QUANTITATIVE ANALYSIS OF SPATIAL INDICATORS IN THE CASE OF VARYING ADMINISTRATIVE BOUNDARIES

Janis Paiders[1], Juris Paiders[2]

Latvia University (Latvia)

ABSTRACT
The aim of the research is to develop methods that would allow comparing the statistical parameters of Latvia at differing administrative divisions, thereby solving the problems that arise in the case of discrepancy of borders. As part of the work, the developed method was implemented in order to assess the suitability of the created tool for mathematical analysis of statistical parameters on Latvia. The territory examined in the work is Latvia; the discrepancy of borders was examined by comparing the parish borders of 1935 with the parishes of 2000. Therefore it was necessary to create a transfer matrix in order to link the administrative borders of both periods. The transfer matrix was prepared in GIS environment, combining the parish borders of 1935 and 2000. By conducting transformation of parish parameters at differing administrative divisions and making a transfer from the parish borders of 1935 to 2000, it is possible to obtain relatively accurate value of the relevant parameter within the present-day borders. The obtained value can be used as a quantitative indicator, allowing for quantitative comparison of temporal changes of the relevant parameter at differing administrative divisions, and the value of those changes in the examined territories can be used in subsequent statistical analysis of the results. By transforming the arable land proportion in rural parishes of 1935 in accordance with the parish borders of 2000, it was possible to carry out temporal comparison of this parameter with its value in 2001.
KEY WORDS: *modifiable areal unit problem (MAUP), quantitative indicators, GIS analysis.*

JEL CODES: B160, N540, R110

Introduction

In geographical studies, hypotheses can be tested by using the spatial properties characteristic to a scalar field (Arbia, 2006). While in a vector field a vector is assigned to each point of space, characterized with direction and size of the vector, in a scalar field each point of space is assigned a certain numerical value (scalar) (Timothy Whitten, 1974). When studying the connection between various scalar fields, the goal is often determining causes or factors in order to use them in predicting the changes of other parameters over time (Moore and Carpenter, 1999; Berger, 2001). In spatial analysis, when examining scalar fields of various indicators, the goal is the change of the spatial indicators over time (Turner 1987; Baker, 1989; Irwin, Geoghegan 2012, etc.). If phenomena can be described as scalar fields, then quantitative analysis can be used to study them; it allows determining which indicators are mutually independent, and which form statistically-significant relationships. Such an approach allows accepting or rejecting hypotheses, discovering common trends and isolating territories with anomalies that do not match the general trends (Paiders, 2013).

The problem encountered by researchers in comparing various statistical indicators and conducting analysis of temporal changes is related to administrative reforms that have taken place at various times (Gregory, 2002; Martí-Henneberg, 2005 etc.). Due to these administrative division changes, it is difficult to

[1] Janis Paiders – master study program student of the Geography and Earth Science Department, Latvia University
E-mail: paidersjanis@inbox.lv

[2] Juris Paiders – PhD (geography), associate professor in the Geography and Earth Science Department, Latvia University
E-mail: jpaiders@inbox.lv

analyze various statistical indicators on a detailed (parish) level, making it more complex to compare the First Republic period in the territory of Latvia with the modern day, in spite of large availability of detailed statistics (Maldups, 1937; Salnītis, 1939; Maldups, 1939).

It means temporal studies that quantitatively compare changes over time mainly analyze them on a countrywide scale (Aizsilnieks, 1968; Strods, 1992), or on the scale of such statistical units where changes have been small (e.g., municipalities). In studies that compare parish indicators of the 1930s with the modern day, their changes have been compared in the form of analysis of the prepared cartographic material, however such a method does not allow for quantitatively precise assessment of the value of changes, which would allow using this change value as an indicator in subsequent mathematical analysis.

Comparison of statistical parameters is relatively simple if the problem of different boundaries has arisen due to several administrative territories merging or splitting, e.g., in the different district divisions in the 1930s and 1940s, where the problem can be solved by summing the different statistical units (Paiders, 2013). However, in order to make it possible to compare parameter values with completely different borders, it is necessary to create more complex methodological solutions.

The goal of the work is to develop methods that would allow comparing the statistical parameters of Latvia at differing administrative divisions, thereby solving the problems that arise in the case of discrepancy of borders. As part of the work, the developed method was implemented in order to assess the suitability of the created tool for mathematical analysis of statistical parameters on Latvia.

1. Data and methods

The territory examined in the work is Latvia; the discrepancy of borders was examined by comparing the parish borders of 1935 with the parishes of 2000. Therefore it was necessary to create a transfer matrix in order to link the administrative borders of both periods. The transfer matrix was prepared in GIS environment, combining the parish borders of 1935 and 2000. A transfer model was thereby obtained in which a transfer was made from the 513 rural parishes in 1935 to the 512 parishes and town rural territories in 2000 (city territories were not included in the model).

The obtained model allows transforming any available statistical parameter that was gathered on the parish level in accordance with the borders of 1935 to an approximate value of the parameter for the borders of 2000. As a result of such transfer, it is then possible to compare the parameter with any statistical parameter of the modern parishes, ensuring compatibility of indicators obtained during periods of varying administrative divisions.

Transformation of statistical parameters of Latvia during the First Republic period to the modern parish borders was conducted with the following formula:

$$X = \sum_{i=1}^{n} \left(Z_{yi} \left(\frac{y_i}{Y_i} \right) \right),$$

where: X – value of the territorial parameter in the parish borders of 2000;

Z_y – value of the parameter in the parish of 1935 as part of the parish of 2000;

y – the area taken by the parish part of 1935 in the parish territory of 2000;

Y – the area taken by the parish part of 1935.

By using this method, it is possible to calculate the approximate value of a 1935 indicator within the parish borders of 2000. The value should be considered approximate, because there are several factors limiting the accuracy of results obtained with such a method. Firstly, such mathematical transformation is based on the assumption that the relevant 1935 parish indicator is internally homogenous. As many indicators exhibit great variation of the parameter also on the parish level (e.g., for distribution of population), it lowers the accuracy of the obtained results. However, this problem cannot be solved, because most statistical parameters of the First Republic period of the Republic of Latvia have not been gathered on a more detailed level. However,

the internal heterogeneity of an indicator is an issue for only some parishes, because for a large part of the parishes, most of the territory of 1935 is part of one parish in 2000, making the heterogeneity of a parish parameter largely unimportant.

Secondly, some inaccuracy is associated with the scale of the parish border materials analyzed in GIS environment. The borders of 1935 have been analyzed on the scale of 1 : 600,000, while the 2000 scale is 1 : 400,000. The analyzed scale slightly reduces the accuracy of the obtained results; however, since the inaccuracies associated with the scale of the cartographic material are too small in the context of the size of the parish territory, their impact can be considered insignificant.

However, in spite of the limitations associated with the selected method, it generally provides accurate results if the encompassed scale of territory is large. When encompassing 512 parishes, which were transferred from the parish borders of 1935 to 2000, the number of values for which the limitations of the applied method provide a potentially approximate value is very low.

The parameter for which, as part of the work, the value has been transformed in accordance with the modern day parish borders, is the proportion of arable land (%) in total agricultural land in 1935 (Maldups, 1937). The examined parameter has been compared with the proportion of arable land in the total area of farm land in 2001 (LR CSP, 2003).

Linear regression analysis has been used as the mathematical analysis method to determine the interrelation of the parameters. A correlation and determination coefficient has been used to characterize the closeness of the relationships. As part of the regression analysis, the significance of the respective interrelation was determined at 95 % probability ($p < 0.05$).

GIS analysis, by establishing a base of transfer of the parishes of 1935 to the borders of 2000, was conducted in GRASS 6.4.3. All working calculations were performed in Microsoft Excel. In addition, as part of the work, cartographic materials were prepared by using QGIS 1.7.2. The cartographic materials and spatial analysis were prepared in the LKS92 / Latvia TM coordinate system (EPSG: 3059).

2. Results

In the case of Latvia, there are two problematic issues. The first: the borders of the administrative unit (parish) do not change, but the division of higher administrative units (districts) does. The second is a more difficult issue: change of the borders of the smallest administrative unit (parish).

Let us first examine the former case. It applies to data on Latvia from 1920 to 1946 and from 1991 to the present day. In order to conduct statistical analysis of data in a period where change of the administrative division has occurred, data calibration must be carried out; they are reduced to the numerically lowest administrative unit system. For example, in order to include in a single statistical analysis statistical indicators of Latvia in the period up to 1940 and statistics of the Soviet period, it is more convenient to transform all Soviet period data in accordance with the district structure as on June 1940. At that time, the administrative reform was carried out by dividing larger districts into smaller ones, or into even smaller areas (administrative division of 1950). Data of administrative units is summed in such a way as to ensure that the indicator matches the district structure of 1940. All data after 1945 were transformed as if the old district borders (up to 1940) were still in effect during the Soviet period. In this example, all data, including newer data, are transformed to an older system of administrative division in order to make them intercomparable. On the other hand, to conduct quantitative analysis regarding the period from 1991 to 2012, the reverse procedure must be performed. At that time, the administrative reform was taking place by merging parishes or adding parishes to municipalities, ending with establishment of the municipality structure. To make the data comparable, the absolute values of parish and town data must be summed in accordance with their association with a municipality as in 2012. As a result, absolute and relative indicators are obtained regarding municipalities for any period of time – also for the time when the municipalities had not been established. In this case, older data are transformed and calibrated in accordance with the new system of administrative division.

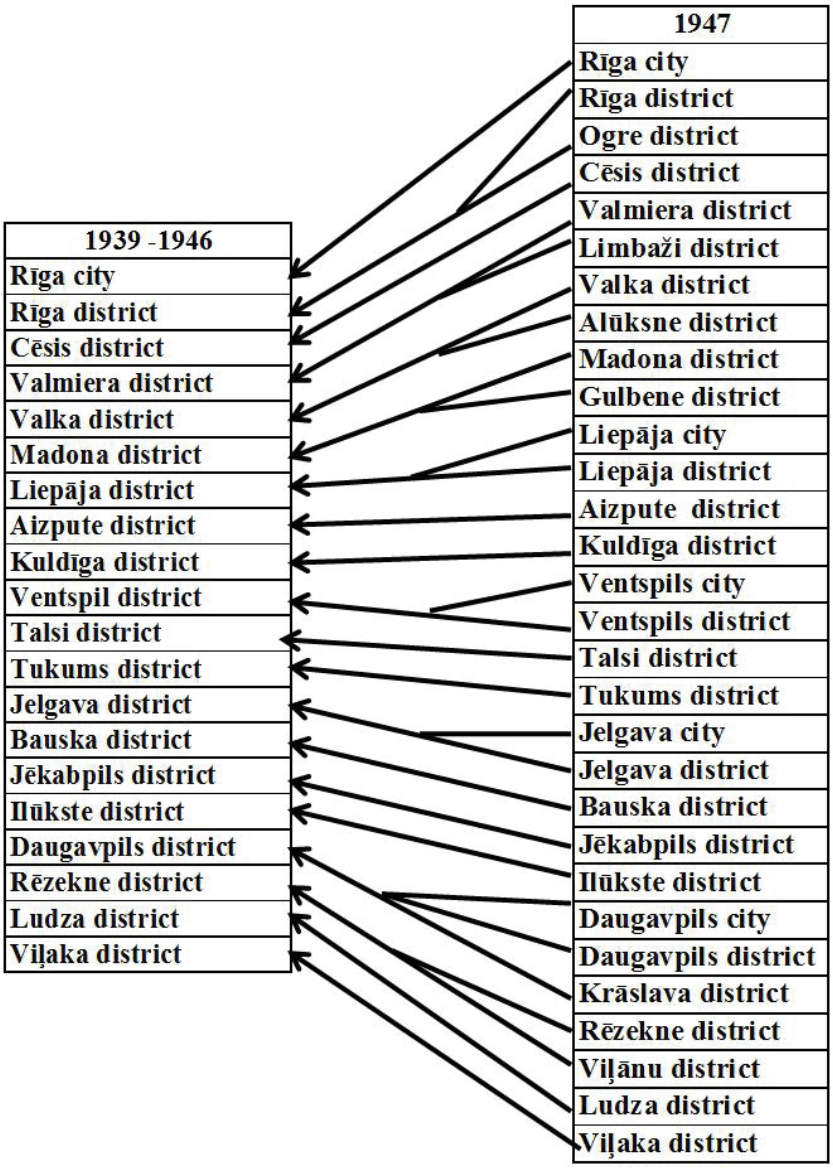

Figure 1. Transformation methodology to combine the district borders of 1939–1946 with the district borders of 1949

Source: author's calculations

The figure shows the transformation principle, demonstrating how to combine district borders of differing periods (Figure 1). In this example, the transformation principle is simple, as the number of districts has changed at various comparable times, but the changes have taken black through districts splitting and the old borders remaining unchanged. In such a case, changes of the parameter over time are compared by going back to the old borders and adding together the values of the relevant parameters.

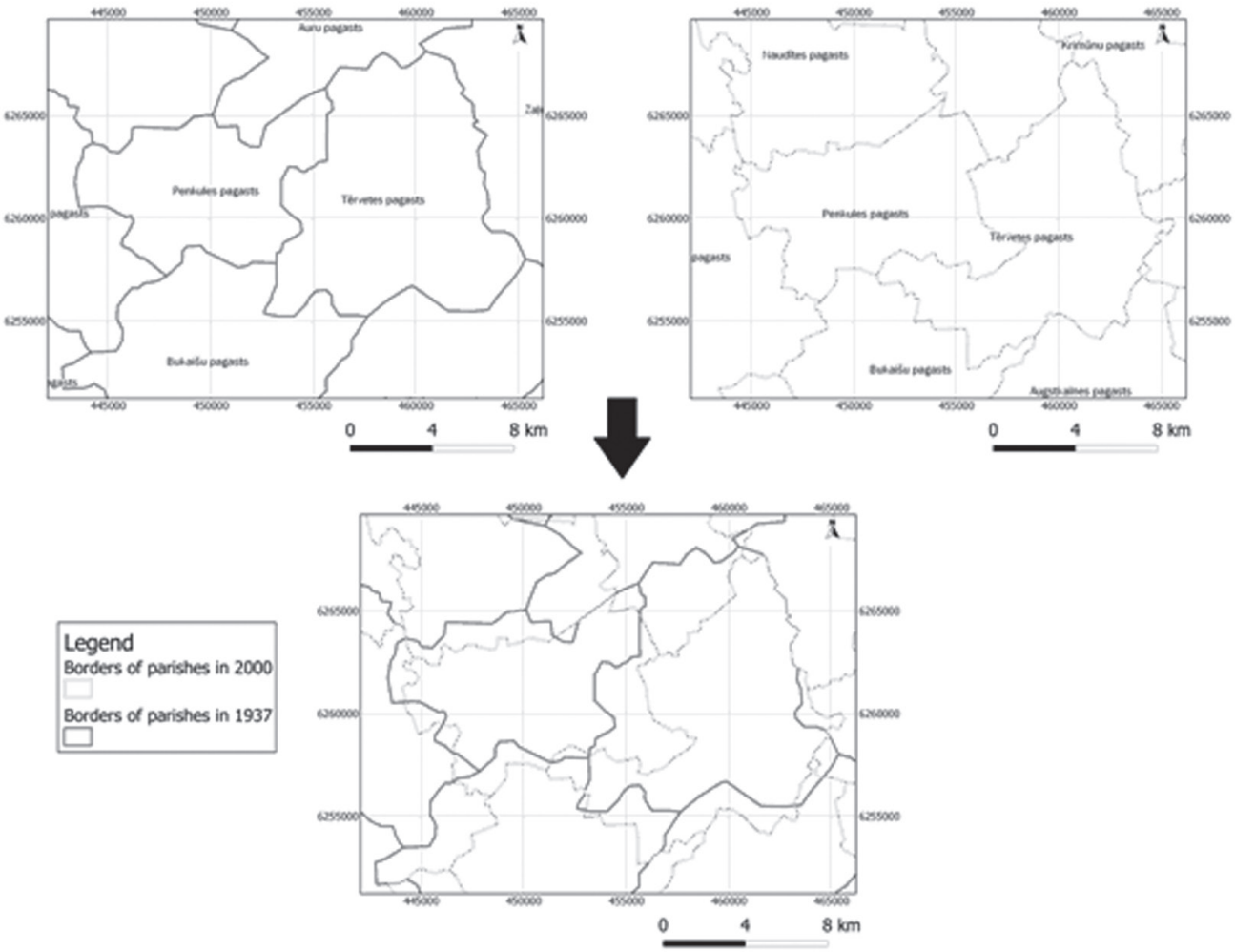

Figure 2. Transformation methodology to combine the parish borders of 1935 with the parish borders of 2000

Source: author's calculations

Carrying out transformation from the parish borders of 1935 to 2000 is much more complicated than in the case of districts (Figure 1), due to entirely changed borders. When examining a fragment of these border differences (Figure 2), it can be seen that only a very small part of the borders of 1935 are unchanged in 2000. Unchanged borders, when comparing both periods, were most often found in places where the parish border has followed along a natural obstacle (mostly various watercourses).

However, it can be seen in the presented fragment of parish borders that the central two parishes (Tērvete parish in Tērvete municipality and Penkule parish in Dobele municipality), in spite of changed parish borders, have very large overlap of the parish territory in the compared periods. In the Tērvete parish of 2000, 78 % of the total area is comprised of the Tērvete parish of 1935. In the Penkule parish of 2000, 68 % of the total area consists of the same parish of 1935. Therefore, when carrying out transformation to the borders of 2000, most of the values of the obtained transformation parameter for both of these parishes will be determined in the value of one parish. It means that, in such a case, the potential internal heterogeneity of a particular transformed parameter will be relatively insignificant.

It must be added that the applied methodology reduces the total range of the original parish indicator values. As a result of the transformation, the size of extreme parameter values is reduced, as in the new borders they slightly converge with neighboring territories. However, the amount of such convergence can also be very small if most of the parish area of 2000 is formed by a single parish of 1935.

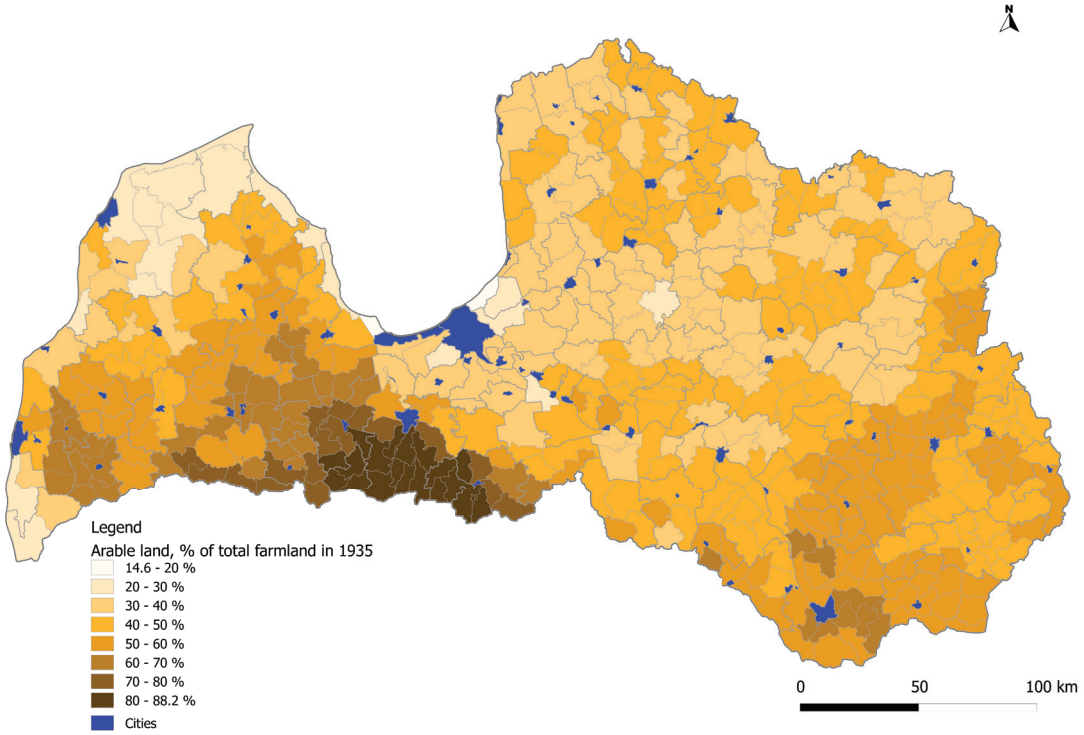

Figure 3. The proportion of arable land in total farm land in 1935 in accordance with the borders of 2000

Source: Agricultural census of 1935, author's calculations

When looking at the layout of arable land in 1935 in accordance with the parish borders of 2000 (Figure 3), it is possible to observe significant clusterization of rural parishes with a high proportion of arable land. Parishes with the highest proportion of arable land are concentrated in Zemgale, in the southern part of the former Jelgava and Bauska districts, along the Lithuanian border. In those parishes, the proportion of arable land exceeds 80 % of the total farm land, which signifies pronounced specialization of the territories in agriculture. The parishes in question also contain the most fertile soil in Latvia that is best suited for agriculture.

The lowest arable land usage intensity in 1935 is observed in parishes along the seashore, as well as in the vicinity of Riga, where it is lower than 30 % of the total farm land – significantly below the average usage level of arable land in the territory of Latvia (47 %). In these coastal parishes, as well as in the vicinity of Riga, the soil is infertile and sandy, with poor suitability for agriculture. Under such conditions, cattle-breeding has a greater role, and consequently part of the total land area is taken by pastures and meadows.

A significant role in the usage intensity of arable land in 1935 was also played by population density. For example, Latgale Upland is comparatively unsuitable for agriculture in terms of natural factors; however, the large population density dictated the need for large arable land areas. The population density of rural parishes located in Latgale Upland was, on average, twice as high as in the rest of the rural territories of Latvia, exceeding 50 inhabitants per one square kilometer of land in certain parishes.

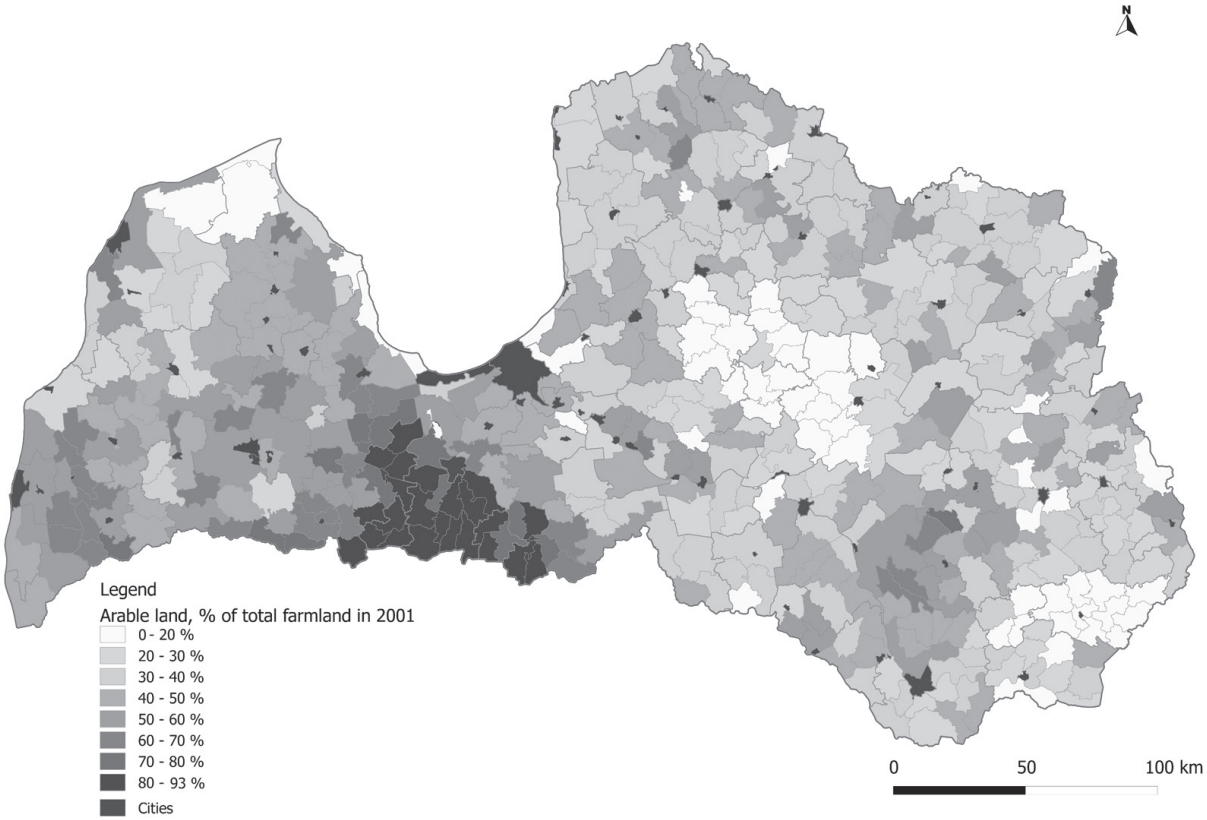

Figure 4. Proportion of arable land in total farm land in 2001

Source: Agricultural census of 2001, author's calculations

In 2001, the proportion of arable land in total farm land has very large value dispersion across the territory of Latvia (Figure 4). The usage intensity of arable land in parishes of Jelgava, Bauska, Auce and Dobele municipalities is very high, exceeding 80 % of the total farm land. This area with very high arable land usage intensity has largely remained unchanged since 1935, with some expansion in the western direction, in the territory of Dobele and Tukums municipalities. The reasons for the great role of agriculture in these parishes have also remained unchanged, mainly based on the natural factors (soil fertility).

Parishes with low usage intensity of arable land have formed mainly in Vidzeme Upland, Latgale Upland, as well in some of the coastal parishes (except along the coast of Kurzeme). In these parishes, arable land comprises less than 20 % of the total farm land, more than twice as low as the average level of Latvia in 2011 (42 %). The reasons dictating the low agricultural land usage intensity in these territories are also connected with the relative unsuitability of the areas for agriculture (poor soil fertility).

It can be noted that the value range of the arable land proportion has significantly increased in the period between 1935 and 2001. In 1935 this indicator was within 14.6–88.2 %, while in 2001 the range has increased to 0–93 %. The reason for the increased polarization between territories suitable and unsuitable for agriculture is largely connected with the fact that the transport costs of produce in the period between 1935 and 2001 have very significantly decreased, which reduces the need for agricultural production in territories unsuitable for such purpose.

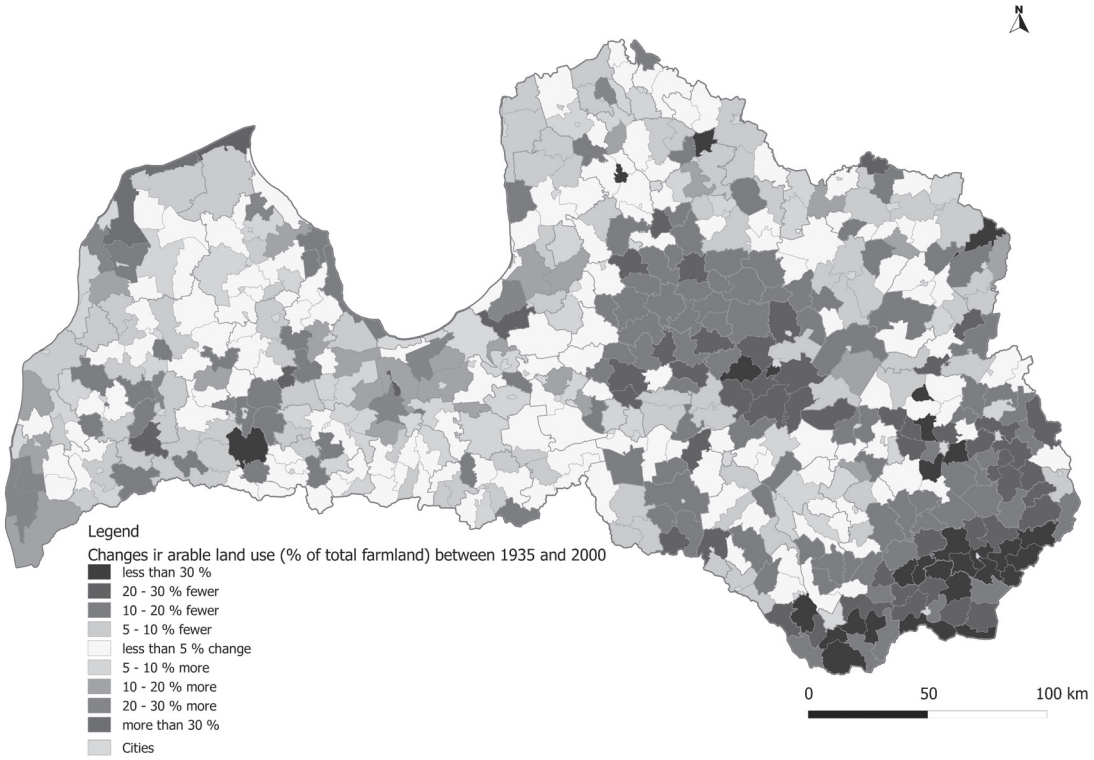

Figure 5. Variations in the proportion of arable land in total farm land between 1935 and 2001

Source: Agricultural censuses of 1935 and 2001, author's calculations

When looking at the variations in arable land usage intensity between 1935 and 2001 (Figure 5), it can be noted that the greatest arable land decline has been observed in a large part of Latgale parishes, as well as in the Vidzeme Upland area. The average decline in arable land in these parishes has been greater than 20 %, even exceeding 40 % in some parishes of Daugavpils and Dagda municipality. The significant decline in the role of agriculture in these parishes is largely connected with the gradual decline of rural populations in the territory of Latvia during the examined period, which has particularly affected the rural parishes of Latgale. Under such circumstances, with the number of rural inhabitants significantly falling, there is no more need for large arable land areas, particularly if they are not well-suited for agriculture.

The biggest increase of the arable land proportion has been noted in parishes around Riga, Liepāja, Jelgava and Ventspils. It is very likely that the significant increase of arable land usage intensity in these parishes is connected with the increasing population of the respective cities between 1935 and 2001 (particularly in Riga) and consequent increasing consumption of agricultural produce. It must be added that the overall arable land proportion between 1935 and 2001 has fallen from 47 % to 42 % of the total farm land.

Discussion

Geographical research with the application of scalar field properties is limited by several factors. One of them is the modifiable areal unit problem (MAUP). Gehlke and Biehl (1934) discovered that the correlation coefficient is sensitive to scale changes of the examined territories. Openshaw and Taylor (1979) began using the term MAUP for the investigation and assessment of this problem. The focus has largely been placed on research of how spatial models are affected by scale changes (Fotheringham and Wong, 1991, Briant et al., 2010, etc.).

The authors look at a very specific MAUP case – what methods can be used in cases where spatial statistical data are available in accordance with the administrative division, under conditions where

significant variation of administrative divisions is observed. Sensitivity to scale (Fotheringham & Wong, 1991) Although small units with high spatial accuracy can present unreliable rates, large spatial units may remove relevant geographical variation (Nakaya, 2000).

Conclusions

By conducting transformation of parish parameters at differing administrative divisions and making a transfer from the parish borders of 1935 to 2000, it is possible to obtain relatively accurate value of the relevant parameter within the present-day borders.

The obtained value can be used as a quantitative indicator, allowing for quantitative comparison of temporal changes of the relevant parameter at differing administrative divisions, and the value of those changes in the examined territories can be used in subsequent statistical analysis of the results.

By transforming the arable land proportion in rural parishes of 1935 in accordance with the parish borders of 2000, it was possible to carry out temporal comparison of this parameter with its value in 2001.

The territories were found in which the most significant changes of arable land usage intensity have occurred, with very significant decline in arable land use found in parishes of Latgale and Vidzeme Upland, while a very notable increase was observed in parishes near Riga and other largest cities of Latvia.

Belonging to one cluster of scalar fields does not mean direct causality and impact, but may usually mean that the respective indicators are different dimensions of one common scalar (or vector) field.

References

Aizsilnieks, A. (1968). *Latvijas saimniecības vēsture 1914−1945 (Latvian economic history 1914–1945)*. Stokholma, Daugava, 983.lpp.

Arbia, G. (2006). *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*. New York, Springer, 207 p.

Baker, W. L. 1989. A review of models of landscape change. *Landscape Ecology*, Vol. 2 (2), p. 111–133.

Berger, T. (2001). Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. *Agricultural Economics*, Vol. 25 (2–3), p. 245–260.

Briant, A., Combes, P. P., Lafourcade, M. (2010). Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*, Vol. 67 (3), May, p. 287–302.

Fotheringham, A. S., Wong, D. W. S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, Vol. 23(7), p. 1025–1044.

Gehlke, C., Biehl, K. (1934). Certain Effects on Grouping upon the Size of the Correlation Coefficient in Census Tract Material. *Journal of the American Statistical Association*, Vol. 29 (185), p. 169–170.

Gregory, I. N. (2002). Time-variant GIS Databases of Changing Historical Administrative Boundaries: A European Comparison. *Transactions in GIS*, Vol. 6 (2), p. 161–178.

Irwin, E. G., Geoghegan, J. (2012). Theory, data, methods: developing spatially explicit economic models of land use change. *Agriculture, Ecosystems & Environment*, Vol. 85 (1–3), p. 7–24.

LR CSP. (2003). *Latvijas 2001. gada lauksaimniecības skaitīšanas rezultāti (Results of the 2001 agricultural census in Latvia)*. Latvijas Centrālā Statistikas Pārvalde.

Maldups, A. (1937). Apriņķu un pagastu apraksti (Descriptions of districts and parishes). Valsts Statistikas Pārvaldes izdevums. Rīga, 654. lpp.

Maldups, A. (1940). *Lauksaimniecības skaitīšana Latvijā 1939.gadā (Agriculture census in Latvia in 1939)*. Valsts Statistikas Pārvaldes izdevums. Rīga, 128. lpp.

Martí-Henneberg J. (2005). The Map of Europe: Continuity and Change in Administrative Boundaries (1850–2000). *Geopolitics*, Vol. 10 (4), p. 791–815.

Moore, D. A., Carpenter, T. E. (1999). Spatial Analytical Methods and Geographic Information Systems: Use in Health Research and Epidemiology. *Epidemiologic Reviews*, Vol. 21 (2), p. 143–161.

Nakaya, T. (2000). An information statistical approach to the modifiable areal unit problem in incidence rate maps. *Environment and Planning A*, Vol. 32(1), p. 91–109.

Paiders, J. (2013). *Skalāra lauka regresijas modeļu interpretācijas problēmas un lietojuma ierobežojumi. Telpisko datu kalibrēšana, mainoties administratīvam iedalījumam. Ģeogrāfija. Ģeoloģija.* Vides Zinātnes. Referātu tēzes. Latvijas Universitāte, 494. lpp.

Salnītis, V. (1939). *Ceturtā tautas skaitīšana Latvijā 1935. gadā (The fourth census in Latvia in 1935)*. Valsts Statistikas Pārvaldes izdevums. Rīga, 672. lpp.
Strods, H. (1992). *Latvijas lauksamniecības vesture (Agricultural history of Latvia)*. Rīga: Zvaigzne, 287. lpp.
Timothy Whitten, E. H. (1974). Scalar and directional field and analytical data for spatial variability studies. *Journal of the International Association for Mathematical Geology*, Vol. 6 (2), p. 183–198.
Turner, M. G. (1987). Spatial simulation of landscape changes in Georgia: A comparison of 3 transition models. *Landscape Ecology*, Vol. 1 (1), p. 29–36.
William, L. B. (1989). A review of models of landscape change. *Landscape Ecology*, Vol. 2(2), p. 111−133.

ERDVINIŲ RODIKLIŲ KIEKYBINĖ ANALIZĖ ADMINISTRACINIO SUSKIRSTYMO TERITORIJOSE

Janis Paiders, Juris Paiders
Latvijos universitetas (Latvija)

Santrauka

Šiame straipsnyje analizuojamas metodo, kurio pagrindu lyginami statistiniai duomenys Latvijos administracinio suskirstymo teritorijose, taikymas. Taip siekiama nustatyti problemas, kurios kilo dėl administracinių teritorijų ribų kaitos. Latvijos administracinių teritorijų ribų kaita tiriama lyginant seniūnijų ribų duomenis 1935 ir 2000 m., todėl tyrimo metu gauti tikslūs rezultatai atitinka šiandienos situaciją. Tyrimo rezultatai gali būti naudojami kaip kiekybiniai rodikliai, kurių pagrindu atliekamas kiekybinis skirtingų statistinių parametrų palyginimas skirtingose administracinėse teritorijose, nustatytų pokyčių reikšmės gali būti taikomos tyrimo teritorijose siekiant atlikti gautų rezultatų statistinę analizę.

PAGRINDINIAI ŽODŽIAI: *kintančio arealo vieneto problema (angl.* MAUP*), kiekybiniai rodikliai, GIS analizė.*

JEL KLASIFIKACIJA: B16, C0, C46, R1.