On the Electoral Behaviour of Russians

Michael G. Sadovskya* and Ekaterina B. Loginovab

aInstitute of computational modelling SB RAS
50-44, Akademgorodok, Krasnoyarsk, 660036 Russia
bSiberian Federal University
79 Svobodny, Krasnoyarsk, 660041 Russia

Received 18.04.2012, received in revised form 18.04.2012, accepted 10.03.2013

Some preliminary results towards the clusterization of voters inhabiting various regions of Russia are described; the clusterization is based on the homology of the electoral behaviour. The clusters were identified with the help of both K-means unsupervised classification and elastic map technique implementation. The factors affecting the electoral behaviour of the voters from different regions are enlisted.

Keywords: elections, distribution, clusterization, entropy, control mechanisms.

1 Introduction

Elections of representatives are the most widely spread worldwide practice making a core of social or political control and public management. An efficiency of this social mechanism is heavily influenced by several factors. Apparently, comprehensive and reciprocally directed political and social communications seem to be the most essential ones among other types. They must be based on distinct, objective and detailed knowledge of the political views of voters, and their social status, or self-esteem.

Obviously, voters elsewhere do not form a homogeneous body; they differ in a number of characteristics, including biological (say, gender difference), demographic, anthropological figures, as well as in social, political, etc. issues. This diversity is rather important when a choice of a political agent is made. Electoral behaviour is a set of phenomena determining the development of the intention to choose some specific political figure.

Russia is a huge country, from the point of view of geography; in addition, it is inhabited by a great number of ethnic groups with sounding difference in abundance. These groups differ in a number of issues, being at times polar in religion or social, linguistics, etc. values. Thus, an identification and careful description of the territories where the residents display relatively homogeneous electoral behaviour is very important, especially for Russia. Vice versa, one may derive a lot from study of the homologies in the electoral behaviour in geographically discrete sites, when studying the social and economic processes in the country.

Some remarks should be made on the territory expected to yield a homogeneous electoral behaviour. Firstly, there exists formal administrative structure in Russia: it consists of
83 regions with political, legal and administrative semblance to state system in the USA, department system in France, or land system in Germany. These regions differ drastically in size, population density, infrastructure level, etc. For example, the cities of Moscow and Saint-Petersburg make the regions themselves, with the number of dwellers of 18.4 and 3.7 billion residents respectively. Meanwhile, the geographic area of these entities is extremely small, in comparison to the area of other regions. On the contrary, the Krasnoyarsk Territory has slightly less than 2.95 billion residents, while the area is equal to 2 339 700 square kilometers. What is more, this region is not the largest one in Russia.

Secondly, the elections are organized through administrative structure in Russia. General governance and control of the elections is carried out by the Central Electoral Committee; then each region has the Regional Election Committee managing the elections within the region. Each region is subdivided into a number of territories where elections are organized by the Territorial Election Committee (hereinafter referred to as TEC). The borders of each TEC are set in a manner to justify the number of voters inhabiting that region in comparison with other TECs. Finally, the “ground” level in electoral process is embodied by Site Election Committees (hereinafter referred to as SEC) which are just the polling stations; again, the borders of each SEC area are drawn to equalize to the maximum possible extent the number of voters inhabiting each specific site.

An identification of the clusters of territories with homogeneous (similar) electoral behaviour is important both for fundamental and applied studies. Indeed, the differences in electoral behaviour result in the methods and channels of the spread of propaganda or other political issues, with due respect to the specific peculiarities of the voters inhabiting a given site.

The fundamental value of such studies is based on deeper comprehension of structure of a society; besides this, such studies explicate the political patterns standing behind the observed electoral behaviour. Several theories of electoral behaviour are in the focus of attention (Zakharov, 2008; Akhrimenko, 2007). A theory of rational vote, valency theory, cleavage theory and some others are among them. The latter explains the electoral choice through attribution of a person to some social group; such groups yield a cleavage. The cleavage may manifest itself in income, residence area (urban area vs. rural one), religion, etc. Also, these papers provide an attempt to classify the Russian voters through the analysis of VTsIOM (All-Russian Center for the Study of Public Opinion) pollings.

A thorough paper (Norris, 2004) shows the results of regression analysis over 34 countries; and the impact of “classical” cleavages (such as ethnic, linguistic, religious ones) on the political preferences of residents is proven, see also (Enyedi, 2005). The paper by K. Reif and H. Schmitt (1980) also supports these ideas and brings more evidences of the feasibility of the theory under consideration. Also, the impact of income differences and social class identification is also shown; some more up-to-date details on that point are comprehensively discussed in (Blais, 2006, Dean, Croft, 2009).

The geographic and social issues in electoral behaviour of Russian voters are discussed in (Zhidkin, 2002; Luzanov, 2007; Turovsky, 2003; Grishin, 2008). The basic difference results from the residence area: urban vs. rural one. Urbanization level together with the density of urban population leads to forming six clusters [4]: 1) rural TECs; 2) mixed TECs; 3) urbanized TECs; 4) urban TECs in small and average cities; 5) urban TECs in larger cities, and 6) urban TECs in megacities. Some evidences and theories explaining the divergence of electoral behaviour
of voters are presented in (Blais, 2006; Fowler, 2006; Scappini, 2006; Bowler, Hanneman, 2006).

Here we present and analyze the differences in political behaviour expressed in polling stations' attendance and electoral preferences observed during the elections run in December, 2011, when the deputies for State Duma were elected.

2 Material and methods

We have analyzed the electoral behaviour of Russian voters at the TEC level. Russia has 2717 TECs for the elections under consideration, and the number of TECs within a region strongly depends on the population density of that latter. The data to analyze the electoral behaviour were taken from the official website of the Central Electoral Committee. There have been analyzed the results of the State Duma elections held on December 4, 2011. Not discussing here all the issues of that electoral campaign, we deal only with the official data records.

The votes gathered by each party during the electoral race were calculated for each TEC, as it is stated by the official records; obviously, the sum of all the parts is equal to one. These pretreated data were used for the further analysis. To analyze the electoral behaviour two techniques have been used: the first is unsupervised K-mean classification and the latter is elastic mapping.

2.1 Unsupervised classification

Standard K-mean technique has been used to develop an unsupervised classification of TECs. The parts of votes gathered by each party, which had taken part in the elections, were the variables. Let us now describe the procedure, in brief. At the starting point we separated randomly the set of TECs into N classes (we tried 2 ≤ N ≤ 6). Then the centroid for each class was determined; that latter was the arithmetic mean of the parts of ballots gathered by parties within the TEC in a given class. Next, for each point (namely, for each TEC) the distance to every centroid was calculated; we used Euclidian distance. A point remained in the same class, if it yielded the least distance to the centroid of the class; otherwise, a point (that is TEC) changed the class for another one, so that the centroid of that other class yielded the minimal (among all the centroids) distance. The class composition was upgraded, and then the procedure ran again. The cycle of the centroid determination and class upgrade runs while a point changes a class, due to the procedure. The procedure always stops for a rough data set.

The K-mean classification described above keeps the number of classes; thus, a distinguishability of the obtained classes should be checked out. The indistinguishable classes must be merged, and the procedure should run de novo. Yet, we checked no class distinguishability. Further details on K-mean unsupervised classification implementation could be found in (Gorban, Rossiev, 1996).

To verify the stability of the classification obtained through the implementation of unsupervised classification, we carried out a number of computations (i.e., classification implementations), in order to figure out the most proper number of classes for this one classification. It was found that the clusterization with three classes yielded the stability of 85 % in a series of the experiments, while the classification with two or four classes got significantly worse results.

2.2 Elastic map technique

Elastic map technique has been used to develop a nonlinear clusterization of TECs. This approach is a powerful tool of nonlinear statistical analysis. In short, it consists of the following steps:
Step 1. Determine the principal components of the original data set. Still, this is the
linear statistics treatment technique. The principal components are the directions of the greatest variability of the data. Mathematically, they are defined as the correlation matrix eigenvectors corresponding to the greatest eigenvalues. Two greatest ones must be defined;

Step 2. Expand a plane on these two eigenvectors, as on the axes. Then each point must be projected on the plane.

Step 3. Each data point must be bounded to the projection point with an elastic spring. Then, the originally rigid plane is allowed to be elastic, as well. Here are two basic parameters determining the procedure: the first is spring rigidity, and the latter is plane elasticity. The system is able to reach the minimum of the potential energy.

Step 4. Next, each data point is rearranged on the elastic map. The most proximal point on the surface must be found for each data set. These points represent the original data points.

Step 5. The elastic map is (nonlinearly) transformed into a plane, so that it looks like a rectangular grid. The data are represented in this nonlinearly transformed surface, and various types of cluster identification might be applied; an average local density of the points, for example. Here the parameter called the correlation radius is the key factor for efficient cluster pruning.

ViDaExpert software was used to do that; see details of the procedure and technique in (http://bioinfo-out.curie.fr/projects/vidaexpert/).

3 Results

Now we turn to the results in more detail. Paragraph 3.1 presents the results of implementation of unsupervised classification; Paragraph 3.2 shows the results of the clusterization due to elastic map technique.

3.1 Unsupervised classification

A stability of implemented K-mean clusterization is always doubtful from the point of view of the stability. Since an implementation takes start from a random distribution of points into N classes, then one may observe a significant lability of a great number of TECs between the classes. The situation is getting worse, when such redefinition of TECs shows no regularity, resulting in a random, completely disordered cluster composition.

A series of computational experiments let us conclude that a good clusterization is of three classes, whereas an implementation of a classification with two classes, or four (and more) exhibits significantly less stability. A part of the mobile TECs approached one.

Table 1 shows the distribution of TECs over three clusters. For each cluster the part of rural TECs vs. the urban ones is shown. It should be borne in mind that Table 1 shows the parts of urban vs. rural TECs in each cluster separately, but not the distribution portions of the territories between the clusters.

Fig. 1 and 2 show the TECs distribution in the space of parts of votes; each point is TEC. Fig. 1 shows the distribution in principal components, while Fig. 2 shows the same distribution in natural coordinates (i.e., in parts of votes). Here the axes are the ratios of Fair Russia, Liberal Democrats and Patriots of Russia parties. Fig. 3 shows the distribution of TECs into three classes developed by K-means classification on the corresponding elastic map, in inner coordinates. A soft 16×16 elastic map has been developed; here and further “soft” means a grade from the standard software options.

Geographic diversity of Russia results in a significant variation in a number of voters in
Table 1. The distribution of TECs in three clusters

<table>
<thead>
<tr>
<th>Type of a settlement</th>
<th>Cluster I</th>
<th></th>
<th>Cluster II</th>
<th></th>
<th>Cluster III</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>%</td>
<td>$N$</td>
<td>%</td>
<td>$N$</td>
<td>%</td>
</tr>
<tr>
<td>Urban</td>
<td>312</td>
<td>18.7</td>
<td>107</td>
<td>17.3</td>
<td>375</td>
<td>86.0</td>
</tr>
<tr>
<td>Rural</td>
<td>1355</td>
<td>81.3</td>
<td>510</td>
<td>82.7</td>
<td>61</td>
<td>14.0</td>
</tr>
</tbody>
</table>

Fig. 1. Distribution of TECs in the space of vote parts. Each point is TEC. Here the clusterization into three classes (identified in colour) is shown as developed through $K$-mean classification. Greater balls represent the class centroid; the ball size is proportional to the weight. The distribution is shown in principal components coordinates.

Fig. 2. Same distribution as in Fig. 1, shown in natural coordinates.

Fig 3. Same distribution as in Fig. 1, shown on elastic map (soft map, 16×16).

A majority of the regions fall into the same cluster; in other words, TECs belonging to the same region occupy the same class, that is true for greater part of the regions. There are some regions, where TECs are split into two or three clusters. Thus, approximately 30% of TECs for Tula Oblast, Bryansk Oblast and Arkhangelsk Oblast fall into the second cluster, while the majority of those regions occupy the first class. Similar situation is observed for the Krasnodar Territory and the Republic of Altai.

Analysis of geographical proximity of TECs from the second cluster reveals a group of TECs a site to provide accessibility while going to the polling station on foot; obviously, a region of a large physical size together with the very low population density observed in it may have a decrease of the abundance of voters registered within the site of a given SEC (polling station). Similar pattern holds true for TECs.
belonging to Tula Oblast and Bryansk Oblast situated close to each other. Moreover, they are close to some TECs from other regions (belonging to the first cluster). Similar situation is observed in the Republic of Altai, Saratov Oblast, Penza Oblast, Belgorod Oblast, the Republic of Adygea republic, and Yamalo-Nenets Autonomous Okrug, where TECs established in these regions are spread between the first and the second clusters; it should be noticed that the second cluster bears greater part of TECs (up to 60 %).

Probably, the most intriguing is the list of regions that are split into all three clusters (Table 2). The analysis of the “spatial conjunction” of different clusters within the same region shows that the second and the third clusters seem to be the most different in the light of electoral behaviour of their residents. It should be noticed that the city of Voronezh (member of the third cluster) is surrounded on all sides by TECs belonging to the second cluster (Novousmanskaya, Semilukskaya, Khokholskaya, Repyevskya, Ramonskaya, Paninskaya TECs, etc.). Rostov oblast exhibits pretty close pattern: the territories belonging to the third cluster are located at the southern end of oblast, while the territories from the second cluster are arranged into several dense groups.

Table 3 shows the distribution of TECs observed for larger cities. Here cluster II includes the minimal number of TECs from larger cities.
It is remarkable that two larger cities with mainly Islamic population are in the second cluster (Ufa and Kazan). Finally, that is the third cluster that grabs the larger cities: TECs located in the larger cities occupy this one cluster, as a rule. Again, two cities that are capitals of economically developed and relatively reach regions (Bashkiria with Ufa as a capital, and Tatarstan with Kazan as a capital) do not correspond to this pattern.

3.2 Elastic map clusterization

Fig. 4 shows the distribution of TECs in elastic map. The map parameters are indicated in the legend to the Figure. A number of clusters are evident in this Figure. The clusters differ in abundance of TECs belonging to a cluster. Thus, different clusters differ in their explication through the average local density. It should be mentioned that the number of clusters depends on the correlation radius \( r \): if radius is large enough, then very few (maybe, a single one) cluster is identified. The number of clusters grows up, as \( r \to 0 \). Ultimately, the number of clusters becomes equal to the number of the points, when \( r = 0 \). The choice of \( r \) value is a matter of expertise.

Let now consider the structure of the clusters shown in Fig. 4 in more detail. The most evident one is located in the left down corner of the map and includes TECs from seven regions: the Chechen Republic (completely), the Tyva Republic (5 TECs), the Republic of Tatarstan (23 TECs), the Republic of Mordoviya (12 TECs), the Karachay-Cherkess Republic (3 TECs), the Republic of Dagestan (2 TECs) and a single TEC from the Republic of Bashkortostan. Obviously, the composition of this cluster exhibits a very good correlation to geography, religion and social characteristics of the regions grouped into this cluster. Similar situation is observed for other clusters, too.

The cluster which follows the previous one according to the average local density is located at the center of the map. It is rather diverse from the point of view of composition and includes 24 regions. Actually, these regions seem to be rather typical for Russia, in spite of a wide geographical diversity. Indeed, the cluster includes a single TEC from Altai Krai and Zabaykalsky Krai, 5 TECs (of 74 total) from the Krasnoyarsk Territory, from one to three TECs located in Irkutsk, Kemerovo, Novosibirsk, Omsk and Tomsk Oblasts; 9 TECs representing the Udmurt Reublic, 7 from Stavropol Krai, 4 and 2 entities from Kurgan and Rostov Oblasts, respectively. The Jewish Autonomous Oblast, Khanty-Mansi Autonomous Okrug, Kaliningrad, Kaluga, Magadan, Ryazan, Sakhalin and Ulyanovsk Oblasts are present with a single TEC each in the cluster, as well as the Krasnodar Territory, Perm and Khabarovsk Krais do.

The third cluster ranked in local average density is located slightly higher the center of the map and consists of 19 TECs: 9 TECs belong to the Udmurt Republic, 5 TECs belong to Stavropol Krai, 4 TECs belong to the Krasnoyarsk Territory. Then, Amursk Oblast, Kemerovo Oblast and
Kurgan Oblast are present with three TECs each, while Tomsk Oblast is represented with a couple of TECs. The Jewish Autonomous Oblast, Irkutsk Oblast, Kaliningrad Oblast, Kaluga Oblast, Novosibirsk Oblast, Omsk Oblast, Rostov Oblast, Sakhalin and Tumen Oblasts are present with a single TEC each in that cluster; the same situation is observed for Khanty-Mansi Autonomous Okrug, Khabarovsk Krai and Perm Krai. It should be noted, that the level of support of United Russia was close to 50 – 60 % in all these regions.

The fourth cluster consists of 38 TECs and occupy the very center of the map and contains 14 TECs from the Sakha (Yakutia) Republic, and 8, 5, 3, 2 and 2 TECs from Volgograd Oblast, Voronezh Oblast, Chelyabinsk Oblast, Ulyanovsk Oblast and Tver Oblasts, respectively. Also, there are one TEC from Altai Krai, and one TEC from Rostov Oblast in the cluster. United Russia gathered here from 42 % to 64 % of votes.

Finally, the fifth cluster is the most abundant containing 55 TECs; it is located in the center of the upper part of the map. It has 13 and 6 TECs from Altai Krai and the Krasnoyarsk Territory, respectively. Also, it has 10 TECs from Novosibirsk Oblast, 5 TECs from Orenburg Oblast, 3 TECs from Irkutsk Oblast, 2 TECs from Sakhalin Oblast, and a single TEC from Samara, Ryazan, Rostov, Omsk, Kurgan, Kirov and Ivanovo Oblasts each. The level of United Russia support was quite close to the average one observed over Russia entirely.

4 Discussion

K-mean classification explicitly identifies three classes covering the territories of the Russian Federation, with respect to the electoral behaviour of the residents of those territories. The composition of the classes is apparently non-random. The territories composing a class possess similar characteristics in ethnicity, anthropology, culture, geography (to some extent), language, etc. The problem of the spatial vs. political and other social issues stratification in Russia was also discussed in (Wilson, 2005). While this paper focuses mainly on the á priori analysis of social, economical and cultural differences in the regions, we have taken quite an opposite approach; such approach seems to be less subjective and, furthermore, it proves the sounding correlation between electoral behaviour and social issues, not contrariwise.

An analysis of the class structure reveals quite intriguing fact that regions entitled with the word republic occupy the same class. This fact (the name of a region) seems to be the leading factor for class attribution. Surely, there are some exceptions from this pattern: say, the Republic of Karelia yields drastically different electoral behaviour. Obviously, the ethnic, social, language, cultural and other relevant features of the population of this republic manifest themselves in the electoral behaviour and affect the classification stronger that the legal status of republic. This fact resembles, to some extent, the situation in Ireland (Marsh, 2006, 2007), where various groups of voters have been found, and this separation also correlates to social, economical and religious issues together with the geographical aspects. It should be said, that any analogy between the patterns observed elsewhere in Western Europe and in Russia must be carefully analyzed: the point is that the social mechanisms determining the electoral behaviour differ strongly for these two entities. Some more ideas concerning a feasibility of the cleavage theory to explain the electoral behaviour are discussed in (Enyedi, 2005); further studies should be made to verify this theory in Russian conditions. Yet, all these papers analyze the electoral behaviour on the sociological background rather than the pure behaviour observations based on the voting
records. Also, some useful ideas may come from a comparative study of the electoral behaviour of French voters (Bélanger et al., 2006).

Analysis of electoral behaviour with the help of elastic maps reveals significantly greater number of clusters; yet, the cluster identification here could not be completely formal. Correlation radius \( r \) is the key parameter for the cluster identification, while it could not be determined formally and is matter of taste of a researcher. We chose \( r = 0.05 \) that obviously reveals at least fifteen clusters, and two of them are extremely dense. The cluster composition is obviously non-random. Yet, the structure of the clusters seems to be more complex, in comparison to the classes obtained through K-mean classification.

The most obvious fact is that the clusters are definitely quite complex from the point of view of the set of TECs composing them. This is neither mistake, nor discrepancy: since the elastic map technique is more sensitive and specific in clusterization, the clusters may include geographically discrete territories, if the electoral behaviour of the voters inhabiting these territories looks similar. Here the researcher has to be rather careful in understanding the reasons standing behind such cluster construction: whether the similitudes in economy, social and cultural issues really take place in these distant territories, or this is just an arbitrary coincidence of the voters’ choice resulted in the same cluster occupation, and such cases must be studied individually.

One of these two dense clusters enlists the republics of the Northern Caucasus and some other territories with mainly Islamic population. Residence area (a type of a settlement, to be exact) is another important factor of the split of voters into several patterns of electoral behaviour.

It should be stressed that larger cities, in turn, are not homogeneous, if to consider the electoral behaviour of the residents. Table 3 shows the distribution of TECs of larger cities over three classes developed due to unsupervised classification.

Evidently, these larger cities are separated into two groups: the first consists of the cities whose TECs occupy the first and the third classes, and the latter consists of the cities with TECs occupying the first and the second classes. It should be stressed that the second group consists of two greater cities only, and they both are the capitals of great economically developed republics populated with mainly Islamic people, they follow quite explicit ideology, and both entities are located in Middle Volga Region. The results of the elections of 2003 to State Duma with respect to the political behaviour issues is considered in (Wegren, Konitzer, 2006). Yet, it should be said that the up-to-date situation in Russia, both in political, economy and social dimensions has nothing to do with those conditions observed in 2003; thus, the observations and theories discussed in this paper are hardly feasible and current. Further analysis of these issues could be found in (Bowler, Hanneman, 2006). Mathematical modelling of voters’ behaviour is present in (Lo Schiavo, 2005, 2006). The models studied in these papers may bring some more understanding of the peculiarities of electoral behaviour, especially under the implementation of fine non-linear techniques of statistical analysis of data records. Finally, more narrative model of electoral behaviour is discussed by W. Claggett and P.H. Pollock III (2006).

In conclusion, let figure out some further problems and questions to be answered, when studying the electoral behaviour of the population of the Russian Federation.

**Scaling problem.** Do separate (but with abundant enough population) regions exhibit a similar pattern of electoral behaviour of their residents, as the whole country does? Definitely,
the answer to this question heavily depends on the choice of a particular region, while one may try to figure out typical scaling figures for several (typical) regions, for example, situated at the centers of relevant clusters or classes.

**Time stability problem.** What would happen to the structure of the clusters and/or classes, if one traces the dynamics of electoral behaviour through a series of elections? It is a common place, that the countries with a long electoral history exhibit rather stable patterns of electoral behaviour. Probably, Russia has too short and too perturbed latest political history, so no reasonable stability could be expected. Indeed, the list of parties taking part in elections changes significantly each electoral cycle.

**Spatial inhomogeneity problem.** Russia, again, is a huge country with extremely inhomogeneous population spatial distribution. It results in a strong diversity of TECs from the point of view of the voters’ presence registered within the site. Federal Law requests to keep the difference in the abundance of voters in different TECs below 10 %, while there is no way to follow the Law. Thus, one should “purify” the original database eliminating some TECs which fall significantly beyond an average figure of registered voters.

Finally, it is necessary to carry out a comparative study of this style, to compare Russia and some other countries with electoral and political systems resembling the Russian one.

**References**


Michael G. Sadovsky and Ekaterina B. Loginova. On the Electoral Behaviour of Russians


**Об электоральном поведении избирателей России**

М.Г. Садовский, Е.Б. Логинова

Институт вычислительного моделирования СО РАН
Россия 660036, Красноярск, Академгородок, 50, стр. 44
Сибирский федеральный университет
Россия 660041, Красноярск, пр. Свободный, 79

В статье представлены предварительные результаты по изучению стратификации избирателей России по признакам электорального поведения. Кластеры были определены с помощью метода динамических ядер и упругих карт. Также обсуждены некоторые факторы, влияющие на электоральное поведение избирателей.

Ключевые слова: выборы, распределение, кластеризация, энтропия, механизм управления.