UDC 004.85: 004.89: 519.25:519.651: 532.5.013.12: 627.152

Yaroslav V. Khodnevych, PhD (Engineering), Research Scientist
ORCID ID: 0000-0002-5510-1154 ***e-mail:*** *ya.v.khodnevych@gmail.com*

Dmytro V. Stefanyshyn, D. S. (Engineering), Senior Research Scientist
ORCID ID: 0000-0002-7620-1613 ***e-mail:*** *d.v.stefanyshyn@gmail.com*

Institute of Telecommunications and Global Information Space of NASU, Kyiv, Ukraine

DATA ARRANGEMENTS TO TRAIN AN ARTIFICIAL NEURAL NETWORK WITHIN SOLVING THE TASKS FOR CALCULATING THE CHÉZY ROUGHNESS COEFFICIENT UNDER UNCERTAINTY OF PARAMETERS DETERMINING THE HYDRAULIC RESISTANCE TO FLOW IN RIVER CHANNELS

Abstract. *Hydraulic calculations and mathematical modelling of open flows in river channels keep still being among the most topical hydro-engineering today's problems in terms of practice. While solving them, independently on the research topic and purpose, and methods used, a number of simplifications and assumptions are usually accepted and applied. Moreover, there is a range of methodological, structural, and parametric uncertainties, which to be overcome require complex empirical pre-researches. First of all, these uncertainties relate to assessing hydraulic resistances and establishing numerical characteristics of them, which depend on many factors varying spatially and temporally.*

One of the most frequently used integral empirical characteristics expressing the hydraulic resistance to open flows in river channels is the Chézy roughness coefficient C. However, despite a large number of empirical and semi-empirical formulas and dependencies to calculate the Chézy coefficient, there is no ideal way or method to determine this empirical characteristic unambiguously. On the one hand, while opting for an appropriate formula to calculate the Chézy coefficient, we need to take into account practical experience based on comprehensive options analysis considering different empirical equations used alternatively to represent the hydraulic resistance to open flows. On the other hand, the fullness and comprehensiveness of field researches of numerous hydro-morphological factors and parameters characterizing various aspects of the hydraulic resistance to open flows can also have an essential role. In particular, the accuracy assessment of the Chézy coefficient computing based on field data, despite methods and formulas, indicates that the accuracy of field measurements of the parameters included in selected formulas largely determines the relative error of such calculations.

This paper deals with the problem of data arrangements and the development of general rules for the formation of training and test samples of data to train artificial neural networks being elaborated to compute the Chézy coefficient taking into account the parametric uncertainty of data on the hydro-morphological factors and parameters characterizing the hydraulic resistance in river channels. The problem is solved on the example of an artificial neural network of direct propagation with one hidden layer and a sigmoid logistic activation function.

Keywords: *artificial neural networks; Chézy's roughness coefficient; data arrangements; hydraulic resistance in river channels; parametric uncertainty*

DOI: https://doi.org/10.32347/2411-4049.2022.2.59-85

1. Introduction

The concept of hydraulic resistance is widely used to solve numerous practical hydro-engineering and fluid mechanics tasks, in particular, those relating to open flows in river channels [1-4]. Among them, first of all, we need to note common hydraulic calculations of the river channels' capacity and the position of the free water surface of open streams [1, 3, 4]. The notion of hydraulic resistance is also used while solving the special tasks of mathematical modelling of streams with free surface in canals and river channels within one (1D) and two-dimensional (2D) flow models of shallow water (non-linear de Saint-Venant equations) describing unsteady open channel flow [5-10]. For instance, the above-mentioned models are applied in numerous modern computational modelling systems, such as the HEC-RAS River Analysis System supporting steady and unsteady flow water surface profile calculations, sediment transport computations, and water quality analyses, etc. [11]. The shallow water models keep successfully competing with more advanced today's hydrodynamic solutions based on the Navier-Stokes equations of the real fluid motion and Reynolds' averaged equations of turbulent water flow, which describe the behaviour of an unsteady three-dimensional flow. Applying them, we avoid assumptions and simplifications that are connected with the hydraulic resistance concept usage. However, results obtained from traditional hydraulic calculations and flow modelling due to 1D and 2D shallow water models may be used as boundary conditions for the next computations based on the Reynolds and Navier-Stokes equations [12-14]. Such an approach can essentially simplify solutions to complex real-world case study tasks of hydrodynamics [13]. Some additional examples of recent pieces of literature relating to the use of 1D and 2D flow models of shallow water are also highlighted in [15-18].

Relying on hydraulic resistance concept when performing traditional hydraulic calculations and mathematical modelling of open flows in river channels, we keep repeatedly dealing with the complex challenge relating to determining numerical characteristics of hydraulic resistance in spite of this problem has long been considered by hydraulic scientists and engineers. Regarding practice, it has still been discussed even in terms of a friction factor (namely, the Darcy-Weisbach friction factor) [19-23] or a roughness coefficient [24-33] usage as appropriate hydraulic resistance numerical characteristics. In the last case, in term of a roughness coefficient, there are also two options, namely, which of them, the Manning roughness coefficient or the Chézy roughness coefficient might fit better.

Admittedly, there are three practically equivalent empirical equations (or appropriate empirical models) linking mean flow velocity V to the hydraulic resistance numerical characteristics. They are the Chézy, Manning (Gauckler-Manning or Gauckler-Manning-Strickler), and Darcy-Weisbach equations, which may be summarized as [1, 3, 4, 30-33]:

$$V = C\sqrt{R \cdot S_f} = \frac{1}{n} R^{\frac{2}{3}} \sqrt{S_f} = \sqrt{\frac{8g \cdot R \cdot S_f}{\lambda}}\,, \tag{1}$$

where C is the Chézy roughness coefficient ($m^{1/2}/s$), n is the Manning (Gauckler-Manning) roughness coefficient ($s/m^{1/3}$), and λ is the Darcy-Weisbach friction factor; $V = Q/A$ is the depth-averaged or cross-sectional averaged velocity (m/s),

Q is the water discharge (m^3/s), $A = B \cdot h$ is the cross-sectional area of the flow (m^2), B is the average flow width (m), h is the average flow depth (m), $R = A/P$ is the hydraulic radius (m), P is the wetted perimeter (m), S_f is the energy grade line slope (or the water surface slope); g is the gravitational acceleration (m/s^2).

Usually, the Chézy roughness coefficient C and the Manning roughness coefficient n are used in calculating the averaged velocity of open flows; the Darcy-Weisbach friction factor λ is more common in calculating the averaged velocity of water movement in pipelines [1, 3, 4]. However, there are no formal restrictions on the use of one or another numerical characteristic of hydraulic resistance, independently on whether it is an open flow or a water movement in a pipeline. The Darcy-Weisbach formulation of hydraulic resistance is used for open flows either [6, 9, 19-23, 26].

Formally, if we take into account the equivalence of equations (1) the following simple relationships between the roughness coefficients C, n, and the friction factor λ may be established:

the Chézy roughness coefficient C will relate to the Darcy-Weisbach friction factor λ as [3, 4, 33]:

$$C = \sqrt{\frac{8g}{\lambda}}\ ,\ \lambda = \frac{8g}{C^2}; \tag{2}$$

the Gauckler-Manning roughness coefficient n will relate to the Darcy-Weisbach friction factor λ as [15, 19, 33]:

$$n = \sqrt{\frac{\lambda \cdot R^{1/3}}{8g}}\ ,\ \lambda = \frac{8g \cdot n^2}{R^{1/3}}; \tag{3}$$

and, eventually, the Chézy roughness coefficient C will relate to the Manning (Gauckler-Manning) roughness coefficient n as:

$$C = \frac{1}{n} R^{1/6},\ n = \frac{C}{R^{1/6}}. \tag{4}$$

Practice shows, it does not matter what the kind of characteristic of the hydraulic resistance to open flow in river channels we use, whether it is the friction factor λ or one of the two roughness coefficients, either the roughness coefficient C or n. More important is how fully it can characterize the hydraulic resistance in a real-world case study, as well as how accurately we can calculate numerical values of the appropriate characteristic relating to this case study.

It should be noted that historically the first empirical equation linking mean flow velocity V to the hydraulic resistance was the Chézy formula, which was obtained by the famous French hydraulic engineer Antoine de Chézy in 1775. It concerned the velocity of pipe flows, but in the modified form $V = C\sqrt{R \cdot S_f}$ Chézy proposed to use this dependence for open channel flows as well. In general, Chézy's equation can be considered as the most generalized empirical model in open-channel hydraulics [1, 3, 4]. We tried showing it in Fig. 1.

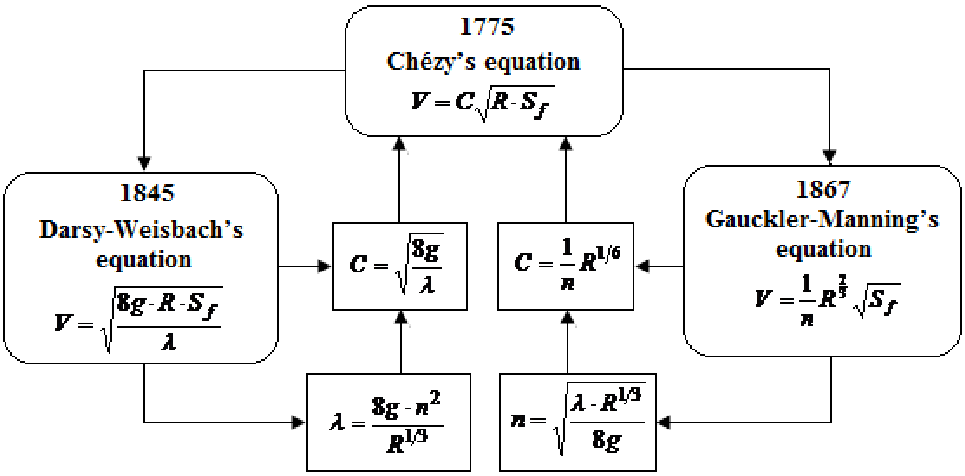


Fig. 1. Flow-chart to explain relationships between the Chézy, Gauckler-Manning, and Darcy-Weisbach equations in the context of empirical models of the hydraulic resistance to open flows in river channels

The next empirical equation was the Darcy-Weisbach formula. It was first proposed by Julius Weisbach in 1845 and relates the head loss Δh due to friction along a given length L of pipe with diameter D to the average velocity V of the fluid flow for an incompressible fluid as [1, 3, 4, 23, 34]:

$$\Delta h = \lambda \cdot \frac{L}{D} \cdot \frac{V^2}{2g}, \tag{5}$$

where the dimensionless friction factor λ is regarded as a function of relative roughness and the Reynolds number (Re) characterizing flow regime.

At present, in hydraulics, there seems to be no formula more accurate or universally applicable than the Darcy-Weisbach equation (5) supplemented by the Moody diagram or Colebrook equation [34]. Therefore, a lot of modern formulas and dependencies proposed to calculate the Chézy resistance coefficient are derived from the relationship (2), which links the Chézy coefficient with Darcy-Weisbach's friction factor [19, 20, 23, 33]. The latter one, in turn, is also determined due to various empirical formulas [20, 22, 28, 29, 35].

The third, Gauckler-Manning's formula was first presented by the French engineer Philippe Gauckler in 1867 and later re-developed by the Irish engineer Robert Manning in 1890 [1, 3, 4]. This equation can be obtained by use of dimensional analysis. Moreover, in the 2000s the Gauckler-Manning formula was derived theoretically using the phenomenological theory of turbulence [36, 37].

The Gauckler-Manning formula is not so universal one as the Chézy and Darcy-Weisbach equations. It can only be applied to streams that have a free surface, such as an open channel, etc. This formula can be considered as a kind of approximation of the Chézy formula, namely, as a partial case of Chézy's equation, when the Chézy coefficient values relate to the Gauckler-Manning roughness coefficient n values according to the relationship (4). In addition, in contrast to the Darcy-Weisbach friction factor λ, the Gauckler-Manning roughness coefficient n does not depend on the Reynolds number and flow regime, which needs being taken into account while analysing hydraulic resistance.

To sum, in term of quantitative presentation of the hydraulic resistance to open flows in river channels, the Chézy roughness coefficient C can be thought to be the most complete empiric numerical characteristic compared with the Darcy-Weisbach friction factor λ and Gauckler-Manning roughness coefficient n. This is because natural watercourses like rivers are characterized by a significant variety of hydro-morphology conditions changing in space and time. Usually, the hydro-morphological changes occur constantly, although stochastically, seasonally, and regularly. Sometimes, they occur sporadically on large scales. The hydraulic resistance to open currents in river channels can depend on manifold elements of roughness such as bottom ridges, dunes, and riffles, turns and bends of a channel, heterogeneity of size and shape of a river channel along its length, including suspended sediments and bottom deposits, vegetation, ice, and others. In some parts of a river, at local scales, essential hydro-morphological changes can occur due to compressions of river channels and floodplains because of temporary formations, such as ice gorges, rubbish of logging, recent alluvial deposits, etc. As a result, similar water levels in rivers can be observed at different water discharges, and vice versa [38]. Human activity can also change dramatically the hydraulic resistance within a river channel and within its floodplain. Herewith, errors, oversights, and flaws in determining the hydraulic resistance, especially when it comes to forecasting flood danger, can result in catastrophic consequences (Fig. 2).

The Halych town inundation

A flooded solar power plant

Fig. 2. Consequences of the June flood of 2020 on the Dniester River near the Halych town (from www.pravda.com.ua)

Thereby, the comprehensive pre-studies relating to estimation of the hydraulic resistances in river channels can be thought an urgent need. In particular, the current comprehensive research of integral empirical numerical characteristics of the hydraulic resistances would open up significant opportunities to prompt flood risk management on rivers. Considering the variety of hydro-morphology and hydrology of rivers, the Chézy roughness coefficient C seems the fittest numerical characteristic to present the hydraulic resistance to open flows in river channels comparing with other integral empirical characteristics, namely, the Darcy-Weisbach friction factor λ and Gauckler-Manning roughness coefficient n. The Chézy coefficient seems to be the most holistic and dynamic numerical empirical characteristic comparing with others. The Chézy coefficient enables to control more factors and parameters determining the hydraulic resistance to open flows in river channels. Using it, we can take into account features of individual river sections and their hydraulic regimes. Eventually, the friction factor λ and roughness coefficient n are often included to formulas and dependencies to calculate the Chézy roughness coefficient C just as needed components.

2. Generalization of the problem. The research aim and objectives

Currently, there are a lot of different empirical and semi-empirical formulas and dependencies in order to calculate the Chézy roughness coefficient values within solving real-world tasks of hydro-engineering calculations and mathematical modelling of open flows in river channels. In the previous study [33], we examined and systematised some well-known of those presented in the literature on open-channel hydraulics, in reference books, tutorials, manuals, and articles highlighting the results of original research on analyzing the hydraulic resistance to flow in open channels. Moreover, numerous publications on mathematical modelling of uniform and non-uniform water flow within 1D and 2D flow models of shallow water were reviewed. In total, we gathered 43 empirical dependencies to compute the Chézy coefficient values, as well as 13 empirical dependencies that can be used to calculate the Gauckler-Manning roughness coefficient values. Based on these dependencies, near 250 empirical equations can be compiled to compute the Chézy coefficient values taking into account hydro-morphology peculiarities of river channels, various flow regimes, specific application limits of the formulas, etc.

We divided all examined empirical formulas to compute the Chézy coefficient into five groups [33]. The four groups are represented with explicit dependencies, which allow calculating the Chézy coefficient values directly due to values of the parameters included in those formulas. The fifth group consists of implicit formulas, which need applying a trial-and-error procedure (iterative calculation).

The first group of explicit formulas, those we analysed and systemised [33], consists of thirteen dependencies the Chézy coefficient C on the Gauckler-Manning roughness coefficient n and the hydraulic radius R:

$$C = f\left(n, R\right), \tag{6}$$

where the Gauckler-Manning roughness coefficient n characterises the roughness of the banks and bottom of river channels and floodplains; the roughness coefficient n values can be obtained in different ways, in particular, due to selecting them from published in the literature on open-channel hydraulics n-value tables [1, 3, 4, 24, 25], or using special empirical formulas; some of them (thirteen dependences) we gave in [33].

Provided that the average flow width $B >> h$ and $R \cong h$, instead of the hydraulic radius R, the dependencies (6) may include the average flow depth h: $C = f\left(n, h\right)$. Some dependencies entering the group (6) may also include the water surface slope S_f as an additional parameter.

Among the formulas of the type of (6) used to calculate the Chézy roughness coefficient values we have to mention the well-known and frequently cited in literature pieces the Manning, Guanguillet-Kutter, Forchheimer, and Pavlovskii formulas [1, 3, 4, 29, 33]. These formulas are now the most common to compute the Chézy coefficient C in various applications. They can be applied for both mountain and plain rivers, both small and medium or large rivers, rivers with earthen or indelible channels. Most of them are considered to be acceptable when values of the Gauckler-Manning roughness coefficient n range from 0.011 (for example, these are closed conduits flowing partly full uncoated or concrete culvert with bends, connections, and some debris, lined or built-up channels with a smooth concrete trowel finish, etc. [1]) to 0.04 (including excavated or dredged and not maintained channels with a clean bottom and brush on the side, as well as natural plain streams, mostly clean, but with some weeds and stones, including floodplains with light brush and trees in summer, and mountain streams, no vegetation in channels with gravels, cobbles, and few boulders in their bottom [1]), and values of the hydraulic radius R range from 0.1 to 5.0 m [33]. Some dependencies may be used even in cases where the roughness coefficient n values reach as much as 0.2 (for example, these are mountain rivers with extremely high resistance with channels composed of boulders or floodplains with trees, dense willows in summer [1]), and the hydraulic radius R values up to 20.0 m [33]. In turn, when estimating the roughness coefficient values, if necessary, it may be taken into account river channel geometry features including meandering and cross-section shape; water-surface profile; roughness because of friction within river bed and due to bank sediments, debris and sediment transport; roughness attributable to vegetation, ice cover, natural and artificial obstructions, and other flow-retarding factors in channels and floodplains.

The second group of explicit formulas, those we defined and systemised, consists of fourteen dependencies in which the Chézy roughness coefficient values are determined based on the relationship (2) between the Chézy coefficient C and the Darcy-Weisbach friction factor λ.

In general, there can be two sorts of roughness influencing the Darcy-Weisbach friction factor λ as the integral characteristic of hydraulic resistance to flows in river channels [33]. The first sort of roughness in terms of the hydraulic resistance to flows in river channels relates to the micro-roughness characterised by the height of protrusions of roughness Δ depending on the size of the bottom fractions of sediments. With taking into account this sort of roughness, the Chézy roughness coefficient C_Δ is established, which is determined as a function of the height of protrusions of roughness Δ and hydraulic radius R:

$$C_\Delta = f(\Delta, R). \tag{7}$$

In particular, among the formulas of the type of (7) used to calculate the Chézy roughness coefficient C_Δ values it should be noted the Strickler, Colebrook-White and Williamson formulas [19, 20, 23, 28, 33]. The formulas of the type of (7) are usually used to compute the Chézy roughness coefficient for mountain and foothill, mostly small and medium-sized rivers, which have practically non-erosion gravel-pebble, pebble-boulder channels.

The second sort of roughness in terms of the hydraulic resistance to flows in river channels relates to the macro-roughness characterized by the size (the height h_r and length l_r) of the bottom ridges (riffles, dunes, etc.) and other similar channel formations. With taking into account this sort of roughness, the Chézy roughness coefficient C_r is established, which is determined as a function of the height h_r and length l_r of the bottom ridges, and hydraulic radius R:

$$C_r = f\left(h_r, l_r, R\right). \tag{8}$$

Among the formulas of the type of (8) used to calculate the coefficient C_r values it can be noted the Knoroz, Snischenko, and Sterenlicht-Polad-zade formulas [33, 39-42]. However, the formulas of the type of (8) have been developing mainly for large canals and plain rivers, where there are conditions to exist of the bottom ridge phase of sediment movement.

In exceptional cases, if we need to take into account the micro- and macro-roughness simultaneously the Chézy coefficient can be written as [33, 39, 40, 42]:

$$\frac{1}{C^2} = \frac{1}{C_\Delta^2} + \frac{1}{C_r^2}. \tag{9}$$

Instead of the hydraulic radius R, the dependencies (7), (8) may include the average flow depth h: $C_\Delta = f\left(\Delta, h\right)$, $C_r = f\left(h_r, l_r, h\right)$. The height of protrusions of the roughness Δ of a channel is usually equated to the average diameter d of soil particles making up the bottom and banks of a river channel: $\Delta = d$ [33]. It should also be noted that formulas of the types of (7) and (8) do not include the hydraulic slope S_f; however, some formulas of the type of (7) include the Reynolds number Re.

In general, the dependencies of the type of (7) and especially ones of (8) do not have wide applications in practice. However, they can be successfully used for estimating hydraulic resistance within gauged river sections, where detailed field research is conducted on a regular basis [33, 45].

The third group of explicit formulas, those we singled out and analysed [33], involves six special dependencies to compute the Chézy coefficient values taking into account the effect of the water surface slope S_f and hydraulic radius R:

$$C = f\left(S_f, R\right). \tag{10}$$

Practice shows that it is especially important to pay attention to the hydraulic slope while assessing the hydraulic resistance characteristics in the case of unstable river channels and variability of floodplain morphological characteristics. Usually, the hydraulic resistance to flow in open river channels keeps changing due to the variability of seasonal hydraulic and hydro-morphological conditions. Sometimes, these changes occur unpredictably. At the same time, the purposeful monitoring of the water surface slope enables taking into account the influence of various hydro-morphological factors on hydraulic resistance within river sections where conditions change dynamically. Altogether, the hydraulic slope may be considered a kind of an indirect integral hydraulic resistance characteristic. In particular, the main feature of formulas of the type of (10) is that they contain neither the Gauckler-Manning roughness coefficient n, nor micro- and macro-roughness parameters used to define the Darcy-Weisbach friction factor λ. On the contrary, the water surface slope S_f is often used as one of the key parameters in some empirical formulas to calculate the roughness coefficient n (Bray's, Jarrett's, Sauer's formulas [25, 33]).

Among the formulas of the type of (10) used to compute the Chézy coefficient values it should be noted Matachievitch's, Winkel's, and Altshuhl-U-Van Thein's formulas [33, 44]. As well, provided that the average flow width $B \gg h$ and $R \cong h$, instead of the hydraulic radius R, the dependencies (10) may include the average flow depth h: $C = f\left(S_f, h\right)$.

Currently, formulas of the type of (10) do not have wide applications in practice. Their usage scope is limited to partial cases, mostly such as earthen canals and canalized rivers, small and medium, foothill and plain rivers with relatively stable self-regulating channels. There are also some restrictions relating to values of the hydraulic slope S_f, hydraulic radius R (or average flow depth h), and average flow width B. However, we consider using the dependencies of the type of (10) as a promising approach to computing the Chézy coefficient in gauged rivers, including monitored rivers by means of modern GNSS technology applications. One of key advantages of the approach seems that the accuracy of the water surface slope determination depends mostly on the accuracy of water level measurements being carried out instrumentally and, usually, with relatively high accuracy [33]. This can minimise the influence of human errors while monitoring the water surface slope. Accordingly, being the simplest element in terms of direct measurements of the river flow [33], the ongoing water level measurements can provide a quite reliable underpin to compute the Chézy coefficient properly with using dependencies of the type of (10).

The fourth group of explicit formulas we generalized as:

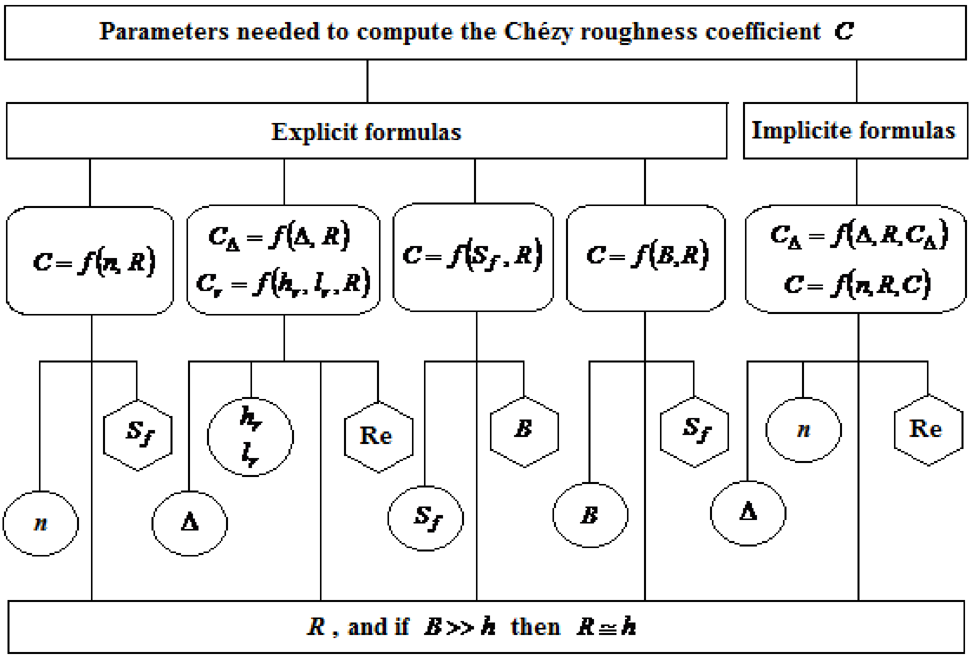
$$C = f\left(B, R\right), \tag{11}$$

where R is the hydraulic radius, B is the average flow width.

These are formulas, where the ratio $\left(B/R\right)$ is used to take into account the shape of a river channel cross-section in terms of determining the hydraulic resistance to open flow [39, 42]. Therefore, these formulas can also be summarised as [33]: $C = f\left(B/R\right)$. We found only four similar formulas that can be attributed to the group (11). Often, in practical applications, a uniform open flow with an arbitrary cross-sectional shape is reduced to a flat flow with the average flow depth h. Then, instead of the hydraulic radius R, the dependencies (11) may include the average flow depth h: $C = f\left(B, h\right)$.

Among the implicit formulas, those we analysed and systemised [33], the most common are formulas of the type of $C_\Delta = f\left(\Delta, R, C_\Delta\right)$. In particular, these are the Colebrook-White [23, 46], Thijse [47], and Powell [1, 29] formulas. Being derived from the Darcy-Weisbach friction factor λ, these dependencies involve also the Reynolds number Re as an additional parameter. Among the implicit formulas, where the Gauckler-Manning roughness coefficient n and hydraulic radius R are used, we might recommend the Agroskin-Zheleznyakov equation [39]. It is a formula of the type of $C = f(n, R, C)$. Moreover, provided that the average flow width $B >> h$ and $R \cong h$, instead of the hydraulic radius R, the average flow depth h may also be used: $C_\Delta = f(\Delta, h, C_\Delta)$, $C = f(n, h, C)$.

Below, Fig. 3 shows the results of our examination of relationships between the Chézy roughness coefficient C and main parameters needed to compute it.



Parameters: n is the Gauckler-Manning roughness coefficient; S_f is the water surface slope; R is the hydraulic radius; B is the average flow width; h is the average flow depth; Δ is the height of protrusions of roughness; h_r is the height and l_r is the length of the bottom ridges (riffles, etc.); Re is the Reynolds number

Fig. 3. Flow-chart showing the relationships between the Chézy roughness coefficient C and main hydro-morphological parameters needed to compute it

According to Fig. 3, the different parameters needed to compute the Chézy roughness coefficient may be divided into two characteristic groups. The first group consists of special parameters presented in formulas of a certain type. These are, for example, the Gauckler-Manning roughness coefficient n, height of protrusions of roughness Δ, water surface slope S_f, and the average flow width B. In the flow-chart (Fig. 3), the circles highlight the key parameters presented in all formulas of a certain type, the pentagons – additional ones, which are only used in some formulas of a certain type. The second group includes the hydraulic radius R or, provided that

$B >> h$, the average flow depth h. One of these parameters is required in all formulas to compute the Chézy coefficient, regardless of their type.

While choosing an appropriate formula for calculating the Chézy roughness coefficient, we should take into account the availability and quality of information about all parameters and focus on the formulas with special parameters whose values are less questionable. Further, depending on flow conditions and factors affecting hydraulic resistance in the river channel section under study, we can choose the best formula of a certain type. If necessary, we can also focus on more detailed research on a special parameter fitting most to solve the task.

In general, as practice shows, regardless of area research, methods and tasks, in modelling and making decisions under data uncertainty, it is important to consider all available information [7, 48-53]. This allows you to implement a comprehensive and holistic approach to solving the problem. Taking into account all available information can be especially useful at the stage of preliminary research, when priorities have not been yet sufficiently identified. Engaging all available information can be useful in data analysis and their arrangements, as well as in modelling, computing of model parameters, and in decision-making processes, including final decision-making stages.

Supporting the comprehensive and holistic approach to hydraulic resistance research, we propose performing the Chézy coefficient calculations using an artificial neural network (ANN). Among the priority tasks needed preliminary solving to achieve that we consider the problem of correct data arrangements to train an ANN. In this research, we tried solving the problem of correct data arrangements to train ANNs being elaborated to calculate the Chézy coefficient on the example of an ANN of direct propagation with one hidden layer and a sigmoid logistic activation function. The main purpose of the study was to develop general rules for the formation of training and test data samples when creating ANNs to compute the Chézy coefficient under parametric uncertainty. To achieve the aim of the study, the following main objectives were set and carried out: (1) generalization of the problem relating to computing the Chézy roughness coefficient, including defining and studying of the subject area; (2) data processing and analysis relating to key parameters defining the Chézy roughness coefficient values; (3) modelling of the ANN to compute the Chézy roughness coefficient, opting of the ANN components and its structure; (4) supervised learning (training and testing) of the proposed ANN with processing examples based on using sets of paired inputs and desired outputs learning; (5) analysis of obtained results with detecting challenges and difficulties relating to computing the Chézy roughness coefficient values by means of the proposed ANN, and outlining ways of their overcoming.

3. Materials, methods, main assumptions, and constrictions of the study

This study keeps a continuation of our previous work [33]. There, on the basis of different pieces of literature, we reviewed, analysed, and systematised a wide set of the well-known and frequently cited empirical and semi-empirical formulas and dependencies, which might be used to compute the Chézy roughness coefficient in cases of open river channels taking into account application limits in term of hydro-morphological conditions. In order to clarify some of the problematic issues related to the use of the reviewed formulas, we have more carefully revised pieces of classical literature on open channel hydraulics [1, 3, 4], reference books, tutorials, and manuals [24, 25, 39-42], articles highlighting the results of original research on estimating hydraulic resistance [2, 19-23, 26, 27, 30-32, 34, 36, 37, 43, 44], including materials of articles devoted to the peculiarities of computing the Chézy roughness coefficient values [28, 29, 35, 46, 47], as well as publications on mathematical modelling of uniform and non-uniform water flow in open channels [7-14, 16-18]. In total, we analysed more than 40 different formulas that can be used while the Chézy roughness coefficient calculating, and revealed the specifics of the use of different formulas depending on the available data, limitations and conditions of their practical applications. In addition, there was studied the problem of the origin, propagation, estimation, and overcoming of the uncertainty of the parameters included in these formulas and hydraulic models [48, 49, 51, 53, 54]. In particular, this research allowed clarifying the tasks of defining and studying the subject area (1), and data processing and analysis relating to key parameters defining the Chézy roughness coefficient values (2).

When researching, we used different methods within the holistic approach to the problem under study [55-59]: historical method; method of dialectical cognition and generalised scientific methods of theoretical and empirical research; heuristic methods; methods of analysis and synthesis; methods of expert evaluation and comparison; methods of formalization and modelling. Moreover, we used modern methods of intelligent data analysis [60, 61], methods of decision making under uncertainty [7, 49, 50, 52, 62], as well as methods and models of artificial intelligence, including ones relating to development and application of ANNs to solve various application problems [63-72].

This article is devoted to solving the problem of correct data arrangements and the development of general rules for the formation of training and test samples of data to train ANNs being planned to be elaborated to compute the Chézy roughness coefficient taking into account the parametric uncertainty of data on the hydro-morphological factors and parameters characterizing the hydraulic resistance in open river channels. The problem is solved on the example of an ANN of direct propagation with one hidden layer and a sigmoid logistic activation function. The training of the ANN and its testing is planned to be carried out taking into account the following hydro-morphological parameters: the Gauckler-Manning roughness coefficient n and water surface slope S_f; the average flow width B and depth h; the height of protrusions of roughness Δ and hydraulic radius R. It is assumed that multicollinearity between the parameters n, S_f, B, Δ, h, and R is absent or can be neglected. Taking into account the relationships between the Chézy coefficient and the defining parameters (Fig. 3), samples of input variables (x_1, x_2) are prepared. Using them, the ANN, according to the algorithm shown below in Fig. 4, calculates the Chézy coefficient $C = f(x_1, x_2)$ as a dependent variable, where, $x_1 \in \{n, \Delta, S_f, B\}$ and $x_2 \in \{h, R\}$ are considered as independent variables. As a result, the ANN of direct propagation with one hidden layer and a sigmoid logistic activation function performs approximation of continuous $C = f(x_1, x_2)$ functions. The training of the ANN is carried out on the learning samples (x_1, x_2, C) using the method of inverse error propagation [65].

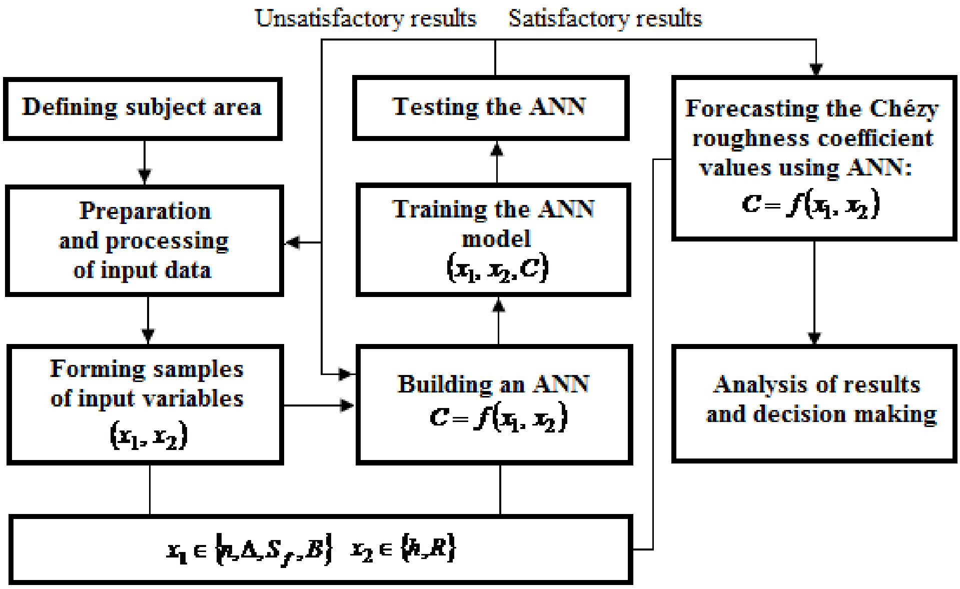


Fig. 4. Flow-chart showing the algorithm of computing the Chézy roughness coefficient C values by means of the ANN under study

To train and test the ANN, a limited amount of field data on hydro-morphological characteristics was used. They related to two channel sections on the Dnieper River (within the city of Kyiv and downstream of Kyiv), the Desna River section near Chernihiv, and the Pripyat River section near the town of Turiv. These areas are characterized by a straight earthen channel with a simple cross-sectional shape and calm current (the Froude number, Fr << 1). Training and testing the neural network was carried out within the following limits of change of hydro-morphological parameters: the water discharge Q = 48.8 ÷ 3665.0 m^3/s; the average flow velocity $V = Q / A$ = 0.336 ÷ 0.968 m/s, where A is the cross-sectional area of the flow (m^2); the water surface slope S_f = 0.000036 ÷ 0.00016; the average flow depth h = 1.0 ÷ 6.2 m; the average flow width B = 122.0 ÷ 611.0 m; the Gauckler-Manning roughness coefficient n = 0.027 ÷ 0.045; the Chézy roughness coefficient C = 27.0 ÷ 48.1 (Table 1). Field data regarding these hydro-morphological parameters were taken from the Hydrological Yearbooks of the Central Geophysical Observatory named after Boris Sreznevsky [73].

Table 1 – Hydro-morphological data used in the ANN training and testing

Rivers, channel sections	Q (m^3/s)	A (m^2)	B (m)	h (m)	$S_f \cdot 10^3$	n (s/m$^{1/3}$)	C (m$^{1/2}$/s)
Dnieper, Kyiv (training)	545.1	1125	375	3.0	0.045	0.029	41.8
	1433	1768	393	4.5	0.067	0.028	46.7
	1842	1988	398	5.0	0.074	0.027	48.1
Dnieper, Kyiv (testing)	1082	1551	388	4.0	0.060	0.028	45.2
	787.2	1336	382	3.5	0.052	0.028	43.6

Rivers, channel sections	Q (m^3/s)	A (m^2)	B (m)	h (m)	$S_f \cdot 10^3$	n ($s/m^{1/3}$)	C ($m^{1/2}/s$)
Dnieper, downstream of Kyiv (training)	657.4	1956	575	3.4	0.046	0.045	27.0
	1123	2403	586	4.1	0.054	0.040	31.4
	3665	3787	611	6.2	0.079	0.031	43.7
Dnieper, downstream of Kyiv (testing)	1763	2858	595	4.8	0.063	0.036	35.6
	2601	3320	604	5.5	0.071	0.033	39.7
Desna, Chernihiv (training)	188.0	501.8	125	4.0	0.036	0.041	31.1
	249.4	580	129	4.5	0.040	0.040	32.2
	403.7	742.4	135	5.5	0.046	0.039	34.2
	497.5	826.3	138	6.0	0.049	0.038	35.1
Desna, Chernihiv (testing)	321.2	660.3	132	5.0	0.043	0.039	33.3
Pripyat, Turiv (training)	48.8	122	122	1.0	0.16	0.032	31.6
	89.0	195.4	130	1.5	0.128	0.033	32.9
	248.6	437.3	146	3.0	0.087	0.034	35.1
Pripyat, Turiv (testing)	136,3	273	136	2.0	0.109	0.033	33.8
	189.7	353.8	142	2.5	0.097	0.034	34.5

In order to correctly use the actual data for training and testing the ANN, they were normalised. Numerical data were converted in such a way as to obtain their model values varying in the range between 0 and 1. In particular, for the purpose of normalization, the parameter B was replaced with a ratio $S_f \cdot B \cdot h^{-1}$ [74]. Instead of parameters V, h, C, model characteristics $V \cdot 10^{-2}$, $h \cdot 10^{-2}$, $C \cdot 10^{-2}$ were considered. Parameters S_f and n remained unchanged. Training data samples consisted of normalized values of the characteristics obtained with uniform linear interpolation in the vicinity of the observed values of parameters. The observed values that were used as test examples were not included in the training samples.

4. The ANN used in the study

4.1. The ANN architecture

Usually, a multilayer direct propagation ANN (a multilayer perceptron) with a nonlinear activation function to approximate continuous functions is used. Such a neural network is considered as a hierarchical structure in which neurons are structured in layers. In a fully connected ANN, each neuron in one its layer is connected to all neurons in its next layer. Neurons of the input layer in such networks transmit input signals to the first hidden layer without converting them. In hidden neurons, sequentially, layer by layer, there is a nonlinear conversion of signals. Each network neuron produces a weighted sum of its inputs, passes this value through the activation function and gives the output value. Signals from the last hidden layer arrive at the neurons of the output layer, which eventually form the ANN response [63-65, 67, 71, 72].

In practice, it is often used one or two hidden layers [64, 65]. It is known that the perceptron with even one hidden layer is a very powerful computing system [63]. In turn, an additional hidden layer can significantly increase the complexity of calculations, processing time, and the risk of the ANN retraining [71, 72, 75].

The ANN we used to compute the Chézy roughness coefficient $C = f(x_1, x_2)$ is a fully connected direct propagation neural network with one hidden layer. The network has 2 inputs, 4 neurons in its hidden layer and 1 neuron in its output layer. Examples of similar networks are given in [63-65, 71, 72, 75].

Fig. 5 shows the neural network architecture as a set of such blocks: the input layer that receives the parameters (x_1, x_2) values and transmitting them (without conversion) to the next layer neurons; the weight matrix $W^1 = \left\{w_{ij}^{(1)}, i = \overline{1,2}, j = \overline{1,4}\right\}$ containing the weight values of inputs for the all hidden layer neurons; the hidden layer containing four neurons, each of which calculates the weighted sum $s_j = \sum_{i=1}^{2} x_i \cdot w_{ij}^{(1)}$, $j = \overline{1,4}$ of its inputs, conducts the sum value through the activation function $F(s_j)$, and transmits the resulting value to the next layer; the weight matrix $W^2 = \left\{w_{ij}^{(2)}, i = \overline{1,4}, j = 1\right\}$ containing the weight values of relationships of the each hidden layer neuron with the output neuron; the output layer containing one neuron in which the weighted sum of its inputs is calculated and the Chézy roughness coefficient value is determined: $C = \sum_{i=1}^{4} F(s_i) \cdot w_{ij}^{(2)}$, $j = 1$.

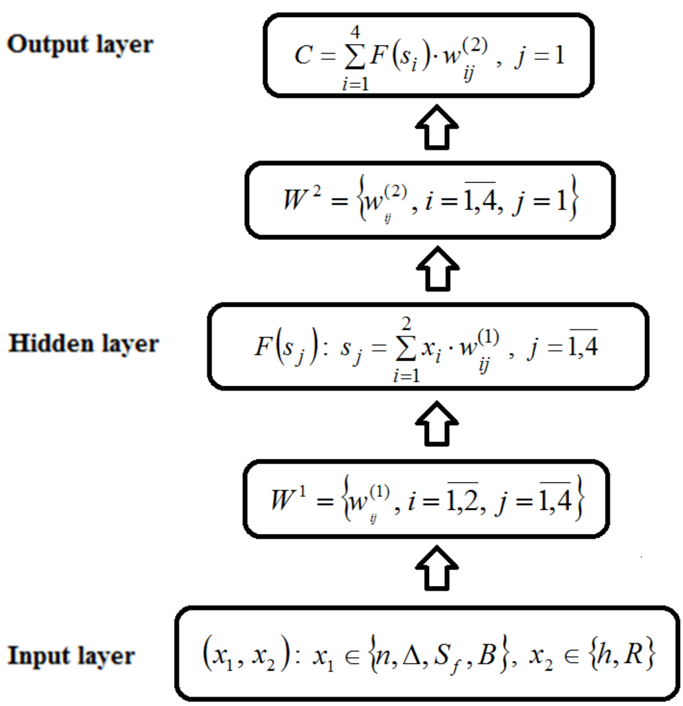


Fig. 5. Flow-chart showing the architecture of the ANN under study

The ANN uses a direct propagation network model with a linear source neuron; the sigmoid (logistic) activation function is applied to the neurons of the hidden layer [64, 65]:

$$F(s) = \frac{1}{1 + e^{-\beta \cdot s}}, \tag{12}$$

where the parameter β influences the steepness of the transition.

The advantage of the function (12) is a quite convenient expression of its first derivative [64, 65]:

$$F'(s) = \beta \cdot F(s)(1 - F(s)), \tag{13}$$

that allows effectively using different algorithms for the ANN learning, where, in turn, parameter β (in our case study β = 1) allows amplifying weak signals and adjusting the learning speed of the network.

4.2. The ANN algorithm to compute the Chézy roughness coefficient values

According to the proposed ANN architecture (Fig. 5) and the recommendations [65], the following algorithm for computing the Chézy roughness coefficient C was developed. It consists in gradually calculations of the outputs of all neurons $y_j^{(1)}$ and $y^{(2)}$ in the network in the direction from the first to the last layer of neurons performed by parameter (x_1, x_2) values and weight matrices W^1, W^2; the neuron of the output layer forms the result of the network work as [65]:

$$y_j^{(1)} = F(s_j),\ s_j = \sum_{i=1}^{2} x_i \cdot w_{ij}^{(1)},\ j = \overline{1,4}, \tag{14}$$

$$y^{(2)} = \sum_{i=1}^{4} y_i^{(1)} \cdot w_{ij}^{(2)},\ j = 1,\ C(x_1, x_2) = y^{(2)}, \tag{15}$$

where $y_j^{(1)}$, $j = \overline{1,4}$ are output values of the hidden layer neurons, $F(s_j)$ is the neuron activation function (12), x_i is an input parameter, $w_{ij}^{(1)}$ are weight coefficients of connections of each input layer neuron with all neurons of the hidden layer, $w_{ij}^{(2)}$ are weight coefficients of connections of each hidden layer neuron with the neuron of the output layer, $y^{(2)}$ is the ANN output value.

4.3. The ANN training

The Python object-oriented programming environment [70-72, 76] was applied to build and train the neuron network (Fig. 5). The ANN training was carried out by adjusting the weights of connections between neurons of all its layers using the inverse error propagation algorithm [64, 65]. The learning factor was assumed to be 0.01. The software implementation of the computational algorithm for learning of the ANN being created to predict the Chézy coefficient C is given in [77].

The ANN weight coefficients were adjusted on a series of real case examples of the parameters (x_1, x_2) values, where $x_1 \in \{n, \Delta, S_f, B\}$, and $x_2 \in \{h, R\}$, in such a way as to achieve a reduction in the error between the predicted (computed) C_p and observed (reference) C_o values of the Chézy coefficient C. The Chézy roughness coefficient reference C_o values were calculated on actual data of hydrological observations with the Chézy formula as:

$$C_o = \frac{Q_o}{A \cdot \sqrt{R \cdot S_f}}, \tag{16}$$

where Q_o is the observed water discharge (m^3/s), A is the cross-sectional area of the flow (m^2), R is the hydraulic radius: $R \cong h$ as $B \gg h$, B is the average flow width (m), h is the average flow depth (m), S_f is the water surface slope.

The initial values of the weight coefficients were set randomly, near to zero. At each iterative step (epoch in learning), at the ANN entrance, in turn, training examples were input and the output values of the neural network were computed, which were further compared with the reference values with error estimating. The network error was also calculated for the hidden layer neurons. The obtained error values were used to recalculate weight coefficients according to the inverse error propagation algorithm [64, 65, 71, 72]. Then, the transition to the next learning epoch was made or, when the required number of epochs was performed or the computational error amounted to an acceptable value the algorithm was stopped.

The ANN learning algorithm includes the following sequence of steps implemented in the program code of the Python module *C_ANN_training.py* [77]:

1. Initialization of the ANN parameters.
2. Direct course of calculations.
3. The reverse course of calculations.
4. Checking the ANN training adequacy.
5. Storage of the ANN learning outcomes.

The proposed computational algorithm implements the classical approach to learning multilayer neural networks using the method of inverse error propagation [63-65, 67, 72, 78]. Herein, the number of learning epochs, the value of learning speed, the amount of learning error, the number of neurons in the hidden layers of the network, etc. are selected by the developer (expert) empirically applying a trial-and-error procedure. Further, the ANN training results are used in predicting the values of the Chézy coefficient C by means of a computer program developed on Python (See, the Python module *C_ANN_calculating.py* [77]).

4.4. The ANN testing

The ANN testing was carried out according to the actual data of hydro-morphological observations (Table 1), which were not used in the network training.

The testing procedure consisted of a comparison of the actual (observed, gauged) Q_o and forecasted (predicted) Q_p water discharge values:

$$Q_p = C_p \cdot A\sqrt{R \cdot S_f}\ , \tag{17}$$

where C_p is the predicted (computed by means of the ANN) Chézy roughness coefficient value (Table 2).

Table 2 – The identified (observed) and computed (predicted) Chézy coefficient C values

Rivers, channel sections	The Chézy roughness coefficient values ($m^{1/2}/s$):	
	C_o, calculated (identified) with the formula (16)	C_p, computed by means of the ANN
Pripyat, Turiv	45.0308	36.8622
Pripyat, Turiv	35.4731	36.8624
Desna, Chernihiv	33.1753	36.8632
Dnieper, Kyiv	33.8147	36.8623
Dnieper, Kyiv	39.6453	36.8624
Dnieper, downstream of Kyiv	43.6760	36.8622
Dnieper, downstream of Kyiv	34.4313	36.8623

Table 3 shows the actual (observed, gauged) and forecasted (predicted) water discharge values, as well as absolute and relative errors of discharge forecasts performed using the tested ANN.

Table 3 – The actual (observed) and forecasted (predicted) water discharges

Rivers, channel sections	Discharges (m^3/s)		Absolute errors (m^3/s)	Relative errors (per cent)
	observed, Q_o	predicted, Q_p		
Pripyat, Turiv	136.3	148.0	11.7	8.6
Pripyat, Turiv	189.7	203.8	14.1	7.4
Desna, Chernihiv	321.2	356.7	35.5	11.1
Dnieper, Kyiv	787.2	664.9	122.3	15.5
Dnieper, Kyiv	1082.0	886.3	195.7	18.1
Dnieper, downstream of Kyiv	1763.0	1830.8	67.8	3.8
Dnieper, downstream of Kyiv	2601.0	2419.9	181.1	7.0

To assess the forecast (predictive) skill of the ANN the Nash-Sutcliffe model efficiency coefficient (*NSE*) was used [79]. At present, this criterion is widely applied for assessment of the predictive power of hydrological models [80].

In our case study, the *NSE* coefficient was calculated as:

$$NSE = 1 - \frac{\sum_{i=1}^{7}(Q_{o,i} - Q_{p,i})^2}{\sum_{i=1}^{7}(Q_{o,i} - \overline{Q}_o)}, \qquad (18)$$

where $Q_{o,i}$, $Q_{p,i}$ are observed and predicted values of water discharges for a river channel section i, $i = \overline{1,7}$; $\overline{Q}_o$ is the mean of the observed discharges $Q_{o,i}$.

It is thought, in the situation of a perfect model with an estimation error variance equal to zero, the resulting Nash-Sutcliffe Efficiency equals 1 (NSE = 1). Values of the NSE nearer to 1, suggest a model with more predictive skill.

In our case study the NSE = 0.9818. It can signify the forecast (predictive) skill of the ANN is quite high.

Moreover, we used an application of NSE in regression procedures (i.e. when the total sum of squares can be partitioned into error and regression components). Herein, the Nash-Sutcliffe efficiency (NSE_r) is considered equivalent to the coefficient of determination (R^2). Fig. 6 shows the graphical illustration of the NSE_r assessing. Its value confirms the high predictive skill of the ANN as well.

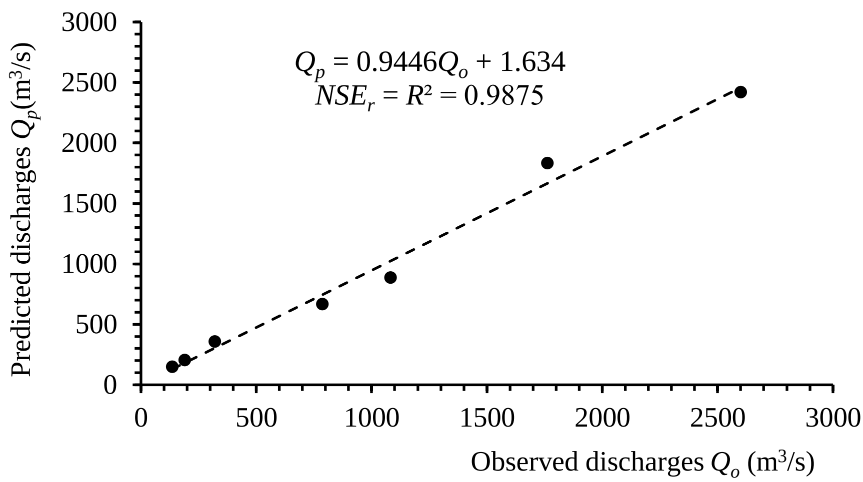


Fig. 6. The graphical illustration of the Nash-Sutcliff model efficiency criterion NSE_r assessment

It should be noted, the NSE coefficient is sensitive to extreme values and might give sub-optimal results when the dataset contains large outliers. Usually, some subjective threshold indicating that the model can be objectively accepted or rejected based on the Nash-Sutcliffe model efficiency coefficient is the NSE = 0.3.

5. Discussion

Predicting the Chézy roughness coefficient values to present the hydraulic resistance to open flows in river channels is a complex challenge burdened by parametric uncertainty of data relating to hydro-morphological parameters. There is also the methodological uncertainty connecting with opting for an appropriate empirical dependency among lots of different empirical and semi-empirical formulas involving various parameters. One of the most promising approaches to overcome such a kind of uncertainty might be the use of artificial neural networks (ANNs). However, this approach needs preliminary solving of the problem on correct data arrangements to train ANNs being developed, taking into account all available information on influential parameters determining the hydraulic resistance, especially in cases as the necessary data are not reliable enough, as well as incomplete, and ambiguous ones.

To study the problem of applying neural networks to predict the Chézy coefficient values using empirical and semi-empirical formulas and dependencies taking into account the parametrical uncertainty of the necessary data, we examined an ANN of direct propagation with one hidden layer and a sigmoid logistic activation function. The ANN testing results indicate the prospects of using such networks in predicting the empirical characteristics of hydraulic resistance to open flows in river channels within certain limitations and applications.

First of all, they confirm that the proposed (not overtrained) neural network of direct propagation with one hidden layer and a sigmoid logistic activation function computing the Chézy coefficient $C = f(x_1, x_2)$ in terms of input variables (x_1, x_2), where $x_1 \in \{n, \Delta, S_f, B\}$, and $x_2 \in \{h, R\}$, can forecast (predict) only some averaged values of the C within a chosen subject area (including rivers, channel sections, and variation of hydro-morphological parameter values) (Fig. 7).

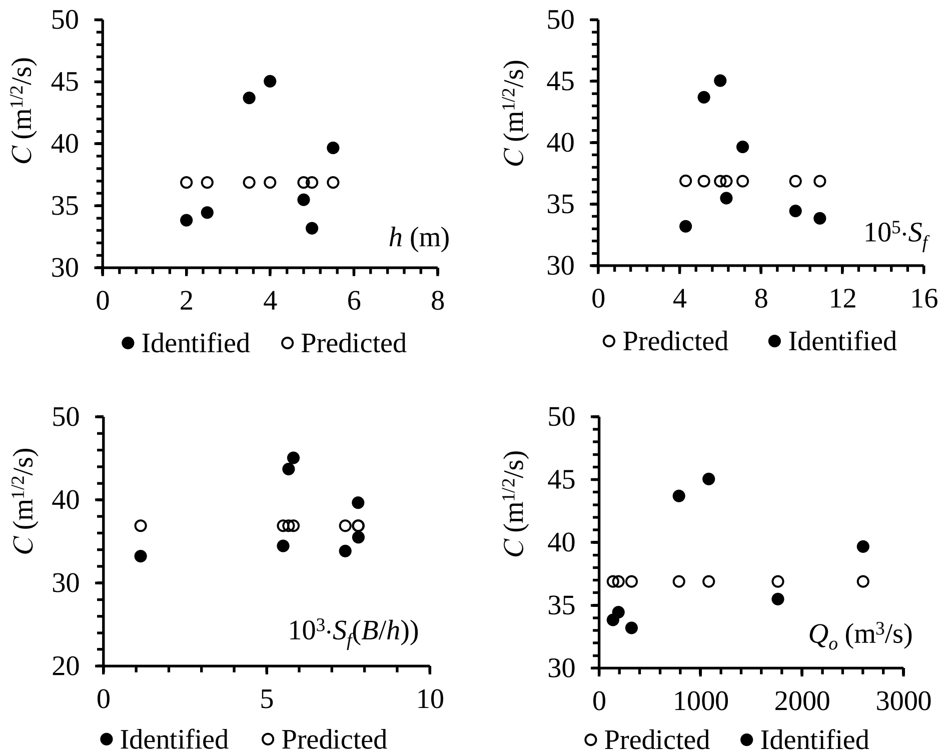


Fig. 7. The Chézy coefficient C values: predicted (computed) by means of the ANN, and identified (calculated) with the formula (16) according to data of hydrological observations

Herein, according to the results of the ANN testing, the prediction relative error of the Chézy coefficient C values depending on the river and the channel section varied from 3.9 to 18.1 percent. Its average value was 10.3 percent.

It should be noted, similar values of the prediction relative error were also recorded for water discharges computed by means of the ANN (See above Table 3). It ranged from 3.8 to 18.1 per cent, and on average was 10.2 per cent. At the same time, the predictive skill of the ANN checked according to the Nash-Sutcliffe model efficiency coefficient (*NSE*), occurred to be high enough (*NSE* = 0.9818). In general, this may indicate that the proposed neural network is able to predict the Chézy coefficient values with sufficient accuracy in terms of practice, provided the correct arrangements of data relating to the subject area. For example, an appropriate hierarchy of data arrangements may include a certain type of a river, its separate channel section with a certain type of fluvial-morphological process [74] and other hydro-morphological peculiarities of river channels [1-4], as well as acceptable variation of hydro-morphological parameter values including seasons, etc. There are also wide opportunities to advance to further improve the neural network. For example, it is possible to apply an ANN with more hidden layers [65-68, 75], as well as combinations of different activation functions for different layers of neurons, including genetic algorithms for approximate computing of initial values of weight matrices, etc. [64-72, 75]. In addition, an ensemble of neural networks can be used to increase the approximation accuracy of the Chézy roughness coefficient output values in conditions of parametric uncertainty of data relating to hydro-morphological characteristics [65, 67], combining several separate neural networks into an ANN with a common architecture.

Conclusions

1. Developed were general rules to the arrangements and the formation of training and test samples of data needed to create ANNs aimed to compute the Chézy roughness coefficient taking into account the parametric uncertainty of findings regarding the hydro-morphological factors and parameters characterizing the hydraulic resistance to flows in river channels. The study was performed on the example of an ANN of direct propagation with one hidden layer and a sigmoid logistic activation function. The training of the ANN and its testing was carried out on the actual data related to several channel sections on the Dnieper River (within the city of Kyiv and downstream of Kyiv), the Desna River channel section near Chernihiv, and the Pripyat River channel section near the town of Turiv.

2. To achieve the aim of the study, the following objectives were set and carried out: (1) generalization of the problem relating to computing the Chézy roughness coefficient, including defining and studying of the subject area; (2) data processing and analysis to provide correct their arrangements in computing the Chézy roughness coefficient values by means of ANNs; (3) modelling of an ANN to compute the Chézy roughness coefficient, including opting of the ANN components and its structure; (4) supervised learning (training and testing) of the proposed ANN by processing examples based on using a set of paired inputs and desired outputs learning; (5) analysis of obtained results with detecting challenges and difficulties relating to computing the Chézy roughness coefficient values by means of the proposed ANN and outlining ways of their overcoming.

3. The algorithm of calculating the Chézy coefficient C as a dependent variable $C = f(x_1, x_2)$ was developed, where, $x_1 \in \{n, \Delta, S_f, B\}$ and $x_2 \in \{h, R\}$ were considered as independent variables (predictors) representing such parameters as the Gauckler-Manning roughness coefficient n, height of protrusions of roughness Δ, water surface slope S_f, average flow width B, average flow depth h, and hydraulic radius R. The training of the ANN was carried out using the method of inverse error propagation.

4. The ANN testing was performed on a comparison of the observed (gauged) Q_o and computed (predicted) Q_p water discharges. To assess the forecast (predictive) skill of the ANN, the Nash-Sutcliffe model efficiency coefficient (*NSE*) was used. We suppose the value of *NSE* = 0.9818 might signify the quite high predictive skill of the ANN. The prediction relative error of water discharges ranged from 3.8 to 18.1 per cent; on average, it was 10.2 per cent.

5. According to the results of the ANN testing, the prediction relative error of the Chézy coefficient C depending on rivers and channel sections varied from 3.9 to 18.1 percent. Its average value was 10.3 percent. Thereby, it was established that the proposed (not overtrained) neural network of direct propagation with one hidden layer and a sigmoid logistic activation function computing the Chézy coefficient $C = f(x_1, x_2)$ can forecast (predict) only some averaged values of the C within a chosen subject area (including different rivers, channel sections, and variation of hydro-morphological parameters). It was shown that the proposed ANN is able to predict the Chézy coefficients with sufficient accuracy for practice, provided the correct arrangements of data relating to the subject area.

REFERENCES

1. Chow, V.T. (1959). Open-channel hydraulics. N.Y., McGraw-Hill, 680 p.
2. Leopold, L.B., Bagnold, R.A., Wolman, M.G., and Brush, L.M., jr. (1960). Flow resistance in sinuous or irregular channels. Physiographic and hydraulic studies of rivers. Geological survey professional paper 283-D. https://pubs.usgs.gov/pp/0282d/report.pdf.
3. French, R.H. (1986). Open-channel hydraulics. N.Y., McGraw-Hill, 705 p.
4. Sturm, T.W. (2001). Open Channel Hydraulics. McGraw-Hill, N.Y., 493 p.
5. Mujumdar, P.P. (2001). Flood wave propagation. Reson 6, 66–73; doi: https://doi.org/10.1007/BF02839085.
6. Cao, Z. and Carling, P. A. (2002). Mathematical modelling of alluvial rivers: reality and myth. Part I: General review. Proc. of the Institution of Civil Engineers Water & Maritime Engineering, 154, Issue 3, 207–219.
7. Veksler, A.B., Ivashintsov, D.A., and Stefanishin, D.V. (2002). Reliability, social and environmental safety of hydraulic structures: risk assessment and decision making. St. Petersburg: VNIIG B.E. Vedeneeva, 591 p. [in Russian].
8. Liang, D., Falconer, R.A., and Lin, B. (2007). Linking one- and two-dimensional models for free surface flows. Water Management, Vol. 160, Issue 3, 145–151; doi: https://doi.org/10.1680/wama.2007.160.3.145.
9. Nikishov, V.I. (2007). From hydraulics of open streams to hydromechanics of river systems. Applied hydromechanics, Vol. 9, No. 2-3, 103–121 [in Russian].
10. Churuksaeva, V., Starchenko, A. (2015). Mathematical modeling of a river stream based on a shallow water approach. Procedia Computer Science, 4th Int. Young Scientists Conf. on Computational Science,Vol. 66, 200–209; doi: https://doi.org/10.1016/j.procs.2015.11.024.

11. HEC-RAS River Analysis System. User's Manual V. 6.0. (2021). US Army Corps of Engineers. Inst. for Water Resources. Hydrologic Eng. Center. Available from https://www.hec.usace.army.mil/software/hec-ras/documentation/HEC-RAS_6.0_Users_Manual.pdf.
12. Kasvi, E., Alho, P., Lotsari, E., Wang, Y., Kukko, A., Hyyppä, H., and Hyyppä, Yu. (2014). Two-dimensional and three-dimensional computational models in hydrodynamic and morphodynamic reconstructions of a river bend: sensitivity and functionality. Hydrological Processes, Pub. online in Wiley Online Library; doi: https://doi.org/10.1002/hyp.10277.
13. Khodnevich, Y.V., and Stefanyshyn, D.V. (2014). Mathematical modelling the conditions of intensification of the riverbed local erosion behind of obstacle that deviates from the shore downstream. Zeszyty Naukowe Inżynieria Lądowa i Wodna w Kształtowaniu Środowiska, Nr 10, Kalisz, 7–18.
14. Qin, X., Motley, M., LeVeque, R., Gonzalez, F., and Mueller, K. (2018). A comparison of a two-dimensional depth-averaged flow model and a three-dimensional RANS model for predicting tsunami inundation and fluid forces. Natural Hazards Earth System Sci., 18, 2489–2506; doi: https://doi.org/10.5194/nhess-18-2489-2018.
15. Dyakonova, T., and Khoperskov, A. (2018). Bottom friction models for shallow water equations: Manning's roughness coefficient and small-scale bottom heterogeneity. IOP Conf. Series: Journal of Physics: Conf. Series 973; doi: https://doi.org/10.1088/1742-6596/973/1/012032.
16. Two-Dimensional Hydraulic Modeling for Highways in the River Environment. Ref. Document. (2019). Publ. No. FHWA-HIF-19-061, U.S. Department of Transportation, FHWA, 301 p. Available from https://portal.ct.gov/-/media/DOT/documents/ddrainage/2-D-Hydraulic-Modeling-Reference-Document.pdf.
17. Gordillo, G., Morales-Hernández, M., and García-Navarro, P. (2019). Finite volume model for the simulation of 1D unsteady river flow and water quality based on the WASP. Journal of Hydroinformatics, 22.2, 327–345; doi: https://doi.org/10.2166/hydro.2019.080.
18. Delis, A.I., and Nikolos, I.K. (2021). Shallow Water Equations in Hydraulics: Modeling, Numerics and Applications. Water 2021, 13, 3598; doi: https:// doi.org/10.3390/w13243598.
19. Fenton, J.D. (2010). Calculating resistance to flow in open channels. Alternative Hydraulics Paper 2. Available from http://johndfenton.com/Alternative-Hydraulics.html.
20. Zidan, A.R.A. (2015). A review of friction formulae in open channel flow. International Water Technology Journal, Vol. 5, No. 1, 43–57.
21. Wang, W.-J., Peng, W.-Q., Huai, W.-X., Katul, G.G., Liu, X.-B., Qu, X.-D., and Dong, F. (2019). Friction factor for turbulent open channel flow covered by vegetation. Scientific reports. 9:5178; doi: https://doi.org/10.1038/s41598-019-41477-7.
22. Zeghadnia, L., Robert, J. L., and Achour, B. (2019). Explicit solutions for turbulent flow friction factor: A review, assessment and approaches classification. Ain Shams Engineering Jour., 10, 243–252; doi: https://doi.org/10.1016/j.asej.2018.10.007.
23. Bacoţiu, C. (2020). The quest for the ideal Darcy-Weisbach friction factor equation from the perspective of a building services engineer. Ovidius University Annals Series: Civil Engineering, Vol.21, No.1, 65–73; doi: https://doi.org/10.2478/ouacsce-2019-0008.
24. Arcement, G.J., and Schneider, V.R. (1989). Guide for Selecting Manning's Roughness Coefficients for Natural Channels and Flood Plains. United States Geological Survey, Water-supply Paper 2339; doi: https://doi.org/10.3133/wsp2339.
25. Coon, W.F. (1998). Estimation of roughness coefficients for natural stream channels with vegetated banks. Prepared in cooperation with the New York State Department of Transportation, 133 p.
26. Machiels, O., Erpicum, S., Archambeau, P., Dewals, B., and Pirotton, M. (2009). Bottom friction formulations for free surface flow modeling, Proc. of the 8th NCTAM Congress, Brussels. Available from http://hdl.handle.net/2268/28208.
27. Dash Saine S., Khatua, K.K., Naik, B., and Mohanty, P.K. (2013). Energy loss for a highly Meandering open Channel Flow. Research J. of Eng. Sciences. Vol. 2 (4), 22–27.
28. Jiménez-Medina, O. (2015). Fórmulas generales para los coeficientes de Chézy y de Manning. Water Technology and Sciences, 6 (3), 33–38 [in Spanish].

29. Achour, B. (2015). Chezy's Resistance Coefficient in a Rectangular Channel. Journal of Scientific Research & Reports, 7(5), 338–347.
30. Stewart, M.T., Cameron, S.M., Nikora, V.I., Zampiron, A., and Marusic, I. (2019). Hydraulic resistance in open-channel flows over self-affine rough beds. J. of Hydraulic Research, 57:2, 183–196; doi: https://doi.org/10.1080/00221686.2018.1473296.
31. Kalinin, A.V. (2019). Dependence of the Chezy coefficient on Froude number. Bulletin of Science and Education of North-West Russia, Vol. 5, №3, 19 p. [in Russian].
32. Zhu, X., Liu, B., and Liu, Y. (2020). New Method for Estimating Roughness Coefficient for Debris Flows. Water, 12, 2341; doi: https://doi.org/10.3390/w12092341.
33. Stefanyshyn, D.V., Khodnevich, Y.V., Korbutiak, V.M. (2021). Estimating the Chezy roughness coefficient as a characteristic of hydraulic resistance to flow in river channels: a general overview, existing challenges, and ways of their overcoming. Env. safety and natural resources, 39(3), 16–43; doi: https://doi.org/10.32347/2411-4049.2021.3.16-43.
34. Jones, G.M., Sanks, R.L., Tchobanoglous, G., and Bosserman, B.E. (2006). Pumping Station Design: Revised 3rd Edition. Burlington, MA: Butterworth-Heinemann, 2855 p.
35. Strupczewski, W.G., Szymkiewicz, R. (1996). Analysis of paradoxes arising from the Chezy formula with constant roughness: I. Depth-discharge curve. Hydrological Sciences Jour., 41:5, 659–673; doi: https://doi.org/10.1080/02626669609491537.
36. Gioia, G., and Bombardelli, F.A. (2001). Scaling and Similarity in Rough Channel Flows. Physical Review Letters, 88 (1): 014501; doi: https://doi.org/10.1103/PhysRevLett.88.014501.
37. Gioia, G., and Chakraborty, Pinaki. (2006). Turbulent Friction in Rough Pipes and the Energy Spectrum of the Phenomenological Theory. Physical Review Letters. 96 (4): 044502; doi: https://doi.org/10.1103/PhysRevLett.96.044502.
38. Stefanyshyn, D.V., Korbutiak, V.M., Stefanyshyna-Gavryliuk, Y.D. (2019). Situational predictive modelling of the flood hazard in the Dniester river valley near the town of Halych. Environmental safety and natural resources, Issue 1 (29), 16–27; doi: https://doi.org/10.32347/2411-4049.2019.1.16-27.
39. Zheleznyakov, G.V. (1981). Throughput ability of channels of canals and rivers. Leningrad, Gidrometeoizdat, 311 p. [in Russian].
40. Shterenlicht, D.V. (1984). Hydraulics. Moscow, Energoatomizdat, 640 p. [in Russian].
41. Baryshnikov, N.B. (2003). Hydraulic resistance of river channels. Tutorial. St. Petersburg, Russian State Hydrometeorological University, 147 p. [in Russian].
42. Baryshnikov, N.B. (2016). Channel flow dynamics. St. Petersburg, Russian State Hydrometeorological University, 342 p. [in Russian].
43. Khatua, K.K., Patra, K.C., and Nayak, P. (2011). Meandering effect for evaluation of roughness coefficients in open channel flow. River Basin Management VI. WIT Trans. on Ecology and the Environment, Vol. 146, 213–224; doi: https://doi.org/10.2495/RM110191.
44. Altshul, A.D., U-Van Thein. (1973). Comparison of formulas without roughness coefficient for determining the average water flow velocity in rivers. Hydrotechnical construction, No 1, 41–42 [in Russian].
45. Korbutiak, V., Stefanyshyn, D., Lahodniuk, O., and Lahodniuk, A. (2020). The combined approach to solving issues of the flood hazard assessment using water gauge records and spatial data. Acta Sci. Pol. Architectura 19 (1), 111–118; doi: https://doi.org/10.22630/ASPA.2020.19.1.12.
46. Giustolisi, O. (2004). Using genetic programming to determine Chézy resistance coefficient in corrugated channels. Journal of Hydroinformatics, 06.3, 157–173.
47. Loukam, I., Achour, B., Djemili, L. (2018). Chezy's resistance coefficient in an egg-shaped conduit. Journal of Water and Land Development, No. 37, 87–96; doi: https://doi.org/10.2478/jwld-2018-0028.
48. Peggy A. Johnson, P.A. (1996). Uncertainty of Hydraulic Parameters. Journal of Hydraulic Engineering, Vol. 122, Is. 2; doi: https://doi.org/10.1061/(ASCE)0733-9429(1996)122:2(112).
49. De Rocquigny, E. (2012). Modelling Under Risk and Uncertainty: An Introduction to Statistical, Phenomenological and Computational Methods. Wiley series in probability and statistics, 484 p.

50. Kuhn, M., Johnson, K. (2013). Applied Predictive Modeling. Springer Science + Business Media: New York, 600 p.
51. Oubennaceur, H., Chokmani, K., Nastev, M., Tanguy, M., and Raymond, S. (2018). Uncertainty Analysis of a Two-Dimensional Hydraulic Model. Water, 10, 272; doi: https://doi.org/10.3390/w10030272.
52. Trofymchuk, O.M, Bidiuk, P.I., Prosiankina-Zharova, T.I., Terentiev, O.M. (2019). Decision support systems for modelling, forecasting and risk estimation. Riga: LAP LAMBERT Academic Publishing, 176 p.
53. Ahmadi, A., Nasseri, M., and Solomatine, D.P. (2019). Parametric uncertainty assessment of hydrological models: coupling UNEEC-P and a fuzzy general regression neural network, Hydrological Sciences Journal, 64:9, 1080–1094, doi: https://doi.org/10.1080/02626667.2019.1610565.
54. McIntyre, N., Wheater, H., and Lees, M. (2002). Estimation and propagation of parametric uncertainty in environmental models. Jour. of Hydroinformatics, 04.3, 177–198.
55. Amini, A. (2001).The Holistic Approach to Development Research. Journal of Interdisciplinary Economics. 2001; 12 (4): 375–394; doi: https://doi.org/10.1177/02601079X00001200404.
56. Brunner, W., and Urenje, S. (2012). The Parts and the Whole: A Holistic Approach to Environmental and Sustainability Education. Visby: Swedish Int. Centre of Education for Sustainable Development. Ecoprint, Tartu, Estonia, 73 p.
57. Vojinovic, Z. (2015). Flood Risk: The Holistic Perspective (Urban Hydroinformatics). IWA Publishing; 1st edition, 270 p.
58. Shwetzer, C., Maritz, A. and Nguyen, Q. (2019). Entrepreneurial ecosystems: a holistic and dynamic approach. Journal of Industry – University Collaboration, Vol. 1, No. 2, P.P. 79-95. doi: https://doi.org/10.1108/JIUC-03-2019-0007.
59. Altman, M. (2020). A holistic approach to empirical analysis: The insignificance of P, hypothesis testing and statistical significance. In D.H. Bailey, N.S. Borwein, R.P. Brent, R.S. Burachik, J.H. Osborn, B. Sims, and Q.J. Zhu (Eds.). From Analysis to Visualization: A Celebration of the Life and Legacy of J.M. Borwein, Callaghan, Australia, September 2017. Springer Verlag. Vol. 313, 233–253; doi: https://doi.org/10.1007/978-3-030-36568-4_16.
60. Berthold, M.R., Borgelt, Ch., Höppner. F., and Klawonn, F. (2010). Guide to Intelligent Data Analysis: How to Intelligently Make Sense of Real Data. London: Springer-Verlag, 407 p; doi: https://doi.org/10.1007/978-1-84882-260-3
61. Combs, J.P., and Onwuegbuzie, A.J. (2010). Describing and illustrating data analysis in mixed research. International Journal of Education, Vol. 2, No. 2: E13; https://hdl.handle.net/20.500.11875/2951.
62. Kochenderfer, M.J. (2015). Decision-making under uncertainty. Theory and Application. With Ch. Amato, G. Chowdhary, J.P. How, H.J. Davison Reynolds, J.R.Thornton, P.A. Torres-Carrasquillo, N. Kemal Üre, and J.Vian. Massachusetts Institute of Technology, The MIT Press, Cambridge, Massachusetts, London, England, 323 p.
63. Cybenko, G.V. (1989). Approximation by Superpositions of a Sigmoidal function. Mathematics of Control, Signals and Systems. Vol. 2, No. 4, 303–314.
64. Rutkovskaya, D., Pilinsky, M., Rutkovsky, L. (2004). Neural networks, genetic algorithms and fuzzy systems. Moscow, Telecom, 452 p. [In Russian].
65. Haikin, S. (2008). Neural Networks and Learning Machines (3rd Edition), Prentice Hall, 906 p.
66. Russell, S.J., and Norvig, P. (2010). Artificial Intelligence: A Modern Approach. 3rd ed. Pearson Education, Inc.: Upper Saddle River, New Jersey, 1132 p.
67. Adamenko, V. and Mirskykh, G. (2012). Artificial neural networks in problems of material objects implementation. Part 2. Design and application specifics, Bulletin of the National Technical University of Ukraine “KPI”, Series – Radio Engineering, Kiev, № 48, 213–221 [In Ukrainian].
68. Artificial Neural Networks – Architectures and Applications. (2013). Edited by Kenji Suzuki. Janeza Trdine 9, 51000 Rijeka, Croatia. 256 p; doi: http://dx.doi.org/10.5772/3409.

69. Piotrowski, A.P., Osuch, M., Napiorkowski, M.J., Rowinski, P.M., and Napiorkowski, Ja.J. (2014). Comparing large number of metaheuristics for artificial neural networks training to predict water temperature in a natural river. Computers & Geosciences, Vol. 64, 36–151; doi: https://doi.org/10.1016/j.cageo.2013.12.013.
70. Muller, A., and Guido, S. (2016). Introduction to Machine Learning with Python, Published by O'Reilly Media, 378 p.
71. Chollet, F. (2018). Deep Learning with Python. Manning Publications Co., 384 p.
72. Trask, A.W. (2019). Grokking Deep Learning. Manning Publications Co, 336 p.
73. Central Geophysical Observatory named after Boris Sreznevsky. Available from http://cgo-sreznevskyi.kyiv.ua/index.php?lang=en&dv=main.
74. Khodnevych, Y.V., Stefanyshyn, D.V., Korbutiak, V.M. (2017). Simulation modeling of condition for initiation of local riverbed erosion. Environmental safety and natural resources. Issue 24 (3-4), 112–123 [In Ukrainian].
75. Keim, R. (2020). How Many Hidden Layers and Hidden Nodes Does a Neural Network Need? Available from https://www.allaboutcircuits.com/technical-articles/how-many-hidden-layers-and-hidden-nodes-does-a-neural-network-need/
76. Lutz, M. (2013). Learning Python. Fifth Edition. Printed in the United States of America. Published by O'Reilly Media, Inc., 1540 p.
77. Khodnevych, Ya. (2022). The software implementation of a neural network computational algorithm for predicting the Chézy roughness coefficient. Available from https://github.com/yakhodnevych/ANN_approximation_C.
78. Novotarskyi, M.A., and Nesterenko, B.B. (2004). Artificial neural networks: calculations. Proceedings of the Institute of Mathematics of the NAS of Ukraine, Vol. 50. 408 p. [In Ukrainian].
79. Nash, J.E., and Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I – A discussion of principles, Journal of Hydrology, 10 (3), 282–290; doi: 10.1016/0022-1694(70)90255-6.
80. Ritter, A., Muñoz-Carpena, R. (2013). Performance evaluation of hydrological models: statistical significance for reducing subjectivity in goodness-of-fit assessments. Journal of Hydrology, 480 (1), 33–45; doi: https://doi.org/10.1016/j.jhydrol.2012.12.004.

The article was received 21.02.2022 and was accepted after revision 31.05.2022

Я.В. Ходневич, Д.В. Стефанишин

ПІДГОТОВКА ДАНИХ ДЛЯ НАВЧАННЯ ШТУЧНОЇ НЕЙРОННОЇ МЕРЕЖІ ПРИ РОЗВ'ЯЗАННІ ЗАДАЧ РОЗРАХУНКУ КОЕФІЦІЄНТА ШОРСТКОСТІ ШЕЗІ ЗА УМОВИ НЕВИЗНАЧЕНОСТІ ПАРАМЕТРІВ, ЩО ВИЗНАЧАЮТЬ ГІДРАВЛІЧНИЙ ОПІР ТЕЧІЇ В РУСЛАХ РІЧОК

Анотація. *Гідравлічні розрахунки та математичне моделювання відкритих течій у руслах річок досі залишаються одними з найактуальніших гідротехнічних задач сучасності з точки зору практики. При їх розв'язуванні, незалежно від теми та мети дослідження, використаних методів тощо, зазвичай приймається та застосовується ряд спрощень та припущень. Крім того, існує низка методологічних, структурних і параметричних невизначеностей, подолання яких вимагає складних емпіричних попередніх досліджень. Перш за все, ці невизначеності стосуються оцінки гідравлічних опорів та встановлення їх чисельних характеристик, які залежать від багатьох факторів, що змінюються в просторі та в часі.*

Однією з найбільш популярних інтегральних емпіричних характеристик, що виражають гідравлічний опір відкритим потокам у руслах річок, є коефіцієнт шорсткості Шезі С. На даний момент існує велика кількість емпіричних і напівемпіричних формул і залежностей для розрахунку коефіцієнта Шезі. Однак, незважаючи на велику кількість емпіричних і напівемпіричних формул і залежностей для його розрахунку, ідеального способу чи методу для однозначного визначення цієї емпіричної характеристики не існує. З одного боку, щоб вибрати відповідну формулу *для розрахунку коефіцієнта Шезі, ми повинні приймати до уваги практичний досвід, заснований на комплексному аналізі варіантів, розглядати різні емпіричні рівняння, які альтернативно використовуються для представлення гідравлічного опору відкритим потокам. З іншого боку, суттєву роль може відігравати повнота та комплексність польових досліджень численних гідроморфологічних факторів і параметрів, що характеризують різні аспекти гідравлічного опору відкритим потокам. Зокрема, оцінка точності обчислення коефіцієнта Шезі за польовими даними, незважаючи на методи та формули, свідчить про те, що точність польових вимірювань параметрів, що входять до обраних формул, значною мірою визначає відносну похибку таких розрахунків.*

У цій статті розглядається проблема упорядкування даних та розробки загальних правил формування навчальних і тестових вибірок даних для навчання штучних нейронних мереж, які розробляються для обчислення коефіцієнта Шезі з урахуванням параметричної невизначеності даних про гідроморфологічні фактори та параметри, що характеризують гідравлічний опір у руслах річок. Задача вирішується на прикладі штучної нейронної мережі прямого поширення з одним прихованим шаром і сигмоподібною логістичною функцією активації.

Ключові слова: *штучні нейронні мережі; коефіцієнт шорсткості Шезі; підготовка даних; гідравлічний опір у руслах річок; параметрична невизначеність*

Ходневич Ярослав Васильович
кандидат технічних наук, науковий співробітник Інституту телекомунікацій і глобального інформаційного простору НАНУ
Адреса робоча: 03186 Україна, м. Київ, Чоколівський бульвар, 13
ORCID: https://orcid.org/0000-0002-5510-1154 ***e-mail:*** ya.v.khodnevych@gmail.com

Стефанишин Дмитро Володимирович
доктор технічних наук, провідний науковий співробітник Інституту телекомунікацій і глобального інформаційного простору НАНУ
Адреса робоча: 03186 Україна, м. Київ, Чоколівський бульвар, 13
ORCID: https://orcid.org/0000-0002-7620-1613 ***e-mail:*** d.v.stefanyshyn@gmail.com