WAVELET-BASED CONFIRMATORY FACTOR ANALYSIS OF MONITORING DATA FOR REVEALING LATENT FACTOR INFLUENCES ON EVOLUTION OF A SYSTEM

Lev S. Kuravsky
Department of Computer Science
Moscow State University of Psychology and Education
Moscow, Russia
l.s.kuravsky@gmail.com

Sergey N. Baranov
Russian Aviation Co
Moscow, Russia

Paul A. Marmalyuk
Department of Computer Science
Moscow State University of Psychology and Education
Moscow, Russia
ykk.mail@gmail.com

ABSTRACT – A new technology for monitoring of factors responsible for evolution of technical and other systems is under consideration. It combines capabilities of wavelet transforms and trained factor structures. According to the proposed approach, the samples of coefficients resulted from discrete wavelet transform of initial parameter time series under study and responsible for different observation periods are considered as values of observed variables in the subsequent confirmatory factor analysis to reveal time history of factor influences and estimates of factor interaction. Identification of free factor model parameters (factor variances and covariances) is carried out by a direct (noniterative) procedure based on the maximum likelihood method, which is an alternative to traditional local iterative solution of optimization problems. A statistical method to check significance of factor model components is also discussed. Presented are advantages of the given approach over the traditional simplex method, a set of approaches to development of factor models represented by path diagrams as well as their comparison and software implementation on the base of a graphical programming environment. In addition, a new statistical criterion to estimate applied models' goodness-of-fit measure which is based on Kohonen's self-organizing maps and doesn't require multivariate normality testing of parameters under study is given in details.

KEYWORDS
System condition monitoring, longitudinal data, repeated measures, confirmatory factor analysis, factor model's goodness-of-fit, discrete wavelet transform.

I. INTRODUCTION

As a rule, available parameters measured for condition monitoring do not represent characteristics of a system under study in the mode that is suitable directly for understanding system status and formulating reliable conclusions sufficient for proper diagnostics. For multivariate measurements, which condition monitoring usually deals with, it is important to reveal some latent factors responsible for joint variability of observed measurable parameters, determine their nature and scope of influences, and use the obtained information to identify system condition.

It is desirable to replace the parameters those are easy to measure by the parameters those are easy to interpret and understand the system behavior, with minimal information losses being expected during this data mining. Functional relationships between revealed factors and observed parameters are also to be determined for further analysis. As a result of this study, a researcher should get the structure of causal connections between revealed factors and observed variables as well as immediate factor values to differentiate system status, if necessary.

To meet all the indicated requirements, empirical mathematical models and corresponding methods of multivariate statistical analysis were developed [3-4, 7, 16-17]. The most appropriate in the discussed situation are exploratory and confirmatory factor models and methods of their analysis. Both approaches are based on the analysis of sample covariance or correlation matrices of the observed parameters under study. The exploratory analysis assumes unknown number of uncorrelated factors with a priori undetermined interpretation[1], whereas the confirmatory factor analysis (CFA) assumes the factors, their interpretation, causal connections with observed variables and correlation connections between latent factors to be known beforehand.

Confirmatory models provide a convenient technique for estimating statistical significance of each their component. Since substantial hypotheses of the reasons of possible influences on the observed variables are usually available in practice, the confirmatory approach is preferable.

Condition monitoring usually needs to take into account time dynamics of observed parameters, with their magnitudes for different time points being formally interpreted as different quantities to be analyzed. To comply with this demand, the simplex method of the confirmatory factor analysis was developed [8]. However, it has serious inherent limitations, which frequently make its practical applications questionable, viz.: capacity of studying factor interaction for adjacent checkpoints only, impossibility of associating factors

[1] Factors are usually interpreted using variables, which they are connected with: to identify a factor it is necessary to assign it a name generalizing the meanings of relevant variables.

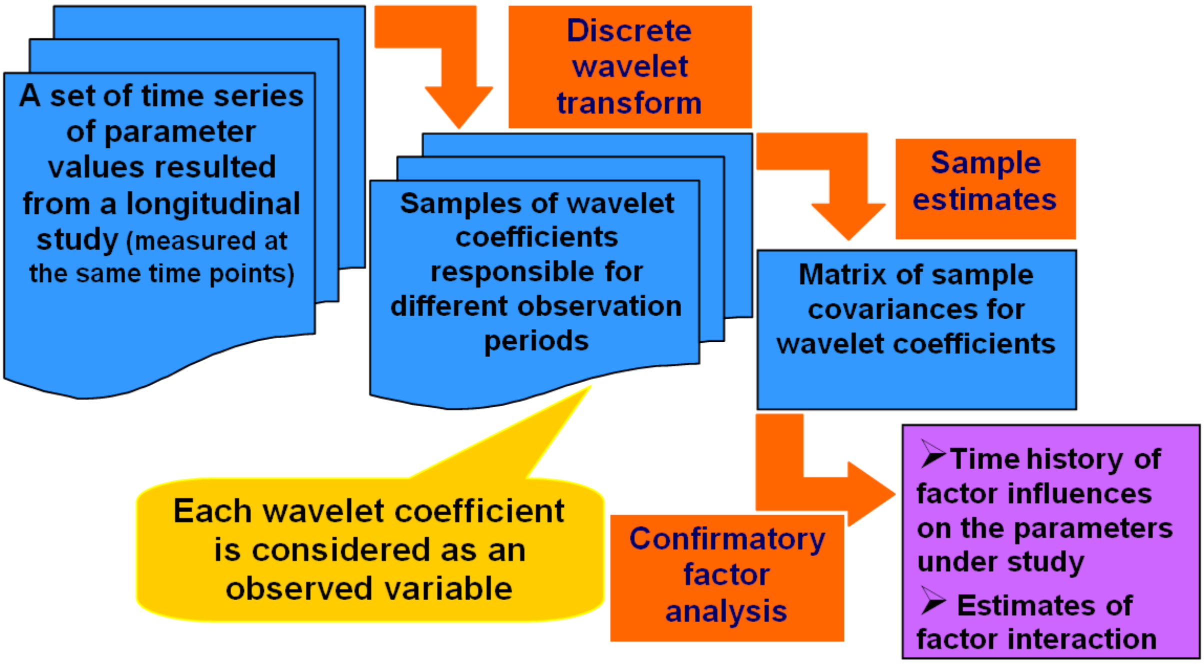

Figure 1. Principal stages of the analysis.

with time periods, acceptability for the analysis of covariance matrices with simplex structure merely, etc. Besides, the traditional CFA has its own intrinsic defect. It needs solution of the laborious local multivariate optimization problem to estimate the values of free model parameters that brings about impossibility of the global minimum estimation and ambiguous solution.

To overcome these problems, a new approach combining capabilities of both wavelet transforms and trained confirmatory factor structures was developed [9]. Its features and advantages, including the possibility of finding the values of free model parameters by direct (noniterative) methods ensuring an unambiguous optimal solution, flexible capacity of studying factor interaction, applicability for the analysis of arbitrary covariance matrices et al., are presented in this paper.

II. Principal Stages of the Wavelet-Based Confirmatory Factor Analysis

Principal stages of the suggested wavelet-based CFA are presented in **Figure 1**. This technology combines capabilities of wavelet transforms and trained factor structures. According to the proposed approach, the samples of coefficients resulted from discrete wavelet transform of initial parameter time series under study and responsible for different observation periods are considered as values of observed variables in the subsequent confirmatory factor analysis (see **Figure 2**) to reveal time history of factor influences and estimates of factor interaction.

Data representation created with the aid of wavelet transforms makes it possible to reveal differences in process characteristics for diverse scales, to filter analyzed parameters time series from noise components and greatly reduces a number of observed variables without significant empirical information losses.

Identification of free factor model parameters (factor variances and covariances) is carried out in the subsequent alternative variant of the CFA by a direct (noniterative) procedure based on the maximum likelihood method, which is an alternative to traditional local iterative solution of optimization problems.

A. Principal Components of the Technology: Wavelet Transforms

Monitoring process representation to be analyzed is created with the aid of wavelet transforms. These transforms make it possible to reveal differences in process characteristics for diverse scales, with the process features being available for analysis in different time points of some interval under study. If the dependence under test is a usual one-variable function, resulting wavelet-spectrum is the function of two arguments, viz.: scale parameter characterizes oscillation time cycles whereas shift parameter – time displacements. Wavelet-spectra are calculated using wavelets, which are special functions in the form of short waves with both zero integral value and localization along the axis of the independent variable, which are able to shift along this axis as well as to scaling (stretching/contraction).

Wavelet-analysis has clear superiority over the traditional spectral analysis since it yields correct representation in case of transition (non-stationary) processes and keeps more useful information about the object behavior under study. Its discrete variant is used here to represent initial parameters time series in the form of points of a certain metric functional space with a wavelet-basis.

B. Principal Components of the Technology: Alternative variant of the CFA

General Principles of the Approach

The traditional CFA expects decision of the laborious multivariate nonlinear optimization problem to estimate the values of free model parameters and, therefore, it results in impossibility of the global minimum estimation and solution ambiguities.

Proposed alternative variant of the CFA allows us to find the values of free model parameters by direct (noniterative) methods ensuring an unambiguous optimal solution.

Hereinafter, each observed variance and covariance is associated with an equation that expresses their expected value via free model parameters (variances and covariances of latent factors). In particular, special tracing rules of the path analysis may be used for that [4, 17].

So, in the alternative variant of the confirmatory factor analysis one has to:

– compose an overdetermined set of the equations each of which expresses observed variances and covariances via free factor variances and covariances with the aid of a factor model;
– solve them by a direct (noniterative) method using a certain form of the maximum likelihood approach, which is different from the one used in the confirmatory factor analysis [12-13];
– examine for the adequacy of the obtained equation sets to observations with the aid of statistical goodness-of-fit tests.

For correct use of the maximum likelihood approach it is necessary to keep certain conditions:

1. Multivariate normalcy of observed variables must be traced.
2. The number of linear equations in the set under study must be equal to a number of observed variances and covariances, and must be greater than the number of free model parameters.

To avoid solving non-linear equation sets as respects to free correlation coefficients and factor loadings the variance components model [19] in which path coefficients (factor loadings) equal to unity is in use (see **Figure 3**).

Hereinafter, each observed variance and covariance is associated with an equation that expresses analytically their expected value via free variances and covariances of latent variables and equates it with the corresponding sample estimation. The set of the equations is obtained, in which number of the equations equals to the number of observed variances and covariances. If this number of equations exceeds the number of free model parameters, the overdetermined set of equations is the case. It is last situation that is necessary for the further decision. The method under consideration needs also multivariate normalcy of observed variables.

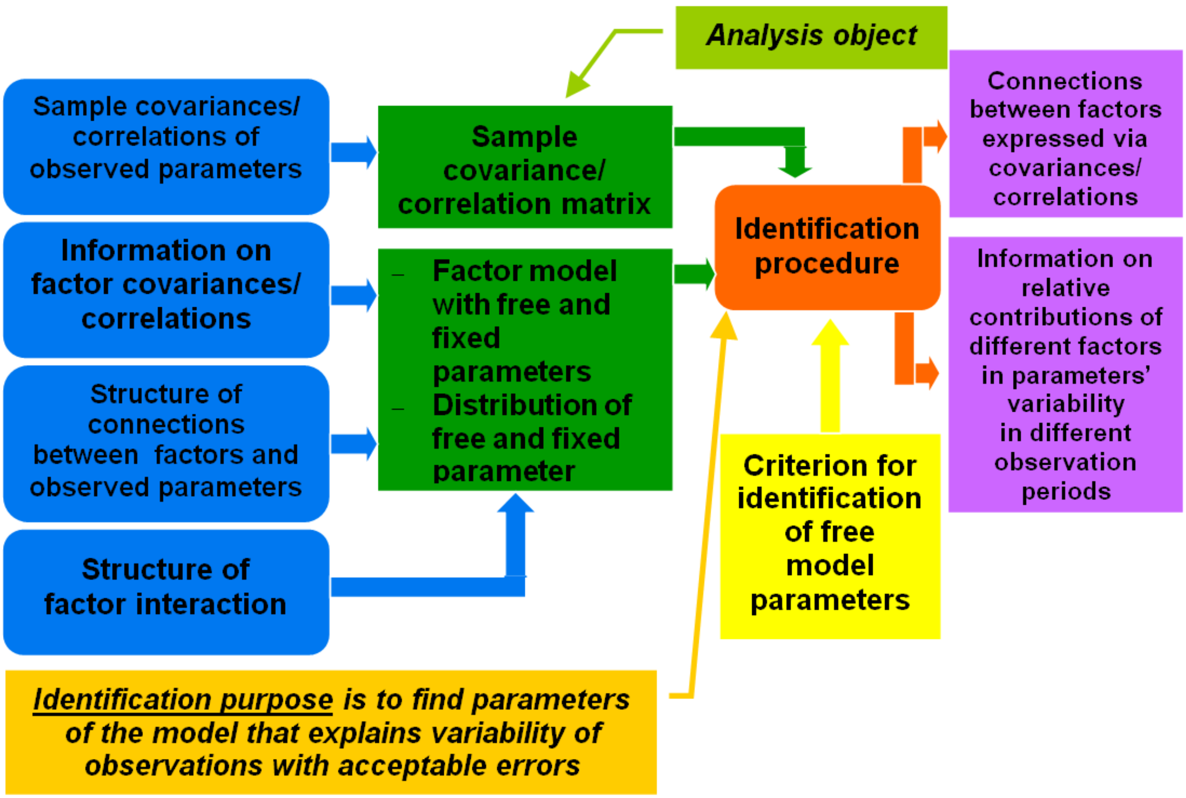

Figure 2. Principal components of the confirmatory factor analysis.

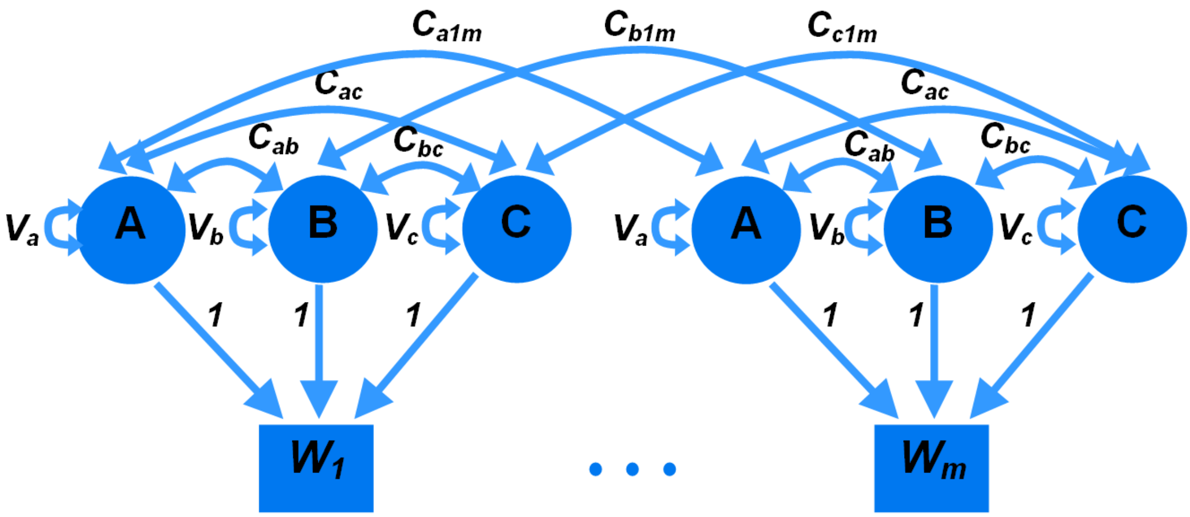

Figure 3. Variance components factor model represented by a path diagram.

Provided that the variance components path model is used, the obtained overdetermined set of linear equations can be represented in a matrix notation:

$$\mathbf{Ax}=\mathbf{b},$$

where **A** - system nxm matrix, which coefficients are determined using the factor model (path diagram) under consideration; **b** - column vector of n variance and covariance sample estimates, which are determined using observation results; **x** - column vector of m unknown free model parameters of interest (viz.: variances and covariances for latent variables).

The vector $\boldsymbol{\varepsilon}=\mathbf{Ax_*}-\mathbf{b}$ represents residual of the given set pseudosolution $\mathbf{x_*}=(\mathbf{A^TV^{-1}A})^{-1}\mathbf{A^TV^{-1}b}$ obtained by the least-squares method. Assuming in the general case that components of the residual vector are correlated let us express their nonsingular covariance matrix as $\boldsymbol{\sigma}^2\mathbf{V}$.

If

– the equation set matrix is nonsingular ($rank\mathbf{A}=m$)

– the transformed residual vector $\mathbf{V}^{-½}\boldsymbol{\varepsilon}$ has multivariate normal distribution

– $\mathbf{x_*}=(\mathbf{A^TV^{-1}A})^{-1}\mathbf{A^TV^{-1}b}$ is pseudosolution,

then this pseudosolution is a maximum likelihood estimate and statistics

$$X^2=(\mathbf{b}-\mathbf{A\,x_*})^T\mathbf{V}^{-1}(\mathbf{b}-\mathbf{A\,x_*})/\sigma^2$$

has χ^2-distribution with n-m degrees of freedom.

Last statistics makes it possible to evaluate the model validity level. Under the assumptions indicated above, the presented statistics $\mathbf{X}^2$ makes it possible to test the hypothesis of representability of sample variances and covariances constituting the vector **b** with the aid of variances and covariances of latent variables contained in the model under study. Acceptance region is $\mathbf{X}^2 \le \chi^2_{n-m;\alpha}$ where $\boldsymbol{\alpha}$ is criterion significance level.

Advantages of the suggested technique are:

– the problem solution is not reduced to the local multivariate optimization;

– since this method is direct there is no multiplicity of solutions;

– no need in search of global minima.

As in the traditional confirmatory factor analysis, the considered model also allows making conclusions on statistical significance of different model components and judge about the importance of the model components under study using goodness-of-fit tests.

To do this one should compare X^2 statistics for two models: saturated model containing the component of interest and simplified model where this component is absent (equals to zero). Let's denote hypothesis that the saturated model coincides with observation results as H_f. Significance level of the component of interest is revealed if there is no grounds to discard hypothesis H_f. At first one should estimate free parameters of the simplified model. The obtained value for X^2 statistics is compared with similar characteristics for the saturated model.

Since the difference in these statistics is asymptotically distributed as χ^2 with the number of degrees of freedom equal to the difference in degrees of freedom of saturated and simplified models, this difference is used to verify zero hypothesis H_r that the simplified model coincides with the observation results against alternative hypothesis H_f.

If H_r hypothesis is not discarded at the given significance level then the component under study is treated as statistically insignificant and the conclusion is made that the available data do not evidence the influence of the studied model part on the observed characteristic under consideration. If H_r hypothesis is discarded (and H_f hypothesis is accepted), then one can talk about the influence of the studied component on the given characteristic.

Typical example of variance components factor model is shown in **Figure 3**. In case of these models, expressions for covariances and variances of wavelet coefficients W_i are linear:

$$Cov(W_i,W_j)=\sum_k C_{kij}\;;$$

$$Var(W_i)=\sum_k V_k+\sum_k\sum_l C_{kl}\;,$$

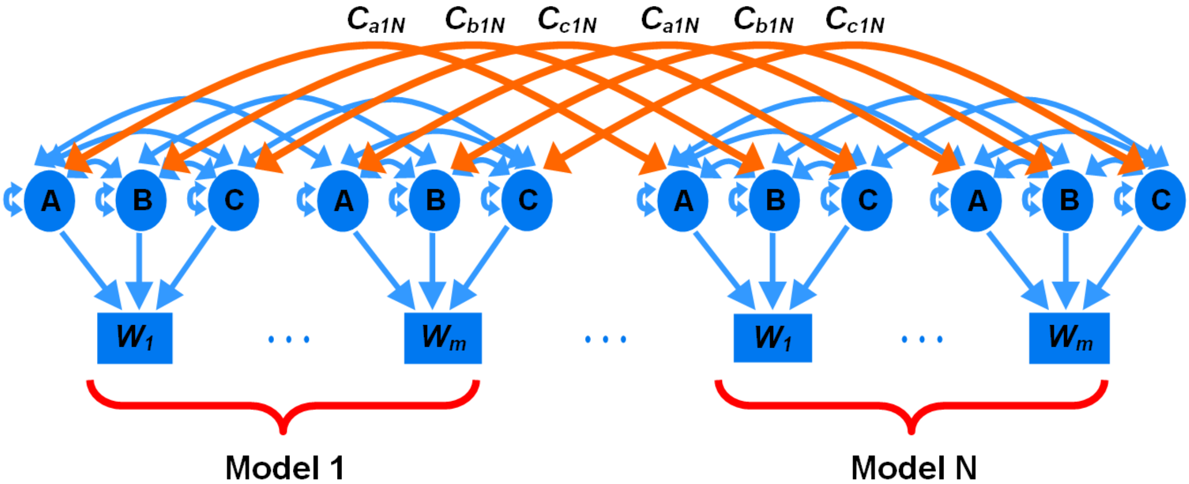

Figure 4. Studying factor influences in case of several variants of observation conditions: simultaneous analysis of model groups. **C*****$_{***}$ are covariances between factors.

where $\boldsymbol{k}$ and $\boldsymbol{l}$ are factor numbers, $\boldsymbol{V}_*$ - variances, $\boldsymbol{C}_{**}$ and $\boldsymbol{C}_{***}$ - covariances between factors. This fact makes it possible to obtain direct estimations of free model parameters using the alternative variant of the confirmatory factor analysis described hereinbefore. Thus, it is the model type that may be used for solution of application problems in reality.

In practical situations, the basic variance components factor model generates a set of particular modifications representing problem peculiarities that are important for solution. For example, simultaneous analysis of different model groups can be useful for studying factor influences in case of several variants of observation conditions (see **Figure 4**).

Typical representation of the wavelet-based confirmatory factor analysis results destined for further interpretation includes:

– factor variances and covariances estimated as free model parameters;

– estimated correlations between different factors relevant to the same time points;

– estimated correlations between the same factors relevant to different time points;

– statistical significance estimations for different model components.

Comparison of the wavelet-based confirmatory factor analysis and the simplex method yields the list of corresponding advantages and disadvantages presented in **Table 1**.

C. Model Overdetermination Reserve

To determine how a factor model fits the accumulated observation results statistics $\boldsymbol{X}^2$ described by the χ^2-distribution is calculated. This operation is possible if the number of observed statistics in use is greater than the number of free model parameters. Corresponding difference forms the number of model degrees of freedom (d.o.f.). This characteristic represents the model overdetermination reserve showing the quantity of additional free parameters that may be included in the model under consideration. This quantity is one of the model capacity characteristics, which are important in practical applications. Expressions of d.o.f. via initial time series sizes and numbers of factors for different model types discussed before are given in **Table 2**.

Table 1. Advantages of the wavelet-based CFA vs. disadvantages of the simplex method

Simplex method: disadvantages ***<u>Destination:</u>*** *the method studies the balance of factor old influences and innovations*	**Wavelet-based CFA: advantages** ***<u>Destination:</u>*** *the method compares the factor influences on different time periods*
Less flexible	More flexible
High-degree nonlinearity of the elements of expected covariance matrices that results in an ambiguous solution	In case of variance components analysis the elements of expected covariance matrices have linear expressions that results in an unique solution
Acceptable for the covariance/correlation matrices with simplex structure only	Acceptable for the arbitrary covariance/ correlation matrices
Lesser reserve of free model parameters (overdetermination degree)	Greater reserve of free model parameters (overdetermination degree)
No possibility to associate factors with time periods	Possibility to associate factors with time periods
If the number of checkpoints is large, multifactor models are very difficult for analysis	If the number of checkpoints is large, multifactor models are moderately difficult for analysis
Limited capacity for studying factor interaction (only for adjacent checkpoints)	Flexible capacity of studying factor interaction
All the existing covariance connections must be taken into account	No need to indicate all the existing covariance connections (only important ones may be taken into account)

Table 2. Numbers of d.o.f. for different model types (M=2^n – time series size, K – number of factors)

Model type	Number of d.o.f.
Simplex model	$M(M+1)/2 - K(3M-2)$
Alternative path coefficients factor model	$M(M+1)/2 – K(M + K – 1)$
Alternative variance components factor model	$M(M+1)/2 – (K^2+K)log_2M/2 – K(log_2M-1)$

In practice, alternative variance components factor models are usually the best. It necessary to note that sometimes only some of the wavelet coefficients can be used in the model if it is expedient for getting rational problem solution.

To ensure model overdetermination when the number of independent observed statistics is insufficient, one has to reduce the number of free parameters equating them to each other, if possible.

D. Model Singularity

If some model derived from an application domain yields system matrix **A** which rank is less than the number of free model parameters, pseudosolution $\mathbf{x}_*$ = $(\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{V}^{-1}\mathbf{b}$ cannot be calculated properly because of singularity of system matrix $\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A}$. In this case, one should reduce the number of free model parameters eliminating certain dependent variables to transform the given matrix into nonsingular form. Number of eliminated variables equals to the defect of matrix $\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A}$. The following technique can be used to determine redundant quantities subjected to this operation:

1. Solution of the eigenvalue problem for the matrix $\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A}$, which is symmetric and nonnegatively definite, to obtain the proper subspace defined by the eigenvectors corresponding to nonzero eigenvalues of the given matrix.
2. Rotation of the obtained proper subspace basis keeping it within this subspace to attain maximal correspondence between directions of coordinate axes of the proper subspace and ones of the initial basis that formally results in transformation of coordinates of the proper subspace axes into either substantial or negligible values (see **Figure 5**).
3. This standard procedure called Quartimax is usually available in the widespread statistical software packages.

The axes of the initial basis, which are represented in the expressions of all the rotated basis directions by negligible coordinate values only, can be considered as lines that are approximately orthogonal with respect to the calculated nonzero proper subspace. Therefore, these lines approximately determine a subspace corresponding to the zero eigenvalues and, accordingly, define variables to be eliminated from the model to reproduce nonsingulary of the matrix in question. Since these quantities cause matrix singularity, they may be considered as dependent (redundant) ones. Their elimination turns into either expressing these variables via independent ones or assigning constant values to them and usually results in the transformed matrix nonsingularity.

If these transformations result in an obviously unacceptable model, one can keep the initial model representation and calculate an approximation of the pseudosolution using, for example, the Gauss-Seidel iteration method.

III. Estimating Goodness-of-Fit Measure with the Aid of Kohonen Self-Organizing Feature Maps

Correct usage of the maximum likelihood method criteria described above for both the traditional and alternative confirmatory factor analysis to identify the values of free model parameters and estimate the model goodness-of-fit measure needs testing multivariate normalcy of distributions of either observed variables or residual vector components. This procedure is laborious and frequently impossible because of deficiency in observed data. The maximum likelihood criteria in case of the traditional confirmatory factor analysis is also too sensitive to a sample size: small deviations from expected

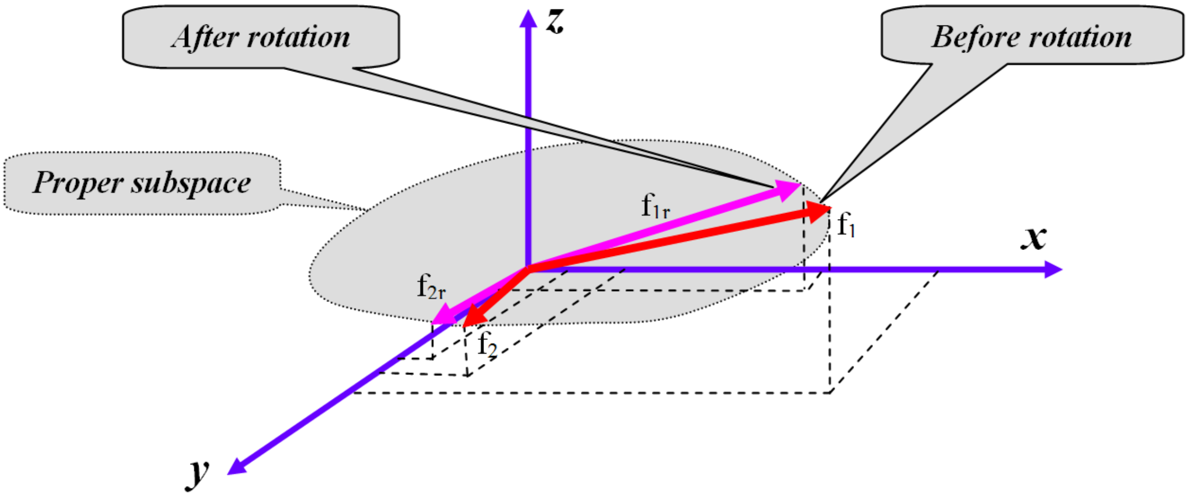

Figure 5. Example: rotation of the 2-D proper subspace basis ($\mathbf{f}_1$, $\mathbf{f}_2$) to attain maximal correspondence between new directions of coordinate axes ($\mathbf{f}_{1r}$, $\mathbf{f}_{2r}$) of the proper subspace and ones of the initial 3-D basis.

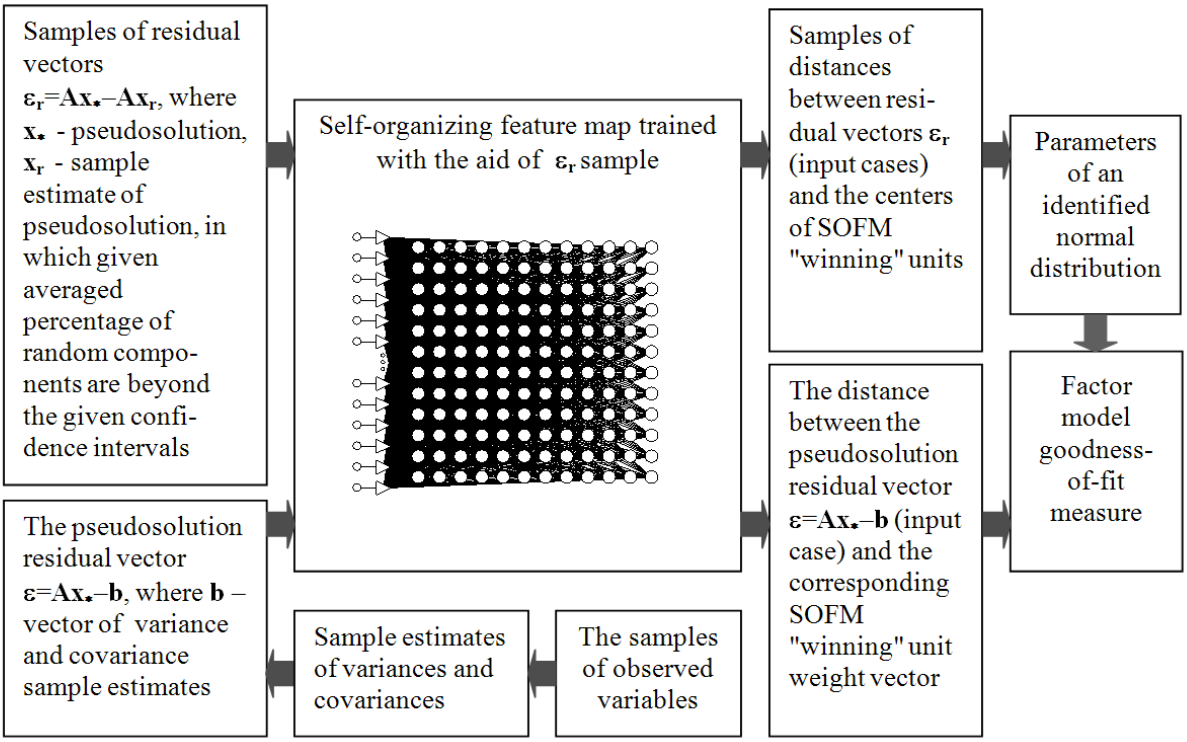

Figure 6. Calculation of factor model goodness-of-fit measure with the aid of the self-organizing feature maps.

characteristics result in considerable goodness-of-fit measures.

To overcome this problem a new technique [18, 10] that uses the capabilities of self-organizing feature maps (SOFM) [6, 11], or Kohonen networks, is proposed. Its framework is presented in **Figure 6**.

Calculation of goodness-of-fit measure is based on comparison of the pseudosolution residual vector $\boldsymbol{\varepsilon}=\mathbf{A}\mathbf{x}_*-\mathbf{b}$ and random samples of residual vectors $\boldsymbol{\varepsilon}_r=\mathbf{A}\mathbf{x}_*-\mathbf{A}\mathbf{x}_r$, where $\mathbf{x}_r$ is sample estimate of pseudosolution, in which given averaged percentage of random components are beyond the given confidence intervals. Residual vectors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}_r$ are both resulted from a factor model under consideration. Random samples of residual vectors $\boldsymbol{\varepsilon}_r$ are used to train self-organizing feature maps of proper dimension and, as a result, to obtain samples of Euclidean distances between residual vectors $\boldsymbol{\varepsilon}_r$ used as input cases and the centers (weight vectors) of SOFM "winning" units. Taking into account the structure of the Euclidian distance and high dimension of residual vectors, which is typical for practical applications, these samples of distances are to be normally distributed.

Estimation of their means and variances identifies the given distributions and yields the opportunity to

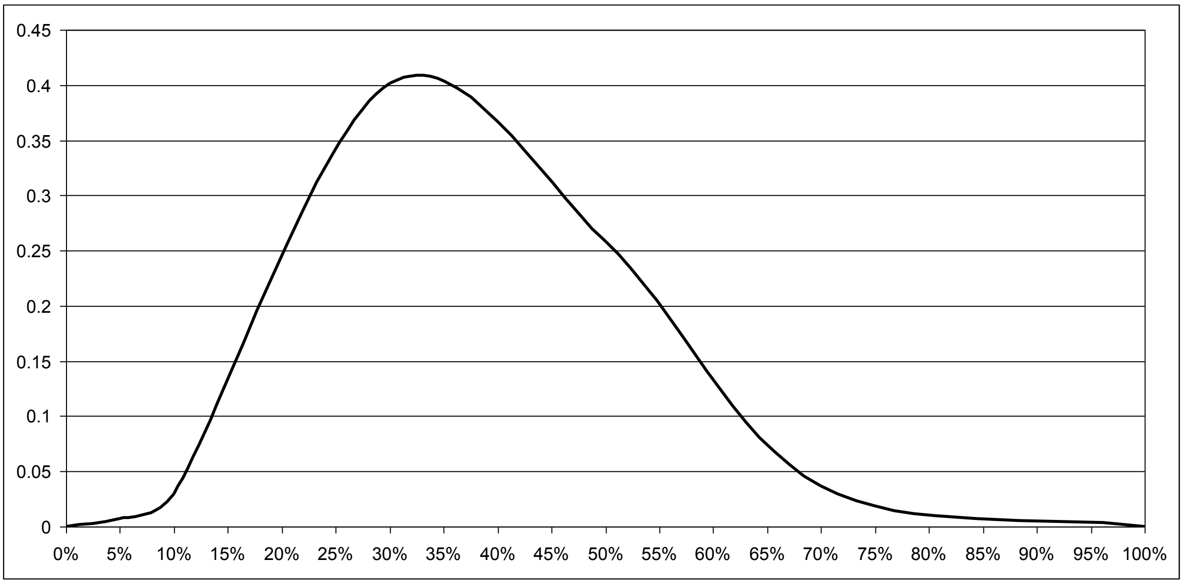

Figure 7. Probability of exceeding the distance between residual vector **ε** and the "winning" unit center as function of averaged percentages of sample pseudosolution component estimates which are going beyond the given confidence intervals.

calculate the probability of exceeding the distance between the pseudosolution residual vector **ε** and its corresponding "winning" unit center that provide one with the factor model goodness-of-fit measure. To get additional information about the structure of deviations of observed variables from their theoretical analogs described by the given factor model, a series of samples with given averaged percentages (e.g. from 0% to 100% by certain intervals) of random components going beyond the given confidence intervals is generated for SOFM training. Comparison of above-stated distance distributions for different percentages makes it possible to reveal the most probable component-wise structure of statistically significant deviations for the pseudosolution residual vector **ε**. As an example, probability of exceeding the distance between residual vector **ε** and the "winning" unit center as a function of these averaged percentages, which was obtained in solving the problem described in paper [1], is shown in **Figure 7**.

According to this approach identified model variances and covariances which compose the pseudosolution are then repeatedly converted to sets of simulated sample estimates of corresponding variances and correlations.

Sample estimates of variances are calculated using the following formula derived from the expression for distribution of sample variance of normally distributed random variable:

$$V_S = \frac{\chi^2_{N-1} V}{N-1},$$

where V_S is variance sample estimate, N is sample size specified for generation, V is an identified variance ingressed in the pseudosolution, χ^2_{N-1} is a random element distributed as χ^2 with $N-1$ degrees of freedom. Elements χ^2_{N-1} are software generated.

Sample estimates of covariances are calculated via corresponding sample estimates of correlations using the fact of approximate distribution normalcy for their Fisher transform, viz.: distribution of the statistics

$$z = \frac{1}{2} ln \frac{1+r}{1-r},$$

where r is sample correlation, can be approximated by the normal distribution with the expectation

$$z = \frac{1}{2} ln \frac{1+\rho}{1-\rho},$$

where ρ is correlation value, and the variance $\frac{1}{N-3}$.

Samples of Fisher transform results are software generated for each covariance ingressed in the pseudosolution, with correlations ρ being substituted for corresponding correlations. Required correlations themselves are restored by means of calculating the inverse Fisher transform for the abovementioned generated values. After that they are converted into the covariances ingressed in the pseudosolution.

Simulated samples of variances and covariances yield required samples of residual vectors $\boldsymbol{\varepsilon}_r = \mathbf{Ax}_* - \mathbf{Ax}_r$ used for SOFM training.

The approach under consideration gives the opportunity to determine easily the sample sizes required for testing hypotheses of equality of the distance between the pseudosolution residual vector **ε** and its corresponding SOFM "winning" unit center to the certain value with both the given significance level and given test power. A formula of interest is derived from the comparison of corresponding acceptance region limits [2]:

$$N = \left(\frac{z_{1-\alpha/2} + z_{1-\beta}}{d_{norm}} \right)^2,$$

where $z_{1-\alpha/2}$ and $z_{1-\beta}$ are standard normal distribution quantiles of orders $1-\alpha/2$ and $1-\beta$, correspondingly; α is significance level; β is probability of type 2 error; d_{norm} is the ratio of deflection of true distance expectation from the tested certain value to the standard deviation of distance distribution.

The given technique was software implemented on the base of the National Instruments LabVIEW graphical programming environment [15]. The work of self-organizing feature maps was simulated with the aid of the STATISTICA Neural Networks software package.

IV. EXAMPLES OF APPLIED MODELS

The approach under consideration was software implemented on the base of the *LabVIEW* graphical programming environment [14] and successfully applied to solution of the following problems:

- Studying influence of maneuvering loads occurrences and climatic conditions of basing on aircraft damage accumulation rate [1]
- Studying of bilingual (Spanish and English) phonological awareness of American schoolchildren [5].
- Testing statistical significance of both national intelligence and economic progress factors influences on gross domestic product in European countries.
- Studying influence of conjuncture factor on the variability of gross added costs in different sectors of national economy in Russian Federation.
- Studying common factor (which has been interpreted as national morality level) impact on social characteristics in European countries [14].

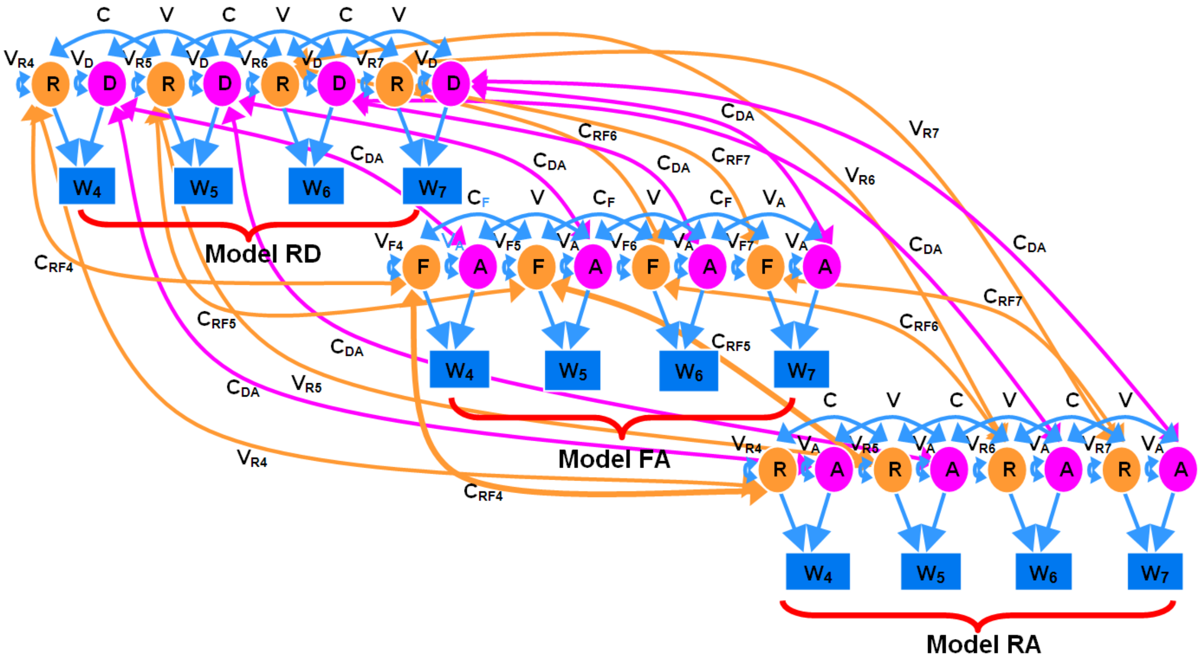

Figure 8. Model to study influence of maneuvering load occurrences and climatic conditions of basing on aircraft damage accumulation rate: influences of national features of pilotage technique are represented by factors **R** and **F**, influences of national environment exploitation – by factors **D** and **A**.

Two examples of factor models used for solution of the above-stated problems are shown in **Figures 8** and **9**.

V. MAIN RESULTS AND CONCLUSIONS

1. Proposed is the wavelet-based confirmatory factor analysis intended for monitoring of factors responsible for evolution of technical and other systems, which combines capabilities of wavelet transforms and trained factor structures. According to the proposed approach, the samples of coefficients resulted from discrete wavelet transform of initial parameter time series under study and responsible for different observation periods are considered as values of observed variables in the subsequent confirmatory factor analysis to reveal time history of factor influences and estimates of factor interaction.

2. Identification of free factor model parameters (usually factor variances and covariances) is carried out by a new direct (noniterative) procedure based on the maximum likelihood method that is an alternative to traditional ambiguous local iterative solutions of multivariate optimization problems, which depend on the initial approximations.

3. Comparison of different sorts of factor structures revealed advantages of variance components models. This fact is conditioned by linearity of their analytical representations, which is convenient for direct estimations of free parameters, and greater overdetermination reserve.

4. A special technique based on the eigenvalue problem solution and rotation of the obtained proper

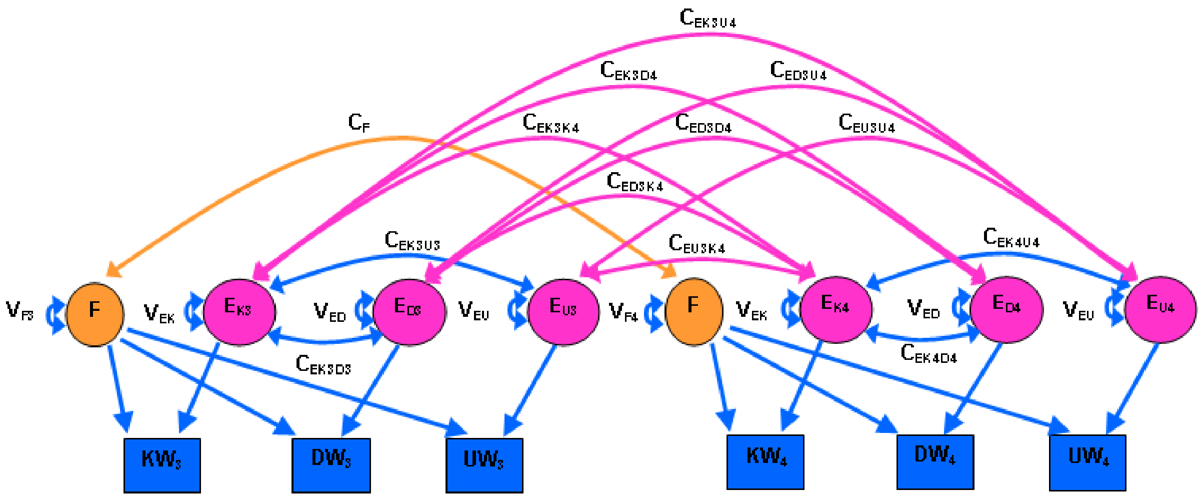

Figure 9. Model to study influences of the common factor **F** (which has been interpreted as national morality level**)** on socio-economical characteristics in European countries: specific factors influencing parameters **K** (corruption index), **D** (Jeany's index) и **U** (murder's index), are represented as $\mathbf{E_{K*}}$, $\mathbf{E_{D*}}$ и $\mathbf{E_{U*}}$; $\mathbf{V_*}$ - variances of factors; $\mathbf{C_*}$ - covariances between factors in different time periods.

subspace basis was developed to find out the dependent model variables, if any, and overcome the resulting singularity of the system under study.

5. A new statistical criterion for estimating factor models' goodness-of-fir measures which doesn't require multivariate normality of observed system parameters under study was developed. It has the following advantages:

a. no need to test multivariate normalcy of distributions of either observed variables or residual vector components;

b. simple procedure of estimating type 2 statistical errors is available;

c. it is possible to reveal the most probable percentage component-wise structure of statistically significant deviations for the pseudosolution residual vector;

d. higher reliability of obtained goodness-of-fit measures because of unrestrictedness of generated random samples of variances and covariances ingressed in the pseudosolution and the following unlimited goodness-of-fit estimation accuracy.

6. The suggested approach has substantial advantages over the simplex method usually used for monitoring of factors responsible for system evolution.

7. The proposed approach was software implemented on the base of a graphical programming environment and applied to solution of technical and other problems.

ACKNOWLEDGMENT

This work is supported by the grant 10-06-00423 from the Russian Foundation for Basic Research.

REFERENCES

[1] Baranov S.N., Kuravsky L.S., Baranov N.I. Studying influence of maneuvering loads occurrences and climatic conditions of basing on aircraft damage accumulation rate with the aid of trained structures. In: Proc. 5th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Edinburgh, Scotland, United Kingdom, 2008.

[2] Bendat, Julius S. and Allan G. Piersol: Random Data - Analysis and Measurement Procedures (2nd edition); New York: Wiley, 1986.

[3] Bishop Y.M.M., Fienberg S.E., Holland P.W. Discrete multivariate analysis: Theory and practice. Cambridge, MA, M I T Press, 1975.

[4] Bollen K.A. Structural equations with latent variables. New York, John Wiley, 1989.

[5] Branum-Martin L., Carlson C.D., Carlo M., Fletcher J.M., Francis D.J., Mehta P.D., Ortiz A. Bilingual phonological awareness: Multilevel construct validation among Spanish-speaking kindergarteners in transitional bilingual education classrooms. Journal of Educational Psychology, Vol. 98, No 1, 2006 pp.170–181.

[6] Galushkin A.I. Theory of neural networks. A tutorial. Moscow: IPRZHR, 2000.

[7] Goldstein H. Multilevel statistical models. 3rd ed., London, Arnold, 2003.

[8] Jöreskog K.G. Estimation and testing of simplex models. British Journal of Mathematical and Statistical Psychology, Vol. 23, pp.121-145, 1970.

[9] Kuravsky L.S. and Baranov S.N. *Development of the wavelet-based confirmatory factor analysis for monitoring of system factors*. – In: Proc. 5th International Conference on Condition Monitoring & Machinery Failure Prevention Technologies, Edinburgh, United Kingdom, pp.818-834.

[10] Kuravsky L.S., Baranov S.N. and Baranov N.I. Wavelet-based confirmatory factor analysis for monitoring of system factors: estimating goodness-of-fit measures with the aid of self-organizing feature maps. – In: Proc. 6th International Conference on Condition Monitoring & Machinery Failure Prevention Technologies, Dublin, Ireland, pp. 224-245.

[11] Kuravsky L.S., Baranov S.N., Malykh S.B. Neural networks in forecast, diagnostics and analysis tasks. Moscow: RUSAVIA, 2003.

[12] Kuravsky L.S., Kornienko P.A. Application of neural networks for identifying quantitative trait loci in behavior genetics. Neyrokomputery: Razrabotka i Primenenie (Neurocomputers: Development and Application), No 4, pp. 15-31, 2007 (in Russian.)

[13] Kuravsky L.S., Kornienko P.A. On the approach to identifying quantitative trait loci in behavior genetics. In: Proc. 2nd World Congress on Engineering Asset Management and 4th International Conference on Condition Monitoring, Harrogate, United Kingdom, pp.1133-1146, June 2007.

[14] Kuravsky L.S., Yurevich A.V., Marmalyuk P.A., Ivanova E.G. Factor analysis of a national morality index in European countries. Psychological science and Education. Moscow: MSUPE, 2010. №1. PP. 94–110.

[15] LabVIEW tutorial for Windows. National Instruments Corp., 1994-2007.

[16] Lawley D.N., Maxwell A.E. Factor analysis as a statistical method. Moscow, Mir, 1967 (in Russian.)

[17] Loehlin J.C. Latent variable models: An introduction to factor, path, and structural analysis. Hillsdale, NJ, Erlbaum, 1987.

[18] Marmalyuk P.A. Estimating factor models goodness-of-fit measures with the aid of Kohonen's self-organizing feature maps; Moscow: Neurocomputers: developing and application, №10 (2010), pp. 53 – 62.

[19] Neale M.C., Cardon L.R. Methodology for genetic studies of twins and families. Dordrecht, the Netherlands, Kluwer Academic Publishers, 1992.

[20] Rubini R., Meneghetti U. Application of the envelope and wavelet transform analyses for the diagnosis of incipient faults in ball bearings. Mechanical Systems and Signal Processing, Vol. 15, pp.287-302, 2001.

[21] Tse P.W., Peng Y.H., Yam R. Wavelet analysis and envelope detection for rolling element bearing fault diagnosis-their effectiveness and flexibilities. Journal of Vibration and Acoustics, Transactions of the ASME, Vol. 123, pp.303-310, 2001.