The use of covariance matrices in laboratory processing of geo-environmental data of field survey for development projects of land use planning

V V Vershinin, A V Fedorinov and A V Dontsov
State University of land use planning, 15, Kazakova str., Moscow, 105064, Russia
E-mail: v.vershinin.v@mail.ru

Abstract. The paper introduces the possibility of using covariance matrices in the factor analysis. The factors affect the effectiveness of issues’ solution, concerning the involvement of unused agricultural lands in the active economic turnover in land use planning projects. The usage of the analysis tool, proposed by the authors, helps to improve the efficiency and objectivity of project decision substantiation in land use planning.

1. Introduction
One of the most important problems of current land management in Russia is the involvement of unused agricultural lands in the active economic turnover.

According to various data [1-3], in Russia such areas of unused land shares are currently about 20 million hectares.

The All-Russian Agricultural Census of 2016 revealed the results, showing the reduction in the total land area belonging to agricultural commodity producers. It is more than 100 million hectares compared to the results of 2006. Moreover, there were only 125 million hectares used by rural producers in 2016 despite the remaining amount of 142.7 million hectares [3].

As our experience shows, there is the most effective way to solve this problem. It is used to identify the unused land plots and involve them into the turnover, based on the special project development of land use planning [4]. The role and importance of land use planning on the whole, and particularly its projects, have been repeatedly emphasized by such specialists in land management as S. N. Volkov, V. N. Khlystun, N. V. Komov and others [5-7].

The development of such projects should be based on the use of an optimal economic and mathematical apparatus.

2. Models and methods
The principal component analysis (PCA) is used for maximum factor consideration which provides the offered measures efficiency for the development of unused land plots. It is one of the main methods allowing the reduction of data dimension with the least information loss.

The principal component calculation can be performed in two ways. The first one is the calculation of singular matrix data decomposition. The second one is the calculation of eigenvectors and eigenvalues of the original data covariance matrix. The usage of the covariance matrix has significantly allowed simplifying and, thus, accelerating the consideration process of factors, affecting the project decision choice.
The objective of principal component analysis has at least four basic versions. We recommend using the data approximation, applying linear manifolds in these works.

3. Results and discussion
The assigned task is to minimize the involvement cost of unused land plots among agricultural lands in active economic turnover. According to the work results, obtained in the Yaroslavl region, the task solution is determined by the influence of 13 key factors (table 1).

| Factor number | Factor name                                      | Factor weight (from 1.1 to 2.1) |
|---------------|--------------------------------------------------|----------------------------------|
| 1             | The distance from a settlement                    | 1.3                              |
| 2             | The distance from an economic center              | 1.5                              |
| 3             | The distance from paved roads                     | 1.6                              |
| 4             | The degree of land plot stocking occupation in the period of 1 – 5 years | 1.2|
| 5             | The degree of land plot stocking occupation in the period of 5 - 15 years | 1.4|
| 6             | The degree of land plot stocking occupation in the period of 15 - 25 years | 1.6|
| 7             | Land plot area                                    | 2.0                              |
| 8             | Land plot form (compactness factor)               | 1.8                              |
| 9             | The degree of land plot peat formation            | 1.5                              |
| 10            | The degree of brushwood of the land plot          | 1.5                              |
| 11            | The species composition of land plot soil         | 1.6                              |
| 12            | The necessity of roads’ repair and (or) their reconstruction | 1.8|
| 13            | The previous use of the land plot                 | 1.1                              |

To evaluate each factor, its significance indicator (“weight” factor) is used (table 1). The weight value is set on the basis of a pair statistical sampling using the covariance matrix. It is a factor cumulative indicator which influences the costs for land plot involvement in economic turnover and the possibility of land plot further usage.

In order to improve the accuracy (objectivity) of territory exploration analysis for each of the listed 13 factors, the factor “internal” weight is set. The process of its setting depends on the characteristic change of the considered factor. The calculation example is shown in table 2.

| Position number | Construction and repair of roads | Value of factor “internal weight” |
|-----------------|---------------------------------|----------------------------------|
| 1               | The minor (small) repair of road surface | 1.1             |
| 2               | The repair of road surface up to 500 m | 1.2             |
| 3               | The repair of road surface from 0.5 to 2 km | 1.3             |
| 4               | The repair of road surface of more than 2 km | 1.4             |
| 5               | The construction of new roads up to 500 m | 1.5             |
| 6               | The repair of the road surface part and new road construction up to 500 m | 1.6             |
| 7               | The construction of new roads from 0.5 to 1 km | 1.7             |
| 8               | The construction of new roads from 1 to 1.5 km | 1.8             |
| 9               | The construction of new roads from 1.5 to 2 km | 1.9             |
| 10              | The construction of new roads of more than 2 km | 2.0             |

Table 3. The results of factor analysis (fragment)
| The land plot number (according to inventory records) | 1   | 2   | 3   | ... | 37  | 38  | The number of participants |
|-----------------------------------------------------|-----|-----|-----|-----|-----|-----|---------------------------|
| from settlement (1.30)* the value                   | 1.30| 1.26| 1.58| ... | 1.62| 1.49| 38                        |
| the value considering the weight                    | 1.69| 1.64| 2.05| ... | 2.11| 1.94|                          |
| from economic center (1.50)* the value               | 1.00| 1.00| 1.02| ... | 1.33| 1.33|                          |
| the value considering the weight                    | 1.50| 1.50| 1.53| ... | 2.00| 2.00| 38                        |
| the distance from paved roads' repair and reconstruction (1.60)* the value | 1.22| 1.23| 1.24| ... | 1.51| 1.68|                          |
| the value considering the weight                    | 1.95| 1.97| 1.98| ... | 2.42| 2.69|                          |
| trees from 1 to 5 years (1.20)* the value            | 0.00| 0.00| 0.00| ... | 0.00| 1.30| 8                         |
| the value considering the weight                    | 0.00| 0.00| 0.00| ... | 0.00| 1.56|                          |
| trees from 5 to 15 years (1.40)* the value           | 1.30| 2.00| 1.20| ... | 2.00| 0.00|                          |
| the value considering the weight                    | 1.82| 2.80| 1.68| ... | 2.80| 0.00| 11                        |
| trees from 15 to 25 years (1.60)* the value          | 0.00| 0.00| 0.00| ... | 0.00| 0.00| 5                         |
| the value considering the weight                    | 0.00| 0.00| 0.00| ... | 0.00| 0.00|                          |
| Land plot area (2.00)* the value                    | 1.60| 2.00| 1.40| ... | 1.20| 1.60| 38                        |
| the value considering the weight                    | 3.20| 4.00| 2.80| ... | 2.40| 3.20|                          |
| Land plot form (1.80)* the value                    | 1.80| 1.80| 1.80| ... | 1.80| 1.80| 38                        |
| the value considering the weight                    | 3.24| 3.24| 3.24| ... | 3.24| 3.24|                          |
| the degree of peat formation (1.50)* the value       | 0.00| 0.00| 0.00| ... | 1.80| 0.00| 6                         |
| the value considering the weight                    | 0.00| 0.00| 0.00| ... | 2.70| 0.00|                          |
| brushwood of the land plot (1.50)* the value         | 0.00| 0.00| 0.00| ... | 0.00| 0.00| 6                         |
| the value considering the weight                    | 0.00| 0.00| 0.00| ... | 0.00| 0.00|                          |
| soils (1.60) the value                               | 1.40| 1.40| 1.40| ... | 1.40| 1.40|                          |
| the value considering the weight                    | 2.24| 2.24| 2.24| ... | 2.24| 2.24| 38                        |
| The necessity of roads' repair and (or) their reconstruction (1.80)* the value | 1.20| 0.00| 0.00| ... | 1.30| 0.00|                          |
| the value considering the weight                    | 2.16| 0.00| 0.00| ... | 2.34| 0.00| 14                        |
| The previous use of land plot (1.10)* the value      | 1.70| 1.80| 1.70| ... | 1.70| 1.80| 38                        |
| the value considering the weight                    | 1.87| 1.98| 1.87| ... | 1.87| 1.98|                          |
| The sum of land plot factor values                  | 19.7| 19.4| 17.4| ... | 21.4| 21.6|                          |
| Land plot location (according to the total factor influence) | 21  | 18  | 10  | ... | 29  | 31  |                          |
| The grouping of land plots                          | 2   | 2   | 2   | ... | 4   | 4   |                          |

*Factor weight (from table 1)
It should be noted that appraisers, as a rule, resort to expert weight distribution, giving the reasons of any kind.

The research procedure is replaced with the use of the mathematical method of weigh coefficient calculation. For this purpose, we have calculated the parameter which is opposite to the specific weight of the corrective sum for each analog in the total sum of analog correctives. The greater the specific weight, the smaller the weight coefficient, and vice versa. The calculation is conducted according to the following formula (figure 1):

\[
C = \frac{(|S_A|+1)\times(|S_1|+1)\times(|S_2|+1)\times\ldots\times(|S_n|+1)}{(|S_1|+1)\times(|S_2|+1)\times\ldots\times(|S_n|+1)}
\]

(1)

where \(C\) – the target coefficient; \(n\) – analog number; \(S_A\) – corrective sum for all analogs; \(S_1\ldots n\) – the sum of analog correctives for which the calculation is conducted; \(S_1; S_2; S_n\) – the sum of the 1\(^{st}\), 2\(^{nd}\) and \(n\)-analog correctives.

It is obvious that the dependence is not linear. Otherwise, the weight distribution would be much easier (in direct proportion) [8].

There is the example of analysis results, received by the authors (table 3). The analysis is conducted to evaluate the possibility of involving unused agricultural lands in the active economic turnover using the covariance matrix.

4. Conclusion
Summation of the weight characteristics of each of the considered factors with their subsequent correction, allows grouping the land plots and ranking them according to the conditions, capabilities and efficiency of their development (table 4).

The solution of such task enables using new indicators, necessary to implement the concept of monitoring development in the Russian Federation, which was repeatedly mentioned in scientific literature [9].

Table 4. The land plot distribution according to development stages with the consideration of cultivated land maximum productivity

| №  | Stage number | The number of participants | The land plot number in inventory records | The total area, ha |
|----|-------------|---------------------------|------------------------------------------|-------------------|
| 1  | The first stage | 12 | 1, 2, 3, 4, 5, 6, 8, 9, 18, 22, 31, 32 | 305.44 |
| 2  | The second stage | 7 | 24, 25, 26, 27, 33, 35, 37 | 278.54 |
| 3  | The third stage | 19 | 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 23, 28, 29, 30, 34, 36, 38 | 281.7 |
|    | Results: | 38 | | 865.68 |

Table 5. The costs for development of unused land plots from among agricultural lands

| Land plot grouping | The land plot number in inventory records | The land plot area, ha | The average development costs for 1 ha, thous. RUB | The total development costs, thous. RUB |
|-------------------|------------------------------------------|-----------------------|-----------------------------------------------|----------------------------------------|
| 1                 | 5, 8, 25, 27, 30, 31, 34, 36 | 244.39 | (24) 25 | 6110 |
| 2                 | 1, 2, 3, 7, 9, 13, 14, 15, 16 | 170.29 | (31) 30 | 5109 |
| 3                 | 4, 11, 17, 18, 22, 23, 24, 26, 28, 29, 32, 33 | 365.98 | (34) 35 | 12809 |
| 4                 | 6, 10, 12, 20, 21, 37, 38 | 81.44 | (39) 40 | 3258 |
| 5                 | 19, 35 | 3.58 | more than 45 | 161 |
| Total amount: | - | 865.68 | - | 27447 |
Each of 13 factors are considered when analyzing the cost value data for the involvement (development) of unused land plot in the turnover. The data usage helps to obtain the average cost values for all participants, belonging to each of 5 selected groups.

The comparison of these data shows the presence of almost identical intervals between the selected groups (figures in brackets; table 5). It indicates the effectiveness and reliability of authors’ method, used to assess the factors of land plot development with the help of the covariance matrix.

References
[1] Volkov S N 2017 The contemporary state of land relations, land management and land use planning in the Russian Federation and the scientific substantiation of their basic regulation directions in agro-industrial complex Materials for the report at the meeting of the Presidium of the Russian Academy of Sciences 28 March 2017 (Moscow: State university of land use planning)
[2] Lipski S A 2017 Integrated approach to the involvement of unused land in agricultural production Nikon readings 371–5
[3] Laikam K E and Fomin A A 2019 About some problems of agricultural land inventory in the Russian Federation Int. Agricultural J. 2(369) 7–12
[4] Vershinin V V and Fedorinov A V 2018 Some systematization aspects of modern land management plans Materials for the Parliamentary hearings of the Council of Federation of the Russian Federation for Stolypin readings at the State university of land use planning of 19 April 2018 eds S N Volkov and A A Fomin (Moscow: State university of land use planning) 54–8
[5] Khlystun V N 2018 Modern land policy and directions of its improvement Economy of agricultural and processing enterprises [in Russian – Economika selskohoziaystvennyh i pererabatvajuschih predprijatij] 88–11
[6] Papaskiri T V, Cherkashina E V, Chepurin E M, Petrova L E and Semochkin V N 2019 The land policy of the Russian Federation under the threat of a new version of the law On Land Management (opinion of experts) Land management, land monitoring and cadaster [in Russian – Zemleustrojstvo, kadastr i monitoring zemel] 4(171) 11–35
[7] Lipski S A 2015 Private ownership for agricultural lands: advantages and disadvantages Experience of two decades Studies on Russian Economic Development 26(1) 63–6
[8] Sleptsov S V 2018 The calculation of weight coefficients in comparative approach Available at: https://www.iop-ocenka.com/ocenka-blog24.html
[9] Volkov S N, Shapovalov D A, Klyushin P V, Shirokova V A and Khutorova A O 2017 Solutions of problems in defining indicators of agricultural land within the framework of activities for the implementation of the concept of development monitoring in the Russian Federation Int. Multidisciplinary Scientific GeoConf. SGEM 17(52) 819–28