

**A. A. Pushkarev**

Ural Federal University (Yekaterinburg, Russia; a.a.pushkarev@urfu.ru)

## CLUSTER ANALYSIS OF REGIONAL INNOVATION ACTIVITY IN RUSSIA IN 2010-2015

*In this article, the indicators of innovation activity in Russian regions are discussed and the regions are divided into five groups, according to their performance in these indicators. Our cluster analysis is based on the recent research and includes several groups of indicators such as innovation activity of enterprises, training of highly qualified personnel, research and development, state support for innovation, and application of innovative technologies. We used the data provided by Rosstat (Federal State Statistics Service) for 83 Russian regions in the period between 2010 and 2015.*

*In terms of their innovation activity, Russian regions can be divided into five groups, two of which are Moscow and St.Petersburg, the two biggest Russian cities that play a special role in Russian economy. Overall, the level of innovation activity in Russia can be assessed as lower middle, although in the given period some regions managed to improve their performance in this sphere. The average level of innovation activity varies considerably across regions, which means that the state innovation policy should be more diversified. Moscow, St.Petersburg, Nizhny Novgorod and Sverdlovsk regions have demonstrated consistent high-level performance and can thus be regarded as prospective centres of innovation. These centres can positively influence the neighbouring areas through the knowledge and technology spillover effect. Although no definitive conclusion can be drawn about the connection between the regions' geographical location and their innovation activity, there is evidence that the most active Russian regions tend to concentrate in the European part of the country. Our findings can be used as guidelines for devising and modifying federal and regional innovation policies.*

**Keywords:** innovation, innovation activity, cluster analysis, regional studies, Russian regions, innovative development

### Introduction

Innovative development is an essential part of the economic development strategy of any country. As the experience of many developed countries show, the right innovation policy and its efficient implementation can provide sustainable and rapid economic growth. A key element of such policy is its region-specific diversification and monitoring of the dynamics of outcome indicators [1].

In modern research literature there is a widely shared view that Russian regions vary significantly both economically and socially. However, there is a lack of consensus regarding the state of innovation in Russian regions: how different or similar the regions are in this respect and how to classify them.

In this paper we analyse the data on innovation and R&D in 83 Russian regions for the period between 2010 and 2015. These data include such indicators as the number of research personnel in the region, the share of R&D spending in the GRP, the overall number of new technologies and the number of these technologies that have been put into practice; the share of companies involved in innovation; the number of students and researchers with Candidate's and Doctor's degrees. We also consider the annual dynamics of the regions' innovation-related indicators, which, apart from the qualitative changes achieved by specific regions, also reflect the overall state of innovation in Russia and the efficiency of the country's innovation policy.

We apply the method of cluster analysis to group Russian regions according to outcome indicators and to compare the results of clusterization with the regions' geographical location. Thus, our research addresses the questions about the connection between the Russian regions' geographical location and their innovation activity: how different are the Western and Eastern Russian regions? What distinguishes Moscow and St.Petersburg from other regions? Are there any regions sharing innovation-related indicators?

The structure of this paper is as follows. After the introduction, we review the existing literature in this field. The next section describes the data and methods used in this research. The fourth section focuses on the cluster analysis and its results. In the final section, the conclusions are drawn. The practical application of our results and the prospects for further research are outlined.

### **Literature review**

The topic of spatial clustering and the knowledge spillover effects it creates arouses significant scholarly interest nowadays.

Spatial clustering creates a widely studied knowledge spillover effect, which appears to be largely a local phenomenon, dependent on the geographical proximity. For example, George Deltas and Sotiris Karkalakos investigate regional patent statistics in the European Union and find that an increase in the distance between the originating and recipient region by 500km reduces the positive effects of spillovers by 55–70% [2]. Similar findings were made by other researchers [3, 4].

Cassandra C. Wang, Cassandra and Aiqi Wu (2015) studied the case of knowledge spillover among Chinese electronic firms and found that the geographical proximity of firms and heterogeneous rather than homogeneous knowledge play an important role in the formation of innovation clusters with Chinese companies tending to concentrate in the same regions of the country [5].

Another study on innovation in China considers the role of spatial factors impeding knowledge spillovers and demonstrates that domestic companies mostly benefit from the positive effects of foreign direct investment (FDI) in their neighbouring regions [6]. Although the effects of FDI are not the main focus of our research, this research model can be transposed onto studying innovation as an independent process.

Luciana Lazeretti and Francesco Capone (2016) study the role of geographical proximity in the creation of innovation network by focusing on the case of high technologies in the agricultural industry of Tuscany. By using stochastic actor-oriented modelling, the authors prove that geographical proximity has a positive impact on innovation dynamics and on the formation of innovation clusters [7].

Doris Läßle and her co-authors (2016) also discuss the spatial aspect of knowledge transfer in agriculture by analyzing the case of agricultural innovation in Ireland and demonstrate the positive effect that the proximity of leaders of innovation has on their neighbours [8].

Yet another study analyzes scientific knowledge networks and technological knowledge networks of China by applying econometric and spatial modelling methods to show the positive correlation between the geographical proximity and the intensity of knowledge spillover effects [9].

Theoretical studies of spatial aspects of innovation diffusion reveal the potential of innovation clusters which comprise closely located regions and territories [10, 11].

To the best of our knowledge, Russian scholars have not yet engaged in the research of regional innovation clusters.

### **Data and methods**

In this research we used the data provided by Rosstat (Federal State Statistics Service) for 83 Russian regions in the period between 2010 and 2015. For clusterization we used sixteen indicators of innovation and research activity. These indicators can be divided into the following groups:

- Innovation activity of enterprises: the number of enterprises involved into R&D; the share of innovative enterprises.
- Training of highly qualified personnel: the number of university students; the number of researchers with Candidate's or Doctor's degrees.
- Research and development: the number of researchers; the number of patent applications; the number of approved patent applications; export of new technologies (mln rbs); import of new technologies (mln rbs);
- State support of innovation: research funding (mln rbs); spending on innovation (mln rbs);

- Application of innovative technologies: the number of new technologies used by manufacturing companies; the volume of innovative products (mln rbs).

These sets of indicators cover the pivotal spheres of innovation, starting from resources to outcomes. These indicators are widely used in a number of other current studies on innovation activities [12, 13, 14, 15, 16, 17].

To avoid incomparability of measurements, we normalized each of the indicators and transformed them into z-scores so that they all lay within the range of (-1;10). This approach allowed us to avoid using additional control variables. The above-mentioned and the following calculations were made with the help of programming language R, version 3.2.2., and its packages.

Table 1 provides the main descriptive statistics for the indicators prior to normalisation.

Table 1.

**Descriptive statistics of the data**

	n	mean	sd	median	min	max	se
Researchers	909	9191,479	29492,39	1711	16	241226	1357,497
Research firms	913	45,97881	92,16995	23	1	811	4,242466
Research spending	494	8511,985	30380,5	1257,05	6,0303	301817,9	1398,376
Number of researchers with Candidate's degrees	912	400,9407	1072,641	181	0	10029	49,37232
Number of researchers with doctoral degrees	901	16,75424	38,93861	8	0	312	1,792295
Patent applications	913	348,3496	1104,047	121	0	12681	50,81786
Patents granted	913	276,053	868,5144	94	0	8699	39,97662
New technologies produced	909	15,56356	34,99215	5	0	259	1,610644
New technologies used	909	2520,561	3166,581	1529,5	0	20021	145,7537
Share of innovative firms	889	9,609534	4,447225	8,8	0,5	34,3	0,2047
Innovation spending	891	11673,17	24100,93	3196,864	0,769	190334,6543	1109,335
Value of innovative goods	908	36179,19	84646,53	8538,125	0	851583,36	3896,172
Technologies exported	913	483,3724	2998,386	3,384516	0	57412,8375	138,0119
Technologies imported	913	914,7791	2276,483	54,11267	0	20183,98079	104,7836
Share of university students	912	12,63136	28,01887	7	0	268	1,289673

In our clustering procedure we applied the K-means clustering algorithm which minimizes the square error:

$$e^2(X, L) = \sum_{j=1}^K \sum_{i=1}^{n_j} \|x_i^{(j)} - c_j\|^2 \tag{1}$$

where X is the vector of characteristics of the given regions; L is the vector of characteristics of cluster centres; and  $C_j$  is the specific cluster's 'centre of masses'.

To measure the distance, we used the standardized Euclidean distance:

$$\rho(x, x') = \sqrt{\sum_{i=1}^n (x_i - x'_i)^2} \tag{2}$$

For preliminary analysis we used five clusters for both theoretical and empirical reasons.

According to the graph below, which shows how the WSS is dependent on the number of clusters, we can see that the WSS falls sharply (2 to 3 clusters) but after the number of clusters reaches 5, it declines at a very slow rate (Fig.1).

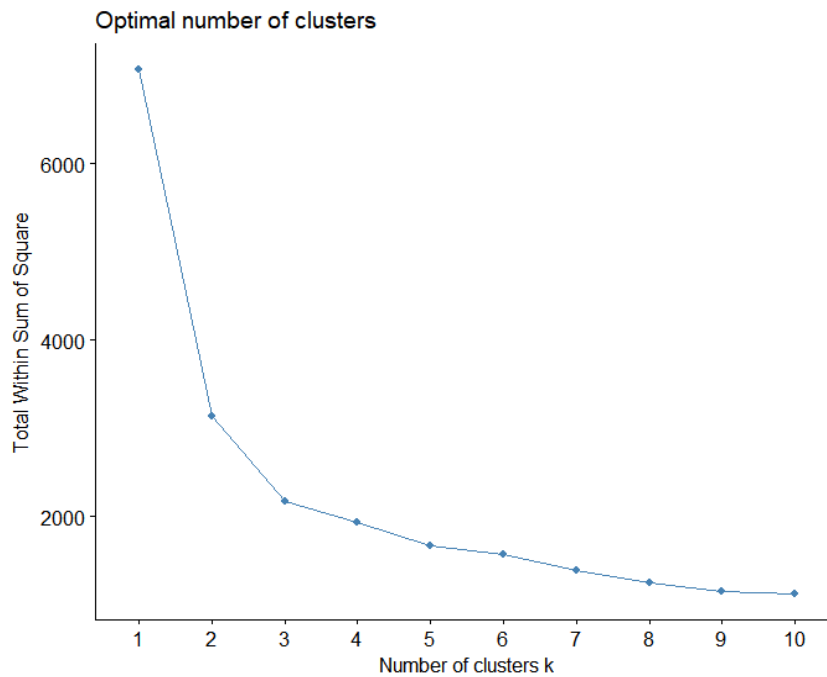


Fig. 1. Optimal number of clusters

Similar results were obtained by using silhouette analysis, which means that if the data are divided into two clusters, it brings more accurate results although the results of division into three, four or five clusters are also quite satisfying (Fig. 2)

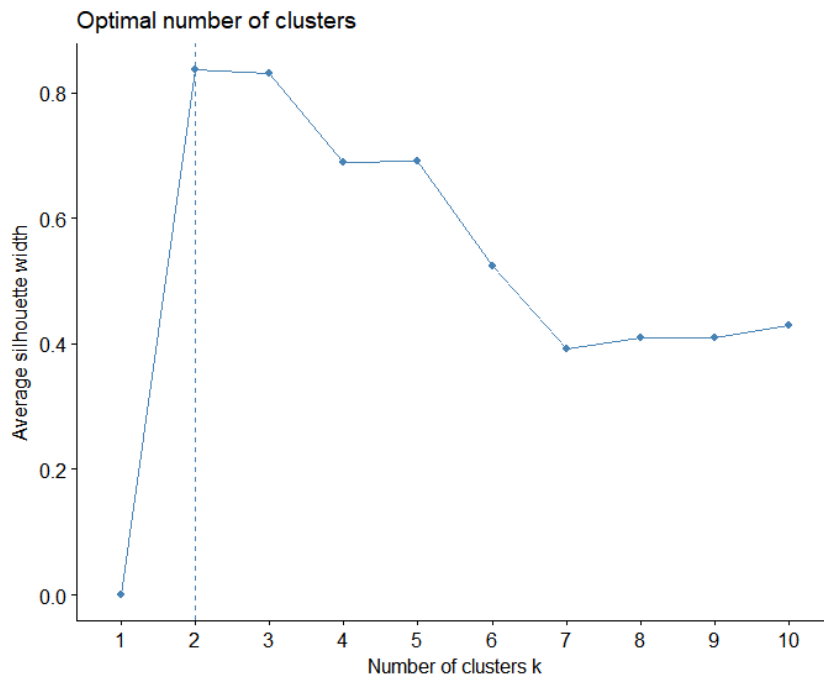


Fig. 2. Silhouette analysis of the optimal number of clusters

The preliminary modelling has also shown that Moscow is significantly different from other regions and that it tends to form a separate cluster regardless of the general number of clusters. Thus, it was decided to

create five clusters for final modelling: one for Moscow and the rest for other leading regions, regions with results above average, regions with middle-level performance, and underperformers.

### Modelling results

Modelling comprised two stages. At the first stage, regions were clustered according to the average values in the given period. Then, to gain a deeper understanding of the innovation dynamics and the effects of the state policy, we considered innovation-related indicators in specific years.

The results of the first stage of modelling are shown in Figure 3 (for Russia in general) and Figure 4 (for the European part of Russia with two specific regions – Moscow and St.Petersburg).

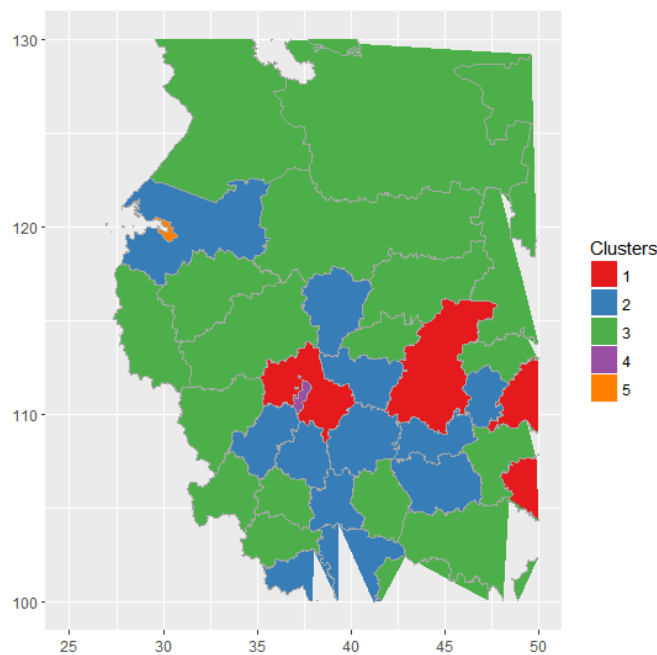


**Fig. 3.** Clusterization of Russian regions according to the average level of their innovation activity in the given period

Apart from Moscow and St.Petersburg, we also observed three specific levels of innovative activity: high, middle, and low (in the map they are indicated with red, blue, and green colours respectively). As Figure 3 illustrates, there are only four highly active regions - Moscow, Sverdlovsk, and Nizhny Novgorod regions.

Other regions have either demonstrated the middle or the low level of innovation activity. It should be noted that the most active regions are concentrated in the European part of Russia, especially around Moscow, which can be seen from the map in Figure 5.

Moscow and St.Petersburg were identified as two separate clusters and were indicated in purple and orange colours respectively. Although these cities have higher levels of innovation than other Russian regions, they significantly differ from each other, which is why we regard them as separate clusters.



**Fig. 4.** Clusterization of Western Russian regions according to their average level of innovation activity in the given period

For Moscow, each of the indicators exceeds those of other Russian regions, even those from the 'red' cluster. In general, such situation is characteristic not only of innovation but of other economic and social spheres. In the areas around Moscow and Moscow region, the level of innovation activity is also quite high, which can serve as an evidence to support the observation that the leading regions stimulate their neighbours' innovative activity.

The innovation-related indicators of St.Petersburg are comparable with other highly innovative regions, except for those indicators that characterize the availability of qualified personnel in the region. In this respect, St.Petersburg is far ahead of other regions.

Therefore, it might be productive to create regional centres specializing in various elements of the innovation process, for example, training of qualified professionals, R&D, implementation of innovations, joint projects with industrial enterprises, and adoption of foreign innovative technologies.

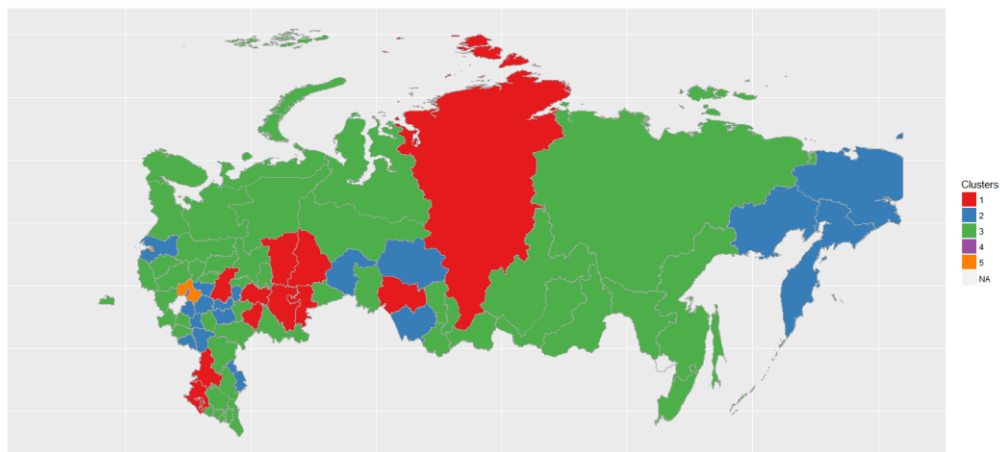
At the second stage of modelling, we focused on the dynamics of innovation in the country. Figures 5 and 6 show the geographical location of the regions' clusters in 2010 and 2015. Figure 5 demonstrates the state of innovation in Russia before the launch of the Innovative Development Strategy 2020.



**Fig. 5.** Clusterization of Russian regions according to their level of innovation activity in 2010

At this stage, the majority of Russian regions were included into the cluster of underperformers. Moreover, we found that in the Asian part of the country, innovative activity is low in almost all the regions.

Figure 6 illustrates the results of clusterization for 2015, the last year in the observation period. These data show the intermediate outcomes of the Innovative Development Strategy 2020.



**Fig. 6.** Clusterization of Russian regions according to their level of innovation activity in 2015

It should be noted that throughout the given period, the regions migrated from one cluster to another although we did not detect any general qualitative growth. The 'centres of mass' of the clusters remained practically the same. Nevertheless, we saw that the regions moved to clusters with a higher level of innovation activity.

Some regions, such as Sverdlovsk and Nizhny Novgorod, unfailingly produce good results. We also noticed that in comparison with 2010, their neighbours have also demonstrated improved performance. A similar trend was observed in the Far Eastern regions, which leads us to the conclusion that there might be a spillover of technologies and innovations from the leaders to their neighbours.

If we analyze the regions' performance in specific years, the majority of Russian regions will be classified as underperformers, which shows the generally low level of innovation in the country. Moreover, only a small number of regions demonstrate the middle level of activity. Therefore, there is a significant discrepancy between the leaders and all the rest.

### Conclusion

Our results confirm that more advanced Russian regions can affect innovation activity of their neighbours through knowledge and technology spillover. This process creates sustainable geographical clusters with high innovation activity around the leading regions. Our findings can thus be used to modify the current innovation policy on the regional and federal levels and to optimize the spending on innovation in the regions.

Moscow and St.Petersburg play a special role in the innovation process as their scores are several times higher than those of other regions. Such situation shows that the economic development of Russia is uneven and that it is necessary to diversify the innovation policy to make it more effective.

Russia has a number of regions that invariably occupy the leading positions. Such regions may become drivers of innovation, maximizing the performance of their neighbours by sharing their knowledge, best practices and technologies with those in proximity. In our analysis, we further focused on specific periods and showed that the innovation policy which has been implemented since 2011 enhances positive dynamics.

Although no definitive conclusion can be drawn about the connection between the regions' geographical location and their innovation activity, there is evidence that in the majority of cases, the most active Russian regions are concentrated in the Western part of the country. At the same time some innovative centres can be also found in Western Siberia and some positive dynamics has been observed in the Far East.

The average levels of innovation, however, differ significantly for different groups of regions, which means that the state policy in this sphere should be more diversified. Our analysis of the clusters' performance in different periods has detected only a slight increase in the clusters' 'centres of mass'. Both of these facts show that although the current innovation policy has brought about some positive changes, it should be modified to ensure a more rapid qualitative growth.

Based on the findings of this study, it can be suggested that further research should be made into such characteristics of Russian regions as their specialization and the available R&D facilities and training centres. Although cluster analysis makes it possible to consider such characteristics, a more precise division of Russian regions into groups will enable us to devise more targeted guidelines for the regional innovation policy.

The results of our cluster analysis can also be used to create an integral innovation-related indicator scheme for assessing Russian regions, comparing them and monitoring their further development.

### Acknowledgements

This research was supported by RFBR, research project 18-010-01190 A ' Models of innovation development factors and comparative advantages analysis in the Russian economy'.

### References

1. World Bank. 2010. Innovation Policy: A Guide for Developing Countries. World Bank. <https://openknowledge.worldbank.org/handle/10986/2460>
2. George Deltas, Sotiris Karkalakos. (2013) Similarity of R&D Activities, Physical Proximity, and R&D Spillovers. *Regional Science and Urban Economics*, Volume 43, Issue 1, January 2013, pp. 124-131.
3. P. Almeida, B. Kogut (1999) Localization of Knowledge Spillovers and the Mobility of Engineers in Regional Networks. *Management Science*, 45 (1999), pp. 905-917
4. L. Bottazzi, G. Peri (2003) Innovation and Spillovers in Regions: Evidence from European Patent Data. *European Economic Review*, 47 (2003), pp. 687-710



5. Cassandra C. Wang, Aiqi Wu (2016) Geographical FDI Knowledge Spillover and Innovation of Indigenous Firms in China. *International Business Review* 25 (2016), pp. 895–906
6. MiLin, Yum K. Kwan (2016) FDI Technology Spillovers, Geography, and Spatial Diffusion. *International Review of Economics & Finance* Volume 43, May 2016, pp. 257-274
7. Luciana Lazzeretti, Francesco Capone (2016) How Proximity Matters in Innovation Networks Dynamics along the Cluster Evolution. A Study of the High Technology Applied to Cultural Goods. *Journal of Business Research* 69 (2016), pp. 5855–5865
8. Doris Läßle et al. (2016) What Drives Innovation in the Agricultural Sector? A Spatial Analysis of Knowledge Spillovers. *Land Use Policy* 56 (2016), pp. 238–250
9. Dandan Li, Yehua Dennis Wei, Tao Wang (2015) Spatial and Temporal Evolution of Urban Innovation Network in China. *Habitat International* 49 (2015), pp. 484 – 496
10. T. Hägerstrand, *Innovation Diffusion as a Spatial Process* (Chicago: University of Chicago Press, 1967)
11. Roger Bivand (2015) Spatial Diffusion and Spatial Statistics: Revisiting Hägerstrand’s Study of Innovation Diffusion. *Procedia Environmental Sciences* 27 ( 2015 ), pp. 106 – 111
12. Ermasova N.B. Factors Affecting Innovation Activities of Organizations// *Ekonomika. Upravlenie. Pravo.* - 2014. - №3. - P. 495-503.
13. Mariev O.S., Savin I.V. Factors of Innovative Activity of Russian Regions: Modelling and Empirical Analysis// *Ekonomika regiona.* - 2010. - № 3.
14. Eriksson, T., Qin, Z., Wang, W. (2014). Firm-Level Innovation Activity, Employee Turnover and HRM Practices — Evidence from Chinese firms. *China Economic Review*, 30, 583-597.
15. García-Quevedo, J., Pellegrino, G., Vivarelli, M. (2014). R&D Drivers and Age: Are Young Firms Different? *Research Policy*, 43, 1544-1556.
16. Guariglia, A., Liu, P. (2014) To What Extent Do Financing Constraints Affect Chinese Firms' Innovation Activities? *International Review of Financial Analysis*, 36, 223-240.
17. Zhang, H. (2015). How does Agglomeration Promote the Product Innovation of Chinese Firms? *China Economic Review*, 35, 105-120.

#### **Author**

**Pushkarev Andrey Aleksandrovich** — Senior lecturer of the Department of Econometrics and Statistics (Ural Federal University, Ekaterinburg; e-mail: a.a.pushkarev@urfu.ru)