Vaneless diffusers characteristics simulating by means of neural networks

A Nikiforov¹, D Avramenko¹, A Kuchumov¹, S Terentev¹, Yu Galerkin² and O Solovyeva²

¹ Department of Mechanization, Smolensk State Agricultural Academy, 214000, Bolshaja Sovetskaja, 10/2, Smolensk, Russia
² R&D Laboratory “Gas dynamics of turbo machines” Peter the Great St. Petersburg Polytechnic University, Polytechnicheskaya, 29, St. Petersburg, Russia

Email: nikiforof@mail.ru

Abstract. The paper presents the results of simulation of loss coefficient and the angle of flow at the outlet of diffuser in centrifugal compressor vaneless diffusers. The calculation was performed in a wide range of design and gas-dynamic parameters by means of neural networks. Also, an analysis performed by CFD (Computational Fluid Dynamics) methods is presented. In order to obtain mathematical models, a data sampling was used for vaneless diffusers with the following characteristics: relative width is $b_2/D_{22} = 0.014 – 0.1$, outlet relative diameter is $D_4/D_2 = 1.4 – 2.0$, inlet flow angle is $\alpha_2 = 10 – 90^\circ$, velocity coefficient is $\lambda_{2} = 0.39 – 0.82$, Reynolds numbers corresponding to them are $Re_{\alpha_2} = 87,500 - 1,030,000$. In order to improve the accuracy of simulating using neural networks, various recommendations on the preparation and processing of initial data were collected and tested: identification of conflict samples and outliers, data normalization, improving the quality of the neural networks training under the insufficient amount of sampling, etc. Application of the listed recommendations and an essential expansion of mathematical models definition significantly improved the accuracy of simulating. A simulation experiment based on neural models for studying the influence of dimensions, diffuser shape, and similarity criteria 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.

1. Introduction

The operation process of the centrifugal compressor cannot be described analytically, since the equations of gas motion are the second-order partial differential equations that cannot be integrated. For gas-dynamic design, approximate methods are used based on mathematical models describing such processes. The actual operation process is schematized using the system of semi-empirical equations. Domestic and foreign universities and manufacturers of compressor equipment possess such mathematical models. Their appearance and features depend on the approaches applied to the schematization of processes. The school of the Nevsky plant created the foundation and made the main contribution to the national compressor science [1, 2]. The Kazan Compressor School has developed its own model based on the calculation of losses in vane channels. The calculation is made by analogy with an equivalent diffuser, taking into account the channel curvature [3, 4]. The resulting loss coefficients...
take into account the influence of Mach and Reynolds numbers using empirical ratios, the finite number of vanes is taken into account using the empirical formula of A. Stodola. The international organizations also develop their own design methods [5, 6, 7]. The Universal Modelling Method, developed by Professor Yu.B. Galerkin, gained the recognition in the practice of gas-dynamic design of industrial centrifugal compressors. The method is the basis for optimal design at the Polytechnical Compressor School, which is now being developed in the research laboratory “Gas dynamics of turbo machines” [8, 9, 10]. The Universal Modelling Method is a set of computer programs for the optimal gas-dynamic design of the flow part of centrifugal compressors [9].

Numerical solutions using computational gas dynamics methods seem to be a suitable tool for gas-dynamic characteristics calculation of a designed compressor [11–14]. But, unfortunately, using the commercial computational gas dynamics programs it is impossible to calculate the stage characteristics with an accuracy acceptable for design practice [15]. However, modelling of stator elements of centrifugal stages gives the correct solution [13]. The model of a vaneless diffuser in version 8 of the Universal Modelling Method is based on the generalization of the results of CFD calculations [16].

The paper [17] presents the results of calculations of the flow structure and the gas-dynamic characteristics of a series of vaneless diffusers with a relative width \( b_2 / D_2 = 0.014 - 0.100 \), with a radial length \( D_4 / D_2 \) of up to 2.0, in the range of inlet flow angles \( \alpha_2 = 10^\circ - 90^\circ \). Calculations were made by the ANSYS CFX program. Similarity criteria varied within \( \lambda_{c2} = 0.39 - 0.82 \) (\( \lambda_{c} \) is velocity coefficient), \( Re_{c2} = 87 500-1 030 000 \) (\( Re_{c} \) is Reynolds number). The paper presents comparison with the theory, which showed the regularity of the flow of gas-dynamic characteristics, and comparison with the well-known experiments, which showed a good agreement by the flow structure.

To calculate the flow parameters in vaneless diffusers (VLD), it is enough to know any two of the following values:

- Efficiency \( \eta = \frac{h_p}{h_i} \);
- Loss coefficient \( \xi = \frac{h_p}{c_s^2 / 2} \);
- Recovery coefficient \( \xi = \frac{h_p}{c_s^2 / 2} \);
- Velocity ratio \( \bar{c} = \frac{c_4}{c_2} \);
- Exit flow angle \( \alpha_4 \).

Here \( h_p \) is polytropic head, \( h_i \) is dynamic head, \( h_s \) is loss of head, \( c \) is absolute flow velocity.

Any dimensionless gas-dynamic characteristic of the VLD is a function of its shape, inlet flow angle, similarity criteria, and relative roughness. The vaneless diffuser configuration with parallel walls is determined by relative width \( b_2 / D_2 \) and relative radial length \( D_4 / D_2 \).

The proposed generalization allows one to calculate the loss coefficient and the output flow angle. The generalization of the numerical experiment results was performed using the EXCEL program, which was described in detail in [17].

The efficiency and the loss coefficient are related by the ratio:

\[
\zeta = (1 - \eta) \left(1 - \left(\frac{c_4}{c_2}\right)^2 \right).
\]  

The diffuser efficiency is calculated by the formula (2):
\[ \eta = \frac{\log(p_1 / p_2)}{\kappa - 1} \log(T_1 / T_2) \]

where \( p \) is pressure, \( T \) is temperature, \( \kappa \) is isentropic coefficient.

The calculated characteristics for \( b_2 / D_2 \) equal to 0.014 and 0.1 are shown in Figure 1.

![Figure 1](image_url)

**Figure 1.** VLD characteristics for \( b_2 / D_2 = 0.014 \) (left) and \( b_2 / D_2 = 0.1 \) (right) [17].

The loss coefficient calculated using the formula (1) is a function of 7 variables:

\[ \zeta = f \left( \beta_2, D_4, \overline{r}_g, \alpha_2, \lambda_{\alpha_2}, Re_{b_2}, k \right), \]

where \( \overline{r}_g \) is relative surface roughness.

The approximate dependence for the loss coefficient in the general form is as follows (4):

\[ \zeta = \frac{A \cdot \alpha^6 \cdot K_{D4}}{K_{Re,\alpha}}. \]

Each formula member is a set of algebraic equations. At the first stage, we carried out approximation of the influence of \( b_2 / D_2, \alpha_2, \lambda_{\alpha_2} \) at fixed value of diffuser length \( D_4 / D_2 = 1.6 \). At the second stage, we introduce a correction taking into account the influence of the VLD relative length \( K_{D4} \). The corrections for the influence of the Reynolds criterion and the relative roughness are made by analogy with the plate drag force coefficient. Computational experiment results were also used.

For a hydraulically smooth surface:

\[ \lambda_{Re} = \frac{\lambda_{CFD}}{\lambda_{MM}} = \frac{0.0032 + \frac{0.221}{8.73 \cdot 10^6 \cdot \beta_2}^{0.237}}{0.0032 + \frac{0.221}{8.73 \cdot 10^6 \cdot \beta_2}^{0.237}} \]

where \( \lambda_{CFD} \) is the friction coefficient, calculated by CFD, \( \lambda_{MM} \) is the friction coefficient, calculated by the Universal Modelling Method.

For a rough surface:

\[ K_{Re} = \frac{\lambda_{CFD}}{\lambda_{MM}} = \left( 0.0032 + \frac{0.221}{8.73 \cdot 10^6 \cdot \beta_2}^{0.237} \right) \left( 21 \log \frac{2 K_{rgh}}{K_{re}} + 1.74 \right)^2. \]

Approximation for angle \( \alpha_4 \) is carried out by the same method as for the loss coefficient.
The system of 32 equations has a good accuracy of 98.6%, Figure 2.

![Image 1](https://via.placeholder.com/150)

**Figure 2.** Comparison of calculated and approximated values of loss coefficients for $b_2 / D_2 = 0.043$, $\lambda_2 = 0.64$ (to the left), and comparison of values $\alpha_1 - \alpha_2$ for $b_2 / D_2 = 0.057$ (to the right)

In the course of performing the studies described above, an extensive data material has been accumulated, which provides an objective basis for processing in order to obtain generalizing dependencies for constructing a mathematical model for calculation the VLD gas-dynamic characteristics from a relatively small number of geometric and gas-dynamic parameters. As a processing method, it is proposed to use neural networks, which, being a universal approximator, make it relatively simple to construct generalized models based on processing a large amount of source data. It is necessary to specifically emphasize that neural network models are a simple tool for use in design or research activities that do not require special preliminary preparation. The main provisions, features, and advantages of the neural network approach in modelling the characteristics of centrifugal compressors are given in [18].

2. Creating a neural network model

In a simplified form, it can be shown that the neural network performs an approximation:

$$Y = f(X),$$  

(7)

where $X$ is the input vector, which is a set of geometric and gas-dynamic parameters, $Y$ is the output vector, that is one or another desired characteristic of the VLD efficiency, $f$ is the transformation performed by the neural network.

The accumulated practice of building and analyzing the use of neural networks for modelling characteristics of centrifugal compressors suggests that an important step in building a neural model is the preliminary preparation and processing of initial data for training neural networks, which can significantly improve the accuracy and the reliability of the desired models.

In general, for processing an initial data sampling and for training of neural networks, you can use the following sequence of steps, which the authors formed empirically:

1. Selection of parameters of the input vector.
   a) Analysis and logic of the subject area;
   b) Analysis of the input neurons weight coefficients;
   c) Perturbation of the input parameters values and analysis of the network response to these perturbations;
4. Improving the quality of neural networks training with an insufficient sampling size (multifold cross-validation, multiple sampling repetition and modification the order of training samplings).
5. Identifying emissions.
6. Data normalization.
7. Selection of neural networks types and activation functions.
8. Network decomposition by the number of output neurons.

Below, we will consider in more detail some of these steps of the algorithm for processing an initial data sampling for constructing a neural network.

The success of creating a neural network model depends largely on the choice of input parameters. In general, the parameters that do not affect the value of the output vector \( Y \) can be included in the initial sampling, and they are assumed to be insignificant. But it is not easy to determine within a huge initial data array what parameters will be significant for the model and what parameters can be definitely excluded. After the creation and the training of a neural network on all the data that the researcher was able to obtain, insignificant parameters can be identified in several ways specified in the algorithm.

For example, it can be carried out by analyzing the input neuron weights values. Since neural networks are self-training systems, during the process of training the parameters, weights that have little effect on the result will weaken and become significantly less than the weights for the other parameters. The analysis of weights of the neural network is given in a previously published paper [18].

After identifying and eliminating insignificant parameters from the training sample, the quality and the accuracy of the neural network model improves due to the reduction of its dimensionality and complexity. However, it is important to remember that an excessive reduction in the number of input parameters and the simplification of a neural network can interfere with the identification of patterns in a particular task. This may also entail the emergence of conflicting (controversial) examples. Examples are called conflicting when they have the same input vectors and different output vectors. Because of the erroneously prepared data, the error will not fall below this value, no matter what training methods we use.

For successful modelling on the basis of neural networks, it is important to provide the necessary volume of the training sample. Partly, the “the more the better” concept is correct, but it is important to remember that the volume of the sample affects the training time and an excessive volume will lead to a large expenditure of machine time for setting up a neural network. In [19], a formula is given with which we can determine the size of a training sample:

\[
Q = 7 \times N_v + 15, \tag{8}
\]

where \( N_v \) is the number of input parameters of the neural network model; \( Q \) is the volume of the training sample.

In practice, it is often not possible to collect a sufficient amount of data for training and there is a need for preliminary manipulations with sampling. We tried such processing methods as multifold cross-validation, and multiple repetitions of the initial sampling supplied to the input of the neural network [20, 21], as well as modification of the order of the training sample. This makes training more stochastic and helps to reduce the probability of hitting local extremes.

The exclusion of emissions in the selection also helps to improve the accuracy of the neural network model before it is created. But in the case of large sampling size, outliers are difficult to detect, so they resort to different outliers search algorithms [22].

It is also desirable to normalize the information prepared for neural network processing by reducing the range of modification values to a common range for all input parameters (for example, [0,1]). This allows obtaining an improved network training. The normalization process is described in detail in [23, 24, 25].

Depending on what problem you need to solve using neural network modelling, you need to select the type of neural network [26]. In order to improve the accuracy of the neural network model in the
context of the task, it is necessary to solve the problem of choosing the activation functions of neurons [27].

The techniques listed above were used to create a neural model of the loss coefficient and of change in the flow angle $\alpha_4 - \alpha_2$. For constructing mathematical models we used data from 24 VLD of various geometries, studied for several values of $\lambda_{d2}$ and $\omega$. Thus, the sampling size of the initial data for building the loss coefficient model was 936 different variants and it was 429 for the VLD flow angle change model.

The data presented above, in addition to its wide range, have a fairly high sampling density of the intermediate values:

- Relative width of the diffuser at the input $b/D_2$ is: 0.014; 0.016; 0.019; 0.022; 0.025; 0.029; 0.033; 0.038; 0.043; 0.05; 0.057; 0.066; 0.074; 0.087; 0.1;
- Relative diameter at the output from the VLD $D_y/D_2$ is: 1.4; 1.6; 1.8; 2.0;
- Flow angle at the input to the VLD $\alpha_2$ is: 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90;
- Velocity coefficient $\lambda_{d2}$, calculated using the input velocity $c_2$ is: 0.39; 0.64; 0.82.

In accordance with the above algorithm for preliminary data processing, the analysis of the subject area and of weights, analysis of the network response to the input perturbations of parameters, elimination of outliers, and normalization of all values were performed.

Based on a pre-processed sampling, the VLD neural model loss coefficient was constructed as a generalized dependence:

$$\zeta = f \left( p_2, D_2, \alpha_2, \lambda_{d2} \right),$$

and models of the flow angle change in a vaneless diffuser are:

$$\alpha_4 - \alpha_2 = f \left( p_2, \alpha_2, \lambda_{d2} \right).$$

Further, in order to select the optimal network architecture of each model, we created neural networks with various numbers of layers and neurons.

To determine the required number of synaptic weights of the neural network, we use the consequence of the Arnold – Kolmogorov – Hecht-Nielsen theorem, expressed as:

$$N_c Q \leq N_w \leq N_x \left( \frac{Q}{N_x} + 1 \right) \left( N_x + N_y + 1 \right) + N_y,$$

where $N_c$ is the number of neurons in the input layer (the number of parameters); $N_x$ is the number of neurons in the output layer (the number of simulated values); $Q$ is the number of elements in the set of training examples, that is, the number of pairs of input and output vectors $X_q$ and $Y_q$; $N_w$ is the required number of synaptic connections.

In accordance with the modelling data, we obtained the range of the number of synaptic connections for the loss coefficient model: $\zeta$ from 70 to 1115 ($70 \leq N_w \leq 1115$), and for the model of VLD flow angle change: $\alpha_4 - \alpha_2$ from 40 to 649 ($40 \leq N_w \leq 649$).

This allows one to determine the required number of neurons in the hidden layers. For example, the number of neurons in the hidden layer of a double-layer perceptron will be equal to [18]:

$$N = \frac{N_w}{N_x + N_y}$$

Calculation performed using the formulas (11), (12) shows that the optimal number of neurons in the hidden layer for a two-layer perceptron for the model of loss coefficient $\zeta$ is in the range from 14 to 223 neurons ($14 \leq N \leq 223$). For the model of VLD flow angle change, $\alpha_4 - \alpha_2$ is in the range from 10 to 162 neurons ($10 \leq N \leq 162$).

At present, there is no strict theory of choosing the optimal number of hidden layers and neurons in hidden layers. In practice, perceptrons with one or two hidden layers are most often used, and the number of neurons in hidden layers usually ranges from $N_c/2$ to $3*N_c$.

To model the loss coefficient $\zeta$, we chose a neural network with the following architecture: two-layer, the number of neurons in the hidden layer is 20, the activation functions is the logical sigmoid for all layers, the learning function is optimized using the Levenberg-Marquardt algorithm.
To model the change in VLD flow angle $\alpha_4 - \alpha_2$, we chose a neural network with the following architecture: two-layer, the number of neurons in the hidden layer is 30, the activation functions is the logical sigmoid for all layers, the learning function is optimized using the Levenberg-Marquardt algorithm.

We choose the two-layer networks as the difference in the errors of the two-layer and three-layer neural networks is insignificant, but the model on two layers of perceptron is more economical in terms of computing power.

The verification of the effect of normalization was carried out when comparing a neural network trained on non-normalized data and a neural network trained on normalized data, Figure 3.

![Figure 3](https://example.com/image.png)

**Figure 3.** Comparison of the modelling results of the normalized and non-normalized sampling with the calculation results obtained using the ANSYS CFX program. NN are values calculated by a neural network using non-normalized data, NN Norm are values calculated by a neural network using normalized data, CFD denotes calculation data obtained using the ANSYS CFX program.

Figures a), b) present loss coefficient; c), d) present change of flow angle.

The average error when using non-normalized sampling for calculated values of $\zeta = f(\beta_2, D_4, \alpha_2, \lambda_{c2})$ is 6.4%, for normalized data it is 2.7%. The average error of non-normalized sampling for the calculated values of $\alpha_4 - \alpha_2 = f(\beta_2, \alpha_2, \lambda_{c2})$ is 8.3%, for normalized data it is 3.9%.
Sampling preprocessing for neural network training is of practical importance, since it allows one to significantly reduce modelling errors compared to neural networks created on unprepared source data.

The training of new network models allowed us to carry out computational studies, on the basis of which it can be concluded that the models are physically adequate. Figure 4 show the results of calculations of the functions $\zeta = f(\bar{b}_2, D_4, \alpha_2, \lambda_2)$ and $\alpha_4 - \alpha_2 = f(\bar{b}_2, \alpha_2, \lambda_2)$.

![Graphs showing the results of calculations for different values of $b_2/D_2$ and $\lambda_2$.](image-url)
Figure 4. a), b), c) Comparison of loss coefficients for VLD calculated by the ANSYS CFX (solid line) and the neural network (dashed line). d), e), f) Comparison of the flow angle change in the VLD calculated by the ANSYS CFX program (solid line) and the neural network (dashed line).

The good match between the calculation results on neural networks and the data of CFD calculations allows us to conclude that the use of neural networks is suitable for modelling vaneless diffusers characteristics.

3. Conclusion
A neural network model was developed for calculations of the gas-dynamic characteristics of vaneless diffusers of centrifugal compressor stages. Recommendations for improving the accuracy of neural network modelling were suggested. Recommendations were formulated into a single algorithm consisting of a sequence of processing steps of the initial sampling. The proposed algorithm was tested when modelling the vaneless diffusers characteristics in centrifugal compressor stages. The use of data rationing made it possible to reduce the error in modelling the loss coefficient by 3.7%, and for the flow angle adjustment value by 4.4% compared with the model trained on non-normalized data. The obtained modelling results and a significant reduction in errors of neural network models show the importance and necessity of preprocessing the training sample.

Acknowledgments
The presented scientific research corresponds to the program of the National technological initiative Centre “New production technologies” of SPbPU and contributes to the formation of competencies in new production technologies in the area of power engineering.

References
[1] Riess V. 1981 Centrifugal compressors. - L: Mechanical Engineering. - p.351
[2] Den G 1960 Study of diffusers of stationary centrifugal compressor machines PhD. Dis.Tech.p. 179
[3] Schneppe V. 1995 Design and calculation of centrifugal compressor machines. - M.: Mechanical Engineering.- p. 240
[4] Khisameev I., Maksimov, V., Batkis, G.and Guzelbaev Y. 2012 Design and operation of industrial centrifugal compressors. – Kazan. - p. 671
[5] Aungier R. 2000 Centrifugal compressors: a stragedy for aerodynamic design and analysis New York. - ASME Press. p.320
[6] Japikse D. and Krivitzky, E. 2016 Radial stages with non-uniform pressures at diffuser inlet. Proceedings of the ASME Turbo Expo. Volume 2D-2016, Seoul; South Korea;
[7] Bianchini A, Ferrari L. and Ferrara G. An academic test rig for industrial centrifugal compressor
stages: a design approach. *Proceedings of the ASME Turbo Expo. Volume 9-2016*, Seoul; South Korea.

[8] Drozdov A and Rekstin A 2018 Analysis of the velocity diagrams of impellers of centrifugal compressor stages after the preliminary design *MATEC Web Conf.* 245.

[9] Galerkin Yu 2010 *Turbocompressors. LTD* information and publishing center KHT, Moscow p. 596

[10] Seleznev K. and Galerkin, Y. 1982 Centrifugal compressors. - L.: mechanical engineering. - p. 271.

[11] Prasad V., Kumar M. and Reddy B. 2011 Centrifugal compressor fluid flow analysis using CFD. *Science Insights: An International Journal.* – №1(1). – p. 6 - 10.

[12] Kabalyk K. and Kryłłowicz W. 2016 Numerical modelling of the performance of a centrifugal compressor impeller with low inlet flow coefficient. *Transactions of the institute of fluid-flow machinery* p. 41 – 53.

[13] Marenina, L. 2016 CFD-modelling and analysis of the characteristics of stationary elements of the flow part of the centrifugal compressor stage. *Compressor technology and pneumatics.* – № 3. - P. 27-35

[14] Gamburger D. 2009 *Numerical modelling of viscous gas flow in a centrifugal compressor stage: methods and results.* PhD. Dis.Tech. – SPbTU. - - p. 190.

[15] Galerkin Y., Voinov I. and Drozdov A. 2017 Comparison of CFD-calcualtions of centrifugal compressor stages by NUMECA Fine/Turbo and ANSYS CFX programs. Conference “Compressors and their Systems”. – London.

[16] Rekstin A., Drozdov A., Solovyeva O. and Galerkin Y. 2018 Two mathematical models centrifugal compressor stage vaneless diffuser comparison. - Oil and Gas Engineering (OGE-2018). – *AIP Conf. Proc.* 2007. - Omsk. – Russia.

[17] Solovyeva O. 2018 *Mathematical model for gas-dynamic characteristics calculation and optimization of vaneless diffusers of centrifugal compressor stages* PhD. Dis.Tech. SPbSPU p. 162.

[18] Nikiforov A., Popova D. and Soldatova K. 2017 A network application for modelling a centrifugal compressor performance map. Conference “Compressors and their Systems”. – London.

[19] Yasnitsky L. 2016 *Intelligent systems.* – M.: Knowledge Lab., – p. 221

[20] Kazantsev P. 2008 Development and research of the route for designing a neural network application with hardware support. PhD. Dis.Tech MIPT. - M. p.117

[21] Dagayeva M., Suleymanov M., Kataseva D., Katasev A. and Kirpichnikov A. 2018 The technology of building fault-tolerant neural network models of handwriting recognition in biometric authentication systems *Bulletin of the Technological University.* -. Vol. 21. No. 2. p. 133-138.

[22] Shkodyrev V., Yagafarov K., Bashtovenko V. and Ilina E. 2017 Overview of methods for detecting anomalies in data streams *Proc.of the Second Conference on Software Engineering and Information Management*, Saint-Petersburg, Russia. Vol. 1864. — p. 7.

[23] Jayalakshmi T., Santhakumaran Dr. 2011 Statistical Normalization and Back Propagation for Classification. *International Journal of Computer Theory and Engineering*, Vol.3, No.1.

[24] Eduardo S., Ogasawara, Leonardo C., Martinez, D., Geraldo Z. and Marta Mattoso. 2010 Adaptive normalization: A Novel Data Normalization Approach for Non-Stationary Time Series. The 2010 International Joint Conference on Neural Networks (IJCNN)

[25] Nayak S., Misra B. and Behera H. 2014 Impact of Data Normalization on Stock Index Forecasting. *International Journal of Computer Information Systems and Industrial Management Applications.*, Vol. 6, p. 257 – 269.

[26] Osovskij S. 2002 *Neural Networks for Information Processing.* – M.:Financy and Statistica. Poland – p. 344.

[27] Borisov V., Fedulov A. and Zernov M. 2017 *Fundamentals of fuzzy models hybridization.* M: Hot Line-Telecom. p. 100.