Theory and practice of neural networks application to building mathematical model of centrifugal compressor vane diffusers

Aleksandr Nikiforov¹, Andrei Rekovetc¹, Yuri Galerkin², Evgeniy Petukhov², Aleksey Rekstin², Vasily Semenovsky², Olga Solovyeva²

¹ Smolensk State Agricultural Academy, Bolshaja Sovetskaja 10/2, Smolensk, 214000 Russian Federation
² Peter the Great St. Petersburg Polytechnic University, Polytechnicheskaya 29, Saint Petersburg, 195251 Russian Federation

Abstract. Optimal design of centrifugal compressor stages needs special computational and experimental methods, both of them could be costly enough. So new advanced design methods which can provide optimal solution faster are needed. Authors developed the set of mathematical models – Universal modeling method - for describing compressor stages characteristics. Its models are being widened and improved. One the most advanced approaches of model building is based on machine learning. A neural network based method for predicting centrifugal compressor vane diffuser characteristics was developed. Input data for network training was obtained from CFD simulations. The resulting model for diffuser loss coefficient shows good approximation quality and can be used for improvement of VD model in Universal modeling method.

Introduction

The use of vane diffusers (VD) in the stages of centrifugal compressors is characterized by the possibility of higher efficiency factor in comparison with vaneless diffusers, but at the expense of the width of the operating area. With the help of vanes the higher slow-down of the flow velocity and pressure build-up is achieved. In these conditions the flow structure in VD after the impeller remains quite complicated with often detachable character. In order to accomplish an objective of optimal design it is necessary to develop a special approach that can take into account peculiarities of every element of centrifugal compressor’s flow channel. 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 (UMM) [1-7] is developed, based on mathematical models that describe diffusers characteristics adequately. The method is being constantly upgraded, according to new research results the models are being specified and completed.

The trials of model stages as well as numerical results from CFD simulations are used for identification. Commercial CFD software seems to be suitable tool for designed compressor characteristics estimation [8-11] but it cannot provide acceptable fidelity for design practices [12-16].
Nevertheless, in case of stationary elements simulations CFD methods provide correct solutions [10, 17, 18]. In 8th version of UMM the model of vaneless diffuser was developed [4, 17-20] based on CFD results summary. Similarly, a methodology for vane diffuser CFD simulation was developed and tested for UMM VD model enhancement [21-23]. Also method for processing simulations data with neural networks (NN) was used, the possibility of its application was shown before [3, 24-26].

In [24] the performances in terms of efficiency, pressure ratio and work coefficient of 39 model stages were modeled with mean squared error 1.5 %. In addition, the loss and friction coefficients of vaneless diffusers of relative widths 0.014-0.10 are modeled with mean squared error 2.45 %.

The paper [25] 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.

A simulation experiment based on neural models for studying the influence of dimensions, diffuser shape, and similarity criteria in [26] made it possible to check the physical adequacy of mathematical models, to obtain new data on energy conversion processes and to establish a number of recommendations on the optimal design of vaneless diffusers.

To improve the model of VD the authors introduce the calculation methods based on data processing of hydrodynamic calculations using neural networks (NN). Based on results of multivariant CFD simulations of VD their characteristics are calculated and together with geometric parameters are used as input data for building NN.

Materials and Methods

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.

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 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.

Base object

Vane diffusers, to a large extent, determine the overall dimensions and energy characteristics of the centrifugal compressor as a whole [27-30]. 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 factor 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 operating area. 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 \( b_3 = b_4 \), in order to simplify the design of the diffuser; the midline of the vane profile is made along an arc. 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 or diameters \( D_3 \) and \( D_4 \) respectively;
- the inlet vane angle \( \alpha_{v3} \);
- the outlet vane angle \( \alpha_{v4} \) or the deviation angle \( \Delta \alpha_v = \alpha_{v4} - \alpha_{v3} \);
- the number of vanes \( z \) or row density \( l/t \);
- the height of the vanes \( b \).

All linear dimensions are assumed as ratio to the impeller diameter \( D_2 \). Figure 1 shows a diagram of the geometry of the investigated vane diffuser, the vane is limited by sections “3” and “4”.

![Figure 1. Scheme of the geometry of the vane diffuser of a centrifugal compressor, (a) meridional section, (b) cross-section.](image)
excluded and the flow is practically incompressible, it is correct to calculate the flow parameters by the ratios as follows:

- VD outlet velocity \( c_4 = \frac{c_5}{D_4/D_5} \);
- outlet VD flow angle \( \alpha_4 = \alpha_5 \);
- temperature \( T_4 = T_{in} - \frac{c_4^2}{2cp} \);
- temperature \( T_5 = T_{in} - \frac{c_5^2}{2cp} \);
- VD outlet static pressure \( p_4 = p_5(T_5/T_4)^{-3.5} \).

Thus, the VD with constant parameters of relative diameters of the beginning (1.1) and the end (1.5) of the vanes, the number of vanes (according to the recommended row density 2 [30]). The investigated parameters and the data ranges are shown in Table 1. The total number of the cases is 522 (for \( \alpha_{v3} = 15^\circ \) are also analyzed \( \alpha_{v3} = 20, 30^\circ \), in some cases \( i=\pm 6^\circ \)).

Table 1. The data ranges

| Parameter | Values |
|-----------|--------|
| \( b \)   | 0.025, 0.045, 0.06, 0.08 |
| \( \alpha_{v3} \) | 15, 20, 25, 30° |
| \( \Delta \alpha_{v} \) | 10, 15, 20° |
| \( i \) | -10..+10° at 2° interval |

Neural network data processing approach

Neural networks as a universal approximator allow to build general models based on significant volume of data. Main ideas, peculiarities and advantages of neural network approach in modeling of characteristics of centrifugal compressors are stated in the papers by A. Nikiforov et al [31-33].

The full sample consisting of 522 controlling cases is analyzed. The neural network is built for Loss coefficient \( \zeta = f(b; \alpha_{v3}; \alpha_{v4}; l/t; i; z) \), and the conflicting samples are ignored. The minimum and maximum values of the analyzed samples are shown in Table 2.

Table 2. Minimum and maximum parameters

| Parameter | \( b \) | \( \alpha_{v3} \) | \( \alpha_{v4} \) | \( l/t \) | \( i \) | \( \zeta \) |
|-----------|-------|--------------|--------------|--------|------|-----|
| Minimum value | 0.025 | 15 | 25 | 14 | 1.96 | -10 | 0.06 |
| Maximum value | 0.08 | 30 | 50 | 26 | 2.05 | 10 | 0.99 |

In order to analyze the sample, a frequency analysis was performed for each of the parameters. For repeated parameters the frequency of coefficients was calculated in the sample for every value. For the parameters with different values the frequency of coefficients was calculated for the ranges. Frequency analysis of the sample allows us to 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. It guarantees the most accurate modeling of the areas described in the initial sample by neural networks.

![Figure 2. Frequency diagram of values of the relative width of VD diffuser.](image)
Based on the previous neural network modeling experience [3] two-layer neural networks with 25 neurons in the first (hidden) layer were created. The simulation accuracy was checked when changing the training functions (trainfunction) as follows: with the BFGS quasi-Newtonian method (trainbfgs), with the Levenberg-Marquardt optimization method (trainlm) and Levenberg-Marquardt optimization with Bayesian regularization (trainbr). In the first case hyperbolic tangent function (tansig) of activation in both layers was used. In the second case neural networks with a logical sigmoid (logsig) and a linear activation function (purelin) for each learning function were used.

As a result of the training the root-mean-square inaccuracies were obtained for every neural network type that are shown in Table 3.

Thus, according to the minimum inaccuracy for neural network model of loss coefficient in the first layer logic sigmoid was chosen as an activation function. In the second layer a linear activation function and training function Levenberg-Marquardt optimization with Bayesian regularization were chosen.
Table 3. The inaccuracies of two-layers neural network with different architecture and different training functions for loss coefficient $\zeta$

| NN training functions | tansig, tansig | logsig, purelin |
|-----------------------|----------------|-----------------|
| trainbfgs             | 0.092          | 0.044           |
| trainbr               | 0.041          | **0.040**       |
| trainlm               | 0.070          | 0.063           |

To choose the network architecture on the chosen training function a calculation experiment was conducted. Two-layer neural networks with 10, 15, 20, 25, 30 neurons in the first (hidden) layer and 1 neuron in the output layer were created. The results are shown in Table 4.

Table 4. The inaccuracies of two-layer neural network with different neurons number in the output layer for loss coefficient $\zeta$

| Neurons number in the hidden layer | 10  | 15  | 20  | 25  | 30  |
|------------------------------------|-----|-----|-----|-----|-----|
| Average error                      | 0.069 | 0.057 | 0.054 | 0.040 | **0.037** |

The least inaccuracy 3.7% in the calculation experiment on neural networks is registered for two-layer neural network with 30 neurons in the hidden layer.

To increase the accuracy of the neural model on the next stage tree-layer networks were investigated. The aim of the study was to calculate the optimal ratio of the neurons number in the first hidden layer 10, 15, 20, 25, 30 and in the second hidden layer 10, 15, 20, 25, 30 by means of selection of different combinations of neurons number in the first and in the second layers. In the third output layer one neuron was used. In the first and the second layers activation functions were logical sigmoid and in the second layer it was linear activation function.

The inaccuracies are defined as the root-mean-square deviations. The results are shown in Table 5.

Table 5. The inaccuracies of three-layer neural network with different neurons number in the first and second hidden layers

| Neurons number | First layer | Second layer |
|----------------|-------------|--------------|
|                | 10          | 15           | 20           | 25           | 30           |
| 10             | **0.053**   | 0.047        | 0.036        | 0.030        | 0.021        |
| 15             | 0.021       | **0.029**    | 0.021        | 0.020        | 0.033        |
| 20             | 0.029       | 0.022        | **0.069**    | 0.021        | 0.025        |
| 25             | 0.033       | 0.020        | 0.037        | 0.018        | 0.030        |
| 30             | **0.222**   | 0.018        | 0.032        | 0.027        | 0.034        |

Results

As a result of training of three-layer neural network a minimum inadequacy 1.8% was achieved if the network architecture was as follows: 30 neurons in the first layer; 15 neurons in the second hidden layer; one neuron in the third layer; in the first and the second layers activation functions were logical sigmoid and in the third layer it was linear activation function with the Levenberg-Marquardt optimization method and Levenberg-Marquardt optimization with Bayesian regularization.

The results of loss coefficient modeling for different vane diffusers are shown on Figures 8-11.
Figure 8. Loss coefficient for $b=0.025; \frac{l}{t}=1.95883; a_{v3}=15; a_{v4}=35; z=20$.

Figure 9. Loss coefficient for $b=0.045; \frac{l}{t}=1.95883; a_{v3}=20; a_{v4}=40; z=20$.

Figure 10. Loss coefficient for $b=0.06; \frac{l}{t}=1.96986; a_{v3}=20; a_{v4}=30; z=17$.

Figure 11. Loss coefficient for $b=0.08; \frac{l}{t}=1.95883; a_{v3}=20; a_{v4}=40; z=20$. 
Conclusions
The article presents the analysis of the existing methods of creating mathematical models and possibilities of their use in engineering. As a result of the analysis, the authors suggest a computer model on the basis of neural networks that is one of the most efficient methods of studying of complex technical systems. This computer model allows to conduct series of calculation experiments with the aim to analyze and compare modeling results with the real object behavior.

The conducted calculation defines an optimal neural network with the acceptable error limit and adequate tracking of the gas-dynamic processes of energy conversion in a vane diffuser of the centrifugal compressor. It also defines the number of layers, activation functions and neurons number on every hidden layer.

To improve the quality of modeling the initial data were processed by the above-mentioned methods. It allowed to increase substantially the modeling quality.

In Creating and Training of Neural Network the authors analyzed different combinations of neural networks, inaccuracies of ready models and the structure of the necessary network. Based on it a neural network with the 2% inaccuracy was created.

The article presents an illustrative example of a mathematical model that is used on the basis of the calculation methods of influence of geometrical parameters on energy characteristics of VD.

The mathematical model of a vane diffuser in the existing Universal Modeling Method consists of a dozen algebraic equations with two dozen empirical coefficients. The results of physical experiments are used to identify the model. The algebraic model works satisfactorily. But in order to improve the accuracy of the calculations, new experiments are needed, which is expensive and time-consuming. As an alternative, the authors carried out a computational experiment. The characteristics of vane diffusers are well simulated by neural network equations satisfactory. The next step of the authors is the inclusion of the VD neural model into the Universal Modeling Method programs in order to find and eliminate the shortcomings.

Nomenclature
\[\begin{align*}
\text{b} & \quad \text{blade height, channel width in the direction of rotor axis;} \\
\text{c} & \quad \text{absolute flow velocity;} \\
\text{c}_p & \quad \text{specific heat capacity at constant pressure;} \\
\text{D} & \quad \text{diameter;} \\
\text{h}_p & \quad \text{polytropic pressure;} \\
\text{h}_d & \quad \text{dynamic pressure;} \\
\text{h}_i & \quad \text{lost pressure in stage flow part;} \\
\text{i} & \quad \text{incidence angle (angle of attack);} \\
\text{k} & \quad \text{isentropic coefficient;} \\
\text{l} & \quad \text{blade length;} \\
\text{p} & \quad \text{pressure;} \\
\text{r} & \quad \text{radius;} \\
\text{t} & \quad \text{distance between the vanes;} \\
\text{T} & \quad \text{temperature;} \\
\text{z} & \quad \text{number of blades;} \\
\alpha & \quad \text{angle between absolute velocity and circumferential direction;} \\
\alpha_v & \quad \text{angle between blade midline tangent and circumferential direction;} \\
\delta_v & \quad \text{blade thickness;} \\
\zeta & \quad \text{loss coefficient;} \\
\eta & \quad \text{polytropic efficiency;} \\
\xi & \quad \text{recovery coefficient;} \\
\rho & \quad \text{gas density;}
\end{align*}\]

Funding: The research was performed by a Grant of the President of the Russian Federation for young PhD MK-1893.2020.8.

Acknowledgments: The results of the work were obtained using computational resources of Peter the Great Saint-Petersburg Polytechnic University Supercomputing Center (www.spbstu.ru).

References
[1] Galerkin Yu, Drozdov A, Solovyeva O and Kabalyk K 2019 E3S Web of Conferences
[2] Galerkin Yu, Rekstin A, Drozdov A, Soldatova K, Solovyeva O, Popova E 2020 E3S Web of Conferences
[3] Galerkin Y B, Nikiforov A G, Solovyeva O A, Popova E Y, Rekovets A V 2020 *BMSTU Journal of Mechanical Engineering* 7 29–42

[4] Solovyeva O, Drozdov A 2020 *E3S Web of Conferences*

[5] Rekstin A F, Drozdov A A, Solovyova O A and Galerkin Yu B 2018 *Proc. Int. Conf. Oil and Gas Engineering* (Omsk)

[6] Rekstin A F, Soldatova K V, Galerkin Yu B 2019 *Proc. Int. Conf. on Compressors and their Systems* (London)

[7] Rekstin A, Popova E, Ucehovscy A 2018 *Proc. Int. Conf. Oil and Gas Engineering* (Omsk)

[8] Prasad V V, Kumar L M, Reddy M B 2011 *Science Insights: An International Journal* 1 6-10.

[9] Kabalyk K, Kryłowicz W 2016 *Transactions of the institute of fluid-flow machinery* 131 41–53.

[10] Marenina L N 2016 *Compressor technology and pneumatics* 3 27-35.

[11] Meduri U K, Selvam K, Nawrocki G 2015 *Proc. ASME Turbo Expo* (Montréal)

[12] Galerkin Y, Voinov I, Drozdov A 2017 *Proc. Int. Conf. Compressors and their Systems* (London)

[13] Kortikov N, Borovkov A, Vojnov I, Kirillov A, Drozdov A 2018 *Proc. Int. Conf. Energy* (St.Petersburg)

[14] Borovkov A I, Voinov I B, Nikitin M A, Galerkin Yu B, Rekstin A F, Drozdov A A 2019 *Proc. Int. Conf. Oil and Gas Engineering* (Omsk)

[15] Borovkov A I, Voinov I B., Galerkin Yu B., Drozdov A A, Soldatova K V 2019 *Proc. Int. Conf. Compressors and their Systems* (London)

[16] Borovkov A, Voinov I, Galerkin Yu, Nikiforov A, Nikitin M, Solovyeva O and Kabalyk K 2019 *E3S Web of Conferences*

[17] Galerkin Yu B and Solovyova O A 2014 *Compressor technology and pneumatics* 3 35-41

[18] Galerkin Yu B and Solovyova O A 2014 *Compressor technology and pneumatics* 4 15-21.

[19] Galerkin Yu B, Rekstin A F, Solovyeva O A 2019 *Proc. Int. Conf Oil and Gas Engineering* (Omsk)

[20] Galerkin Y, Soldatova K, Solovieva O 2015 *Proc. Int. Conf. Compressors and their Systems* (London)

[21] Petukhov E P, Galerkin Yu B, Rekstin A F 2019 *Proc. Int. Conf. Compressors and their Systems* (London)

[22] Petukhov E P, Drozdov A A, Galerkin Yu B 2019 *Compressor technology and pneumatics* 3 37-48

[23] Petukhov E P, Galerkin Y B, Rekstin A F 2019 *Proceedings of Higher Educational Institutions. Machine Building* 8 51–64

[24] Nikiforov A, Popova D, Soldatova K 2017 *Proc. Int. Conf. Compressors and their Systems* (London)

[25] Nikiforov A, Kuchumov A, Terentev S, Petukhov E and Kabalyk K 2019 *E3S Web of Conferences*

[26] Nikiforov A, Avramenko D, Kuchumov A, Terentev S, Galerkin Yu, Solovyeva O 2019 *Proc. Int. Conf. Compressors and their Systems* (London)

[27] Galerkin Yu B 2010 Turbocompressors (Moscow: LTD information and publishing center KHT) p 596

[28] Seleznev K P and Galerkin Yu B 1982 Centrifugal compressors (Leningrad: Mechanical engineering)

[29] Seleznev K P and Galerkin Yu B 1986 Theory and design of turbocompressors (Leningrad: Mechanical engineering)

[30] Ris V F 1981 Centrifugal compressors machines (Leningrad: Mechanical engineering) p 351

[31] Nikiforov A G, Popova D Yu, Soldatova K V, Solovyeva O A 2015 *Compressor technology and pneumatics* 3 18-21

[32] Nikiforov A G, Popova D Yu, Soldatova K V, Solovyeva O A 2015 *Compressor technology and pneumatics* 4 14-16
[33] Nikiforov A G, Popova D Yu, Soldatova K V, Solovyeva O A 2015 *Compressor technology and pneumatics* 6 30-33