Regularized Identification in Engine Models Matching with Measured Data

Sergiy Yepifanov **Roman Zelenskyi**
Zhukovsky National Aerospace University, “Kharkov Aviation Institute”
17, Chkalov St., 61070, Kharkov
UKRAINE

aedlab@gmail.com meerkat_r@ukr.net

Aleksandr Khustochka
SE “Ivchenko-Progress”
1, Ivanova St., 69068, Zaporozhye
UKRAINE

a.khustochka@ivchenko-progress.com

ABSTRACT

The problem of engine model matching with measured data is considered in application to thermodynamic static model used in design, dynamic model used in automatic control and normal state model used in diagnostics. Methods of engine models identification are considerably developed for providing stable, precise and physically adequate solutions. This development is based on the use of a priory information about engine, its parameters and performances directly in the identification procedure. The mathematical basis for this development is the fuzzy sets theory.

NOMENCLATURE

A – turbine nozzle box area;
av – average value;
C – generalized matrix of influence coefficients;
G – air/gas consumption;
H – matrix of influence coefficients;
HPC – high pressure compressor;
HPT – high pressure turbine;
I – unity matrix;
K – number of operation modes;
LPT – low pressure turbine;
m – number of measured parameters;
r – number of estimated parameters;
sim – simulated value;
U – parameters that determine the engine operation conditions;
W – weight diagonal matrix;
Y – measured parameters;
Z – generalized vector of measured parameters;
γ – regularization (weighting) factor;
Δ – parameter’s correction;
δ – relative deviation;
η – efficiency;
σ^2 – dispersion;
θ – component’s map parameter;
$\rightarrow$ – vector;
^ – estimated parameter;
* – measured value

INTRODUCTION

Mathematical models of turbine engines, which are based on working thermo-gas-dynamic process description, are widely used in design of the engine and its automatic control and diagnostic systems. These models calculate the gas path parameters (air and gas temperature and pressure, speeds of rotors, fuel consumption, thrust etc.) as a function of a set steady-state or transient operational mode and ambient conditions (atmospheric temperature and pressure, flight speed and altitude). These relations are non-linear. The considered models decrease material expenses and man hours at design stage, and, sometimes, give new information about the object's behavior.

The models are based on performance of various engine components (compressor and turbine spools, main combustion chamber, afterburner, intake, exhaust system, transition channels, secondary air system etc.). These component performances are known at some degree of confidence. For each engine, they have their own individual deviations because of differences in manufacturing and assembling. Additionally, these component performances degrade in maintenance because of different wear and propagation of faults. These changes are simulated by special parameters (components performance parameters or engine state parameters) that are included in thermo-gas-dynamic models. They serve to shift the initial (average engine) component performances, thus taking into account the current technical condition of the engine.

The models are subjected to a fitting procedure, when the experimentally measured parameters are mated to similar simulated parameters. The identification adjusts the model to make it output parameters as close to experimental as possible. Except the significant improvement in the gas path simulation, the identification has a great diagnostic value because the estimated parameters contain information about the technical condition of each component.

The non-linear thermo-gas-dynamic models of turbine engines and the identification procedures are applied in diagnostic algorithms for more than forty years. Nowadays, due to progress in computing capability, there are no limits for the direct implementation of these models and identification procedures into algorithms of the engine real-time automatic diagnostic systems.

The model calculates parameters of the engine gas path $\vec{Y}$ at steady-state modes depending on a mode, external conditions $\vec{U}$ and component performance $\vec{\Theta}$. Hence, in the general case it is represented as

$$\vec{Y} = F(\vec{U},\ \vec{\Theta})\,. \tag{1}$$

The linear model may be formulated as (1):

$$\delta\vec{Y} = H\delta\vec{\Theta}\,, \tag{2}$$

which relates small deviations of the gas path parameters $\delta\vec{Y}$ and parameters of components' performances $\delta\vec{\Theta}$ at a single operational mode. H is an influence coefficients matrix. The linear model (2) is a component of the identification algorithm for the non-linear model.

The problem of the identification has a long background as the mathematical models are used in the engine design and development. Many scientists tried to solve it. A. Tunakov solved it using the Least Square Method (LSM) in 1979 [1]. His method can be considered as a universal one because it uses information at some off-design modes of the engine operation. S. Yepifanov in 1981 applied correction of the component performances, which provides not only displacement but also turning of these performances at a multi-mode identification; in fact, this corresponds to variable correction coefficients [2, 3]. The similar methods were used by A. Stamatis, K. Matioudakis et. al. [4, 5]

and developed by C. Kong, Y. Li, P. Pilidis et. al. [6, 7, 8], B. Roth, D. Doel [9] and many other researchers. They are based on a minimization of a functional, which is a sum of squared deviations, calculated as differences of calculated and measured values.

The task of the engine mathematical model matching with experimental data is characterized by a presence of multiple parameters, which can be used for the model correction. These parameters can be strongly correlated. At the same time, a quantity of measured parameters is strongly limited. This reasons decrease a stability of correction procedure, which is based on the Least Square Method, and force researchers to check for methods to improve the stability. For this reason, V. Borovick and Ye. Taran applied the Least Modulus Method [10], which is more robust to the data outliers. S. Yepifanov implemented the Marquardt's method [11] and in further methods of the Singular Value Decomposition and ε-structuration [12, 13]. A. Volponi et. al. applied the Kalman filter [14].

There is well known that any regularization reduces to the modification of the above mentioned functional by adding the regularizing component to it. This causes a biased estimates of the model coefficients. B. Roth, D. Doel et. al. demonstrated it clearly [15].

Hence, the engine models matching needs regularization. At that we need to monitor the estimate errors of the model parameters, which are due to the bias. Formal methods of regularization essentially limit the possibility of this monitoring. This paper presents the regularization of the matching task using a priori information about the engine, its mathematical model and proper performance, and also about the measuring system and the measuring procedure. This information is heterogeneous: partially it is presented in a view of statistic parameters and partially in a view of heuristic expert representations. The unified instrument that helps to formalize this information and complete mathematical formulation of the considered task is the Fuzzy Sets Theory [16]. Due to this modification of the functional the identification task becomes non-linear. Therefore, traditional methods of its solution are not applicable. This paper considers the solution based on the genetic algorithm [17], which is adapted to specifics of the engine model matching.

BASIC IDENTIFICATION PROCEDURE

Identification of the non-linear model (1) by measuring results $\vec{Y}^*$ and $\vec{U}^*$ lies in determining estimations $\hat{\vec{\Theta}}$, which are solutions of the following optimization task:

$$\Psi(\vec{\Theta}) = \left\| \vec{Y}^* - \vec{Y}(\vec{U}^*,\ \vec{\Theta}) \right\|, \quad \Psi(\hat{\vec{\Theta}}) = \min \Psi(\vec{\Theta}) , \tag{3}$$

where ‖ ‖ is a norm of vector.

The estimation precision will improve, if a priory information is redundant in relation to a number of unknown parameters. So the identification can be provided by a set of measurements that are done at K different operating modes. In this case, the generalized vector of residuals is minimized

$$\vec{Z}(\vec{\Theta}) = \begin{bmatrix} \vec{Y}_1^* - Y(\vec{U}_1^*, \vec{\Theta}) \\ \vec{Y}_2^* - Y(\vec{U}_2^*, \vec{\Theta}) \\ \dots\dots\dots\dots\dots \\ \vec{Y}_K^* - Y(\vec{U}_K^*, \vec{\Theta}) \end{bmatrix}. \tag{4}$$

The model (1), which is included into (4), is numerical. Therefore, minimization of residuals (4) is numerical iterative procedure, which in each iteration determines solution as a sum of previous solution and current correction:

$$\vec{\Theta}^{n+1} = \vec{\Theta}^n + \Delta\vec{\Theta}^{n+1}, \tag{5}$$

and iterations continue till corrections become negligible.

The correction $\Delta\vec{\Theta}^{n+1}$ is determined as a solution of overdetermined system of linear algebraic equations

$$C(\vec{\Theta}^n)\Delta\vec{\Theta}^{n+1} = Z(\vec{\Theta}^n)\,, \tag{6}$$

where

$$C(\vec{\Theta}^n) = \begin{bmatrix} H_1(\vec{\Theta}^n) \\ H_2(\vec{\Theta}^n) \\ \dots\dots \\ H_K(\vec{\Theta}^n) \end{bmatrix} \tag{7}$$

is generalized matrix of influence coefficients, which is composed of elementary matrixes $H_i(\vec{\Theta}^n)$ determined for each operating mode. In accordance with system of equations (6), for each iteration the such correction $\Delta\vec{\Theta}^{n+1}$ of required parameters is find, which narrows residuals (4) $\vec{Z}(\vec{\Theta})$ down.

The known solution of linear system (6) by the least square method (LSM-solution) has a view

$$\Delta\vec{\Theta}^{n+1} = A^{-1}C^T W\vec{Z}\,, \tag{8}$$

where $A = C^T WC$ – information matrix of Fisher, W – weight diagonal matrix, which elements are inverse to dispersions of measuring errors $\sigma^2_{Y_I}$.

Unfortunately, the least square method is sensitive to outliers in the right part of the system (6), which can be related with faults in experimental data. So practical application of the LSM had demonstrated instability of estimations $\Delta\vec{\Theta}^{n+1}$ and poor convergence of the identification algorithm as a whole. In some cases, the estimations $\hat{\vec{\Theta}}$ are far from expected values. The reasons of such results may be lack of empirical information, excess number of estimated parameters or correlation between two or more state parameters. The least causes ill-conditioning of the Fisher Matrix and excessive deviations of estimations. These estimations loose a physical sense (are out of range of possible components' performances parameters variation – for example efficiency is more than one) and can cause the model calculation program crash.

Hence, the LSM identification procedure needs modification for providing its stability and physically adequate estimations.

REGULARIZED IDENTIFICATION PROCEDURE

In the proposed procedure, the generalized functional is minimized, which besides the residuals by measured parameters includes a norm of the finding parameter vector with a weighting factor γ. The identification task takes a view

$$\Psi'(\vec{\Theta}) = \left\{ \left\| \vec{Y}^* - Y(\vec{U}^*,\ \vec{\Theta}) \right\| + \gamma \left\| \vec{\Theta} \right\| \right\}, \quad \Psi'(\hat{\vec{\Theta}}) = \min \Psi'(\vec{\Theta}) . \tag{9}$$

The new identification procedure is based on the above described procedure. As it is understand from relations (9), the minimized functional is extended by elements, which are related with parameters to be find $\vec{\Theta}$. So the main changes in the identification procedure are extension of the generalized vector of residuals $\vec{Z}(\vec{\Theta})$ and generalized matrix of influence coefficients $C(\vec{\Theta})$. The modified values have the following structure:

$$\vec{Z}'(\vec{\Theta}) = \begin{bmatrix} \vec{Z}(\vec{\Theta}) \\ \gamma\vec{\Theta} \end{bmatrix} \tag{10}$$

and

$$C'(\vec{\Theta}) = \begin{bmatrix} C(\vec{\Theta}) \\ \gamma I \end{bmatrix}, \tag{11}$$

where *I* is unity matrix.

It is understand that influence of coefficient γ will depend on proportions between components $\left\| \vec{Y}^* - Y(\vec{U}^*,\ \vec{\Theta}) \right\|$ and $\|\Theta\|$ of the generalized functional $\Psi'(\Theta)$, hence on conditions of identification: number of measured parameters *m*, number of modes *K*, measuring errors *W* and number of estimated parameters *r*. The effect of regularization will become more depending on conditions of identification, if the regularization coefficient α is entered using the relation

$$\gamma = f(W, m, K, r)\alpha . \tag{12}$$

It is clear that the proposed identification procedure will save possibility to obtain the former non-regularized solution, if the regularization coefficient is zero.

As any new software, the regularized identification procedure need careful testing. Therefore, we performed two stages of testing, using the turbofan engine model: a) initial checking on model information without measuring noise; b) random testing on model information with measuring noise simulation.

Calculations were done for different values of α, which are varied in a range from 0 to 1; corresponding diagrams are shown in fig.1, which represent variation of estimations $\hat{\vec{\Theta}}$ and $\hat{\vec{Y}} = F(\vec{U}, \hat{\vec{\Theta}})$. The following conditions were used:

- – 12 design modes, the mode parameter is fuel flow;
- – simulated changes in the gas path: $\Theta_{1\ sim}=\Delta G_{HPC}=-0.03$ (high pressure compressor (HPC) performance deviation by a flow rate); $\Theta_{2\ sim}=\Delta\eta_{HPC}=-0.04$ (HPC performance deviation by efficiency);
- – 5 gas path parameters Y_i, which are initial data for identification: HPC discharge pressure, high pressure turbine (HPT) discharge temperature, low pressure turbine (LPT) discharge temperature, rotation speeds of both rotors;
- – 4 estimated parameters, including two above mentioned simulated parameters $\Theta_1=\Delta G_{HPC}$ and $\Theta_2=\Delta\eta_{HPC}$, and also $\Theta_3=\Delta A_{HPT}$ – deviation of HPT performance by a flow, $\Theta_4=\Delta\eta_{HPT}$ – deviation of HPT performance by efficiency.

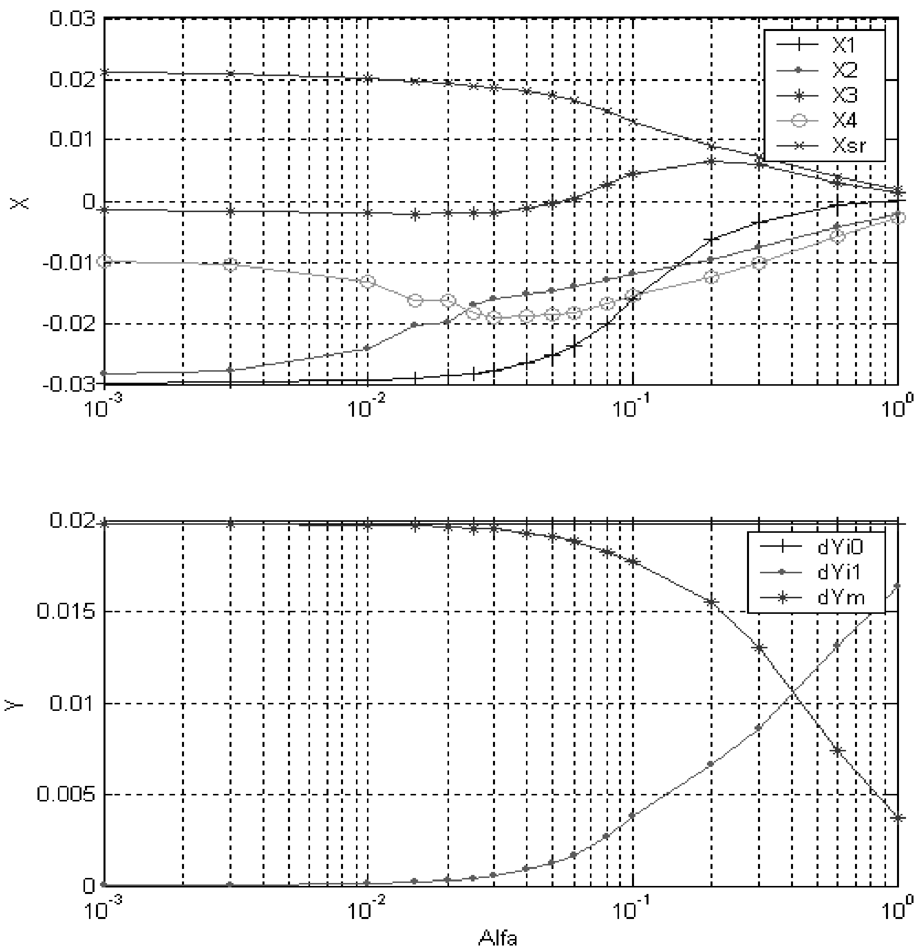

Fig.1 - Influence of the regularization coefficient on estimations when noise is absent

Fig. 1 shows that despite of initial growth of Θ_3 and Θ_4 absolute values, the average value Θ_{av} decreases monotonously and the residual between measurements and corrected model grows. This corresponds to theoretical concept about influence of regularization on the identification process. Diagrams in fig. 1 make possible estimating of the range of the regularization coefficient. If deviation of parameters $\vec{Y}$ must be no more than 5%, and deviation of parameters Θ – 10%, then the value of α must not exceed 0.03.

To obtain precise values of dispersions and to analyze character of estimations $\hat{\hat{\Theta}}$ distribution, cycle of random measuring errors setting with dispersions $\sigma^2_{Y_i}$ and identification was repeated 1000 times. That provided average precision about 1% of initial residuals for parameters Y and about 0.5% of simulated deviation 0.03 for estimated parameters $\hat{\hat{\Theta}}$.

Fig. 2 represents distributions of errors for calculations with two estimated parameters at $\gamma = 0$ and $\gamma = 0.04$. It is seen that as non-regularized as regularized estimations have distribution close to a normal one. Total scatter of estimations is also saved, but centers of regularized estimations displaced essentially from their true values on -0.03 and -0.04 respectively.

The estimations changed behavior in calculations with four unknown parameters. Estimations $\hat{\Theta}_1$ and $\hat{\Theta}_3$ saved the above mentioned properties of estimation with two unknown parameters; at same time estimations $\hat{\Theta}_2$ and $\hat{\Theta}_4$ have significant differences, which are illustrated in fig. 3. At low regularization, centers of

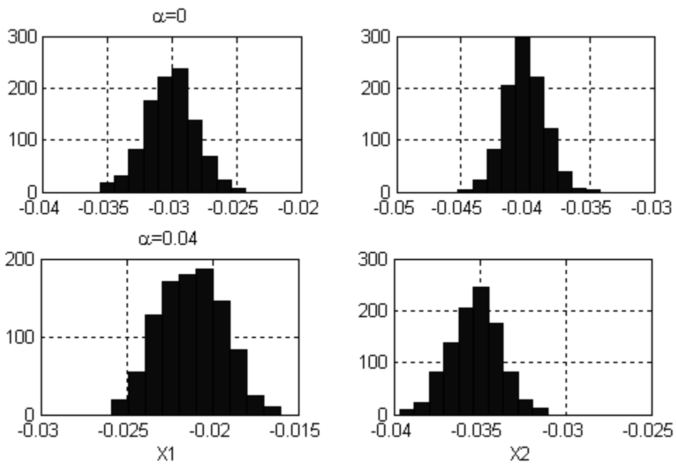

Fig. 2 - Distributions of estimations when two parameters are estimated

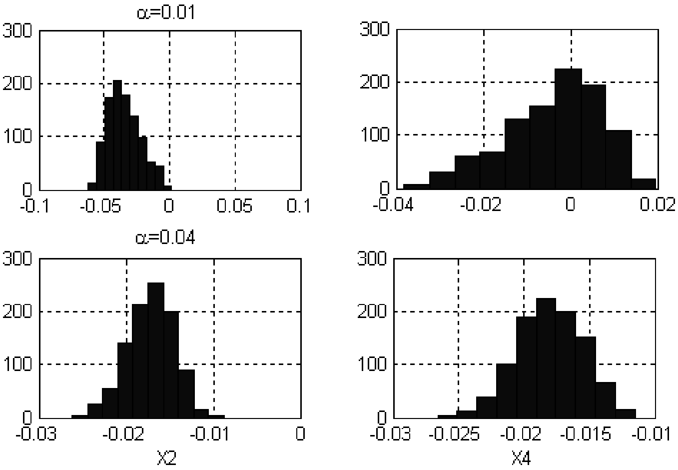

Fig. 3 - Distributions of estimations when four parameters are estimated

distributions are close to true values, but scatter is high and distributions are strongly non-symmetrical.

When α increases then the mean values displace, scatter decreases and distribution shape makes close to a normal one. Such abnormal behavior of parameters Θ_2 and Θ_4 is explained by their correlation.

REGULARIZED IDENTIFICATION PROCEDURES DEVELOPMENT USING A PRIORY INFORMATION

The above represented regularization procedure is formal. The values of the regularization coefficient are set appropriately; therefore, this procedure needs preliminary adjustment and results checking. The searching for the sources to improve the precision and stability led to an idea of implementation of a theoretical information about the engine, its parameters and performances. The main difficulty is in the diversity of this information, which is represented in one of the following forms:

- exact statement (for example, a part-load performance in a determined area is smooth);
- statement in a form of limitations of the area of acceptable solutions (for example, efficiency of individual compressor cannot differ for more than 3% from the efficiency of an “average” compressor, which performance is used in the initial model);
- statement in a form of fuzzy information (for example, the gas temperature in turbine will rather grow with the engine life);
- statistical form (for example probability density functions of parameters).

The next difficulty is in formalization of parameters that characterize the model quality. These parameters are set on a base of subjective preferences of decision makers (DM). The same difficulties appear at ranking of partial criteria and limitations according to their significance for the model quality estimation. Analysis showed that the main problem is that theoretical-probabilistic methods are hardly applicable for operation with uncertainties, which are related to subjective preferences that nature is not statistical. Actually, choice of the model’s structure is the DM procedure, which in multi-criteria case inevitably contains elements of subjectivity. If complexity of the task increases, the role of quality factors will grow. Therefore, there is possible to take into account all criteria using proper mathematical tool.

We propose to use the fuzzy sets theory as this tool. This theory makes the uniform base for description of information, which is given in all above listed forms, thus providing correct mathematical definition of the identification task.

This work includes examples of practical application of this approach to the engine models matching to experimental data. These models, which adequacy is provided at initial stage of maintenance, are used in further as reference models to determine normal values of parameters to be checked in current flight conditions. Deviations of the engine parameters to be checked from normal values are diagnostic characters, which are used in the engine technical state analysis.

CONCLUSIONS

Regularization of the engine's thermo-gas-dynamic model identification procedure will be useful, if initial (non-regularized) procedure is unstable or number of estimated parameters is high or estimations are strongly correlated.

However, if the regularized procedure is applied, one must be careful. The non-regularized procedure gives non-biased estimations. So, if the procedure is stable, we can repeat experiment and reach required precision. The bias, which is fed-in by the regularization, is difficult for practical determination. Therefore, the small regularization coefficients will be preferable, if the bias of estimations is negligibly small.

The optimum regularization coefficient must be selected and precision estimation must be executed prior to the regularization procedure (for example in automated diagnosing system).

Methods of the engine models identification are considerably developed for providing stable, precise and physically adequate solutions. This development is based on the usage of a-priory known information about the engine, its parameters and performances directly in the identification procedure. The mathematical base for this development is the fuzzy sets theory.

ACKNOWLEDGMENT

This work is performed with the support of the AERO-UA project, which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 724034.

REFEFENCES

[1] Tunakov, A.P., 1979. Optimization Methods in Turbine Engine Development and Designing. – Moscow: Mechanical Engineering Publishing.

[2] Simbirskyi, D.F., Yepifanov, S.V., Galchenko, A.L., 1979. Turbofan engine model Designing and Researching for Health Management. – Proc. of the scientific conference "Development Prospect of Aero Engineering Maintenance Methods", Kiev.

[3] Yepifanov, S.V., Simbirskyi, D.F., Kaplun, S.A., 1985. Precision Improvement of Turbomachine Mathematical Model Using Identification Methods. – Kharkov: – Proc. of the scientific conference "Simulation of Energetic Turbomachines Processes and Structures in Computer Aided Systems", Kiev.

[4] Stamatis, A., Mathioudakis, K., Papailiou, K.D., 1990. Adaptive Simulation of Gas Turbine Performance. – ASME Paper 90-GT-205, Journal of Gas Turbine and Power, Vol. 112, No. 2.

[5] Lambiris, B., Mathioudakis, K., Papailiou, K. D., 1991. Adaptive Modeling of Jet Engine Performance with Application to Condition Monitoring. – ISABE 91-7058.

[6] Kong, C. D., Ki, J. Y., 2003. A New Scaling Method for Component Maps of Gas Turbine Using System Identification. – ASME J. Eng. Gas Turbines Power, Vol. 125, No. 4.

[7] Li, Y. G., Pilidis, P., Newby, M., 2006. An Adaptation Approach for Gas Turbine Design-Point Performance Simulation. – ASME Transactions: Journal of Engineering for Gas Turbine and Power, Vol. 128, pp.789-795.

[8] Li, Y.G., Abdul Ghafir, M.F., Wang, L., Sing, R., Huang, K., Feng, X., 2010. Non–linear Multiple Points Gas Turbine Off-design Performance Adaptation Using a Genetic Algorithm. – ASME Paper GT2010-22285.

[9] Roth, B., Doel, D., Cissel, J., 2005. Probablistic Matching of Turbofan Engine Performance Model to Test Data. – ASME Paper 2005-68201.

[10] Litvinov, Yu, Borovick, V. Performances and Maintenance Properties of Aircraft Turbine Engines. – Moscow: Mechanical Engineering Publishing.

[11] Yepifanov, S., 1983. Turbine Engine Parameters Optimum Estimation on Experimental Data. – Proc. of the scientific conference “Gas Turbine and Combined Power Plants”. – Moscow.

[12] Yepifanov, S., Kuznetsov, B., Bogayenko, I. at al., 1998. Synthesis of Turbine Engine Control and Diagnosing Systems. – Kiev: Technical Publishing.

[13] Vapnik, V., 1979. Dependences Reconstruction by Empirical Data. – Moscow: Science Publishing.

[14] Volponi, A.J., Depold, H., Ganguli, H., Chen, D.G., 2000. The Use of Kalman Filter and Neural Network Methodologies in Gas Turbine Performance Diagnostics: A Comparative Study. – ASME Paper GT2000-547.

[15] Roth, B., Doel, D., Mavris, D., Beeson, D., 2003. High-accuracy Matching of Engine Performance Models to Test Data. – ASME Paper 2003-38784.

[16] Zadeh, L.A., 1965. Fuzzy sets. – Information and Control, Vol.8.

[17] Meniailov, I, Khustochka, A., Ugryumova, K., Chernysh, S., Yepifanov, S., Ugryumov, M., 2018. Mathematical Models and Methods of Effective Estimation in Multi-Objective Optimization Problems under Uncertainties. – Advances in Structural and Multidisciplinary Optimization: Proceedings of the 12th World Congress of Structural and Multidisciplinary Optimization (WCSMO12) (Braunschweig, Germany). – SpringerLink.