14 regions

Another reason of the reform is to provide higher-quality and easier accessible services of public administration to the citizens. This request is associated with a development of information society in the Czech Republic and Europe in the framework of eEurope initiatives - the Czech Republic has become an EU member in 2004.

3.3 Data Collection and Pre-processing

Evaluation of development was expected to be done at the regional level so it was necessary to have data about all 14 regions. At the same time it was necessary to have at least short time series to be able to evaluate development of the information society in the country.

There are no such detailed data available at international level, e.g. from Eurostat. In the Czech Republic there are two organizations collecting statistical data concerning ICT utilization and state of information society – Czech Statistical Office and STEM/MARK a.s. Agency STEM/MARK a.s. made same measurements for the former Ministry of Informatics of the Czech Republic but there are no complete data series because regions had to pay for the measurements [19]. Thus, the only source of data about all regions in at least short time series is the Czech Statistical Office (CSO).

Czech Statistical Office has done measurements in two important topics: utilization of ICT by households and utilization of ICT by public administration bodies. The maximum length of data sets was 5 years but only in a few cases [4]. Due to this fact it was decided to use data sets which were collected at least for three years (with one exception). Data from CSO are normalized, statistically pre-processed and without seasonal dependency [5] so they can be used for next analyses without any other special pre-processing. Data are mostly calculated as a percentage share of population so they can be used for comparison without any additional pre-processing.

Following indicators were selected for next analyses:

- Number of owners of personal computer (percentage share)
- Number of Internet users (percentage share)
- Number of regular Internet users (percentage share)
- Interest of Internet users in utilization of eGovernment services (percentage share), available for 2006 only
- Number of online purchases done by citizens
- Number of online purchases done public administration bodies (percentage share), municipalities above 500 inhabitants [10].

3.4 Evaluation of Information Society Development in the Czech Republic

At first, correlation of some of the selected indicators was tested by means of Pearson correlation coefficient $r_{x,y}$ calculated in MS Excel – see formula (1) [15].

$$r_{xy} = \frac{n\left(\sum XY\right) - \left(\sum X\right)\left(\sum Y\right)}{\sqrt{\left(n\sum X^2 - \left(\sum X\right)^2\right)\left(n\sum Y^2 - \left(\sum Y\right)^2\right)}} \quad (1)$$

where:
X, Y – variables,
n – number of values.

It was found that number of owners of personal computer and number of Internet users is highly correlated in all regions of the republic ($r_{x,y}$ = <0,927;0,999>) so in the next analyses only data about users of Internet will be used.

Concerning the growth of the number of Internet users, it was found that the highest growth in the whole period 2003 – 2007 was 25.9 % in Jihocesky Region (from 19.4 % to 45.3 %). The lowest growth in the whole period was 13.5 % in Karlovarsky Region (from 29.5 % to 45.0 %). The highest interannual growth was 15.1 % in Pardubicky Region in period 2004 – 2005, the second highest one was 12.9 % in Olomoucky region in period 2005 – 2006. Prague (the Capital) is a leading region in absolute numbers, e.g. 60.6 % of its inhabitants were Internet users in 2007 but interannual growth is slower then in almost all regions of the republic.

Quite high correlation was found in the case of Internet users and regular Internet users ($r_{x,y}$ = <0,897;0,999>).

3.4.1 Interest of Users in eGovernment services

In this case data about interest of users in utilization of eGovernment services were available only for year 2006 but because of its importance data were treated although there was no time series available so development could not be evaluated. Number of citizens interested in eGovernment services is expressed as a percentage share of the number of the citizens of each region.

At first, data were visualised by means of cartogram (see Fig. 2) containing graph showing percentage share of Internet users in population of each region as well.

To classify regions in according to the number of users interested in eGovernment services, average value was used to split data into the first two classes. Each interval was then split in two more intervals. Because 7 regions still fitted into one interval, the interval was once more split. The resulting map is in Fig. 2. Interest of the users in eGovernment services was measured in the end of 2006 so the most comparable values describing number of Internet users are values for the same year or year 2007.

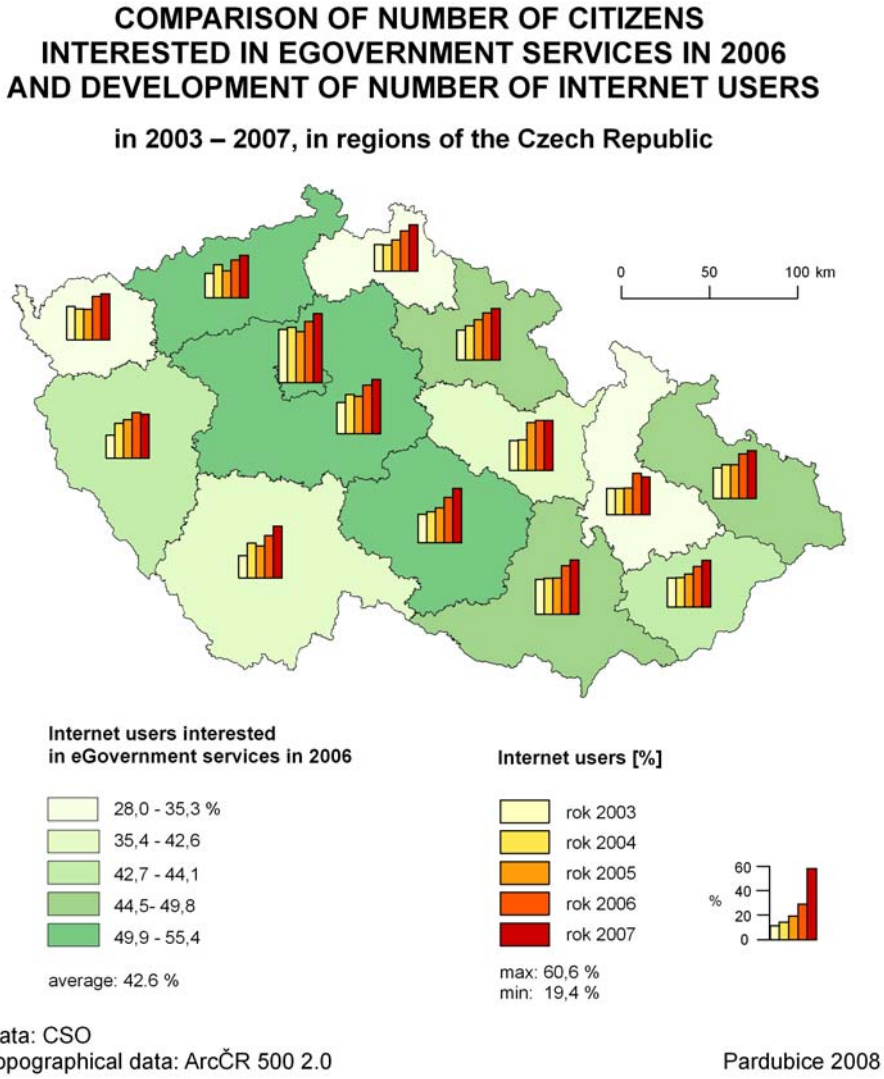

Fig. 2 – Comparison of interest of Internet users in utilization of eGovernment in regions of the Czech Republic, (source: authors – based on [10])

To be able to evaluate directional trends, i.e. orientation of spreading of Internet users and citizens interested in utilization of eGovernment services across the country, standard deviational ellipses were created. Three standard deviations were used to cover majority of features (see Fig. 3). Ellipses can be compared with the ellipse representing geographical orientation of the Czech Republic (not weighted by any attribute value). As it can be seen from the map, all ellipses are very similar. It means both phenomena follow geographic orientation of the Czech Republic, there cannot be observed any special directional trend.

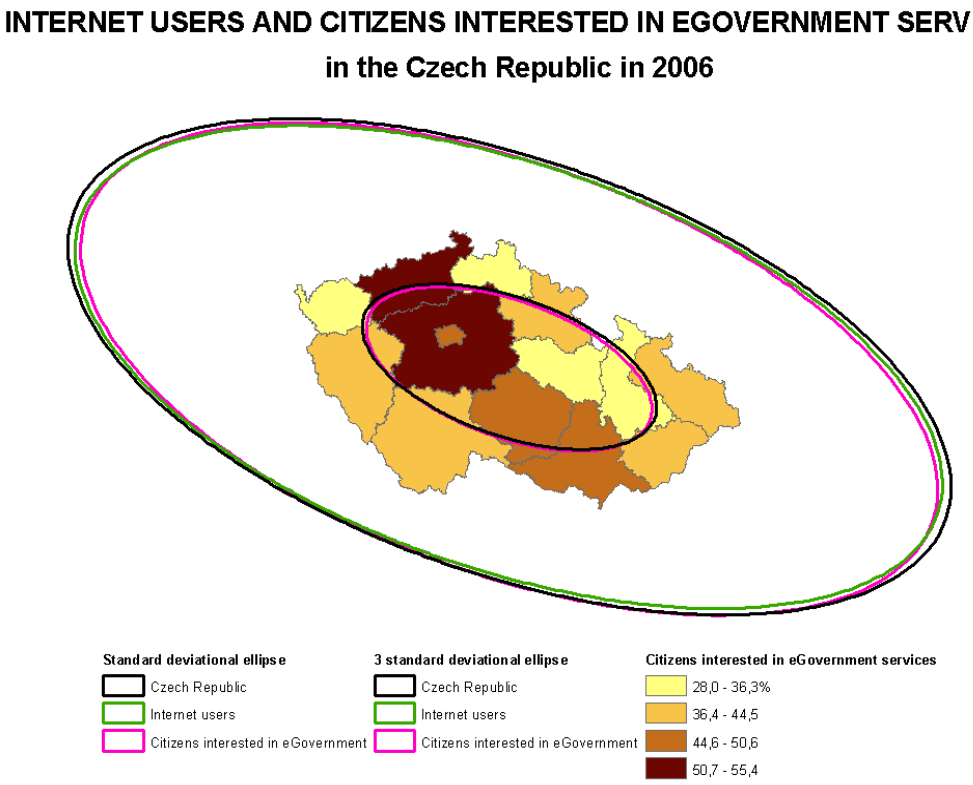

Fig. 3 – Directional trends of spreading of Internet users and citizens interested in eGovernment services across the Czech Republic evaluated by means of standard deviational ellipses, (source: authors)

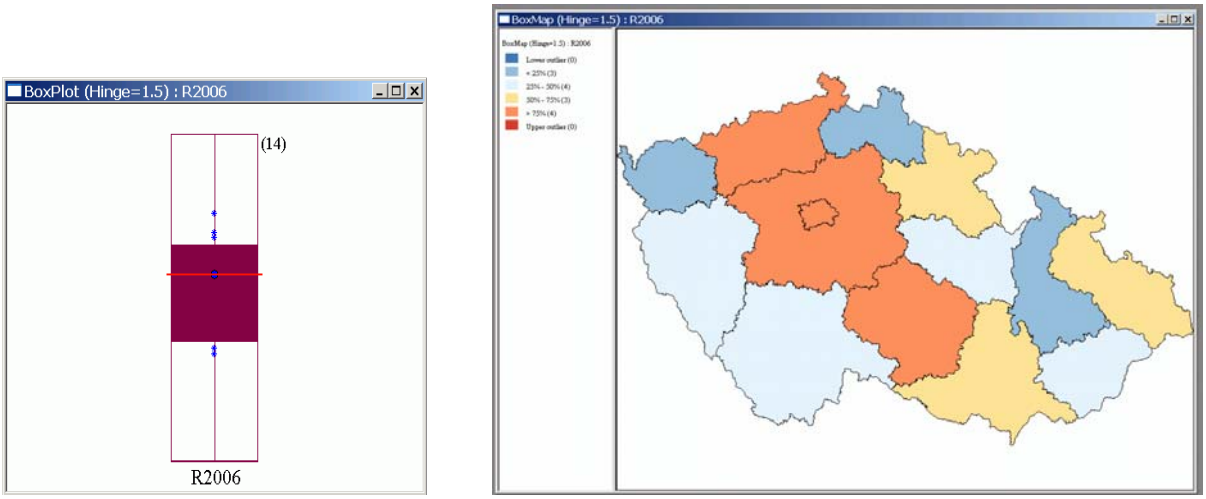

Fig. 4 – Box plot and box map - citizens interested in eGovernment services in the Czech Republic in 2006, (source: authors)

The next step is utilization of EDA and ESDA to describe and analyze data in more detailed way.

At first, box plot and box plot maps were created to describe the distribution of citizens interested in eGovernment services. Results are in the Fig. 4 – no outliers were identified by this way, all values are within interquartile range (1.5 times boundaries). GeoDa was used as a software tool in this case.

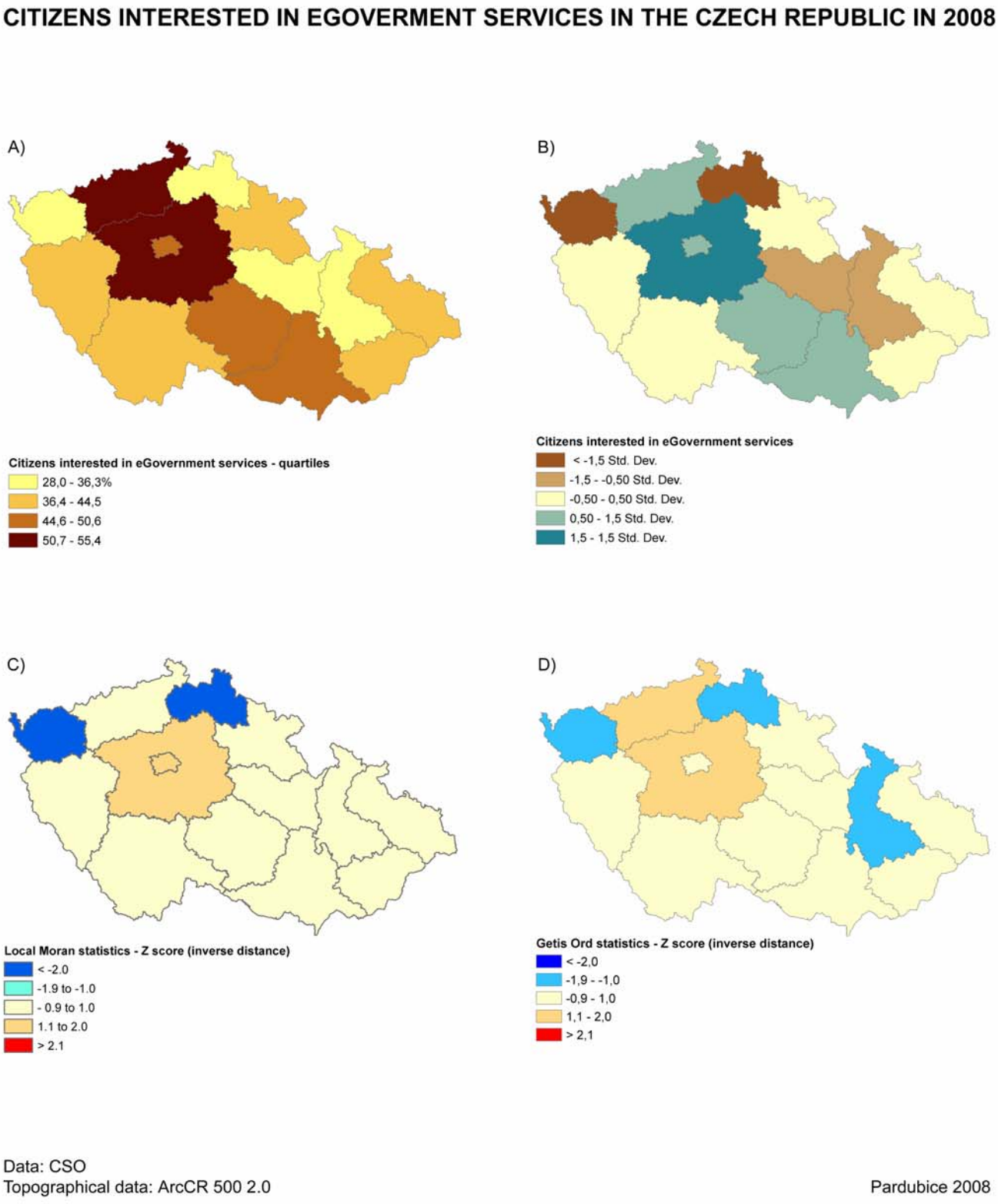

Fig. 5 – Citizens interested in eGovernment services in the Czech Republic in 2006: A) Classification by means of quartiles; B) Classification by means of standard deviation; C) Local Moran's statistics – Z score; D) Getis Ord Gi* statistics – Z score, (source: authors)

Global Moran's I statistics provided following results:

- Moran's Index = 0,064909
- Expected Index = -0,076923
- Variance = 0,015624
- Z Score = 1,134681 standard deviations

Getis Ord General G statistics provided following results:

- Observed General G = 0,000009
- Expected General G = 0,000009
- General G Variance = 0,000000
- Z Score = 1,877301 Standard Deviations

In both cases inverse distance was used to take into account influence of neighbours. In according to the Moran's I null hypothesis of random spatial distribution can be confirmed. In according to General G high values may tend to cluster.

More detailed views on this attribute are in Fig. 5 showing at first regions classified into 4 quartiles (Fig. 5 A)). Then data are classified in according to standard deviation (Fig. 5 B)). Then, two local indicators of spatial association are shown: Anselin Local Moran's I (Fig. 5 C)) and Getis Ord Gi* (Fig. 5 D)). Two regions were identified as local spatial outliers (Liberecky and Karlovarsky). Only 29.2% (Liberecky), 28% (Karlovarsky) of the citizens are interested in eGovernment services. These regions neighbour with regions with quite high values: Ustecky (51.6%), Stredocesky (55.4%). Prague (49.9%) is not far away too. These areas were identified as cold/hot spots by Getis Ord Gi* statistics too.

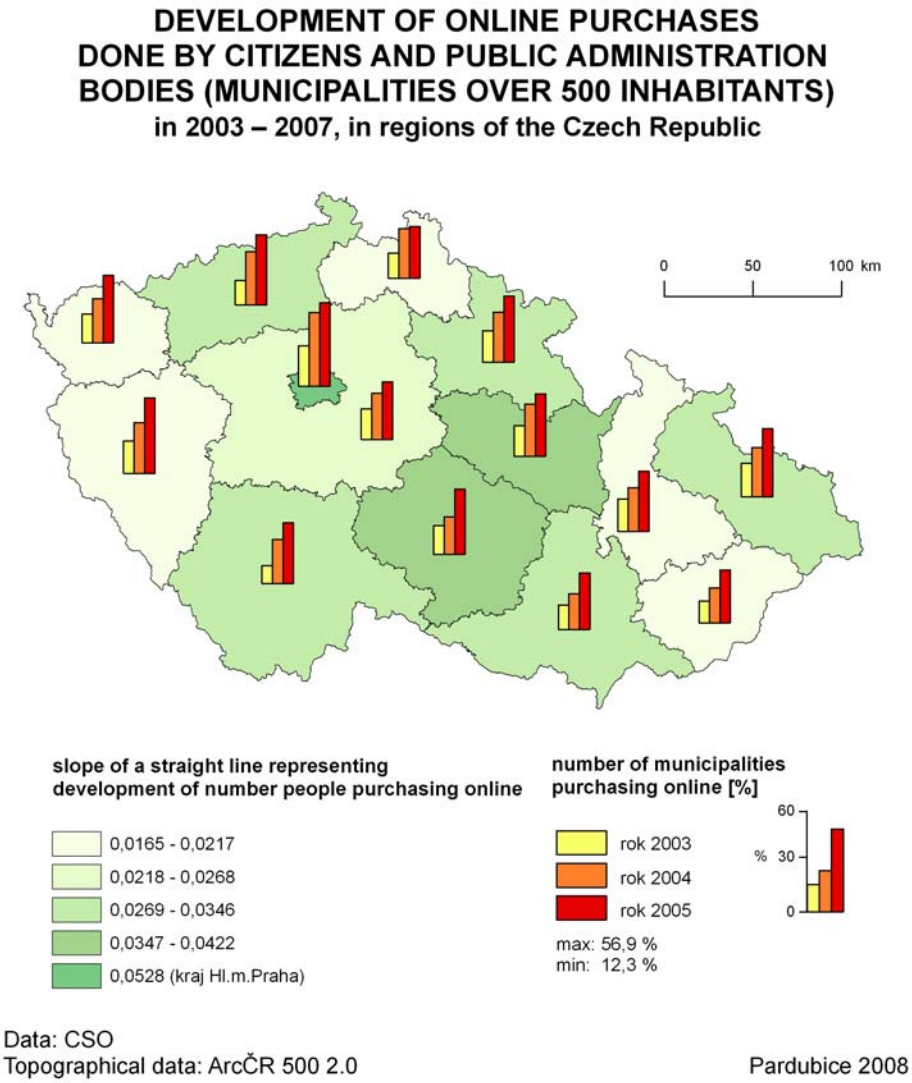

Fig. 6 – Development of online purchases done by the citizens and public administration bodies, (source: authors – based on [10])

3.4.2 Development of Utilization Internet for Online Transactions and Shopping

There are almost no data (data series) available about transactions realized within eGovernment services. So, number of online purchases was used as a substitutional indicator. In this case development of number of people purchasing online is compared with number of municipalities (public administration bodies) purchasing online. All data provide percentage share of the whole. The problem is that only municipalities with more then 500 inhabitants are included which decreases a little bit information capability of obtained results. To be able to show both indicators, development of citizens buying online was expressed as a slope of a straight line. Results are shown on the two following maps (see Fig. 6 and Fig. 7).

On Fig. 6 there can be seen that the Capital Prague (Praha) is an outlying value concerning percentage share of citizens buying online so for classifying purposes in this cartogram it was excluded. In Prague this value increased from 5.8 % to 25.7 % in period 2003 – 2007 what is the best value. Average slope for the other regions is 0.0268, it was used as a border of intervals.

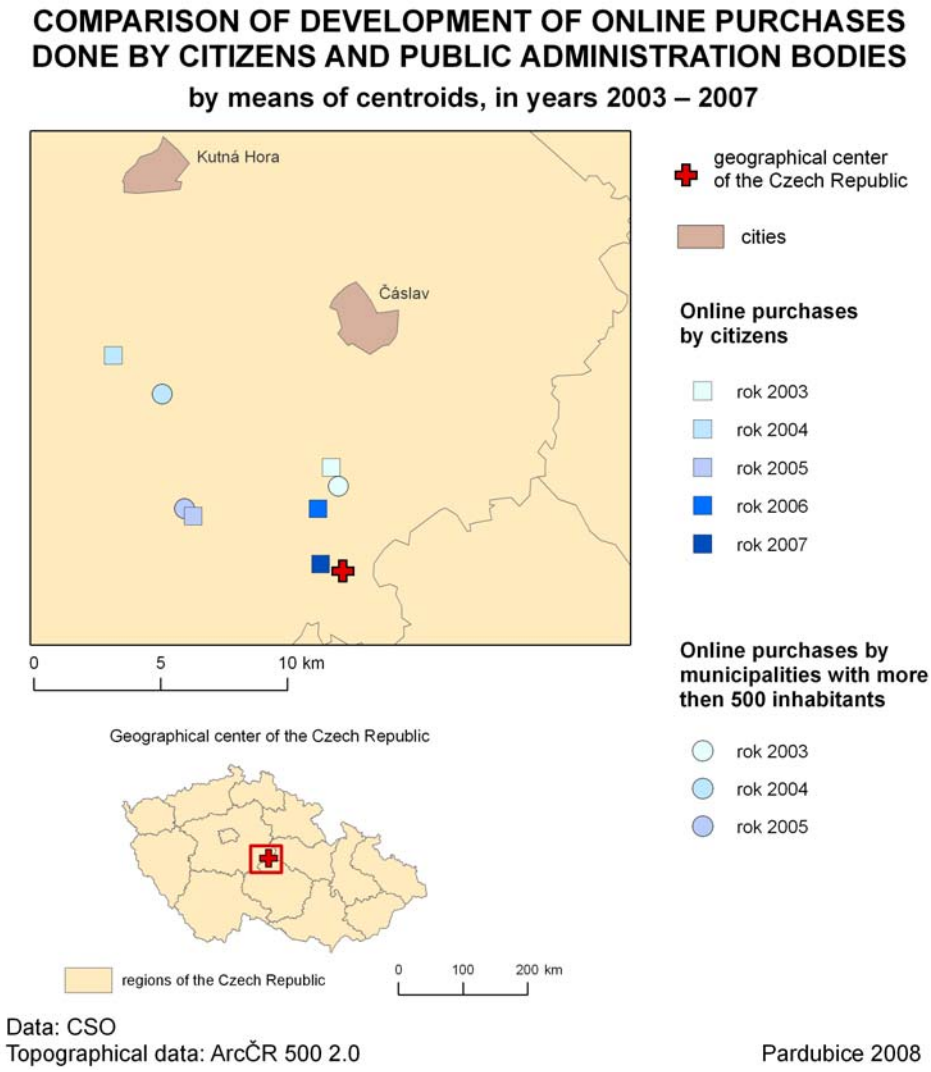

Fig. 7 – Centroids showing development of online purchases done by the citizens and public administration authorities (municipalities with over 500 inhabitants), (source: authors – based on [10])

There can be seen on Fig. 7 that on the beginning centre of gravity of both phenomena was moved to north-west. This position is influence mostly by Prague which lies to north-west from the geographical centre of the country. During the following years centres of gravity of both phenomena moved closer to the

geographical centre of the country as the other regions followed development in the Capital and decrease the difference.

The next step is utilization of EDA and ESDA to describe and analyze data in more detailed way.

At first, box plot and box plot maps were created to describe the distribution of citizens interested in eGovernment services. Results are in the Fig. 8 – one outlier was identified by this way - Prague fell outside interquartile range (1.5 times boundaries). After increasing boundaries to 3 multiple, there was no outlier identified (graph not shown in the paper). GeoDa was used as a software tool in this case.

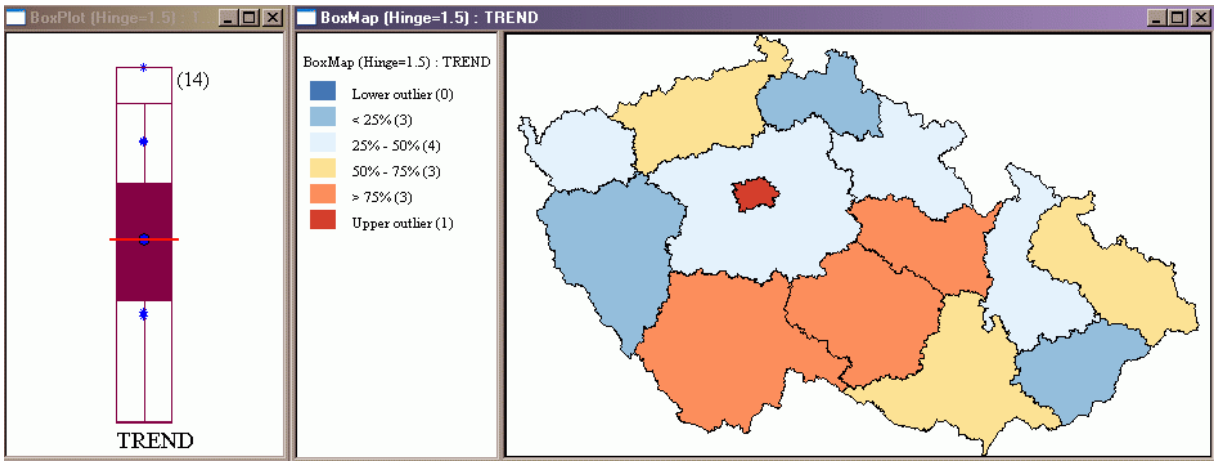

Fig. 8 – Box plot and box map - development of online purchases done by the citizens in the Czech Republic in 2003 – 2007, expressed by means of a slope, (source: authors)

Global Moran's I statistics provided following results:

- Moran's Index = -0,228719
- Expected Index = -0,076923
- Variance = 0,014610
- Z Score = -1,255860

Getis Ord General G statistics provided following results:

- Observed General G = 0,000009
- Expected General G = 0,000009
- General G Variance = 0,000000
- Z Score = 1,383964 Standard Deviations

In both cases inverse distance was used to take into account influence of neighbours. In according to the both indexes null hypothesis of random spatial distribution can be considered.

More detailed views on this attribute are in Fig. 9 showing at first regions classified into 4 quartiles (Fig. 9 A)). Then data are classified in according to standard deviation (Fig. 9 B)). Then, two local indicators of spatial association are shown: Anselin Local Moran's I (Fig. 9 C)) and Getis Ord Gi* (Fig. 9 D)). Two neighbouring regions (Pardubicky and Vysocina, slope 0.0421, 0.0418, respectively) were identified as an area with the highest growth of number people purchasing online, of course besides Prague (slope 0.0528). The regions were identified by Getis Ord Gi* as a hot spot because they are surrounded by regions with much more lover value of this attribute (e.g. Olomoucky, slope 0.0206). Central part of Bohemia (regions Prague, Stredocesky and Liberecky) were identified by Anselin Local Moran statistics as regions which are surrounded by dissimilar values.

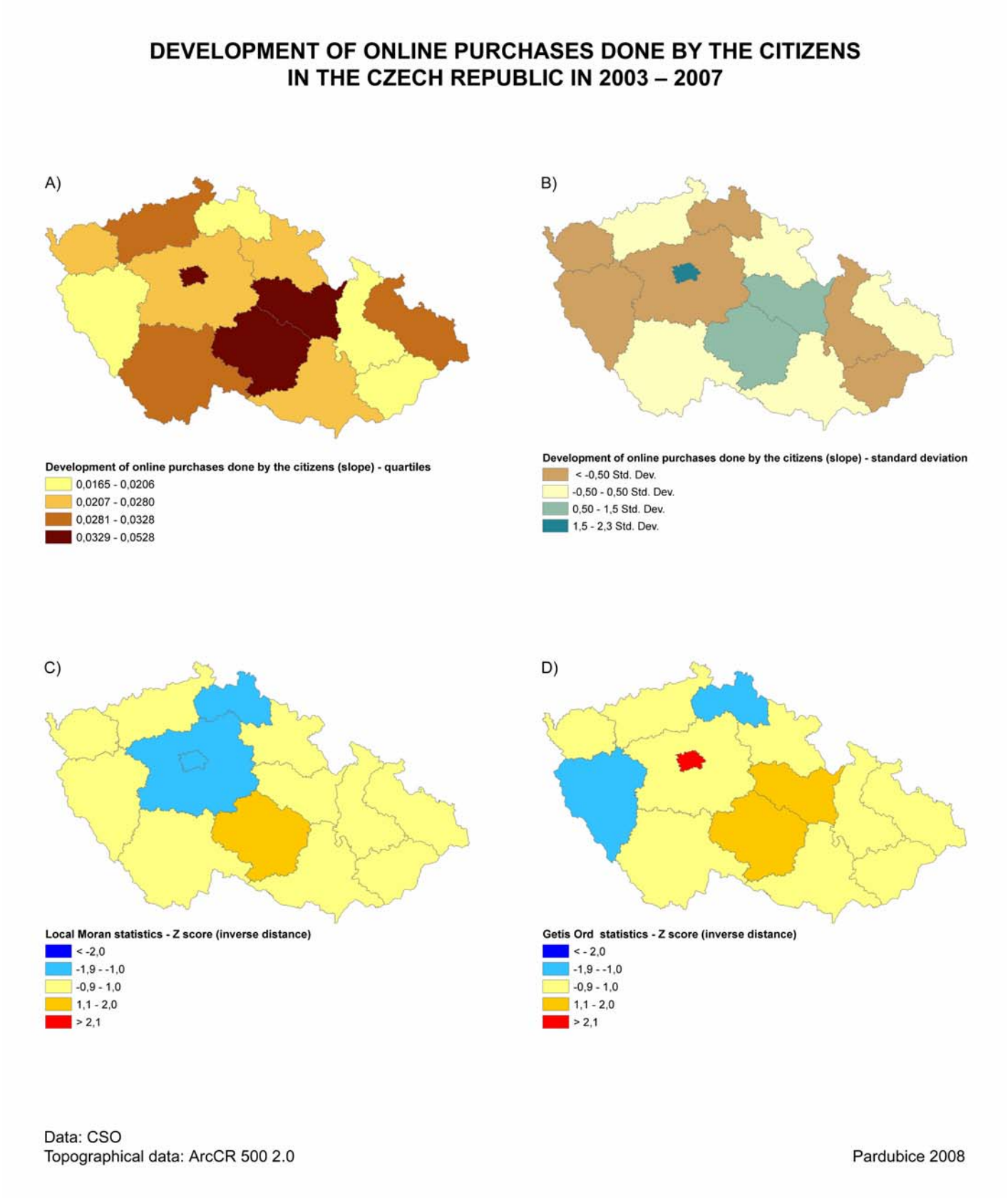

Fig. 9 – Development of online purchases done by the citizens in the Czech Republic in 2003 – 2007 period, expressed as a slope of the straight line: A) Classification by means of quartiles; B) Classification by means of standard deviation; C) Local Moran's statistics – Z score; D) Getis Ord Gi* statistics – Z score, (source: authors)

4 Conclusions and Future Work

The Czech Republic, along with all European countries and EU itself, has strongly supported development of information society so it has invested high amounts of money to support higher utilization of ICT in public administration and introducing eGovernment services. One of the most important reasons is to make participation of the citizens in public deals easier (e.g. to let them participate in electronic way) and consequently higher. Next reason is to make public administration more transparent and cheaper.

Many methodologies of ICT implementation have been developed. On the other side, it is still not very easy to evaluate impact of these investments, their efficiency and compare countries or regions between themselves. The paper proposes utilization of spatial analyses and cartographic visualisation methods as one possible way for evaluation of information society development and comparison of regions.

Case study for the Czech Republic was done and all regions of the republic were compared in according to the available data. Probably the most interesting output is fact that the Capital Prague is a leading region in absolute numbers but its development is becoming slower that development of almost all other regions. This was validated by means of centroids in GIS. As it was shown, people tend to use Internet for online transactions increasingly. It can be understood as a good prerequisite for future utilization of eGovornment services.

Data is a problem in case of this type of studies. It is quite difficult to find relevant and enough detailed data in time series so evaluation of development is still quite a problem. More detailed data and longer time series would allow doing much more precise exploratory spatial data analysis to identify spatial patterns and associations. By this way some problematic or prosperous regions could be identified and lessons could be learned from their experience.

Acknowledgments. This research and paper was created with a kind support of the DEMO-net Network of Excellence project (FP6-2004-IST-4-027219) and Research Project of Faculty of Economics and Administration, University of Pardubice.

References:

[1] Aichholzer, G., Allhutter, D. Evaluation Perspectives and Key Criteria in ePparticipation [CD-ROM]. In *Best Practice and Innovation: 6th International Eastern European e|Gov Days*, Prague, 23th – 25th April 2008, pp. 117-124.

[2] Anselin, L., Bao, S. Exploratory Spatial Data Analysis Linking SpaceStat and ArcView. In Fischer, M. M., Getis, A. (Eds). *Recent Developments in Spatial Analysis: Spatial Statistics, Behavioural Modelling, and Computational Intelligence*. Springer. ISBN 3-540-63180-1, 1997.

[3] ArcGIS Desktop Help [software]. Ver. 9.2. Redlands: ESRI. [cit. 2008-08-28].

[4] ČESKÝ STATISTICKÝ ÚŘAD. *Informační technologie* [online]. [cit.2008-04-19]. Available in Czech only. (Information technologies). URL: <http://www.czso.cz/csu/redakce.nsf/i/informacni_technologie_pm>.

[5] ČESKÝ STATISTICKÝ ÚŘAD. *Sezónně očištěná data* [online]. [cit.2008-03-25]. Available in Czech only. (Seasonal dependency). URL: <http://www.czso.cz/csu/redakce.nsf/i/sezonne_ocistena_data>.

[6] El-Kiki, T., Lawrence, E., Towards a Better Understanding of Efficiency in mGovernment Service. In *Proceedings of the 4th WSEAS Int. Conf. on Information Security, Communications and Computers*, Tenerife, Spain, December 16-18, 2005, pp. 510-519.

[7] Hansen, H.S., Reinau, K.H. The Citizens in E-Participation. In M.A. Wimmer et al. (Eds.): *EGOV 2006*, LNCS, Vol. 4084. Springer-Verlag, Berlin Heidelberg New York, 2006, pp. 70-82.

[8] "*i2010 – A European Information Society for growth and employment*". (2005). [on-line]. Brussels: Commission of the European Communities. [Cit. 2008-03-20]. URL: <http://ec.europa.eu/information_society/eeurope/i2010/i2010/index_en.htm>.

[9] Kopáčková, H., Kopáček, L. Text-based Decision Making with Artificial Immune Systems. *WSEAS TRANSACTIONS on BUSINESS AND ECONOMICS*, Vol. 3, Issue 5, 2006, pp. 428-433.

[10] Langrova, K. *Rozvoj účasti občanů na veřejném dění a využívání služeb informační společnosti v ČR*. Diploma thesis, Pardubice, University of Pardubice, 2008. Supervisor: Jitka Komarkova. Available in Czech only. (The devolvement of the citizens´ participation on public deals and utilization of information society services in the Czech Republic).

[11] Macintosh, A. eParticipation in Policy-making: the Research and the Challenges. In Paul Cunningham and Miriam Cunningham (Eds): *Exploiting the Knowledge Economy: Issues, Applications, Case Studies*. Amsterdam: IOS Press. ISBN: 1-58603-682-3, 2006.

[12] Macpherson, M. *Citizen participation in politics and the new systems of communication* [online]. [cit. 2007-09-01]. URL: <http://home.snafu.de/mjm/CP/cp.html>.

[13] Messner, S.F. et al. The Spatial Patterning of County Homicide Rates: An Application of

Exploratory Spatial Data Analysis. *Journal of Quantitative Criminology*, Vol. 15, No. 4, 1999, pp. 423-449.

[14] MINISTERSTVO VNITRA ČESKÉ REPUBLIKY. *Archív - e-Government* [online]. c2005, [cit.2008-04-19]. Available in Czech only. (Ministry of Interior of the Czech Republic). URL: <http://www.mvcr.cz/micr/egovernment/default.htm>.

[15] *Microsoft Excel 2002: Help: PEARSON* [electronic help]. Microsoft Corporation.

[16] OEMA . *AOEMA : Stages/Phases of e-government* [online]. c2007, [cit.2008-01-02]. Dostupný z: <http://www.aoema.org/E-Government/Stages-Phases_of_e-government.htm>.

[17] Spremic, M., Hrvoje, B. Comparative Analysis of e-Government Implementation Models and Progressive Services. *WSEAS TRANSACTIONS on BUSINESS and ECONOMICS,* Vol. 5, Issue 5, 2008, pp. 260-269.

[18] Spremic, M., Zmirak, Z., Kraljevic, K. Evolving IT Governance Model – Research Study on Croatian Large Companies. *WSEAS TRANSACTIONS on BUSINESS and ECONOMICS*, Vol. 5, Issue 5, 2008, pp. 250-259.

[19] STEM/MARK. *Výzkum informační společnosti-prezentace hlavních výsledků výzkumu. Tisková konference - 25.8. 2005 Hilton Praha* [online]. [Cit.2008-04-19]. Available in Czech only. (Investigation of information society – presentation of main results of investigation) URL: <http://www.dialogin.cz/main.php/>.