Simulation of gas-dynamic characteristics of a centrifugal compressor vane diffuser using neural networks

Aleksandr Nikiforov[1], *Alexei* Kuchumov[1], *Sergei* Terentev[1], *Evgeniy* Petukhov[2,*], and *Kirill* Kabalyk[3]

[1]Smolensk state agricultural academy, Bolshaja Sovetskaja 10/2, Smolensk, 214000 Russia
[2]Peter the Great St. Petersburg Polytechnic University, Polytechnicheskaya 29, Saint Petersburg, 195251 Russia
[3]Lodz University of Technology, Institute of Turbomachinery Wolczanska 219/223 Lodz, Poland 90-924

Abstract. The paper presents the results of mathematical simulation of the characteristics of a vane diffuser of a centrifugal compressor intermediate stage, such as the loss coefficient and the deviation angle versus the outlet vane angle of the diffuser. The simulation of these characteristics was made on the basis of processing the results of studies performed by the Research Laboratory "Gas Dynamics of Turbomachines" of Peter the Great St.Petersburg Polytechnic University at the model characteristics of vane diffusers. Given the almost complete absence of recommendations in the literature, the paper describes the technology for constructing neural network models, which includes preparing a sample of input data and determining the optimal structure of the neural network. Based on the obtained mathematical models, a computational experiment was carried out in order to determine the influence of the main geometric and gas-dynamic parameters on the efficiency of vane diffusers. The results of the computational experiment on neural models of the efficiency of a vane diffuser are analyzed according to the existing ideas about the physics of the processes of energy conversion in a vane diffuser.

1 Introduction

A small number of desired geometrical dimensions and a relatively simple design of the vane diffuser (VD) slightly reduce the problems of optimal design, because the main influence on the losses in the diffuser is provided by the impeller. The spatial complex structure of the viscous compressible flow behind the impeller creates a complex flow structure in the diffuser itself. Previously absent methods for calculating a viscous compressible flow left researchers with only one way out which were experimental studies. In [1 - 12], the results of experimental and theoretical studies of vane diffusers were analyzed and, based on their generalization, a number of recommendations for optimal design were given: the optimal nature of the velocity distribution along the pressure and suction sides of vanes were suggested; the influence of various laws of load distribution over the vane on the efficiency of diffusers was analyzed; some estimates of changes in the minimum values of the loss coefficients for various lattice densities were given; some analytical dependences were given for choosing the number of vanes according to the optimal solidity and to the average cone angle of the diffuser; the dependences of the loss and recovery coefficients on the solidity of the gratings at the optimum mode and the relative change in

the loss coefficient on the incidence angle were obtained; noise characteristics were determined, etc. But the results of these and of a number of other works related, as a rule, to a narrow spectrum of geometric and gas-dynamic characteristics of vane diffusers. Recently, a number of computational software systems have appeared that allow the calculation of a viscous compressible three-dimensional flow in the flow part of a centrifugal compressor stage, which made it possible to transfer a substantial part of the research to a virtual computer experiment [1, 13]. Such systematic and numerous studies of various elements and stages of a centrifugal compressor as a whole were carried out at the Research Laboratory Gas Dynamics of Turbomachines at the St. Petersburg Polytechnic University, under the guidance of Professor Yu.B. Galerkin. The Universal modeling method [14-20], developed in this laboratory and repeatedly tested in theoretical studies of the working processes of turbochargers and in the practical design of new machines made it possible to obtain an extensive database of energy characteristics of centrifugal compressors stages and of their elements. The created Universal Modeling Method and the corresponding software package suggest the use of a designer with large computing power and significant time spent in virtual research or finding the optimal design using variant calculations. These difficulties can be circumvented using the results of studies of the Gas

---
[*] Corresponding author: eugene@lamm.spbstu.ru

Dynamics of Turbomachines laboratory for constructing mathematical models suitable for simple practical use in design or in carrying out research of compressors. The present work is devoted to the method of constructing such models using the sample of a vane diffuser and discussion of the simulation results.

2 Methods

In order to achieve the aim of work, the most modern method of processing the input data, which is the neural modeling, was used. The main task is the methodology of training a neural network (NS) on a sample of input data. As separate stages of training, it is envisaged to provide analysis and selection of types of neural networks and activation functions; creation of several test neural networks of different architecture; analysis of main coefficients of input neurons; perturbation of input parameter values and analysis of the response of the neural network to these perturbations; sequential exclusion of input neurons and observation of a network generalization error; testing trial neural networks of different architectures; the choice of the ultimate neural network architecture. Also, one of the essential tasks for achieving the aim of the work is to develop methods and to implement them by preliminary preparing a sample of initial data for training a neural model of the energy characteristics of a vane diffuser, which includes determining the vector of output and input parameters of the model; identification of conflicting examples; determination of the required minimum sample size to create a neural network; sample conversion in order to improve the quality of training of neural networks with insufficient sample size (multiple cross-validation, multiple sampling and changing the order of training examples); identification of outliers in the input data; excluding outliers from the training set; rationing of input data; adding noise to training samples.

2.1 The aim of the work

The aim of the work is to build mathematical models of the gas-dynamic characteristics of the vane diffuser at the intermediate stage of a centrifugal compressor, which allow performing a computational study and searching for the optimal design of the VD for the given parameters in the process of developing new designs of centrifugal compressors.

2.2 Object under study

Vane diffusers, to a large extent, determine the overall dimensions and energy characteristics of the centrifugal compressor as a whole [1 - 7]. The use of vane diffusers in centrifugal compressors allows to obtain a greater deceleration of the gas flow and, accordingly, to reduce losses in the rotary elbow and the return guide apparatus. The coordination of the optimal modes of the impeller and of the vane diffuser is achieved due to the special installation of the diffuser vanes, which allows to increase the efficiency by 2-4% in the design mode with

a certain narrowing of the stage operation zone. Along with this, we know designs of stages with vaneless diffusers providing high efficiency in a wide area of work. On the other hand, the use of vane diffusers in low-flow stages did not lead to a significant narrowing of the stage operation zone. Thus, the choice of the type of diffuser and the determination of its geometric parameters requires special analysis.

The main structural elements of the vane diffuser are shown in Figure 1. As a rule, the height of the vanes is taken to be constant along the length of the diffuser $b_3 = b_4$ (where b is vane height, 3, 4 - indices of control sections), in order to simplify the design of the diffuser; the midline of the vane profile is made along an arc of a circle with radius R_v from the radius of the centers R_c (v – value for diffuser vane):

$$R_v = 0,5\ (r_4{}^2 - r_3{}^2)/(r_4\ cos\ \alpha_{v4} - r_3\ cos\ \alpha_{v3})\qquad(1)$$
$$R_c{}^2 = R_v{}^2 + r_4{}^2 - 2\ R_v r_4 cos\alpha_{v4}\,.\qquad(2)$$

In this case, the task of determining the optimal design of the vane diffuser is to search for the main structural dimensions: the radii of the input r_3 and of the outlet r_4 from the diffuser, the inlet vane angles α_{v3} and at the output α_{v4}, the number of vanes z and the height of the vanes b.

Figure 1 shows a diagram of the geometry of the investigated vane diffuser.
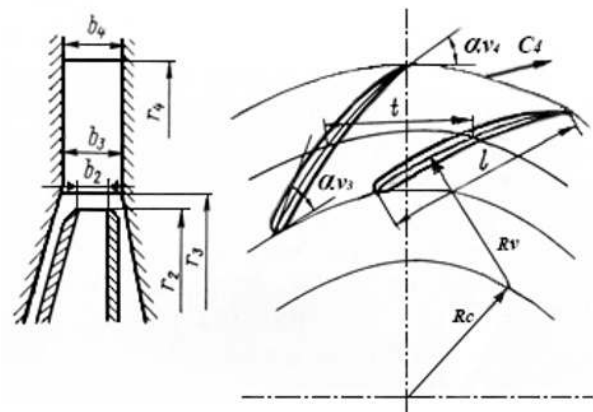

Fig.1. Scheme of the geometry of the vane diffuser of a centrifugal compressor.

3 Results

3.1 Processing a sample of initial data for training a neural network of the characteristics of a vane diffuser

Neural networks, as a universal approximator, allow us building generalized models based on a large data amount. The main provisions, features and advantages of the neural network approach when modeling the characteristics of centrifugal compressors are given in [21]. In a simplified form, it can be shown that the neural network performs the following approximation:

$$Y=f(X),\qquad(3)$$

where *X* is the input vector, *Y* is the output vector, *f* is the transformation performed by the neural network.

Practice and analysis of the use of neural networks for simulation the characteristics of centrifugal compressors allow us to conclude that the analysis, selection and preliminary preparation and processing of input data (formation of a training sample) before training neural networks can significantly improve the accuracy and reliability of models.

In the general case, for processing a sample of input data, we can use the following sequence of stages, which was formed empirically by the authors of this paper, using their own experience in creating neural models [21-24]:

1. Selection of input parameters (logic and analysis of the subject area; analysis of the main coefficients of the input neurons; perturbation of the values of the input parameters and analysis of the response of the network to these perturbations; sequential exclusion of input neurons and observation of the network generalization error).

2. Identification of conflicting samples.

3. Determination of the required number of samples.

4. Improving the quality of training of neural networks with insufficient sample size (multiple cross-validation, multiple sample repetition and changing the order of the training samples).

5. Identification of outliers.

6. Normalization of data.

7. Adding noise to training samples.

8. Selection of types of neural networks and of activation functions.

9. Network decomposition according to the number of output neurons.

Let us dwell in more detail on the points completed in the framework of this work.

3.2 Sampling of initial data and determination of the vector of output and input parameters of the model

The initial total amount of the input data sample describing the VD geometry and gas-dynamic parameters was of 603 samples.

Using the research results provided by the research laboratory "Gas Dynamics of Turbomachines" of the St. Petersburg Polytechnic University, the following parameters were accepted as input ones (arguments of the neural model):

b_3/D_3 is relative width of the diffuser (only the input data of diffusers of constant width were used); l/t is cascade solidity (where l – vane length, t – vane installation stage); α_{v3} is inlet vane angle of VD; $\Delta\alpha_v$ is vane blade-camber angle; D_4/D_3 is relative vaneless diffuser exit diameter; α_4 is outlet VD flow angle; i_3 is incidence angle.

Neural models were built for two parameters:

ζ is VD loss coefficient;

$\Delta\alpha_4$ is deviation angle at the VD outlet.

The ranges of changes in geometric and operational parameters in the sample of input data for constructing a mathematical model are shown in table 1.

Table 1. The ranges of input and output parameters in the original data sample.

Parameter	b_3/D_3	l/t	α_{v3}	$\Delta\alpha_v$	D_4/D_3	i_3	α_4	ζ	$\Delta\alpha_4$
Minimum value	0.0227	1.273	15°	10°	1.3636	-8°	18.651°	0.075	-1.611°
Maximum value	0.0727	2.449	20°	30°	1.3636	6°	38.215°	0.581	22.564°

In order to analyze the sample, a frequency analysis was performed for each of the parameters. In an ideal sample of input data, the input parameters should be evenly distributed in the studied range of their changes with high density. In reality, due to the reduction in the cost of the experiment, this, as a rule, cannot be achieved. Frequency analysis of the sample allows us to clearly see in the diagrams which ranges of the changed or other input parameters are covered most fully by the values, and which areas in the sample were represented to a lesser extent. This analysis allows us to describe the domain of definition of the desired models.

The outlet flow angle α_4 is excluded as an insignificant argument in the model, because actually duplicated by the deviation angle at the VD outlet versus the outlet vane angle of the diffuser $\Delta\alpha_4$. The relative diameter of the exit from VD4 D_4/D_3 is excluded from the list of arguments, because the input data sample for modeling was obtained only for vane diffusers of the same length $D_4/D_3 = 1.3636$.

Thus, in a generalized form, the desired models were presented as follows:

$$\zeta = f(b_3/D_3; \alpha_{v3}; \Delta\alpha_v; l/t; i_3); \qquad (4)$$

$$\Delta\alpha_4 = f(b_3/D_3; \alpha_{v3}; \Delta\alpha_v; l/t; i_3). \qquad (5)$$

3.3 Identifying conflicting samples

After identifying and eliminating insignificant parameters from the training set, the quality and accuracy of the neural network model, as a rule, improves, due to a decrease in its dimension and, therefore, simplification of the mathematical dependence. It should be noted here that an excessive decrease in the number of input parameters and simplification of the appearance of a neural network may interfere with revealing patterns in a particular problem. It can also lead to the emergence of conflicting samples.

Samples are called conflicting if, for identical input vectors, they have different output ones. The presence of conflicting samples in the training set may appear due to random errors or in the case of incorrect statement of the problem. The search for such conflicting samples can be carried out using special algorithms that search for matching training samples [24, 31] or by carefully looking at the sample in search of repeated input vectors.

When processing the input sample, several conflicting samples were found, two of them are shown in Table 2 as an illustration:

Table 2. Conflicting samples found in the described sample.

Parameter	b_3/D_3	l/t	α_{v3}	$\Delta\alpha_v$	i_3	ζ	$\Delta\alpha_4$
Sample 1	0.0545	1.379	20°	15°	-2°	0.075	-
Sample 1'	0.0545	1.379	20°	15°	-2°	0.076	-
Sample 2	0.0227	1.861	20°	20°	4°	-	12.73°
Sample 2'	0.0227	1.861	20°	20°	4°	-	12.47°

All identified conflicting samples were removed from the training set. Thus, for further work, 591 training samples were used.

3.4 Determination of the required number of samples for creating a neural network

For successful simulation using neural networks, it is important to dispose of the required number of training samples. In part, the concept "the more the better" is true, but it is important to remember that the number of samples affects the training time and an excessive number of samples will lead to a large expenditure of computer time for training the neural network. In [25], a formula is given with which you can determine the minimum required amount of training samples:

$$Q=7*N_x+15, \qquad (6)$$

where N_x is number of input parameters of the neural network model; Q is number of training set samples.

The number of input parameters in neural models of the VD loss coefficient ζ and the deviation angle $\Delta\alpha_4$ under the study are the same and equal to 5. Then, provided the samples are distributed as evenly as possible over the entire sampling range, the minimum training sample size should be of 50. The sample size in 591 samples repeatedly covers this value. It should be specially noted here that this quantitative well-being does not provide sufficient grounds for reflecting the desired physical laws in the model, because the second defining moment in the process of setting the problem is the uniformity and density of distribution of the model arguments in the input data sample.

3.5 Sampling transformations to improve the quality of training of neural networks with insufficient sample size

In practice, it is often not possible to collect a sufficient amount of data distributed uniformly with a high density for training, and there is a need for more rigorous testing of the neural network or for certain manipulations with the sample. So, in the problem under consideration, with a relatively favorable distribution density of input parameters in the domain of definition of the desired function, only two values of the vane angles at the entrance to the VD are used in order to construct the model.

As one of the possible ways, you can resort to multiple cross-validation (multifold cross-validation). Also, the problem of a small sample can be dealt with by repeating the original sample supplied to the input of a neural network [25, 26]. In this case, the method of

changing the sequence of training samples was applied. This gives the learning course a more stochastic character and helps to reduce the likelihood of getting into local extremes.

The sample was shuffled several times in a random order using the RAND function integrated in the Excel MSOffice spreadsheet editor in order to eliminate the influence of the order of input of samples during training.

3.6 Identification and exclusion of outliers from the training set

For outliers in the input sample, parameter values are taken that, due to random reasons or due to a simple human factor, significantly differ in their values from other similar information. Emissions may appear during the data collection (a comma is absent when entering the data into the computer and as a result an error in the value arises by an order of magnitude) or may be due to other reasons (errors of measuring instruments, malfunctions of the equipment, etc.). Obviously, this does not reflect the physical laws of the influence of parameters in the described subject area.

The presence of outliers adversely affects the accuracy of the created models, as in the case of conflicting samples, the error inherent in the sample before the training stage is difficult to correct by changing the training algorithms. In the simplest cases, outliers can be detected by careful viewing of the sample, in more complex multi-parameter dependencies and large amounts of information for searching for outliers, we can resort to using simple neural networks with a minimum number of neurons in the hidden layer. For example, the perceptron is poorly trained in samples with outliers and, by alternately removing samples and comparing the resulting errors, we can find out those samples that are outliers. In the case of a large sample size, this is difficult to implement; therefore, they resort to various emission search algorithms [27].

In the course of working with the input data, several samples with outliers were found. In order not to reduce the sample size, the detected outliers were corrected in accordance with the general type of dependence.
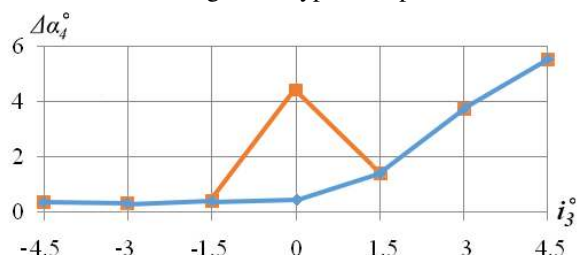

Fig.2. An example of the deviation angle versus the incidence angle with ejection and after correction.

3.7 Rationing of input data

It is desirable to normalize data prepared for neural network processing by aligning the range of variation of the values of the quantities, limiting them to the interval

(for example, [0,1]). The normalization procedure should be subjected to both input and output parameters. The normalization process is described in detail in [25, 34-36].

In this case, data normalization was especially relevant due to the physical nature of data and was carried out for the range [0,1], where 0 corresponds to the minimum and 1 to the maximum value of both the input and the output parameters of the neural network.

After completing all the procedures above for preparing a sample of input data for training a neural network, 591 training samples remained for building the model. It was decided not to introduce noise so as not to distort the simulation results, since the sample size practically coincides with the minimum threshold data volume for introducing noise into the training sample.

Of the total sample size in 591 training samples, 60 samples (10%) were allocated to form a test set that was not involved in the training of the desired models. The test set was used only after training in order to determine the error of the models.

3.8 Neural network training

Depending on which task you need to solve using neural network modeling, you need to select the appropriate type of neural network. So for approximation of functions, perceptron-type neural networks are well suited, Kohonen networks are often used for clustering tasks, and convolutional neural networks are often used for image recognition and classification [29-31].

In this work, in order to create models, two-layer perceptrons were chosen, as the form of neural networks that has proven itself in the approximation of multidimensional functions.

The experience of neural network modeling showed that asymmetric, differentiable functions (for example, such as logical sigmoid, hyperbolic tangent) are efficiently used as activation functions of neurons [29, 30].

For the neural network model of the loss coefficient of the vane diffuser ζ, the logical sigmoid in both layers were taken as activation functions. In order to simulate the deviation angle $\Delta\alpha_4$, the logical sigmoid was used in the first layer and the linear activation function in the second one.

For complex neural network models, where the value of several output parameters is approximated, it is convenient to use network decomposition according to the number of output neurons. So, for example, instead of a neural network with 5 inputs and 3 outputs, create 3 neural networks with 5 inputs and 1 output. The use of this technique allows to reduce the overall error in simulation the output parameters. Each neuron will adjust its weight in accordance with the reduction of error for simulation of one output parameter, and not adjust to several parameters of the output vector at once. The structure of each individual network should be optimized separately, taking into account the minimization of error for each neural network.

That is why in the modeling process in this work two models were constructed – ζ (VD loss coefficient) and $\Delta\alpha_4$ (deviation angle). For each of them, in order to better approximate its dependence in neural networks, different types of activation functions were used in the output models layers.

The selection of suitable network architecture is carried out empirically. Despite the number of existing recommendations on network architecture for solving the problem of approximating dependencies, it is necessary to independently select the architectures that are most suitable for each specific task and the source data used.

In order to select the learning function, two-layer neural networks with 25 neurons in the first (hidden) layer were created. The simulation accuracy was checked when changing the training functions as follows: with the BFGS quasi-Newtonian method, with the Levenberg-Marquardt optimization method and Levenberg-Marquardt optimization with Bayesian regularization. In the first layer, an activation function, a logical sigmoid, was used. In the second layer, neural networks with a logical sigmoid and a linear activation function for each learning function were used.

In accordance with the minimum error, a training function with the Levenberg-Marquardt optimization method with Bayesian regularization was chosen.

Next, a computational experiment was conducted to select the network architecture for the selected learning function. Two-layer models were constructed with the number of neurons in the hidden layer 10, 15, 20, 25, 30 and one neuron in the output layer. As well as three-layer ones with the number of neurons in the first (hidden) layer 10, 15, 20, 25, 30 and in the second (hidden) layer 10, 15, 20, 25, 30 for each variant of the number of neurons in the first layer. There is one neuron in the third output layer. As activation function for all layers a logical sigmoid was used.

It also follows from the Arnold – Kolmogorov – Hecht-Nielsen theorems (Arnold-Kolmogorov-Hecht-Nielsen) [25] that for construction of a neural network model of an arbitrarily complex function, it suffices to use a perceptron with one hidden layer of sigmoid neurons, the number of which is determined by the ratio:

$$\frac{N_y Q}{1 + log_2(Q)} \le N_w \le N_y \left(\frac{Q}{N_x} + 1\right)\left(N_x + N_y + 1\right) + N_y, \ (7)$$

where N_x is number of neurons in the input layer (number of parameters); N_y is number of neurons in the output layer (the number of simulated quantities); Q is number of elements in the set of training samples; N_w is required number of synaptic connections.

In accordance with the initial data for modeling and the range of numerical and synaptic connections for the models of the loss coefficient ζ and for the model of the deviation angle, $\Delta\alpha_4$ is in the range from 58 to 797. This allows us to determine the required number of neurons in hidden layers, so for a two-layer perceptron the number of hidden layer neurons will be [25]:

$$N = \frac{N_w}{N_x + N_y} \qquad (8)$$

For the problem under consideration, the calculation according to the formulas above shows that the optimal number of neurons in the hidden layer for the two-layer perceptron of the model of loss coefficient ζ and deviation angle $\Delta\alpha_4$ lies in the range from 8 to 114 neurons ($8 \leq N \leq 114$).

A rigorous theory of choosing the optimal number of hidden layers and neurons in hidden layers does not currently exist. In the current practice of neural network modeling, perceptrons with one or two hidden layers are most often used, and the number of neurons in hidden layers usually ranges from $N_x/2$ to $3N_x$. A computational experiment on neural networks in the process of constructing models showed that the minimum error was obtained for a network with 30 neurons in a hidden layer, both for models of the loss coefficient ζ and the flow deviation angle $\Delta\alpha_4$. This value is consistent with the recommendations of theoretical studies.

The smallest error in the loss coefficient ζ was obtained for a neural network with the following architecture: two-layer, the number of neurons in the hidden layer is of 30, logical sigmoid activation functions for all layers, learning function with optimization according to the Levenberg-Marquardt algorithm with Bayesian regularization. The average error of the neural model of the loss coefficient ζ was of 5.5%.

The smallest error in the deviation angle $\Delta\alpha_4$ was observed in a neural network with the following architecture: two-layer, the number of neurons in the hidden layer is of 30, logical sigmoid activation functions for the first layer, linear activation function for the second layer, learning function with optimization according to the Levenberg-Marquardt algorithm with Bayesian regularization. The average error of the neural model of the deviation angle $\Delta\alpha_4$ was of 6.4%.

Examples of comparing simulation results for neural models with the input data are shown in Figures 3 and 4.
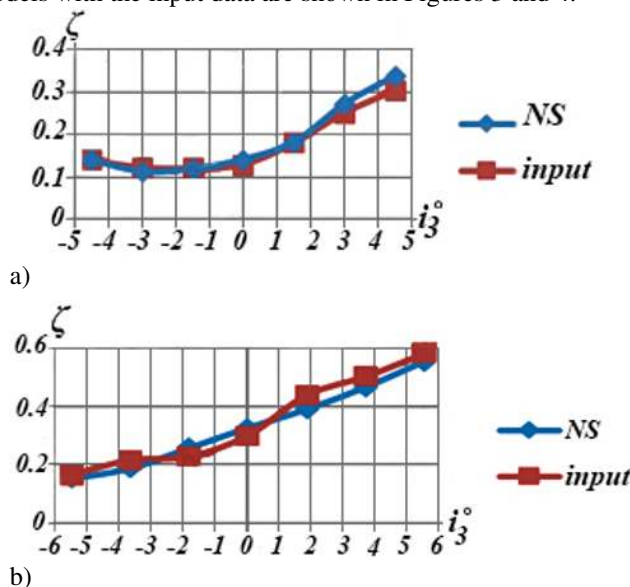

a)

b)

Fig.3. Comparison of simulation results ζ for VD: a) b_3/D_3=0.0545; l/t=1,861; α_{v3}=15°; $\Delta\alpha_v$=10°;
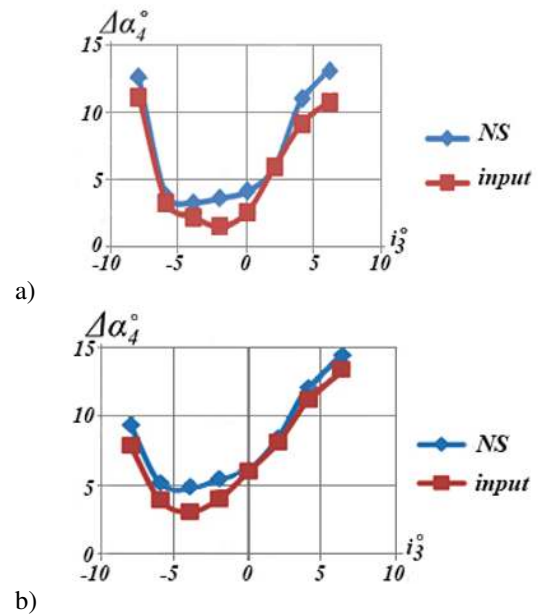b) b_3/D_3=0.0545; l/t=1.763; α_{v3}=15°; $\Delta\alpha_v$=30°.

a)

b)

Fig.4. Comparison of simulation results $\Delta\alpha_4$ for VD: a) b_3/D_3=0.0727; l/t=1,909; α_{v3}=20°; $\Delta\alpha_v$=15°;
b) b_3/D_3=0.0409; l/t=1.909; α_{v3}=20°; $\Delta\alpha_v$=15°.

The results of constructing neural models of the loss coefficient and the deviation angle versus the outlet vane angle of the diffuser show that the models describe well the nature of the change in the studied parameters in the entire domain of definition of models with an accuracy that is satisfactory for the practical use of models.

3.9 Verification of neural network models of loss coefficient and lag angle

In order to verify the conformity of the current understanding of gas-dynamic processes of energy conversion in compressor diffusers with the results of calculations on the obtained models, a computational experiment was carried out to study the laws governing the influence of the VD loss coefficient ζ and the deviation angle versus the outlet vane angle of the diffuser $\Delta\alpha_4$:

- the incidence angle i_3 for various values of the relative width b_3/D_3 , the cascade solidity l/t , the inlet vane angles α_{v3}, the vane blade-camber angle $\Delta\alpha_v$;
- the cascade solidity l/t at various values of the relative width b_3/D_3, the inlet vane angles α_{v3}, the angle of curvature of the vane profile $\Delta\alpha_v$ and incidence angle i_3;
- the relative width of the diffuser b_3/D_3 and at various values of the cascade solidity l/t, the inlet vane angles α_{v3}, the vane blade-camber angle $\Delta\alpha_v$, the incidence angle i_3;
- the vane blade-camber angle $\Delta\alpha_v$ at various values of the relative width b_3/D_3, the cascade solidity l/t, the inlet vane angles α_{v3} and the incidence angle i_3.

When conducting a computational experiment, the ranges of deviation of the listed parameters corresponded to the domain of definition of neural network models shown in table 1.

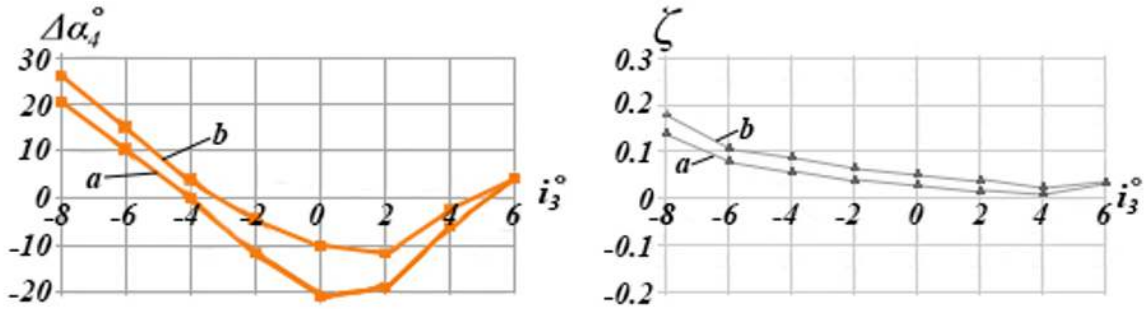Typical examples of the results of a computational experiment are shown in Figures 5–13.

Fig.5. Dependence of the flow delay angle $\Delta\alpha_4$ and the loss coefficient ζ
on the angle of attack i_3 for a BD with the following parameters:
a) b_3/D_3=0.0309; l/t=1.861; α_{v3}=15°; $\Delta\alpha_v$=10°; D_4/D_3=1.3636;
b) b_3/D_3=0.0409; l/t=2.3; α_{v3}=15°; $\Delta\alpha_v$=15°; D_4/D_3=1.3636.

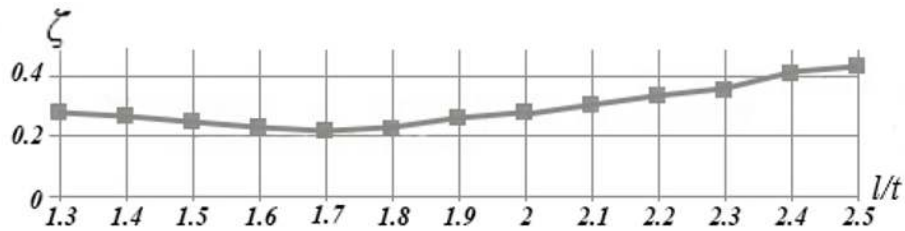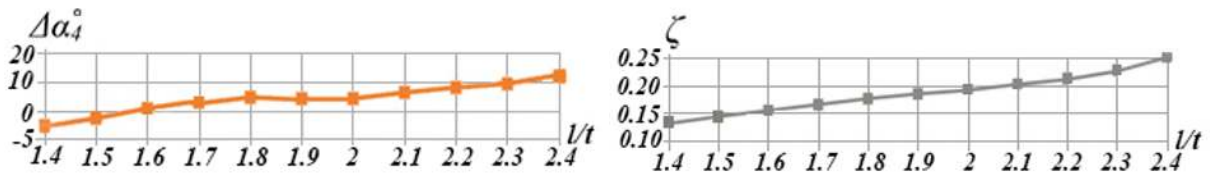**Fig.6.** Dependence of the loss coefficient ζ on the lattice density l/t for a BD with the following parameters:
b_3/D_3=0.0545; α_{v3}=15°; $\Delta\alpha_v$=10°; i_3=-8°; D_4/D_3=1.3636
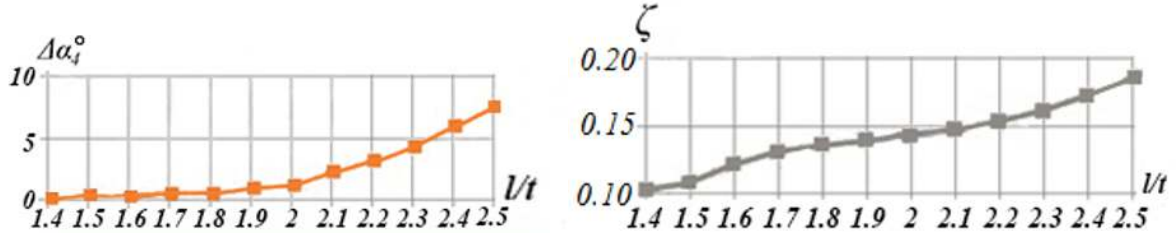

Fig.7. Dependences of the flow delay angle $\Delta\alpha_4$ and the loss coefficient ζ on the lattice density l/t
for a BD with the following parameters: b_3/D_3=0.0309; α_{v3}=20°; $\Delta\alpha_v$=20°; D_4/D_3=1.3636; i_3=-4°.

Fig.8. Dependences of the flow delay angle $\Delta\alpha_4$ and the loss coefficient ζ on the lattice density l/t
for a BD with the following parameters: b_3/D_3=0.0409; α_{v3}=20°; $\Delta\alpha_v$=15°; D_4/D_3=1.3636; i_3=-4°.
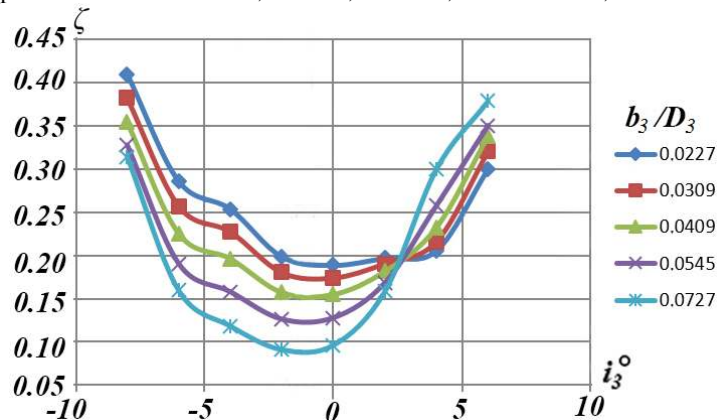

Fig.9. Influence of diffuser width b_3/D_3 on the dependence of the loss coefficient ζ
on the angle of attack i_3 for a BD with the following parameters: α_{v3}=15°; $\Delta\alpha_v$=10°; D_4/D_3=1.3636; l/t=2.291.
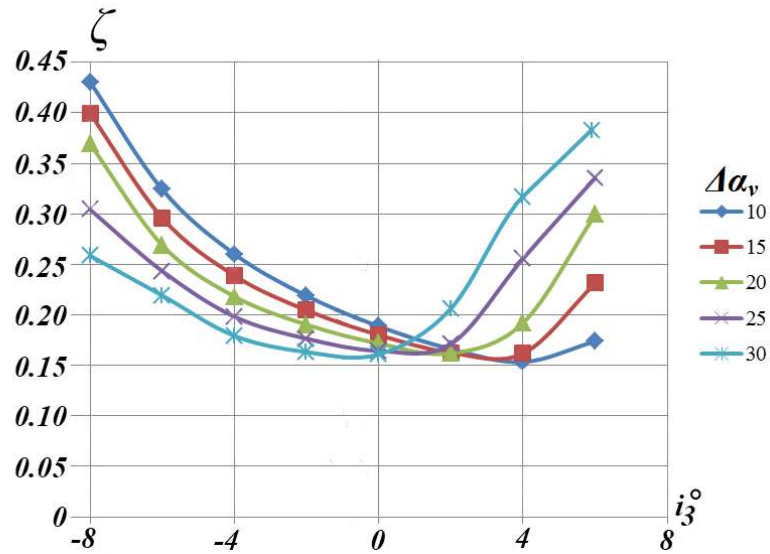
Fig.10. Influence of the bent angle of the diffuser blade $\Delta\alpha_v$ on the dependence of the loss coefficient ζ on the angle of attack i_3 for a BD with the following parameters: b_3/D_3=0.0309; α_{v3}=20°; D_4/D_3=1.3636; l/t=2.449.

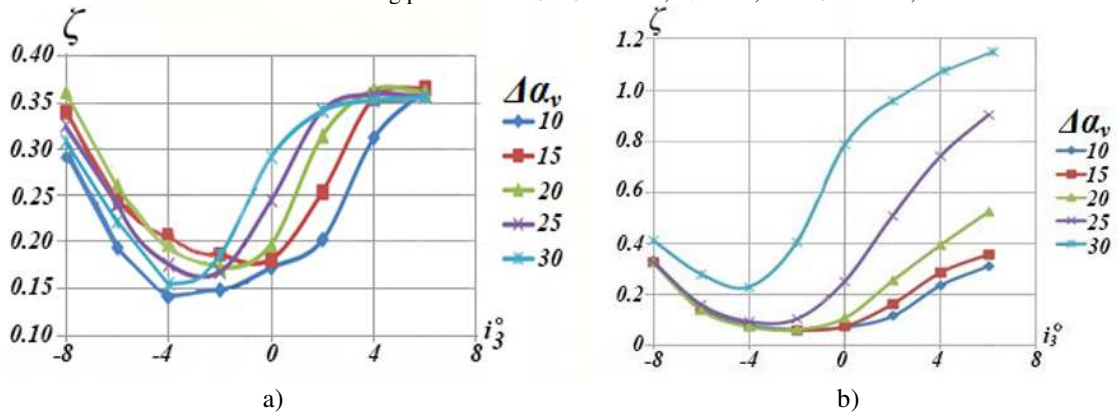a)   b)

Fig.11. Influence of the bent angle of the diffuser blade $\Delta\alpha_v$ on the dependence of the loss coefficient ζ on the angle of attack i_3 for a BD with parameters:
a) b_3/D_3=0.0409; α_{v3}=20°; D_4/D_3=1.3636; l/t=1.92;
b) b_3/D_3=0.0727; α_{v3}=20°; D_4/D_3=1.3636; l/t=1.763.
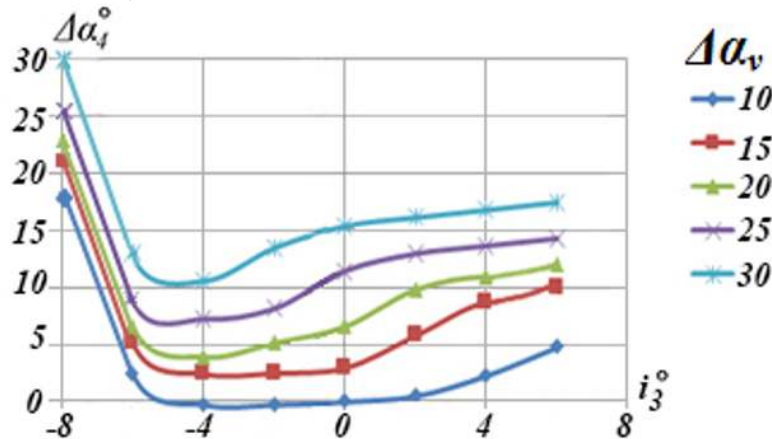

Fig.12. Influence of the bent angle of the diffuser blade $\Delta\alpha_v$ on the dependence of the flow delay angle $\Delta\alpha_4$ on the angle of attack i_3 for a BD with the following parameters: b_3/D_3=0.0227; α_{v3}=15°; D_4/D_3=1.3636; l/t=1.591
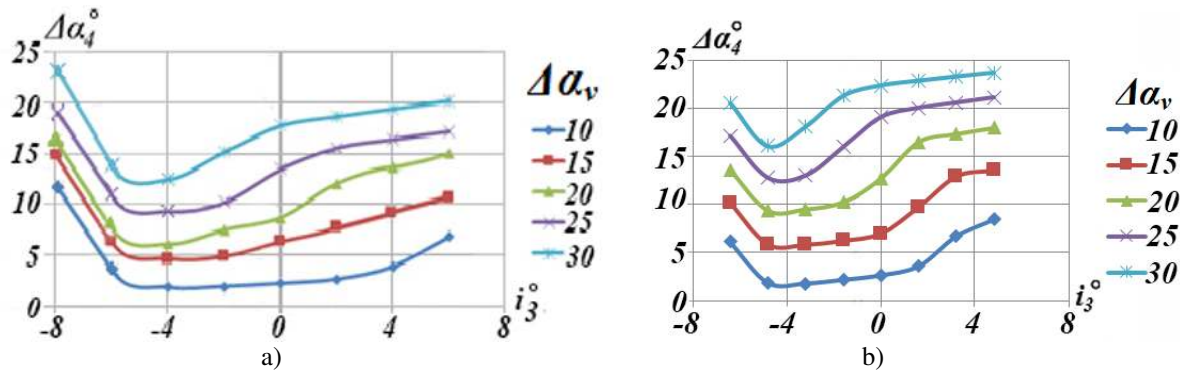
Fig.13. Influence of the bent angle of the diffuser blade $\Delta\alpha_v$ on the dependence of the flow delay angle $\Delta\alpha_4$ on the angle of attack i_3 for an BD with the following parameters:

a) $b_3/D_3=0.0409$; $\alpha_{v3}=15°$; $D_4/D_3=1.3636$; $l/t=1.92$;

b) $b_3/D_3=0.0727$; $\alpha_{v3}=20°$; $D_4/D_3=1.3636$; $l/t=1.763$.

Numerous results obtained during a computational experiment on neural models require special reflection and careful analysis. Here we only note that the results of a virtual experiment do not contradict the existing knowledge about the processes of energy conversion and with sufficient accuracy for engineering calculations coincide with the known quantitative data and recommendations for the design of vane diffusers.

4 Discussion

A number of technological conclusions can be noted that were practically obtained in the search for the optimal structure of neural networks. Obviously, these conclusions can be useful to readers only as conclusions from the results of the above study and cannot be in the nature of comprehensive recommendations:

1. We used as the resulting neural networks with fewer neurons in the hidden layer, which better adapt to the general form of the dependence and respond less to data noise and random outliers that were not detected at the preliminary stage of data processing.

2. The choice in favor of two-layer networks was made due to the insignificant difference in the errors of the two-layer and three-layer neural networks, but it is obviously more economical from the point of view of computing power for structures on two perceptron layers when using the obtained models to carry out research work or construct optimal VD designs.

3. The decrease in the number of neurons in the hidden layer significantly reduces the accuracy of the simulation. So, for the loss coefficient model, the ζ neural network with a similar final architecture: two-layer, logical sigmoid activation functions for all layers, a learning function with optimization according to the Levenberg-Marquardt algorithm, but with the number of neurons in the hidden layer 10, gave an average error of about 9%. A similar result was observed for the model of the deviation angle $\Delta\alpha_4$ (two-layer, the number of neurons in the hidden layer 10, the activation functions of the logical sigmoid for the first layer, the linear activation function for the second layer, the learning function with optimization according to the Levenberg-

Marquardt algorithm) with the number of neurons in the hidden layer of 10, the average error was about 12%.

4. A check of the influence of normalization was carried out when comparing a neural network trained on normalized data and a neural network trained on standardized data, showed a significant increase in accuracy. After normalization, the error decreases on average by 3% compared with the same model, which was only trained on abnormal data. Similar results were obtained in [24].

5. Pre-processing of a sample of input data is a necessary stage for training a neural network, as it can significantly reduce modeling errors.

The developed mathematical model can be applied in calculation programs of gas-dynamic characteristics of centrifugal compressors and compressor stages. The authors have a positive experience of cooperation with the developers of the universal modeling method programs [37, 38].

It should be specially emphasized that the time spent on research on neural models is much less than when using the Universal Modeling Method and is incomparably less than the cost of performing a full-scale experiment.

The results of a computational experiment allow us to conclude about the possible use of the obtained models in practical applications, both for scientific research and in the development of new, more advanced designs of centrifugal compressors with vane diffusers in the stage.

5 Conclusions

This paper summarizes the experience gained in pre-processing data for training neural networks in constructing mathematical models of the energy characteristics of vane diffusers of centrifugal compressors and offers recommendations for improving the accuracy of neural network modeling. The recommendations are framed in a single algorithm, consisting of a sequence of stages of processing the initial sample. The suggested algorithm was tested by simulation of the energy characteristics of vane diffusers of the intermediate stage of a centrifugal compressor.

The obtained simulation results and a significant reduction in the errors of neural network models show the importance and need for preliminary processing of the training sample.

A neural network simulation of the dependences of the deviation angle and the vane diffuser loss coefficient on the incidence angle was obtained with an average error of about 6%.

A computational experiment on the basis of the obtained models allows us to draw conclusions about a satisfactory agreement between the obtained results and the existing ideas about the gas-dynamic processes of energy conversion in a vane diffuser. This gives researchers and designers a new tool for the study of vane diffusers and the search for new, more advanced designs of the centrifugal compressor stage.

References

1. Yu.B. Galerkin, *Turbochargers* (2010)

2. K.P. Seleznev, Yu.B. Galerkin, *Centrifugal compressors* (1982)

3. K.P. Seleznev, Yu.B. Galerkin, *Theory and design of turbochargers* (1986)

4. A.F. Rekstin, K.V. Soldatova, Yu.B. Galerkin, International Conference on Compressors and their Systems, **604**, doi:10.1088/1757-899X/604/1/012045 (2019)

5. Yu.B. Galerkin, V.B. Semenovskiy, K.V. Soldatova, Oil and Gas Engineering **2141**, https://doi.org/10.1063/1.5122076 (2019)

6. Y. Galerkin, A. Drozdov, O. Solovyeva, Proceedings of 13th European Conference on Turbomachinery Fluid dynamics & Thermodynamics **329** (2019)

7. V.F. Ris, Centrifugal compressor machines (1981)

8. A.T. Lunev, R.F. Murtazin, Yu.V. Drozdov, I.F. Khusnutdinov, Compressor Technique and pneumatics **6**, 9-11 (2017)

9. A.M. Akhmetzyanov, A.T. Lunev, D.V. Pashinkin, Gas turbine technologies **7 (134)**, 6-7 (2015)

10. D. Japikse, JANNAF Conference (1998)

11. PCA Engineers Limited, http://www.pcaeng.co.uk/software/

12. PCA Engineers Limited, https://www.pcaeng.co.uk/apps

13. Customer documentation for ANSYS Release 17.2 - © SAS IP, Inc.

14. A.F. Rekstin, A.A. Drozdov, O.A. Solovyeva, Y.B. Galerkin, Oil and Gas Engineering, https://doi.org/10.1063/1.5051896 (2018)

15. Yu.B. Galerkin, A.F. Rekstin, K.V. Soldatova, A.A. Drozdov, Oil and Gas Engineering **2141** https://doi.org/10.1063/1.5122068 (2019)

16. Y. Galerkin, K. Soldatova, A. Drozdov, 10th International Conference on Computer Modeling and Simulation, 187-191 (2018)

17. Y. Galerkin, A. Drozdov, A. Rekstin, K. Soldatova, Oil and Gas Engineering **2007**, https://doi.org/10.1063/1.5051898 (2018)

18. A. Rekstin, Y. Popova, A. Ucehovscy, Oil and Gas Engineering **2007**, https://doi.org/10.1063/1.505189 7 (2018)

19. A. Drozdov, A. Rekstin, International Scientific Conference on Energy, Environmental and Construction Engineering **245**, DOI 10.1051/matecconf/201824504004 (2018)

20. Yu.B. Galerkin, A.F. Rekstin, O.A. Solovyeva, Vaneless diffuser of the centrifugal compressor stage design method, Oil and Gas Engineering **2141**, https://doi.org/10.1063/1.5122057 (2019)

21. A.G. Nikiforov, D.Yu. Popova. K.V. Soldatova, O.A. Solovyeva, Compressor Technique and pneumatics **3**, 18-21 (2015)

22. A.G. Nikiforov, D.Yu. Popova. K.V. Soldatova, O.A. Solovyeva, Compressor Technique and pneumatics **4**, 14-16 (2015)

23. A.G. Nikiforov, D.Yu. Popova. K.V. Soldatova, O.A. Solovyeva, Compressor Technique and pneumatics **6**, 30-33 (2015)

24. A.G. Nikiforov, D.Yu. Avramenko, Scientific and Technical Bulletin of SPbPU. Natural and engineering sciences **24** (**4**), 61-71. DOI: 10.18721/JEST.24406 (2018)

25. L.N. Yasnitskii, *Intellectual systems* (2016)

26. M.V. Dagaeva, M.A. Suleimanov, D.V. Kataseva, A.S. Katasev, A.P. Kirpichnikov, Bulletin of Technological University **21** (**2**), 133-138 (2018)

27. V.P. Shkodyrev, K.I. Yagafarov, V.A. Bashtovenko, E.E. Ilyina, Proc. of the Second Conference on Software Engineering and Information Management (2017)

28. S. Osovskiy, *Neural networks for information processing* (2002)

29. V.V. Kruglov, M.I. Dli, R.Yu. Golunov, *Fuzzy logic and neural networks* (2001)

30. V.V. Borisov, V.V. Kruglov, A.S. Fedulov, *Fuzzy models and networks, 2nd ed.* (2016)

31. V.S. Medvedev, V.G. Potemkin, *Neural networks. Matlab 6* (2002)

32. V.V. Borisov, A.S. Fedulov, M.M. Zernov, *Basics of fuzzy models hybridization* (2017)

33. T. Jayalakshmi, Dr.A. Santhakumaran International Journal of Computer Theory and Engineering **3**(**1**)**,** 89-93 (2011)

34. E.S. Ogasawara, L.C. Martinez, D. Oliveira, G. Zimbrão, G.L. Pappa, M. Mattoso The 2010 International Joint Conference on Neural Networks (*IJCNN*), DOI: 10.1109/IJCNN.2010.5596746 (2010)

35. S. C. Nayak, B. B. Misra and H. S. Behera, International Journal of Computer Information Systems and Industrial Management Applications **6**, 257-269 (2014)

36. K. Levenberg Quart. Appl. Math, **2**, 164—168 (1944)

37. A. Nikiforov, D. Popova, K. Soldatova International Conference on Compressors and their Systems DOI: 10.1088/1757-899X/232/1/012046 (2017)

38. A. Nikiforov, D. Avramenko, A. Kuchumov, S. Terentev, Yu. Galerkin, O. Solovyeva International Conference on Compressors and their Systems, **604**, doi:10.1088/1757-899X/604/1/012046 (2019)