

УДК 519.25: 930

Spatial Approach to the Analysis of the Employment Data in Siberia Based on the 1897 Census (the Experience of Multivariate Statistical Analysis of the District's Data)

**Elena A. Bryukhanova, Sergey V. Dronov
and Oksana I. Chekryzhova***
*Altai State University
61a Lenina pr., Barnaul, 656049, Russia*

Received 12.03.2016, received in revised form 24.05.2016, accepted 18.06.2016

The article presents an original approach to regionalization of Siberia at the turn of 19th -20th centuries. The subdivision into districts is based on the typology of districts (uezds) in Siberia, which allows to make a more detailed spatial analysis of the data on employment of the 1897 census. The key source of the research has become an open information resource developed on the basis of the published materials of the 1897 census by the number of occupational groups, their gender, age and ethnic composition.

Application of mathematical methods such as cluster analysis, post-hoc method and coefficient of cluster differences, allowed not only to identify the socio-demographic and economic types of Siberian districts, but also to rank the importance of the factors for determining the “appearance” of the districts. Further implementation of the cartographic method to analyze the sets of clusters defined spatial boundaries of socio-economic districts in Siberia. This study contributes to comprehensive understanding of the peculiarities of development of Siberian territories and enhancement of modernization processes at the turn of the 19th – 20th centuries.

Keywords: Siberia, employment, database, the 1897 census, cluster analysis, cartographic analysis, regionalization.

The research is supported by the Russian Foundation for Basic Research, project No. 15-06-07553A “Typology of employment in the regions of Russian Empire at the end of 19th – beginning of 20th century (on the basis of multivariate statistical analysis of the first overall census in 1897).

DOI: 10.17516/1997-1370-2016-9-7-1651-1660.

Research area: history.

Spatial analysis and subdivision into districts in Siberia

Spatial analysis in historical research has got its new development in the beginning of the 21st

century, in connection with the use of information technology in data processing and analysis (Vladimirov, 1998). Currently, geographic information systems based on historical material

are actively used. They are implemented for storage, processing and interpretation of data in the field of archeology, land use, migration, history of human settlements, urban space, socio-demographic and economic processes (Frolov, 2013; Konchakov, Baranova, 2010), which also include employment. At the same time, spatial analysis can also be an independent research task, which involves analyzing the characteristics of the objects located in space, acquiring new knowledge (Pashkina 2011: 157), for example, to identify some districts with similar characteristics. In this context, the issues of Siberia historical zoning are of particular interest due to the ambiguity of the decision.

Thus, the pre-revolutionary historiography, as noted by the researcher S.V. Rasskazov (2009), was characterized by the common Siberian approach: from the middle of the 19th until the early 20th century most of regionalization projects (K.I. Arseniev, A.I. Chuprov, D.I. Mendeleev), even if they included Siberia in the map of districts, showed it as a single monolithic region. The division of Siberia into East and West was recorded in the writings of the famous geographers P.P. and V.P. Semenov-Tyen-Shansky. The first attempts of Siberia economic regionalization were undertaken during the World War I, when in the conditions of militarization of the economy, the government discussed the projects on the division of the territory into 4 districts. Practical economic regionalization of Siberia took place in the Soviet period and was associated with significant economic transformation of the region (Ratkovskaya, Timoshenko 2013). In the writings of Soviet and many modern historians, the administrative and territorial approach to the spatial study of the pre-revolutionary Siberia has gained popularity due to the peculiarities of the existing source systems. At the same time, the level of provinces and districts was the most traditional one, which could not reflect

the diversity of economic areas for such a large territory as Siberia.

The approach of geographers to historical-geographical zoning of the territory of Siberia is quite interesting (T.M. Pozdnyakova (2012), S.V. Rasskazov (2009), L.B. Vampilova (2012), B.A. Seredovskih, V.I. Bulatov (2013)). The use a whole range of factors combined in five groups: 1) landscape, 2) paleogeographical, 3) ethnocultural, 4) political and administrative, 5) socio-economic.

The use of the multifactorial approach in historical research also seems to be promising, but in comparison with provinces, at a more detailed level, i.e. at the level of districts and counties. Consistent statistical data about Siberian district appeared due to the 1897 census. Such details will help to identify regions similar in economic development. In this study, we present the spatial approach to the analysis of Siberian districts at the turn of the 19th – 20th centuries, based on multivariate statistical analysis of socio-demographic and occupational data of the 1897 census. At the same time, employment of the district's population in a particular area is considered as the leading socio-economic factor.

Digital sources and mathematical methods

The spatial approach is based on the mathematical analysis of statistical data for certain administrative-territorial units of Siberia, which correspond to districts (districts and counties). A comprehensive statistical source for socio-demographic and economic features of the district's population in Siberia are the materials of the first general census of the Russian Empire in 1897. For the first time it brought together and unified the data in terms of general programme. By the time of the 1897 census in Siberia, there were 55 districts and counties. The most detailed information about Siberian districts was published

in 9 editions on the provinces (*guberniyas*) (Troinitskii, 1898-1905), which included tables with the total number of people, their gender, age and ethnic group, as well as distributed urban and rural residents into 65 groups of activities, including those giving additional income (avocations). Moreover, the characteristics of the employment in Siberian districts were published in the Empire summary of the results of the first national census (T.II) data development and special edition “Population distribution by types of vocations and age groups by specific territorial areas” (T.IV), which included the information about the employees by 390 types of activities.

It is hard to imagine the work with a similar volume of statistical data in the modern study without digital processing. Thus, the published materials of different issues of the 1897 census results, which are cumbersome statistical tables, have been merged into a single thematic database based on MySQL. On the one hand, its linking elements are administrative-territorial units, on the other hand, groups and types of people’s occupations¹. Currently, the database contains information on all regions of the Russian Empire (89 provinces, 800 districts and 931 cities and equivalent administrative units) with the following indicators: number of groups and types of activities, number of employees and their families, men, women, and additional features of professional groups such as age, nationality, avocations (for the persons involved in rural or nomadic farming, fishing and hunting) and literacy of the specific professional category - workers.

On the basis of the database an open information Internet-system “Professions and occupations of the people in the Russian Empire at the end of the 19th – beginning of the 20th centuries” was developed ([Http://hcod.asu.ru/data/](http://hcod.asu.ru/data/)) with partial implementation of the spatial approach to visualization of the data

on employment providing the reflection of the information on the number of employees in a specific occupational group in all provinces of the region on the interactive maps of the region (when the view mode “by administrative division” is chosen) (Bryukhanova, Vladimirov, Ivanov, Chekryzhova 2014).

For the calculation of the majority of socio-demographic indicators (the percent of the male population, the ratio of ethnic groups, the number of people of working age, etc.), the data of historical applications of the project “DemoscopeWeekly” (<http://demoscope.ru/weekly>) were used.

The given digital arrays were the main sources of the data for statistical analysis.

Multivariate statistical analysis had long ago become the tool of historical research aimed at understanding of socio-demographic and economic phenomena, as well as determination of some characteristics of the region’s space. Development of ideas about the methods of multivariate statistical analysis as a special method of historical science is associated with the names of I.D. Kovalchenko and L.I. Borodkin, whose works had become the theoretical and practical basis for the use of these methods for the analysis of historical materials. Modern trends in the use of statistical analysis in historical research are presented in the section on quantitative history of the Information Bulletin “History and Computer” (<http://aik-sng.ru/content>), as well as on the pages of the periodical “Historical Computer Science” (<http://kleio.asu.ru/index.html>).

Methods of multivariate statistical analysis were used for data processing of the 1897 census in the context of agrarian typology of provinces in European Russia (I.D. Kovalchenko, L.I. Borodkin (1979)), typology of cities in the provinces of the Volga region of European Russia (A.V. Konovalov (2003)), the study of the social and class structure of the population of Belarus (A.G. Kokhanovskii, Valegina (2009)). However,

it is the first time when the multivariate statistical analysis is applied for zoning of the territory of Siberia at the turn of the 19th- 20th centuries based on the data of people's occupations.

To solve the given tasks, the method of cluster analysis of socio-demographic and economic (people's occupations) features of Siberian districts was chosen, while the cluster analysis of each set of characteristics was a separate phase of the research. Separate consideration of the factors was aimed at verification of the given districts grouped into regions.

Selection of features for the analysis and evaluation of their significance required special attention. The important socio-demographic factors were: the proportion of men, the proportion of the Russian population (which was calculated as the sum of the data of the following ethnic groups: Great Russians, Belarusians, Little Russians), the proportion of the urban population, the proportion of people of working age (calculated as the ratio of total population to the sum of the age groups in the range of 10-70 years), the proportion of the employed (amateur) population and the employment rate of the district's population (calculated as the ratio of the number of employed people to the number of the people of working age (Ivanov, 2002: 78), the proportion of persons employed in agriculture and with additional sources of income (avocations).

Selection of significant economic factors required preliminary preparation. Firstly, grouping of the data on employment undertaken by the developers of the 1897 census, proved to be very cumbersome and included at least 65 groups, while the number of the smallest group ("Tobacco and its products") in Siberia as a whole was only 27 people, and the most popular one ("Agriculture") was more than 790 thousand people. Secondly, the most common models of the professional structure of the population on the basis of aggregated data of the 1897 census

both for the European Russia (V.I. Lenin, B.N. Mironov) and for the urban population of Siberia (V.A. Skubnevskii, Yu.M. Goncharov) also included groups of activities, which tended to zero for the population of Siberian districts (for example, "Finance"). At the same time, all activities of the agricultural sector fell in the same group "Agriculture" (Bryukhanova, 2011). Ranging groups of activities according to the number of employed persons in Siberia as a whole showed a statistical significance of all agricultural sectors: agriculture, cattle breeding, fishing and forestry; whereas most intellectual jobs in the field of healthcare, education, etc., as well as the number of religious officers turned out to be very small. "Armed Forces" and "Imprisoned people and those serving sentences" were put in specific groups, which occupy 5 and 12 lines of the rating and are identified as significant for the whole Siberia. Thus, 14 groups of occupations had been formed, i.e. the sectors of activities, the number of which was statistically significant (Table 1). The importance of each of the selected factors was analyzed using the analysis of variance - post-hoc method for determining the degree of variability of the average values from cluster to cluster (Dronov, Sazonova, 2014).

Another objective of the study was to compare the socio-demographic and occupational clusters aimed at verifying the objectivity of the sets of Siberian districts. Mathematics does not use the common method of cluster comparisons nowadays. This study used the coefficient of cluster differences previously developed by one of the authors of this article (Dronov, Dementjeva, 2012).

An independent task allowing to implement the spatial analysis of the data on employment, is the localization of the obtained clusters on the map of Siberia at the turn of the 19th – 20th centuries, which is based on the administrative map of the Russian Empire in 1907. The cartographic method

was used not only to visualize the clusters, but also as a tool for spatial analysis of the boundaries of the obtained clusters.

The considered combination of methods seems to be appropriate to detect a similar socio-economic areas of Siberia at the turn of the 19th – 20th centuries.

Interpretation of the results

To determine the optimal number of clusters, the hierarchical agglomerative clustering algorithm was preliminarily applied. Basing on the analysis of its results for each of the groups of characteristics, the optimal number of clusters was decided upon. Following the idea of individual consideration of socio-demographic characteristics and employment factors, two sets of 7 clusters were formed on the basis of each combination of factors: socio-demographic characteristics and numeric factors of population distribution by occupations correspondingly.

It should be noted that by using the post-hoc method for determining the degree of variability in mean values of indicators from cluster to cluster, it was determined that all selected characteristics were significant in both combinations of features. On the basis of the Fischer-Snedecor statistical values (F) the rating of indicators was developed, according to which the formation of clusters by socio-demographic indicators was mainly influenced by ethnic and gender factors, as well as the number of persons with additional sources of income (avocations). Clustering of the districts on the basis of activity types was mainly determined by employment in the field of agriculture, the presence and number of prisons and penal institutions, the number of people employed in fishing and cattle breeding. The last line of the rating is the factor “unearned income”, which was the least to determine the “appearance” of the districts. The rating of indicators is given in Table 1.

The next stage of the study was to compare the sets of the clusters using the coefficient of cluster differences (Dronov, Dementjeva, 2012). As shown in the work cited above, this coefficient is equal to 0 only for a completely matching partitions and the stronger the partitions differ, the greater values it takes. The most differing partitions (and only them) match the value of the coefficient equal to 1. The ratio of cluster differences for clusters by the socio-demographic characteristics and clusters by the employed population had the value of 0.12996, which means a high degree of similarity of the compared sets of clusters. Moreover, it can be concluded that there is an interdependence of socio-demographic characteristics and types of employment in Siberian districts.

Further, two sets of clusters were localized on the model map. The example of the districts’ distribution by clusters basing on the employment is shown at Fig. 1.

The spatial alignment of geographically distributed two sets of clusters showed “isolation” of four groups of districts that are similar both in socio-demographic and economic indicators: the northern, Primorsky, Central Siberia and the southwestern districts adjoined to Pribaikalsky region.

The critical indicator for the northern group of districts is the ethnic composition of population (the proportion of the Russian population was on the average 14%). The northern districts are characterized by a high percentage of persons with an additional source of income (84%) and higher than the average in Siberia (27.5%) employment rate of the population (29%). The northern group of districts incorporated two employment clusters with such leading activities as fishing and timber production (the average indicator for the cluster - 70% of employees) and cattle breeding (66%). The northern group includes the largest areas and the most sparsely populated districts: Berezovsky

Table 1. Analysis of the significance of indicators (Post-hoc method)

Rating of the indicator	Socio-demographic indicators		Rating of the indicator	Occupation groups – fields of population’s activities	
	Name of the indicator	F		Name of the indicator	F
1	The proportion of the Russian population	110.4	1	Agriculture	158.1
2	The proportion of the male population	87.6	2	Imprisoned	128.2
3	The percentage of persons with avocations	80.0	3	Fishing*	77.8
4	The proportion of employed population	55.3	4	Cattle breeding	34.7
5	The employment coefficient	44.7	5	Services*	10.9
6	The percentage of the population of working age	36.7	6	Armed forces	8.3
7	The proportion of the urban population	5.7	7	Construction	7.5
			8	Transport and communications*	7.0
			9	Intellectual and creative jobs, religion*	5.4
			10	Production*	5.0
			11	Trade*	4.8
			12	Unemployed	3.4
			13	State and local administration	2.5
			14	Unearned income*	1.3

*These fields include the following aggregated groups from the source: Fishing – groups 18, 20-21; Services - groups 13, 60-62, 64; Transport and communications - groups 41-45; Intellectual and creative jobs, religion - groups 3, 5-12, 63; Production - 22-37, 39-40; Trade - groups 46-59; Unearned income - 14-15.

and Surgutsky counties of Tobolsk Province, Turukhansky county of Yeniseisk province, all counties (except for Olekminsky) of the Yakutsk region, northern counties (Anadyrsky, Petropavlovsky, Komandorsky, Okhotsky) of Primorsky region.

Primorsky group of districts differs by significant disproportion of the gender composition (70% of men), the highest proportion of the employed population compared to other clusters (59%) and a high percentage of urban residents (25%) with a low percentage of people who have additional sources of income (29%). Primorsky group of districts as well as the previous one includes two clusters by population’s activities, one of which is the military-peripheral (Udinsky, Khabarovsky, Ussuriysky and Yuzhno-Ussuriysky

counties of Primorsky region, Olekminsky is the only county of Yakutsk Province not included in the previous group), is characterized by the highest proportion of military officers if compared to other clusters - the average indicator was 16%; while the second cluster (the whole island Sakhalin) is characterized by the highest number of imprisoned persons and those serving sentences (33% of the total population).

The group of Central Siberian districts is the most complex in composition and includes two clusters by employment and three socio-demographic clusters. In accordance with the social and economic clusters included in the group, the following “dispersed” subgroups can be allocated: “non-agricultural” districts (Amursky district, Akshinsky, Troitskosavsky

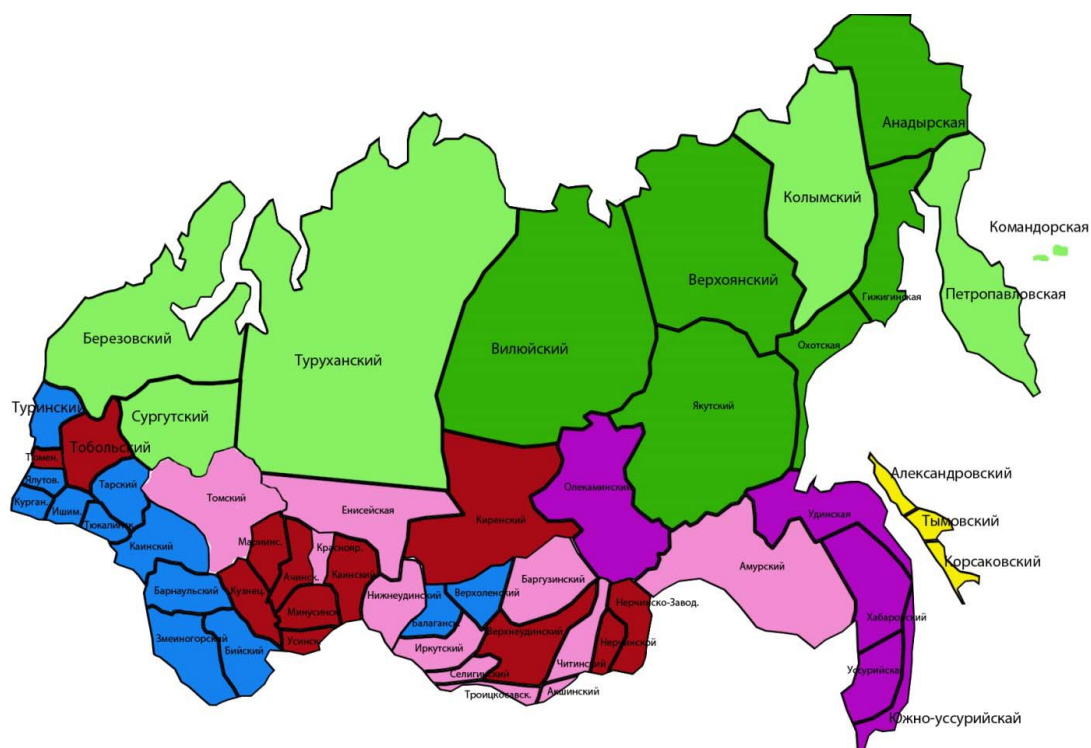


Fig. 1. Distribution of Siberian districts by clusters on the basis of the employment data of the 1897 census

counties of Transbaikal Province, Irkutsky, Nizhneudinsky counties of Irkutsk Province, Yeniseysky, Krasnoyarsky counties of Yeniseisk Province, Tomsk county) characterized by the small proportion of persons employed in agriculture (on the average 35%) with the highest employment rates in the service sector (12%) and production (12%); “semi-agricultural districts (Tobolsky and Tyumensky counties of Tobolsk Province, Mariinsky and Kuznetsky counties of Tomsk Province, Achinsky, Minusinsky, Usinsky, Kansky counties of Yeniseisk Province, Kirensk and Irkutsk Provinces, Nerchinsky and Nerchinsk-Zavodsky counties of Transbaikal Province) with the high proportion of employment in agriculture (52%), which are characterized by relatively high employment rates in the service sector (11%) and production (12%), while in this subgroup there is a concentration of population observed in urban areas (25%) and the highest rate of employment (35%) with a small proportion of avocations

(40%), which confirms the predominance of industrial employment.

The Central Siberian districts include Pribaikalsky districts, which differ due to the ethnic specificity (53%) of the Russian population and peculiarities of employment, for example, in Seleginsky, Barguzinsky, Verkhneudinsky and Chitinsky districts of Trans-Baikal Province a high proportion of individuals engaged in cattle breeding (33%) was recorded, which is not typical for the group of Central Siberian districts. On the whole, the common characteristics of the Central Siberian group is the predominance of the Russian population (on the average, 88%, except for Pribaikalsky districts) and mainly non-agricultural professional activities of the population.

The last group consists of the southwestern districts with two Pribaikalsky counties. This group is characterized by the highest percentage of the Russian population (94%), except for

Balagansky and Verkholsky counties of Irkutsk Province inhabited by 60% of Russians, the low number of urban population (4%), lower employment rate (25%) and a large proportion of persons having additional sources of income (71%). Another feature of this group is a slight disproportion of the gender composition with prevailing female population (51% of women), which is not observed in the other groups. In terms of employment, this group is “monolithic” and includes only one cluster, where the dominant occupation is farming (72%).

The distinguished groups of districts present socio-economic regions of Siberia.

Thus, we can conclude that the multivariate statistical analysis combined with the cartographic method is an effective tool of spatial analysis to identify similar economic regions of Siberia. The most important factors determining the “appearance” of the districts were gender, ethnic characteristics and the number of persons employed in some agricultural field. The research revealed the following socio-economic regions of Siberia: the northern fishing and cattle breeding, Primorsky peripheral-military, the Central Siberian industrial and the southwestern agricultural. The results are preliminary and require further investigation.

¹ Database “Professions and occupations of the people in the Russian Empire at the end of the 19th – beginning of the 20th centuries” (based on the All-Russian census of 1897). Authors: E.A. Bryukhanova, D.N. Ivanov. Certificate of state registration No. 2014621104 dated August 07, 2014.

References

Bryukhnova, E.A. (2011). *Agregirovannye dannye perepisi 1897 g. i modeli professional'noi struktury naseleniia* [Aggregated data of the 1897 census and models of professional structure of the population]. *Sotsial'naia istoriia rossiakoi provintsii: materialy verossiakoi nauchnoi konferentsii* [Social history of Russian province: material of the All-Russian scientific conference]. Yaroslavl, 111-116.

Bryukhanova, E.A., Vladimirov, V.N., Ivanov, D.N. & Chekryzhova, O.I. (2014). Informatsionnaia internet-sistema “Professii i zaniatiia naseleniia Rossiakoi Imperii kontsa XIX — nachala XX v.” kak istochnik dlia professiovedcheskikh issledovanii [Information Internet-system “Professions and occupations of the people in the Russian Empire at the end of the 19th – beginning of the 20th centuries” a source for the study of professions]. *In Istoricheskaiia informatika* [Historical Computer Science], (1), 3-15.

Dronov S.V., Dementjeva E.A. (2012) A new approach to post-hoc problem in cluster analysis, *In Model Assisted Statistics and Applications*, (7), 49 – 55.

Dronov, S.V. & Sazonova, A.S. (2014). Obratnaia post-hoc zadacha klaster'nogo analiza i ee primenie k diskriminatsii dannykh [Reverse post-hoc problem of cluster analysis and its application for discrimination of the data], *In Vestnik Tiimenskogo gosudarstvennogo universiteta, vypusk Fiziko-matematicheskie nauki, Informatika* [Bulletin of Tyumen State University, Physics and Mathematics, Computer Science], (7), 199-204.

Frolov, A.A. (2013). Geoinformatsionnye tekhnologii v sovremennykh istoriko-geograficheskikh issledovaniikh istorikov [Geoinformation technology in modern history and geography studies of historians], *In Voprosy geografii* [Geography Issues]. 136, 447-458.

Ivanov, Yu.N. (2002). *Ekonomicheskaiia statistika* [Economic statistics]. Moscow, 480 p.

Koval'chenko, I.D. & Borodkin, L.I. (1979). Agrarnaia tipologiia gubernii Evropeiskoi Rossii na rubezhe XIX – XX vekov. (Opyt mnogomernogo kolichestvennogo analiza [Agrarian typology of the

provinces in European Russia on the turn of the 19th- 20th centuries (Experience of the multivariate quantitative analysis)], *In Istoriia [History]*, 1, 81-92.

Konovalov, A.V. (2003). *Gubernii Povolzh'ia i Evropeiskaia Rossiia: opyt mnogomernogo statisticheskogo analiza [Provinces of the Volga region and European Russia: the experience of multivariate statistical analysis]*. Pod sen'iu Kliio: Mezhvuzovskii sbornik nauchnykh trudov, posviashchennyi 355-letiiu Simbirskai-Ul'ianovskai I 60-letiiu obrazovaniia Ul'ianovskoi oblasti [In the shadows of Kliio: inter-university collection of research works devoted to 355th anniversary of Simbirsk-Ulyanovsk and the 60th anniversary of Ulyanovsk region foundation]. Ulyanovsk, 52-60.

Konchakov, R.B. & Baranova, E.V. (2010). *Geoinformatsionnye sistemy v istoricheskikh issledovaniakh: sovremennye tendentsii [Geoinformation systems in history research: modern trends]*. XV Derzhvinskii chteniia. Alademiia gumanitarnogo i sotsial'nogo obrazaovaniia: materialy obshcherossiiskoi nauchnoi konferentsii [The 15th Derzhavin' Readings. The Academy of Education in the Humanities and Social Sciences: materials of the All-Russian scientific conference]. Tambov, 154-162.

Pashkina, M.V. (2011). Prostranstvennyi analiz v GIS-sistemakh: sushchnost', napravleniia, vozmozhnosti (Chast' I) [Spatial analysis in GIS- systems: idea, trends, possibilities (Part I)], *In Iaroslavskii pedagogicheskii vestnik, Estesvennye nauki [Yaroslavsky Pedagogical Bulletin, Natural Sciences]*, 1 (III), 156-163.

Troinitskii, N.A. (1899-1905). *Pervaia Vseobshchaia perepis' naseleniia Rossiiskoi imperii 1897 g. [The First All-Russian census of the population in the Russian Empire in 1897]*, LXXII, LXXIII (book 2), LXXIV, LXXV, LXXVI (book 3), LXXVII (book 2), LXXVIII, LXXXIX, . LXXX. Saint Petersburg.

Pozdnyakova, T.M. (2012). *Istoriko-geograficheskoe raionirovanie Aziatskoi chasti Rossii. [Historical and geographical regionalization of Asian part of Russia]*. Author's abstract of the Candidate of Sciences in Geography, Saint-Petersburg, 24 p.

Pochtovo-telegrafnye karty Rossiiskoi imperii po guberniiam i oblastiam [Mail and telegraph maps of the Russian Empire by provinces and regions] (1907). Saint Petersburg, 89 p.

Rasskazov, S.V. (2009). *Iugo-zapadnaia Sibir': evoliutsiia prostranstvennykh struktur obshchestva (s XV v. Do nastoiashchego vremeni) [Southwestern Siberia: evolution of spatial structure of the community (since the 15th century till nowadays)]*. Thesis of the Candidate of Sciences in Geography, Moscow, 24 p.

Ratkovskaya, T.G. & Timoshenko, A.I. (2009). *Ekonomicheskoe raionirovanie [Economic regionalization]*, *In Istoricheskaiia entsiklopediia Sibiri [Historical Encyclopedia of Siberia]*, Novosibirsk. Available at: <http://bsk.nios.ru/enciklopediya/ekonomicheskoe-rayonirovanie> (accessed on 09 December 2015)

Seredovakikh, B.A. & Bulatov, V.I. (2013). Podkhody k istoriko-geograficheskomu raionirovaniu severa Zapadnoi Sibiri [Approaches to historical and geographic regionalization of the north of the West Siberia], *In Vestnik Tiimenskogo gosudarstvennogo universiteta [Bulletin of Tyumen State University]*, 4, 30-40.

Valegina, K.O. (2009). *K voprosu primeneniia statisticheskikh metodov v istoriko-ekonomicheskikh issledovaniakh [Application of statistical methods in historical and economic research]*, *In Global'nyi ekonomicheskii krizis: realii i puti preodoleniia: sbornik statei [Global economic crisis: reality and ways out: collection of works]*, 7, 451-454.

Vampilova, L.B. & Manakov, A.G. (2012). Raionirovanie Rossii: istoriko-geograficheskii podkhod [Regionaliation of Russia: historical and geographic approach], *In Pskovskii regionologicheskii zhurnal [Pskov regionological journal]*, 13, 26-36.

Vladimirov, V.N. (1998). *Nazad k prostranstvennomu analizu? [Back to the spatial analysis?]*. Krug idei: makro- i mikropodkhody v istoricheskoi informatike. Trudy V konferentsii Assotsiatsii "Istoriia i komp'yuter" [Circle of ideas: macro and micro approaches in historical computer science. Works of the 5th conference of the Association "History and Computer"], Minsk, 10-17.

Пространственный подход к анализу данных о занятиях населения Сибири по переписи 1897 г. (опыт многомерного статистического анализа поуездных данных)

**Е.А. Брюханова, С.В. Дронов,
О.И. Чекрыжова**

*Алтайский государственный университет
Россия, 656049, Барнаул, пр-т Ленина, 61а*

В статье представлен оригинальный подход к районированию территории Сибири на рубеже XIX-XX вв. Основой районирования стала типология уездов Сибири, позволившая провести более детальный пространственный анализ данных о занятиях населения переписи 1897 г. Ключевым источником исследования стал открытый информационный ресурс, созданный по опубликованным материалам переписи 1897 г. о численности групп занятий и их гендерном, возрастном, этническом составе.

Применение математических методов, таких как кластерный анализ, post-hoc-методика и коэффициент кластерных различий, позволило не только выявить социально-демографические и экономические типы уездов Сибири, но и ранжировать значимость факторов для определения «облика» уездов. Дальнейшее использование картографического метода анализа полученных наборов кластеров определило пространственные границы социально-экономических районов Сибири. Данное исследование способствует формированию комплексного представления об особенностях освоения территории Сибири и развитии процессов модернизации на рубеже XIX – XX вв.

Ключевые слова: Сибирь, занятия населения, база данных, перепись 1897 г., кластерный анализ, картографический метод, районирование.

Исследование поддержано РФФИ, проект № 15-06-07553А «Типология занятости населения регионов Российской империи конца XIX – начала XX в. (на основе многомерного анализа статистических данных первой всеобщей переписи населения 1897 г.)».

Научная специальность: 07.00.00 – исторические науки.
