# Identification of the Optimal Set of Informative Features for the Problem of Separating of Mixed Production Batch of Semiconductor Devices for the Space Industry

G. Sh. Shkaberina, V. I. Orlov, E. M. Tovbis, and L. A. Kazakovtsev[✉]

Reshetnev Siberian State University of Science and Technology,
prosp. Krasnoyarskiy Rabochiy 31, Krasnoyarsk 660031, Russia
`levk@bk.ru`

Abstract. In this paper, we investigate the problem of separation of a mixed production batch of semiconductor devices of space application into homogeneous production batches. The results of the mandatory testing for each item contain a large number of parameters. Many optimization models and algorithms were developed for solving this clustering problem in the most efficient way. However, due to a rather high data dimensionality, such algorithms take significant computational resources. We analyzed methods of reducing the dimensionality of the data set with the use of factor analysis based on Pearson matrix in order to improve the accuracy of the separation. We investigated efficiency of the proposed method for separating a mixed lot of semiconductor devices which consists of two, three, four and seven homogeneous batches, with various methods of selection and rotation of factors. It was shown on real data that with any orthogonal rotation, with an increasing number of homogeneous batches in the sample, the clustering accuracy decreases. Moreover, it was impossible to identify a universal clustering model with a limited number of factors for dividing a mixed lot composed from an arbitrary number of homogeneous batches. Thus, the use of the multidimensional data was shown to be inevitable.

Keywords: Clustering · Factor analysis · Semiconductor devices

1 Introduction

In order to supply space equipment with highly reliable electronic components, specialized testing centers conduct a variety of tests for each installed semiconductor device. Electronic component base (ECB) designed for installation in spacecraft equipment, along with the input testing is subjected to additional rejection tests, including a selective destructive physical analysis (DPA). DPA allows us to confirm the good quality of the batches of ECB, or to identify the

batches, having defects due to manufacturing technology and are not detected during conventional rejection tests and additional non-destructive testing. In order to be able to transfer the results of DPA of several devices to the entire batch of semiconductor devices, the following requirement is put forward for the ECB intended for installation in space equipment: all devices from the same batch must be made from the same raw materials. Equipment manufacturers for general consumption (not designed solely for use in spacecraft) can not guarantee the implementation of this requirement. Therefore the problem of automatic grouping of semiconductor devices by production batches is very relevant.

It was shown [1], that the problem of allocation of homogeneous batches can be further reduced to a problem of cluster analysis. Authors [1] consider k-means, p-median and other optimization models for solving such a problem. Each group (cluster) must represent a homogeneous batch. To solve the problem of identifying homogeneous batches, in papers [2–4], the application of the clustering optimization algorithm k-means is proposed. In [5], authors consider the clustering method based on the EM algorithm which maximizes the log likelihood function. A model of separation of homogeneous production batches based on a mixture of Gaussian distributions was proposed in [6]. In [7], authors propose using ensembles of optimization models (k-means, k-medoids, k-medians), EM, as well as their optimized versions. In [1], authors consider the application of genetic optimization algorithms with greedy heuristic procedures, in combination with the EM algorithm for the separation of homogeneous batches of electronic devices. The advantage of the new algorithms over classical clustering algorithms for multidimensional data is shown.

In this paper, the initial data are represented by multidimensional sets (arrays) of parameters of electronic radio components (ERC), measured as the results of several hundred mandatory non-destructive tests [8]. In order to reduce the dimensionality of the input parameter sets for clustering devices into homogeneous batches, we propose the application of factor analysis methods. The aim of factor analysis is to find a simple structure that would accurately reflect and reproduce the real dependencies existing in nature [9]. Factor analysis is based on the definition of the factor model

$$X_i = \sum_{j=1}^{m} a_{ij} F_j + u_i \tag{1}$$

where X_i is a vector of values of measured parameter ($i = 1, \ldots, n$), F_j are primary factors ($j = 1, \ldots, m$), a_{ij} are coefficients named factor loadings, u_i are characteristic (specific) factors describing the part of the parameter that is not included in any primary factor. If $m < n$, the reduction of the original problem dimensionality takes place. By reducing the dimension of the data in the article we mean reducing the number of input variables due to the introduction of factors.

The quality improvement is achieved both by more coordinated functioning of radio elements with identical characteristics (from a single production batch), and by improving the quality and reliability of the results of destructive

testing, for which it is possible to select elements from each production batch [1]. This paper is devoted to the problem of reducing the dimension of the original data for the corresponding problems of cluster analysis and attempts to find an optimal set of the informative features used in such cluster analysis optimization problems.

2   Data and Preprocessing

As an example of real data, in this paper we consider a sample consisting of seven different homogeneous batches. The sample is deliberately composed of homogeneous batches, some of which are extremely difficult to separate by known methods of cluster analysis.

One of the largest samplings, which the specialized test center was faced with, is presented in this paper. The total number of all devices in all batches is 3987: batch 1 contains 71 devices, 116 in batch 2, 1867 in batch 3, 1250 in batch 4, 146 in batch 5, 113 in batch 6, 424 in batch 7. Each batch contains information about 205 input measured parameters of the device. Input parameters for which the data vector contains only zero values or for which the number of non-zero values does not exceed 10% were excluded from consideration. For further processing, 67 input parameters remain to be considered.

At the first step, the analysis of the input parameters showed that the considered set of parameters can be divided into three groups:

1. parameters for which the histograms represent the normal Gaussian distribution (In21 - In28, In39 - In46, In92 - In107);
2. parameters for which the histograms represent a Gaussian distribution with frequency gaps (In84 - In91);
3. parameters for which the histogram does not correspond to Gaussian distributions (In 57 - In64, In75 - In82, In10-In20).

For each group, the histograms of observed frequencies and graphs of adjustment of distributions are given on the example of several input parameters (Figs. 1, 2 and 3).
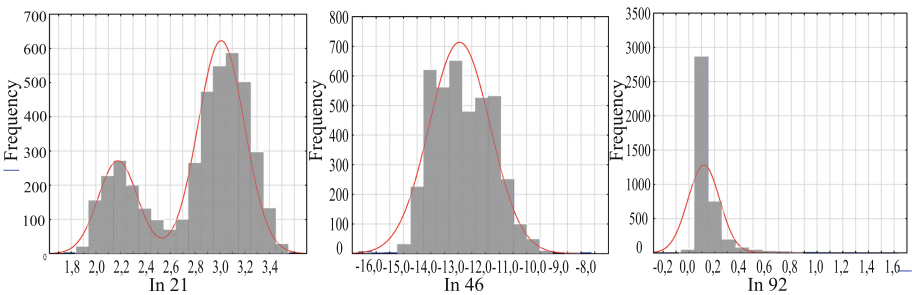

Fig. 1. Histogram of observed frequencies and graphs of the fit of the distributions. Normal Gaussian distribution
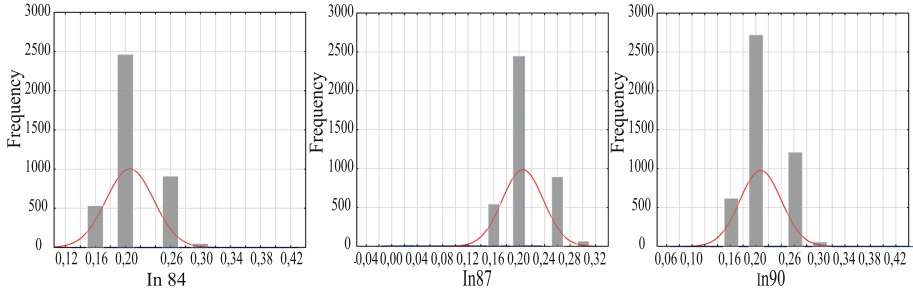
Fig. 2. Histogram of observed frequencies and graphs of the fit of the distributions. Gaussian distribution with frequency gaps

In the second step, the parameters were normalized according to the Eq. (2),

$$a_{i,k} = \frac{a_{i,k}^* - \overline{a_k^*}}{\delta_k^{max} - \delta_k^{min}} \tag{2}$$

where a_{ik}^* is the value of the measuared parameter before normalization, $\overline{a_k^*}$ are average values of the parameter, δ_k^{min} and δ_k^{max} are the lower and upper bounds of the parameter drift, respectively. The drift means the amount of change of parameters of ERC arising during the additional non-destructive testing, simulating extreme operating conditions. This method of normalization by the drift bounds was proposed in [1]. It is shown experimentally that this method of normalization gives a separation by production batches with a much smaller number of errors.
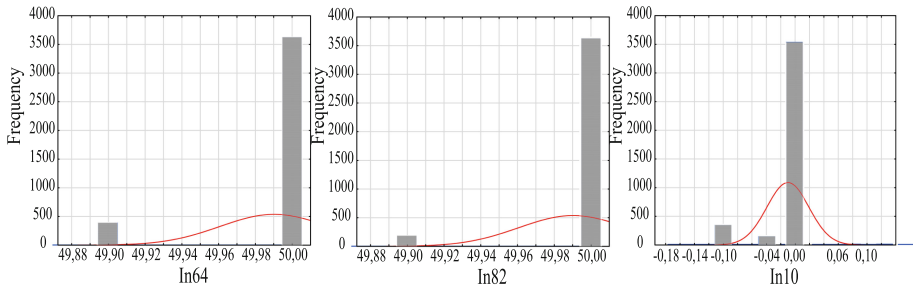

Fig. 3. Histogram of observed frequencies and graphs of the fit of the distributions. The histogram does not correspond to Gaussian distribution

3   Factor Analysis Using Pearson's Correlation Matrix

In the first step, we determine the Pearson correlation coefficient matrix [9] for input parameters. In the second step, we determine the matrix of factor loadings. Assuming the orthogonality of the factors, we obtain

$$R = A \cdot A^T \tag{3}$$

where R is the correlation matrix, A - factor loadings matrix.

The number of factors in the factor model was determined by two criteria. The first of them, the Kaiser criterion [10], selects factors with eigenvalues greater than one. However, the number of sufficient factors also depends on the total share of variance reproduced by these factors. The second of them, Cattel screening criterion [11], selects factors by scree plot based on eigenvalues of factors. The number of factors defined at the point on the chart where the decrease of eigenvalues from left to right slows down as much as possible. Since the Kaiser criterion selects factors with eigenvalues greater than one, and the Cattel screening criterion involves visual observation of the scree plot, there is no need to use any software to calculate these criteria.

Also, to simplify the factor structure, rotation is used to find one of the possible coordinate systems in the space of factors. The consequence of this is the maximization of high correlations and the minimization of low correlations. The problem of rotation is formulated as follows [9]: need to find the transformation matrix T corresponding to:

$$A^{\cdot} = A \cdot T \qquad R = A \cdot A^T = A^{\cdot} \cdot A^{\cdot T} \tag{4}$$

The following methods of orthogonal rotation are used in this paper: the Varimax with Kaiser normalization and the Quartimax with Kaiser normalization [12]. Varimax rotation maximizes the total variance of the loadings squares of the common factors for each input attribute. Quartimax rotation based on the fact that the sum of squares of pairwise products of the matrix A elements will decrease as the values of the loading tend to zero.

Various combinations of parties were subjected to factor analysis: full mixed lot and its subsets lots from four, three and two batches. The full mixed lot consists of seven homogeneous batches. The mixed four-batch lot consists of batch 1, batch 2, batch 5, and batch 6. The mixed three-batch lot consists of batch 1, batch 2, and batch 6. The mixed two-batch lot consists of batch 1 and batch 2.

In this paper, the number of factors was determined by the Kaiser criterion, and the total proportion of variance reproduced by these factors should be at least 70%.

4   Computational Experiments with Various Compositions of the Mixed Lot

To extract factors, we used the principal components method, the principal factor method with multiple R-square, principal axes method, maximum likelihood factors method, iterated communalities method (MINRES) and centroid method [9]. In further consideration, we used principal components method since it describes the maximum variance of input parameters.

For the whole mixed lot, the method based on Cattel criterion recommends to select 4 factors in the model, and this number does not change with any rotation (Fig. 4). According to Kaiser criterion, taking into account the total share of variance of at least 70%, there are five factors selected. Uberla [9] recommends in cases of dispute to select a larger number, therefore we allocate 5 factors for further consideration. Factor 1 corresponds to the highest loadings on the parameters In92-In107. This factor describes 22.779–23.954% of total variance. Factor 2 corresponds to the highest loadings on the parameters In58-In64, In76-In82. This factor describes an additional 19.335–21.265% of the total variance. Factor 3 corresponds to highest loadings on the parameters In39-In46, This factor describes an additional 12.300–14.776% of total variance. Factor 4 (parameters In10, In11, In13, In14, In18) describes 9.003–9.375% of total variance. Factor 5 (parameters In21 - In28) describes 6,781% (unrotated), 11.928% (Varimax) and 11.993% (Quartimax) of total variance. Regardless of the rotation method, the final solution has a cumulative percent of the total variance 75.794% (Table 1).

Table 1. Rotation of factor structure. Full mixed lot

Factor	Eigenvalues			Percent of the total variance (%)			Cumulative percent of the total variance (%)		
	Varimax	Quartimax	Unrotated	Varimax	Quartimax	Unrotated	Varimax	Quartimax	Unrotated
Factor 1	15.26	15.37	16.05	22.78	22.94	23.95	22.78	22.94	23.95
Factor 2	12.95	13.11	14.25	19.34	19.56	21.27	42.12	42.5	45.22
Factor 3	8.29	8.24	9.9	12.38	12.3	14.78	54.49	54.8	60
Factor 4	6.28	6.03	6.04	9.38	9	9.02	63.87	63.8	69.01
Factor 5	7.99	8.04	4.54	11.93	11.99	6.78	75.79	75.79	75.79

The total number of devices in a mixed lot composed of four batches is 446. For further processing 62 input parameters remain. The Cattel criterion, regardless of the rotation, recommends to select 4 factors in the model (Fig. 5), however, according to the Kaiser criterion, taking into account the total percentage of variance at least 70%, we allocate 6 factors. Substantial loadings on the Factor 1 appear for the parameters In21 - In28, In39 - In46. This factor describes 23.304–38.622% of total variance. Factor 2 shows substantial loadings for the parameters In58-In64, it describes in additional 13.220–17.761% of the total variance. Factor 3 has substantial loadings for In91-In107, Factor 4 for In79 - In82, Factor 6 for In57, In78. Regardless of the rotation method, the final solution has a cumulative percent of the total variance equal to 70.364% (Table 2).
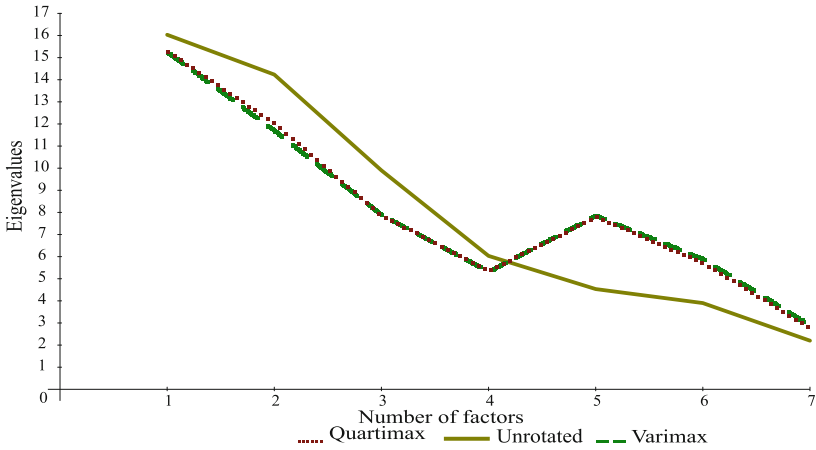
Fig. 4. Scree plot for whole mixed lot. Adv.Grapher

Table 2. Rotation of factor structure. Four-batch mixed lot

| Factor	Eigenvalues			Percent of the total variance (%)			Cumulative percent of the total variance (%)		
	Varimax	Quartimax	Unrotated	Varimax	Quartimax	Unrotated	Varimax	Quartimax	Unrotated
Factor 1	15.07	16.92	23.95	24.3	27.3	38.62	24.3	27.3	38.62
Factor 2	11.01	10.36	8.2	17.76	16.7	13.22	42.07	44	51.84
Factor 3	12.27	11.41	6.96	19.8	18.41	11.23	61.86	62.41	63.07
Factor 4	2.42	2.07	1.9	3.9	3.34	3.06	65.76	65.75	66.13
Factor 5	1.41	1.48	1.38	2.28	2.38	2.22	68.05	68.13	68.35
Factor 6	15.07	16.92	23.95	24.3	27.3	38.62	24.3	27.3	38.62
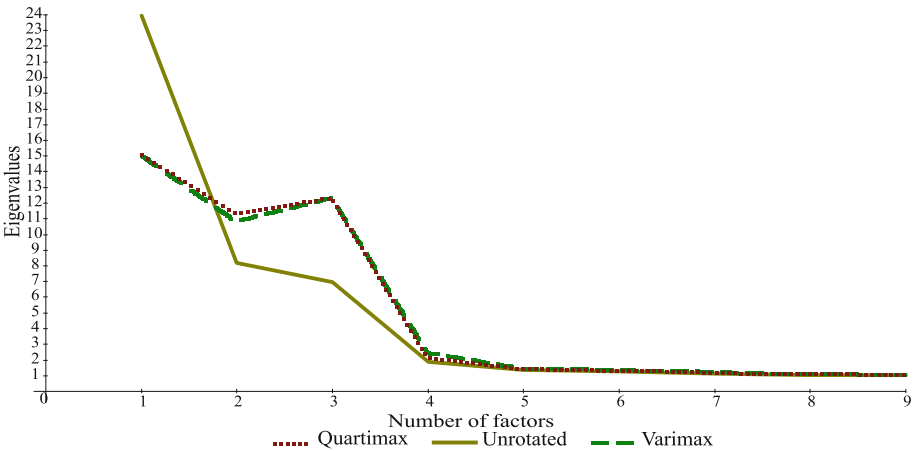
Fig. 5. Scree plot for four-batch mixed lot. Software - Adv.Grapher

The total number of devices in a mix of three batches is 300. The Cattel criterion, regardless of the rotation, recommends selecting 3 factors in the model (Fig. 6). According to the Kaiser criterion, taking into account the total percentage of variance at least 70%, we also allocate 3 factors. Substantial loadings on the Factor 1 appear for the parameters In21-In28, In39-In46. This factor describes 37.09–46.39% of total variance. Factor 2 has substantial loadings for In92-In107 and describes 22.03–26.61% of total variance. Factor 3 has substantial loadings for In84-In91 and describes in addition 9.192-13.905% of the total variance. Regardless of the rotation, the total solution has a cumulative percentage of the total variance 77.61% (Table 3).
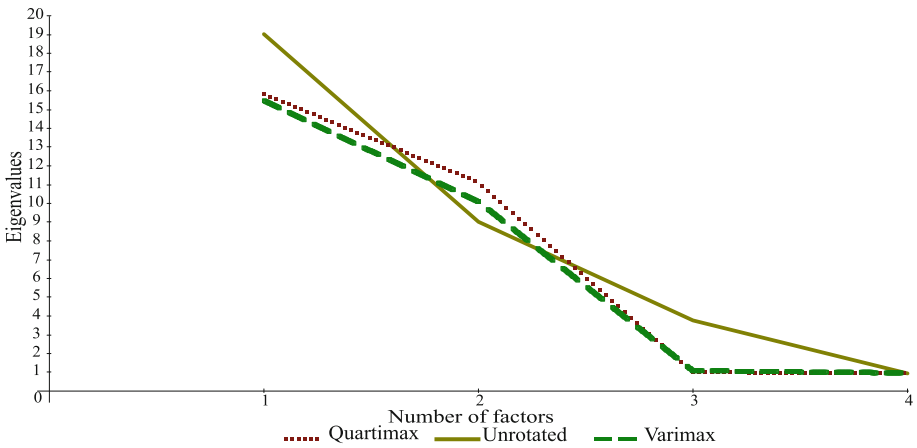

Fig. 6. Scree plot for three-batch mixed lot. Adv.Grapher

Table 3. Rotation of factor structure. Three-batch mixed lot

Factor	Eigenvalues			Percent of the total variance (%)			Cumulative percent of the total variance (%)		
	Varimax	Quartimax	Unrotated	Varimax	Quartimax	Unrotated	Varimax	Quartimax	Unrotated
Factor 1	15.21	15.89	19.02	37.09	38.74	46.38	37.09	38.74	46.38
Factor 2	10.91	10.89	9.03	26.61	26.57	22.03	63.7	65.31	68.41
Factor 3	5.7	5.04	3.77	13.91	12.3	9.19	77.61	77.61	77.61

The number of devices in the simplest mixed lot of two batches is 187. According to the Kaiser criterion, taking into account the total percentage of variance at least 70%, we allocate 2 factors. Factor 1 shows the highest loadings for the parameters In21 - In28, In39 - In46 and describes 45.41–66.20% of the total variance. Factor 2 shows the highest loadings for the In92-In95, In100-In102, In106, and describes in addition 7.28–28.07% of total variance. Regardless of the rotation, the solution has a cumulative percentage of the total variance 73.48% (Table 4).

Table 4. Rotation of factor structure. Two-batch mixed lot

Factor	Eigenvalues			Percent of the total variance (%)			Cumulative percent of the total variance (%)		
	Varimax	Quartimax	Unrotated	Varimax	Quartimax	Unrotated	Varimax	Quartimax	Unrotated
Factor 1	18.62	26.61	27.14	45.41	64.91	66.2	45.41	64.91	66.2
Factor 2	11.51	3.51	2.99	28.07	8.57	7.28	73.48	73.48	73.48

5   Adequacy of the Factor Model

Verification of the factors number sufficiency in the model was performed using The Kaiser and Cattel criteria. Verification the adequacy of the factor model is reduced to checking the achievement of a simple structure. A simple structure is a configuration of vectors that rotates to the state when the vast majority of vectors will be on or near hyperplanes of coordinates [9]. In addition, the simple structure is "contrast": factor loadings are high for variables that determine this factor, and close to zero for all others. To test the significance of a simple structure in various areas of research, modern scientific literature offers the Bargmann test [9], the Lawley-Bartlett's test [9], the Bartlett-Wilks test [9], the Burt's test [9]. In this paper, we use the Bargmann's test [13] due to the ability of this criterion to show that main axis rotation procedure is not completed and control the density of variables positions. It is necessary to calculate the number of zero loadings for each factor:

$$\left| \frac{a_{ij}}{h_i} \right| < 0.1 \tag{5}$$

where a_{ij} are factor loadings on each parameter, h_i - square root of communality (communality refers to the variance of a parameter due to common factors). If the number of zero loadings is not lower than the table value, the simple structure is considered to be achieved.

For the full mixed lot Bargmann test is satisfied for 3 of 5 factors in case of unrotated structure and for all factors in case of rotation with $\alpha <= 0,05$ (where α is a level of significance). For four-batch mixed lot test is satisfied for 3 from 6 factors in case of unrotated structure and for 4 from 6 factors in case of rotated structure with $\alpha <= 0.25$. For three-batch mixed lot test satisfied just for 1 factor in case of unrotated structure and for 2 factors in case of rotated structure with $\alpha <= 0,25$. And for two-batch mixed lot test is satisfied in one case with $\alpha <= 0.25$ (Table 5).

Analysis of the percentage of zero loadings shows, that with increasing the number of batches and at any rotation the number of cases for which test Bargmann is satisfied also increases.

Factor values obtained by orthogonal rotations described above are considered as input data for clustering algorithms. Clustering was performed with Deductor Studio Academic tool. EM algorithm applied with lower bound of likelihood $= 0.2$, level of accuracy $= 10^{-5}$, maximum of iterations $= 100$. Self-organizing Kohonen maps (SOM) [14] applied with linear initialization with

Table 5. Bargmann test

		Factor no.	Table value for $\alpha <= 0,05$	Table value for $\alpha <= 0.25$	The number of zero loading	Percent of zero loading
Full mixed lot (67 parameters)	Unrotated	1	17	14	9	13%
		2			6	8%
		3			18	27%
		4			31	46%
		5			26	39%
	Varimax	1			43	64%
		2			26	39%
		3			33	49%
		4			33	49%
		5			42	63%
	Quartimax	1			41	61%
		2			24	36%
		3			33	49%
		4			33	49%
		5			43	64%
Four-batch mixed lot (62 parameters)	Unrotated	1	20	17	3	5%
		2			8	13%
		3			14	23%
		4			45	73%
		5			36	58%
		6			51	82%
	Varimax	1			14	23%
		2			10	16%
		3			18	29%
		4			44	71%
		5			37	60%
		6			51	82%
	Quartimax	1			14	23%
		2			9	15%
		3			17	27%
		4			52	84%
		5			36	58%
		6			50	81%
Three-batch mixed lot (41 parameters)	Unrotated	1	9	7	0	0%
		2			0	0%
		3			8	20%
	Varimax	1			7	17%
		2			0	0%
		3			10	24%
	Quartimax	1			7	17%
		2			0	0%
		3			15	37%
Two-batch mixed lot (41 parameters)	Unrotated	1	6	4	0	0%
		2			3	7%
	Varimax	1			1	2%
		2			0	0%
	Quartimax	1			0	0%
		2			5	12%

eigenvalues, bubble neighborhood function, significance level $= 0{,}1\%$. The clustering accuracy for considered mixed lots with different orthogonal rotations is presented in Table 6.

The analysis of Table 6 showed that for any orthogonal rotations and clustering algorithms, the clustering accuracy increases with a decrease the number of homogeneous batches in the sample from 39% up to 98%.

Clustering results on three-batch and two-batch mixed lots are shown in Figs. 7 and 8, respectively. Separating batches takes place exclusively on Factor 1 in both cases.

Table 6. Clustering results

|		Unrotated		Varimax		Quartimax	
		EM	SOM	EM	SOM	EM	SOM
Two-batch mixed lot							
Batch 1 n = 71	Number of hits	71 (100%)	71 (100%)	71 (100%)	42 (59%)	71 (100%)	71 (100%)
	Number of errors	0 (0%)	0 (0%)	0 (0%)	29 (41%)	0 (0%)	0 (0%)
Batch 2 n = 116	Number of hits	113 (97%)	113 (97%)	104 (90%)	114 (98%)	116 (100%)	116 (100%)
	Number of errors	3 (3%)	3 (3%)	12 (10%)	2 (2%)	0 (0%)	0 (0%)
Three-batch mixed lot							
Batch 1 n = 71	Number of hits	71 (100%)	36 (51%)	71 (100%)	41 (58%)	71 (100%)	40 (56%)
	Number of errors	0 (0%)	35 (49%)	0 (0%)	30 (42%)	0 (0%)	31 (44%)
Batch 2 n = 116	Number of hits	113 (97%)	110 (95%)	110 (95%)	60 (52%)	111 (96%)	116 (100%)
	Number of errors	3 (3%)	6 (5%)	6 (5%)	56 (48%)	5 (4%)	0 (0%)
Batch 6 n = 113	Number of hits	106 (94%)	102 (90%)	93 (82%)	102 (90%)	98 (87%)	91 (81%)
	Number of errors	7 (6%)	11 (10%)	20 (18%)	11 (10%)	15 (13%)	22 (19%)
Four-batch mixed lot							
Batch 1 n = 71	Number of hits	70 (99%)	71 (100%)	70 (99%)	71 (100%)	70 (99%)	71 (100%)
	Number of errors	1 (1%)	0 (0%)	1 (1%)	0 (0%)	1 (1%)	0 (0%)
Batch 2 n = 116	Number of hits	108 (93%)	108 (93%)	108 (93%)	116 (100%)	108 (93%)	86 (74%)
	Number of errors	8 (7%)	8 (7%)	8 (7%)	0 (0%)	8 (7%)	30 (26%)
Batch 5 n = 146	Number of hits	146 (100%)	68 (41%)	146 (100%)	116 (79%)	146 (100%)	38 (26%)
	Number of errors	0 (0%)	78 (59%)	0 (0%)	20 (21%)	0 (0%)	108 (74%)
Batch 6 n = 113	Number of hits	107 (95%)	107 (95%)	107 (95%)	107 (95%)	108 (96%)	103 (91%)
	Number of errors	6 (5%)	6 (5%)	6 (5%)	6 (5%)	5 (4%)	10 (9%)
Full mixed lot							
Batch 1 n = 71	Number of hits	68 (96%)	71 (100%)	70 (99%)	71 (100%)	63 (89%)	71 (100%)
	Number of errors	3 (4%)	0 (0%)	1 (1%)	0 (0%)	8 (11%)	0 (0%)
Batch 2 n = 116	Number of hits	108 (93%)	106 (91%)	60 (52%)	114 (98%)	81 (70%)	113 (97%)
	Number of errors	8 (7%)	10 (9%)	56 (48%)	2 (2%)	35 (30%)	3 (3%)
Batch 3 n = 1867	Number of hits	487 (35%)	1337 (72%)	618 (33%)	1453 (78%)	699 (37%)	781 (42%)
	Number of errors	1380 (65%)	530 (28%)	1249 (67%)	414 (22%)	1168 (63%)	1086 (58%)
Batch 4 n = 1250	Number of hits	537 (43%)	721 (58%)	462 (37%)	583 (47%)	467 (37%)	571 (46%)
	Number of errors	713 (57%)	529 (42%)	788 (63%)	667 (53%)	783 (63%)	679 (54%)
Batch 5 n = 146	Number of hits	121 (83%)	79 (54%)	135 (92%)	102 (70%)	133 (91%)	73 (50%)
	Number of errors	25 (17%)	67 (46%)	11 (8%)	44 (30%)	13 (9%)	73 (50%)
Batch 6 n = 113	Number of hits	107 (95%)	113 (100%)	107 (95%)	113 (100%)	105 (93%)	113 (100%)
	Number of errors	6 (5%)	0 (0%)	6 (5%)	0 (0%)	8 (7%)	0 (0%)
Batch 7 n = 424	Number of hits	314 (74%)	369 (87%)	256 (60%)	421 (99%)	255 (60%)	284 (70%)
	Number of errors	110 (26%)	55 (13%)	168 (40%)	3 (1%)	169 (40%)	140 (30%)
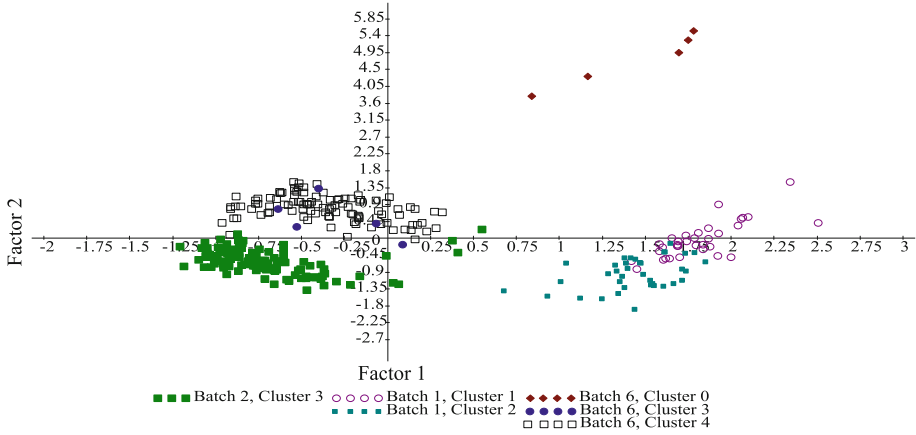
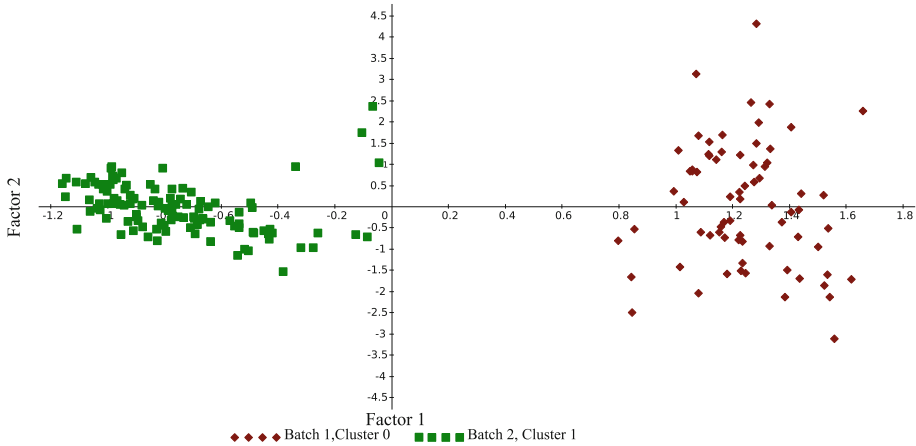**Fig. 7.** Clustering results for three-batch mixed lot

Fig. 8. Clustering results for two-batch mixed lot

6   Conclusions

The possibility of using factor analysis for the separation of a mixed lot, consisting of an arbitrary number of homogeneous batches of electronic radio components, has been proposed and described in the paper. Thus, the use of the factor model is appropriate to improve the accuracy of batch separation, regardless of the clustering algorithm used. It is shown, that the optimal number of the selected factors depends on the number of considered devices in the mixed lot, as well as on the input measured parameters of the device in a given sample. Regardless of the type of orthogonal rotation, the clustering accuracy decreases with the increase of the number of homogeneous batches in the mixed lot. A similar result was shown earlier in [6,7] when using the ensemble approach

of cluster algorithms and [5], where efficiency of EM algorithm at the small volume of input data was demonstrated. At the same time, the considered factor analysis methods do not allow us to obtain a universal set of a small number of features for the separation of mixed lot consisting of an arbitrary number of the homogeneous batches. Thus, despite the fact that the proposed method makes it possible to somewhat reduce the dimensionality of the data, for reliable separation of homogeneous batches with cluster analysis methods, the use of multidimensional data is inevitable.

References

1. Orlov, V.I., Kazakovtsev, L.A., Masich, I.S., Stashkov, D.V.: Algorithmic support of decision-making on selection of microelectronics products for space industry. Siberian State Aerospace University, Krasnoyarsk (2017)
2. Ackermann, M.R., Märtens, M., Raupach, C., Swierkot, K., Lammersen, C., Sohler, C.: J. Exp. Algorithmics **17**, 2.4:2.1–2.4:2.30 (2012). https://doi.org/10.1145/2133803.2184450
3. Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., Wu, A.Y.: Computing nearest neighbors for moving points and applications to clustering. In: Proceedings of the 10th ACM-SIAM Symposium on Discrete Algorithms, Baltimore, MD, pp. S931–S932 (1999)
4. Kazakovtsev, L.A., Stupina, A.A., Orlov, V.I.: Modification of the genetic algorithm withgreedy heuristic for continuous location and classification problems. Sistemy upravleniya i informatsionnye tekhnologii **2**(56), 31–34 (2014)
5. Orlov, V.I., Stashkov, D.V., Kazakovtsev, L.A., Stupina, A.A.: Fuzzy clustering of EEE components for space industry. IOP Conf. Ser.: Mater. Sci. Eng. **155** (2016). Article ID 012026. https://doi.org/10.1088/1757-899x/155/1/012026
6. Kazakovtsev, L.A., Orlov, V.I., Stashkov, D.V., Antamoshkin, A.N., Masich, I.S.: Improved model for detection of homogeneous production batches of electronic components. IOP Conf. Ser.: Mater. Sci. Eng. **255** (2017). https://doi.org/10.1088/1757-899x/255/1/012004
7. Rozhnov, I., Orlov, V., Kazakovtsev, L.: Ensembles of clustering algorithms for problem of detection of homogeneous production batches of semiconductor devices. In: 2018 School-Seminar on Optimization Problems and their Applications, OPTA-SCL 2018, vol. 2098, pp. 338–348. CEUR-WS (2018)
8. Kazakovtsev, L.A., Antamoshkin, A.N.: Greedy heuristics method for location problems. Vestnik SibGAU **16**(2), 317–325 (2015)
9. Uberla, K.: Factorenanalyse. Springer, Heidelberg (1977). https://doi.org/10.1007/978-3-642-61985-4
10. Kaiser, H.F., Dickman, K.: Sample and population score matrices and sample correlation matrices from an arbitrary population correlation matrix. Psychometrika **27**(2), 179–182 (1962). https://doi.org/10.1007/bf02289635
11. Cattel, R.B.: The scree test the number of factors. Multivar. Behav. **1**, 245–276 (1966). https://doi.org/10.1207/s15327906mbr0102_10

12. Harman, H.: Modern Factor Analysis. The University of Chicago Press, Chicago (1967). https://doi.org/10.1002/bimj.19700120119
13. Bargmann, R.: Signifikanz Untersuchungen der einfachenStruktur in der Faktoren-Analyse. Mitteilungsblatt für Mathematische Statistik **1**, 1–24 (1955)
14. Kohonen, T.: Self-organized formation of topologically correct feature maps. Biol. Cybern. **43**(1), 59–69 (1980). https://doi.org/10.1007/bf00337288