NATIONAL ACADEMY OF AGRARIAN SCIENCES OF UKRAINE
INSTITUTE OF IRRIGATED AGRICULTURE
KHERSON STATE AGRARIAN UNIVERSITY

ARTIFICIAL NEURAL NETWORKS AND THEIR IMPLEMENTATION IN AGRICULTURAL SCIENCE AND PRACTICE

Warsaw – 2019

Authors:

Raisa Anatoliivna Vozhehova, Doctor of Agricultural Sciences, Professor, Corresponding Member of the National Academy of Agrarian Sciences of Ukraine, Honored Worker of Science and Technology of Ukraine, Director of the Institute of Irrigated Agriculture of the National Academy of Agrarian Sciences of Ukraine

Pavlo Volodymyrovych Lykhovyd, Candidate (Ph. D.) of Agricultural Sciences, Senior Researcher of the Department of Scientific and Innovative Activity, Transfer of Technologies and Intellectual Property of the Institute of Irrigated Agriculture of the National Academy of Agrarian Sciences of Ukraine

Serhii Vasyliovych Kokovikhin, Doctor of Agricultural Sciences, Professor, Deputy Director of the Institute of Irrigated Agriculture of the National Academy of Agrarian Sciences of Ukraine

Iryna Mykolaivna Biliaieva, Doctor of Agricultural Sciences, Head of the Department of Scientific and Innovative Activity, Transfer of Technologies and Intellectual Property of the Institute of Irrigated Agriculture of the National Academy of Agrarian Sciences of Ukraine

Olena Yevhenivna Markovska, Doctor of Agricultural Sciences, Assistant Professor, Head of the Department of Botany and Plant Protection of Kherson State Agrarian University

Sergiy Olehovych Lavrenko, Candidate (Ph. D.) of Agricultural Sciences, Assistant Professor of the Department of Agriculture of Kherson State Agrarian University

Oleksandr Leonidovych Rudik, Candidate (Ph. D.) of Agricultural Sciences, Assistant Professor of the Department of Agriculture of Kherson State Agrarian University

Reviewers:

Yurii Oleksandrovych Lavrynenko, Doctor of Agricultural Sciences, Professor, Corresponding Member of the National Academy of Agrarian Sciences of Ukraine, Principal Researcher of the Plant Breeding Department of the Institute of Irrigated Agriculture of the National Academy of Agrarian Sciences of Ukraine

Oleksandr Volodymyrovych Averchev, Doctor of Agricultural Sciences, Professor, Honored Worker of Science and Technology of Ukraine, Vice-rector in Scientific Work and International Activities of Kherson State Agrarian University

UDC 57.087: 517: 519.22

Vozhehova R. A., Lykhovyd P. V., Kokovikhin S. V., Biliaieva I. M., Markovska O. Y., Lavrenko S. O., Rudik O. L. Artificial neural networks and their implementation in agricultural science and practice/ Vozhehova R. A., Lykhovyd P. V., Kokovikhin S. V., Biliaieva I. M., Markovska O. Y., Lavrenko S. O., Rudik O. L. - Warsaw: “Diamond trading tour”, 2019. - 108 p..

ISBN 978-83-66401-06-8

ARTIFICIAL NEURAL NETWORKS AND THEIR IMPLEMENTATION IN AGRICULTURAL SCIENCE AND PRACTICE

The monograph is dedicated to the overview of the core and peculiarities of practical use in agricultural science of the newest data processing tool – artificial neural networks. The issue contains necessary systematized information about the essential properties, methodology, key features of artificial neural networks, provides the reader with information about historical path and incipience of the computation technique, its main advantages, drawbacks, and pitfalls, further prospects of development and implementation in different branches of agricultural science and practice. A special attention is paid to the practical usage of artificial neural networks within various computer software applications for solving the tasks of different compacity and difficulty level related to agricultural science. The publication is targeted on the wide range of specialists in the field of agriculture including scientists, students, and agricultural producers.

Keywords: agriculture, agricultural statistics, artificial neural network, classification, forecasting, modeling, simulation.

CONTENTS

INTRODUCTION

We are standing at the threshold of the Fourth Industrial Revolution, which is mainly instigated by the wide implementation of informational technologies in almost all branches of modern science, technology and industry, including agriculture. Agricultural science and practice in the well-developed countries like the United States of America, Canada, leading countries of the European Union, Australia, etc., is strongly connected with development of informational technologies for their purposes. And the positive effect of informational technologies on the agricultural sector of these countries is not a debatable question: implementation of these innovative technologies caused a huge leap of agricultural production in the above-mentioned countries. Joint applying of the achievements of modern agricultural science and informational technologies opens new possibilities for further development of both branches. Ukraine as one of the leading agricultural countries of the world should also attend to the experience of scientific innovations connected with deep percolation of informational technologies, and reap the benefits of the conceptually new way of conducting agriculture.

First of all, informational technologies connected with simulation and forecast of the processes, which take place in agricultural ecosystems, are developed and presented to farmers and scientists in the form of various computer software applications, which are easily to understand and use even for people who are not very good in complicated mathematical statistics. These applications are used for better agricultural resources management directed to promotion of environmental sustainability and preservation of natural resources simultaneously with stable development of agricultural production (for example, water and land management). Besides, various models of simulation and prediction of the natural processes in biosystems under the impact of agricultural activities of mankind are successfully used for better planning of agricultural production, understanding of the current tendencies in environmental changes both on the local and global scale, forecasting and prevention of possible negative phenomena of biogenic and anthropogenic origin, estimation of the economic prospects, etc. Lots of the achievements of modern informational technologies are also used in automation of the technological processes during the crop's cultivation, for example, irrigation management, fertilizers application, light, air humidity and temperature management in green-houses, etc. Informational technologies related to versatile, accurate simulation and prediction of natural processes, crops productivity under the various cultivation technology treatments and weather conditions, efficient resources management and economic prediction are of the most importance and interest to modern agricultural science. One of the most well-developed, widely spread, accurate and easy-in-use tool for the above-mentioned purposes are artificial neural networks that are nowadays realized within a number of software applications, both commercial and free. Artificial neural networks are a new word in mathematical statistics, therefore, the mission of this paper is to introduce the method to wide auditory of people who are interested in modern tendencies of agricultural science and practice and have a wish to learn about the peculiarities of their use in solving different questions related to efficient farming.

CHAPTER 1

ARTIFICIAL NEURAL NETWORKS: THE OVERVIEW

1.1. Theoretical Basis and Historical Narrative of the Artificial Neural Networks

Artificial neural networks (further in the text – ANN) are one of the methods of mathematical statistical computing, which is used for holding and processing complex data inputs, mainly very huge and exuberant. It is believed that the method was inspired by the biological neural networks that are the general way of thinking and taking decisions in animal and human brains. The principle of the brain functioning, which might be summarized to the complex system of comparatively simple interconnected units (neurons) that together can handle and process a number of complicated and versatile operations and tasks with various objects, is the basis for the idea of emulating such system in an artificial environment by the means of mathematical algorithms and computer power. The main distinctive feature of ANN is a possibility of "learning" (or, as some authors claim, "training") to take "the right decision" on the basis of previous "experience" or "knowledge" (which is set up by a number of examples or samples) or on the basis of settled in advance parameters and computation algorithms. For example, you can create an ANN that will differentiate wheat grain with barley grain by taking into account a number of specific characteristics, which are common for the concrete specie of grains by teaching it what kernel type refers to wheat, and what type refers to barley. So, in the example the ANN imitates some kind of "analytical thinking" based on the previously conducted training. The main goal of any ANN is solving problems in the way the human brain does. The structure of an ANN also imitates a simplified one of the human brains. It consists of a number of neural nodes, which are interconnected with each other, a number of inputs that make these nodes activated, and a number of outputs that give the results of the ANN consideration referring the settled task. The interconnections between the neuron nodes in the ANN perform the functions of the synapse connections in the brain. The neurons and neuron nodes are grouped in the ANN layers. Modern ANNs can contain one, two, three or even much more hidden layers of neurons, which are also tightly interconnected. The structure of the simple ANN with ten inputs, four hidden layers with eight – two neural nodes, and one output is graphically represented in the Fig. 1.

The earliest ideas on an ANN development appeared in the end of the XIX century. These theories and scientific findings were related to the study of how the human brain works. W. Jones was the first to publish the paper, which contained the description of the human brain working patterns.

The first scientists to provide an artificial neuron model were Warren McCulloch

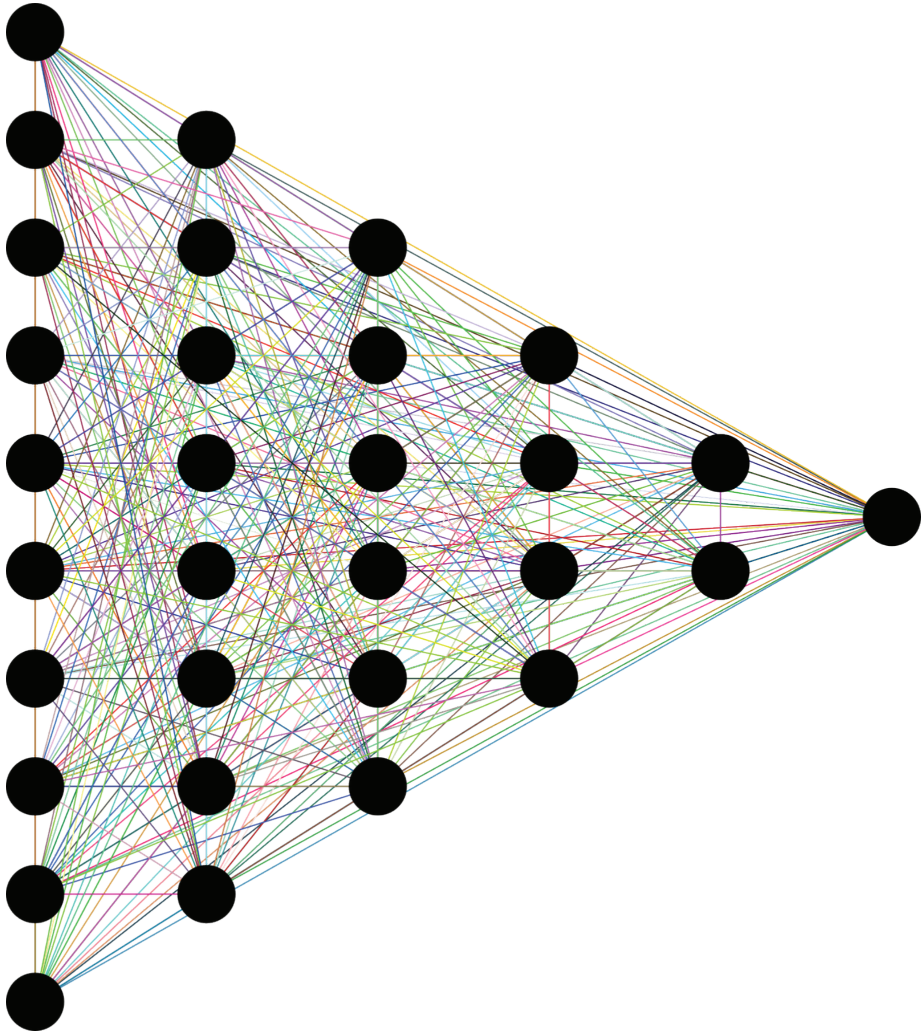


Figure 1. The structure of a simple artificial neural network

and Walter Pitts. They created the functioning model on the basis of mathematical algorithms called threshold logic in 1943 [138]. Their model was the pioneer one, and opened the way for new automatic mathematical computation method. Further their artificial neuron was laid in the basis of Rosenblatt's Mark I Perceptron. McCulloch-Pitts neuron was a product of mutual cooperation of mathematics and neurophysiology. Besides, it stood at the beginning of split in scientific community, which saw the future of computational method quite differently: one part believed that the technology should work on better understanding of the biological processes in the human brain, while another part considered the mission of new method in creating of the artificial intelligence.

Norbert Wiener published a scientific paper, in which he explained an idea of using mathematical models and algorithms for simulation of complex biological processes [145].

In the late 40s of the XX century D. Hebb designed a new learning hypothesis. It was based on the theory of neural plasticity that gave Hebbian's learning algorithm a great advantage – it was not supervised[1] [31]. D.O. Hebb also published the book named "The Organization of Behavior" (1949) where he described the idea of a law for synaptic neuron learning. The Hebb's learning is the first and the simplest learning algorithm for ANN.

The first attempt to put ANNs into life was made by N. Rochester at the IBM research laboratories. However, he did not succeed, and his attempt of running the ANN within the computer environment failed. It is believed that the first ANN-based computer called "Snark" was created by M. Minsky in 1951 at the time of working at Princeton.

Hebbian's ANN learning algorithm was implemented in some calculation machines, including the perceptron[2], which is an algorithm for supervised learning of binary classifiers and is used in prediction functions [36, 109, 110].

In 1958 the book by the famous mathematician J. von Neumann, which represented the earliest thorough scientific study on the interrelation between the mathematics and human brain algorithms functioning, was published. The book had great effect on the development of applied mathematical computing and further improvement of the ANN algorithms used in computer calculations [135].

Rosenblatt created the Mark I Perceptron at Cornell University. The perceptron created by Rosenblatt in 1958 had only one hidden layer of neuron nods, and used a linear algorithm. It was a suitable and quite effective instrument for solving the tasks of classification and weather forecasts [51]. The perceptron used a simple input-output relationship (McCulloch-Pitts neuron) with a step activation function. Besides, Rosenblatt wrote the book devoted to the topic of neurocomputing "Principles of neurodynamic" [149].

---

1 There are several main types of learning used in the ANN: supervised, non-supervised, reinforcement learning. The simplest and the most imperfect way of learning is supervised learning. The learning process is in matching an output to an input by training on the previously set number of examples «input-output» pairs [112]. Non-supervised learning is a learning algorithm when a machine learns and trains on the inputs, which contain unmarked and uncategorized data sets. Reinforcement learning or so-called «approximate dynamic programming» and «neuro-dynamic programming» is different to supervised learning in that there is no need of matching correct inputs to outputs, and learning process is focused on the achievement of balance between using «ready knowledge» and «finding new experience» to provide the best system performance [12, 13, 69]. The basis for the reinforcement learning is a Markov decision process [134].

2 The perceptron is a mathematical model developed to imitate the functioning of a biological neuron. While a biological neuron is activated by electric impulses, an artificial neuron is activated by numeric signals. An artificial neuron receives numeric signals and calculates the output value by summarizing of the input weights. The output numeric is transmitted to other neurons in the network to provide the generalized output of the net.

In 1959 ANN-based model MADALINE (which is an abbreviation of Multiple Adaptive Linear Elements) was developed by B. Widrow and M. Hoff. In 1962 the scientists developed a better learning algorithm. The weight change was estimated as the product of pre-weight value and error, which is divided on the number of inputs. The formula has a drawback: it results in error if the pre-weight value is zero.

In 1963 Petrov conducted a thorough study of the perceptron capacities in solving difficult complicated tasks. The works by Petrov and Bongard were devoted to improvement and correction of the actual at that time perceptron's algorithm [17, 105].

However, the first multi-layer ANN was introduced just in 1965 [65]. Further development of ANN method met an obstacle in low power of computers, which were available in late 60s. Minsky and Papert in 1969 proved that the computers of that time were unable to handle proper functioning of huge ANN, which could lead to getting unfair and inaccurate results of calculations [97]. They also claimed that the perceptron is limited, and it is impossible to use it for solving the tasks connected with invariability and the tasks of non-linear nature. This statement cooled down the interest of the global scientific community to ANNs. However, in 1972 Kohonen and Anderson proposed a conceptually new type of ANNs. Their principal distinction was the feasibility of functioning as a memory [51]. In 1973 Khakimov proposed a non-linear model of an ANN with the synapses based on splines. This ANN was implemented for solving the tasks of medicine, geology, ecology, etc. [71].

But in the late 70s ANN got a new powerful impulse to further development. The new algorithm of backpropagation discovered by Werbos in 1975 solved the "exclusive-or" problem, and made the training of large ANN efficient. Backpropagation (which is an abbreviated version of "the backward propagation of errors") is used to estimate and adjust the weight of neurons in the multi-layer complex ANN by calculating the gradient of the loss function. Backpropagation is considered to be a supervised learning method; however, it might be used in some not supervised ANN, too. It is believed that the algorithm is strongly related to the Gauss-Newton one [52].

The first multi-layer non-supervised perceptron was introduced in 1975. This year Fukusima presented a cognitron – the self-organized ANN, which was developed for the recognition of images. The state of the art of Fukusima's ANN was its capacity to remember all possible states of images and recognition of them on the basis of the ANN's artificial memory.

In 1982 Kohonen presented his self-studying ANN, which was able to solve the tasks of clusterization and visualization. Besides, he developed the self-organizing maps on the basis of the self-studying algorithms. At the same time Hopfield created his own model of an ANN, which was named the Hopfield's ANN. Reilly and Cooper contrived to use the ANN of hybrid nature where every hidden layer uses its own different algorithm for finding decision.

In the middle of 80s, the algorithm of connectionism[3] (parallel distributed processing) grew more and more popular, especially, thanking to the works of Rumelhart and McClelland dated 1986 [111]. In 1986 two groups of scientists, one from the West and another one from the USSR [8, 111], rediscovered and developed the previously proposed algorithm of backpropagation. This event significantly increased the interest to the ANN method of mathematical modeling.

To deal with the issues and researches on the ANN the International Neural Network Society was created in 1987. A year later the Neural Networking Journal started to cover the latest achievements and current problems of the studies related to the topic of ANN.

The first perceptron, which was able to handle and solve non-linear tasks, was the Hecht-Nielsen's one that was created in 1990. The main difference to the Mark I perceptron was in non-linear training algorithm and use of more hidden layers. The Hecht-Nielsen's perceptron was a multi-layer one.

In 1992 the maximized aggregation was introduced to help with invariability related to the least inversion and tolerance to deformation, which apparently occurred in the tasks connected with the recognition of visual images and 3D-objects [142, 143, 144].

Meanwhile, popularity of more common mathematical methods, such as method of support vectors in various modifications, and linear classifiers, considerably increased among the scientists, first of all, because of their simplicity in comparison to ANNs [11, 27]. These modeling methods are widely used in different branches of modern science till nowadays.

In 2006 Hinton created the algorithms of the deep studying of the multi-layer ANNs. In his work he used the restricted Boltzman machine, which is a type of the stochastic recurrent neural network, and might be looked upon as a stochastic generative variant of the Hopfield's ANN [59, 61]. The drawback of the Hinton's ANN is a very slow process of studying. Hinton's ANN used a so called "ancestral pass in training deep multi-layer networks.

Since 2011 a new era of convolution ANNs has been started. The training algorithms realized in the convolution ANN were the first to provide the productivity, which was not far worse from the productivity of human brain [26]. The best examples of modern highly accurate and productive ANNs are evolution networks, viz., developmental networks [141], and their partial case – "where-what networks" [148].

Nowadays, one of the major constraints for further improvement of the ANN technologies is considered to be a hardware imperfection, first of all, limitations of the processors power. Some specialists think that the future of the ANN technology relies upon the introduction of optical chips in computers. Besides, new mathematical algo-

3 Connectionism is an approach in the field of cognitive science that uses ANN to explain the course of mental processes in human brain.

rithms of training the ANNs should also be taken under consideration of the scientific community.

In the end of the historical narrative about development of the ANN technology we should mention the main differences between the conventional (so called, Neumann's) and ANN's computing method. Besides, we should clarify the main feature of the ANN approach and define what do we understand under the term of "learning" in computation.

The key difference lays in the sphere of application of these two methods. While traditional Neumann's computing suits well for solving the tasks, which can be settled by a number of concrete computing algorithms and rules, or for the determined tasks with no possible diversions from the linear path, we cannot successfully use it for more complicated tasks of non-linear nature, which could not be defined by the rules of a number of calculations conducted in the certain sequence. So, if we would like to create an algorithm for so-called "top-down" computation or decision-making method based on Neumann's computing, we would face the problem of impossibility to create the computation procedure, which could take into account all the possible diversions of the linear course while the decision-making process. We cannot put into computation machine all possible force major situations and scenarios for giving it an opportunity of taking correct steps, because this is almost impossible, especially, when we are talking about such complex systems as biological and ecological ones. Therefore, "top-down" based computation program is going to fail every time when it is used for solving the questions of the above-mentioned nature.

The ANN computation uses so-called "bottom-up learning" algorithm. The algorithm is based on the set of attempts to get the necessary result. During the process of these attempts, the computation program is trained to take right steps, and it is getting better from the trial to trial because of gained during this process "experience" in solving the settled tasks. The program turns down the steps that do not lead to the result and tries to find ones that are suitable for achieving the goal. These latter algorithms are "remembered" and fixed in the computation machine's memory. It is interesting that A. Samuel used the "bottom-up" learning method in the programming of his chess computers. As you can see, the method is more similar to the brain's one of taking decisions and primitively imitates human's process of learning. That's why it is more suitable and convenient for dealing with complicated non-linear tasks and computations.

We think that it is quite an obvious thing that the "bottom-up" approach looks more like "learning" process than so-called Neumann's model of "if/else" type. However, we cannot associate the general process of "learning" referring to the human brain neither with "top-down" nor with "bottom-up" model itself. The learning process includes not only abstract operations with some data sets, but it involves interaction with the complex environmental system surrounding the man at the certain circumstances and life situations. So, we see that use of only one methodological approach for creation of the artificial neural network, which would be able to perform learning as the human brain does and take into account the environment, is sentenced to failure.

As you can see, the ANN foresee "learning" in the proper sense of the word more than traditional Neumann's computation, which in its core cannot be trained or taught at all, because it relies upon the rules, which were previously put into the "machine's brain". And here we meet another limitation of the conventional computers: the latter can function only by some logical program on the basis of rules, which are represented by a number of numeric calculations conducted by the settled algorithm. Whereupon ANN-based computers have an ability to "be taught" by the examples or some actions, and then adjust their "activity" with accordance to the "experience" obtained during the learning process. We want to say that the ANN-based machine can be a self-programmed one, while a traditional computer strongly depends on the program support from a human.

Another important thing about ANN and traditional computing is the ability of multi-processing or parallel solving of a number of simultaneous tasks. Conventional computation technique cannot provide such an option, even if it seems to be multitasking nowadays. At the same time, the ANN's architecture is a multi-processing one, parallel working on two and more tasks is natural for it. This feature makes the ANN's architecture more like the human brain. For example, people can look out of the window on what happens in the street and speak on the phone at the same time without any restriction or difficulty. So, the ANN technique does. Of course, it is a great advantage over the conventional computer systems, however, the ANN technology requires development and introduction of special processors.

And the last point we want to stress is that the ANN-based computers are believed to be faster than our conventional ones because of principally different mechanism of functioning and requirements for hardware environment.

1.2. Types of Modern Artificial Neural Networks and Their Key Features

There are several types of modern ANNs. We will stop on each type separately to show its main features, advantages and weak points.

1.2.1. Feed-forward Artificial Neural Networks

It is the first-developed and the simplest type of ANN. The key feature of this ANN is that activation of the neurons goes only in one direction from the inputs of the net to the outputs. A simple scheme of the feed-forward ANN is represented in the Fig. 2.

The main features of the feed-forward ANNs are:

a) Arrangement of the perceptron in layers where the first layer (or so-called "zero-layer") is represented by the inputs, and the last layer represents the outputs of the ANN. The other layers, which are situated in between of the above-mentioned layers, are called "hidden". Every hidden layer is directly connected with the closest two other layers.

b) The neuron nodes, which are placed in the same layer, have no mutual connections. The signals in the ANN are only transported in forward direction. The inter-node connections in the feed-forward ANN do not create a cycle [151].

c) They use a supervised learning algorithm, so their capabilities are limited. Hidden layers of the feed-forward ANN have no connections with an environment, the ANN can be trained only by a number of previously set examples of correct "input-output" pairs.

The most primitive example of the feed-forward ANN is a single-layer perceptron. The neurons in nodes are activated by the sum of the input weights fed to the layer, which is calculated by the chosen mathematical function. If the sum of the weights is more than a set threshold value, the neurons are stressed and transfer the signal farther to the next layer. Otherwise, the neurons remain inactive.

There are several popular activation functions of the artificial neuron. The artificial neuron activation might be descended to the following formula 1:

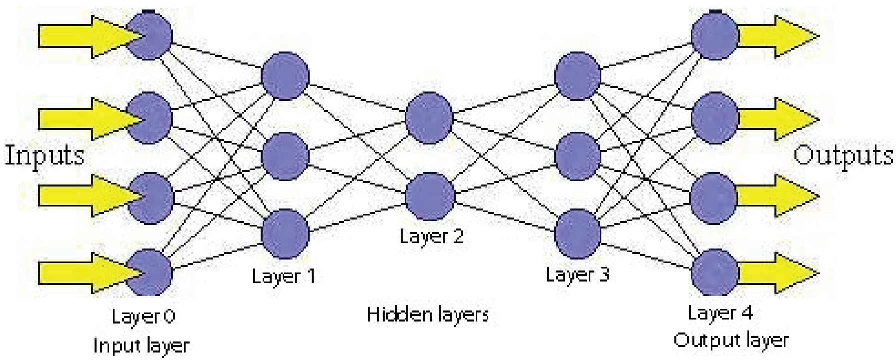


Figure 2. The example of a feed-forward artificial neural network [130]

$$u = \sum_{i=1}^{n} w_i x_i \ (1)$$

where u is the sum of the weights, n is the number of the inputs, w is the vector of the weights, x is the vector of the inputs.

There can be used the step function, which activates the neuron in the case of meeting of the sum of the input weights to the set threshold. There also can be used a sigmoid function in its different variations [57]. In general, the function is represented by the formula 2:

$$S = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1} \ (2)$$

In this case the feed-forward ANN with one hidden layer will be very close to the logistic regression mathematical model.

The feed-forward ANN is more frequently organized in the form of the multi-layer perceptron. There is more than one hidden layer in the structure of such a network. Hidden layers are interconnected with each other in the hierarchic direct way, and have no connections with an outer environment. In the multi-layer perceptron, a sigmoid function is mostly used. The learning algorithm of the multi-layer feed-forward ANN is commonly backpropagation ("the backward propagation of errors"). To put it in plain words, the state of art of backpropagation is use of the gradient descent optimization algorithm with the purpose of adjusting the weights of neurons. It is achieved by calculation of the gradient of the loss function. The errors obtained in the calculation of the output weights are then traced back through the network, and computations are made again and again for as many cycles as is required to obtain such a small error as is only possible.

Backpropagation learning has its imperfections. For example, it is very hard to obtain satisfactory performance in the case of small number of training samples [6]. The speed of convergence and the possibility of ending up in a local minimum of the error function are considered to be other important troubles while implementing backpropagation learning within the ANN technology (en.wikipedia.org).

1.2.2. Competitive Artificial Neural Networks

A feed-forward ANN with its backpropagation learning algorithm is quite helpful in getting predictions and models if we have the sufficient number of training samples, which we can provide to the network, and give it an opportunity of analyzing the considerable amount of correct "input-output" pairs to make the right decision. However, when we are speaking about unorganized raw data sets with no training samples available, first of all, we need to arrange the data. So, we are going to speak about the clusterization issue, which can be successfully solved by using comparatively simple competitive learning algorithm in the ANN. Competitive learning is a type of unsupervised

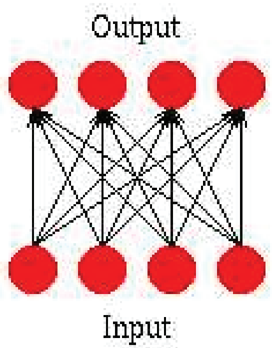


Figure 3. The principal scheme of the Hemming net [130]

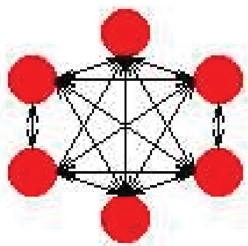


Figure 4. The principal scheme of the Maxnet [130]\

learning algorithm. Its name "competitive" comes from the main principle of the ANN functioning where each neural node "competes" with others for the right to respond to the input activators [111].

A simple competitive ANN is a combination of the Hemming net and the Maxnet. Each net is responsible for performing the concrete task. While the Hemming net estimates the correspondence between the input vector and the weight vector of the perceptron, the Maxnet deals with searching for the perceptron with the maximum value. The Hemming net structural scheme is shown in the Fig. 3.

The Hemming net perceptron calculates the weighted sum of the input vectors and then compares it to the weight vector to conclude whether it is close enough to the input vector. Mathematically we can represent it by the formula 3:

$$\sum wi = |w| \times |i| \times cos\theta \quad (3),$$

where *w* is the weight vector, *i* is an input vector.

The Maxnet network has an interconnected with each other nodes of neurons (Fig. 4).

The Maxnet is that purely competitive neural network where every node struggle to be the "winner" for giving the response to the input. The nodes in the Maxnet translate each other the input weights gradually inhibiting the signals to the point where the signal values of all nodes excluding the one "winner-taking" node will decrease to "0" value. Combining the Hemming net with the Maxnet we are getting typical competitive neural network where the Maxnet connects the top nodes of the Hemming net. This combination gives the highest accuracy in finding the closest weight vector to the input. Learning process of such ANN is based on the moving of the "winner-taking" weight vector towards the input vector. At the same time, all other weight vectors, which are not "winning" the competition, are not activated and remain without any changes. The process is repeated many times until the time when every weight vector reaches the proper cluster. Then the task of clusterization is finished.

Competitive ANNs have several variations. For example, there are ANNs based on the Adaptive resonance theory (ART). The authors of this theory are Carpenter and Grossberg [21, 22, 23, 24]. The ART networks are used for pattern recognition and prediction. The main idea of this ANN is that the process of recognition of objects is the result of interactions between the "top-down" expectations and "bottom-up" sensory information. The "top-down" expectations are represented by the prototypes, which further are compared with the real objects. If the difference between the prototype and real object is below the set threshold, it is considered to belong to the certain category or cluster. Naturally, ART competitive ANNs are not supervised. Besides, there are a number of variations of standard ART ANN:

– ART1 – the simplest variant, which can operate only with binary values [21].
– ART2 – supports continuously fed values [22].
– ART3 – simulates mediator regulation of the synaptic connection [23].
– Fuzzy ART – uses the principles of the fuzzy logic [24].

Another variation of competitive ANN is Kohonen self-organizing map (SOM). This is a non-supervised ANN firstly developed in 1984 for solving the tasks of clusterization. The method is based on the projection of multi-dimensional space into the space with lower dimensions (one or two). Nowadays SOM is widely used for modeling, forecasting, classification, data analysis, computer games development, etc.

Every neural node of the SOM is described by two vectors: weight vector m and coordinate vector r, which is used to mark the node on the map. At first, the primary variant of map is built. Then the learning process starts, and the primary map is reconstructed by the principle of compression because the nodes with a quite similar weights will be placed nearly on the Kohonen map, they will appear in the same cluster. As a result, we will obtain such maps: map of the inputs (the internal structure of the inputs), map of the outputs (the model of the input samples situation), and special maps (the map of clusters and satellite maps) [73].

The work of the SOM includes:

– Initialization.
– Choosing the next vector among the input data set.
– Searching for the best matching unit (BMU) for the chosen vector.
– Computation of the BMU neighbors and training.
– Calculation of the map error.

Kohonen SOM network has its own features and advantages. We can point out the most valuable ones such as tolerance to noisy data, fast learning process, necessity for the ANN to set the number of output clusters. The main drawback is that the final result strongly depends on the settings of the SOM. So, it is almost impossible to study the data without the previously formed [29].

1.2.3. Recurrent Artificial Neural Networks

This is a type of ANN using directional sequence in connections between the elements of the network. This feature makes possible processing of the time series and space links. Recurrent ANNs are used in the tasks of text or speech recognition [53, 84]. The impulse to development of the recurrent ANN was made by Hopfield in 1982,

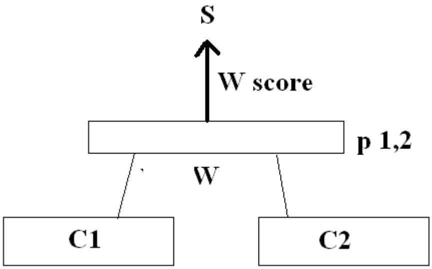


Figure 5. The simplest structure of a recursive neural network

when he introduced his own model of ANN. Now there are a lot of various architectures of the recurrent neural networks. However, nowadays only two of them are widely spread and used: "long-short term memory" (LSTM) networks and "gated recurrent units" networks. LSTM networks are one of the mostly used methods in hand-writing, speech recognition, automated translation and in text-to-speech technologies [54, 127].

The difficulty of the recurrent ANN is that we need to build a separate neural layer for each time period if we are going to take into account every time step. This leads to serious difficulties in calculations and imperfection of the created network, which would be overloaded with neural layers. On the other hand, if we do not take into account all the periods of the time series, we are risking to obtain an inaccurate prediction, especially, if we are speaking about long-term predictions. Various approaches to the recurrent networks try to solve the above-mentioned trouble in different ways.

Completely recurrent neural network is a basic architecture of the ANN of this type. It was developed in the 1980s. The network is constructed with the mutually interconnected neural nodes. The activation threshold of the artificial neurons is changeable in time. All the nodes are divided into input, hidden and output ones. When the network is trained by a supervised learning process, every period of time signal is fed to the input nodes while other nodes finish their activity. The output signals are prepared to be transferred to the nodes of the higher hierarchy. Reinforcement learning of the ANN does not provide any "teacher" sending target signals to the network. The "teacher" is replaced by the reward function or fitness function, which is used to estimate the efficiency of the network performance. Fitness function shows how close is the ANN solution to the aims of the performed task. The function is usually used in genetic algorithms to solve the tasks of finding the optimal solution.

Recursive neural network is a type of ANN, which is developed to work with the data of variable length. They were firstly developed and introduced in the 1990s [41, 123]. A recursive ANN uses hierarchic structures of patterns and samples in the learning process. Artificial neurons with similar weights are activated recursively with accordance to the network structure. The simplest recursive ANN structure could be represented as in the Fig. 5.

Recursive ANNs usually use non-linear activation functions. All the neural nodes of the simple recursive ANN converge to the parental nodes through the matrix of the weights of the hidden layer. The parental nodes weights are calculated as a function 4:

$$p1,2 = \tanh(W[c1c2], \tag{4}$$

where W is the trained matrix of weights ×

This simple type of recursive neural network found its practical use in structuring the sentences [122]. Besides, a number of the improved architectures and approaches were developed for the recursive ANNs, for example, recursive cascade correlation, unsupervised learning network approach, tensor[4] networks, tree echo state networks, etc. [14, 43, 55, 56].

Recursive neural networks are usually trained by using the stochastic gradient descent method. The gradient is determined through the cross-cutting structure of backpropagation of the errors, so, the methodology in its core is a modification of the backpropagation method for time series, which is used in the recurrent neural networks.

Kosko network (Bidirectional association memory) is a type of recurrent ANN developed by Bart Kosko. It was introduced by the author in 1988 [74, 75]. The network is based on two main ideas: adaptive resonance theory of Stephen Grosberg and auto-associative memory of Hopfield. Kosko ANN uses hetero-associative memory: the input vector comes to one neural node, and the output vector comes out from the other node. Same as Hopfield network, bidirectional association memory can generalize and provide adequate reactions even with corrupted data. Adaptive variants of Kosko networks can find the standard sample among the noisy unorganized data set that makes their functioning similar to the latter of the human brain.

The synchronous network consists of two completely interconnected neural layers. The network can be described by the matrix of weights. If the matrix of weights is square and symmetric, the ANN is turning into the auto-associative Hopfield network. All the neurons in the network have their memory and change their status simultaneously.

The continuous variant of bidirectional association memory network has neurons, which can change their status whenever it is needed. The mostly common function for this variant of Kosko network is a sigmoid one.

The adaptive network changes its weights while functioning. Input signals make the network change its energy. After a while short-term memory of the network turns into long-term memory, which makes adjustments of the ANN due to the functioning model.

The competitive Kosko network has the element of competition between the neurons as it was described in the paragraph devoted to competitive ANN.

Jordan network is a type of recurrent ANN, which is obtained by feeding the inputs of the multi-layer perceptron not only with input vectors but also with output vectors

4 Tensor is an object of linear algebra, which transforms the elements of one linear space into the elements of another. Particular cases of tensors are scalars, vectors, bi-linear forms, etc. To put it in other words, tensor is a mathematical tool of putting multi-linear tensors to the resulting tensor.

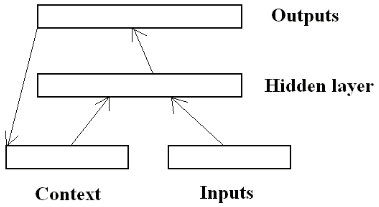


Figure 6. Jordan neural network structure and working cycle

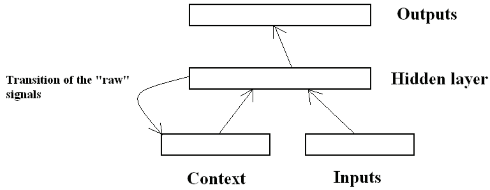


Figure 7. Elman neural network structure and working cycle

with a delay in several tact [68]. The working cycle of Jordan ANN becomes very clear if you take a look at the Fig. 6. As you can see, the information comes to the functioning hidden layer of neurons both from the inputs layer and context layer, which means the layer with the transitional results provided to the outputs from the hidden layer in the previous cycle.

In the mathematical form Jordan network could be represented as follows (formula 5):

$$h_t = s_h(W_h x_t + U_h y_{t-1} + b_h),\ y_t = s_y(W_y h_t + b_y) \tag{5}$$

where x_t is a vector of the input layer; h_t is a vector of the hidden layer; y_t is a vector of the output layer; *W, U, b* are the matrix and parameter vectors, respectively; s_h and s_y are the activation functions used in the perceptron.

Elman network is very alike Jordan ANN with only difference that context layer is formed by the transition of signals from the hidden layer [33]. These ANNs are used in the systems of controlling the moving objects, because their main valuable feature is fixing in the memory of the consequences. The principle of the network functioning is represented in the Fig. 7.

Elman ANN is used in the architecture of more complex recursive auto-associative memory, which is the doubled Elman network. The network uses backpropagation learning algorithm. It is implemented for compression and ciphering the information. Mathematical form of the network is represented below (formula 6):

$$h_t = s_h(W_h x_t + U_h h_{t-1} + b_h),\ y_t = s_y(W_y h_t + b_y) \ (6)$$

where x_t is a vector of the input layer; h_t is a vector of the hidden layer; y_t is a vector of the output layer; W, U, b are the matrix and parameter vectors, respectively; s_h and s_y are the activation functions used in the perceptron.

Echo state networks has the only one hidden reservoir layer with occasional rare interconnections between the neurons. The connections within the reservoir are fixed. The connections with the output layer could be trained. The state of the reservoir is the function of the previous reservoir state and previous state of the input and output signals. Echo state networks are very good in reproduction of the time series [66].

A variation of echo state ANN is a pulsed (spiking) neural network, which is the third-generation network [94]. These networks are the most realistic from the point of view of physiology [49]. There are two major scientific approaches to the pulsed ANNs. The first one sees the profit of the implementation of these type of networks in creation of the computer models, which simulate the real functioning of the human brain for better understanding the principles of its work and diagnostics of the central neural system diseases. The second approach looks upon the spiking ANNs as a powerful tool for analysis of huge amounts of different data.

Pulsed network working cycle consists of several stages. At the first stage, network receives a series of input impulses and generates impulses in the outputs. Every neuron has its value at the period of time. If this value is higher than set threshold, the neuron sends the signal, and then loses its value. The neurons have only two parameters of weight connections: delay time and weight value. There are a number of models of the artificial neurons applied in spiking ANNs [106]. They might be grouped as follows:

– conduction models: Hodgkin-Huksley model [63]; FitzHugh–Nagumo model [39]; Hindmarsh-Rose model [58]; Morris–Lecar model; Wilson–Cowan model [147]; Galves–Löcherbach model [44]; multi-compartment model [50], etc.

– threshold models, viz., method "integrate and work", "integrate and work with leaks".

Spiking ANNs use following ways of information representation:

– Phase (time-depending) – the information about the signal is described by the exact position of impulses in time;

– Synchronous (population) – the information is set by the synchronous activity of different groups of neurons;

– Time period until the first impulse – the information about the signal is set by the time of appearance of the first impulse in any output;

– Serial – the information about the signal is set by the order of impulse obtained at the outputs;

– Interval – the information about the signal is set by the distance between the impulses obtained at the outputs;

– Resonance – the information about the signal is set by the dense consequence of impulses leading to the resonance in the network.

Pulsed ANNs could be learned both in supervised and unsupervised manner [10, 106]. Learning methods used in the spiking neural networks are classified:

1) Unsupervised learning: Spike-timing-dependent plasticity, Growing spiking neural networks, Artola, Bröcher, Singer (ABS) rule, Bienenstock, Cooper, Munro (BCM) rule, Relationship between BCM and STDP rules, General unsupervised learning.

2) Supervised learning: SpikeProp, Deep learning methods, Remote Supervised Method, FreqProp, Local error-driven associative biologically realistic algorithm, Supervised Hebbian Learning.

3) Reinforcement learning: Spiking actor-critic, RL through reward-modulated STDP.

Spiking neural networks have lots of advantages over the networks of previous generations, for example:

1) They are dynamic and they suit best for the tasks with dynamic components.

2) They are multi-tasking.

3) They can provide recognition with prediction.

4) They are trained easily, because it is sufficient to train only output and receptive neurons.

5) They are highly productive in data processing and resistant to noise.

6) They require a smaller number of neurons, because every neuron replaces two neurons (activating and breaking) of conventional neural network.

7) They are fast and have a potential for parallel work.

8) They could be trained in the working process.

However, pulsed ANNs are not as ideal as it might seem at the first sight. They have some drawbacks, such as absence of perfect learning algorithm. Also, it is unreasonable to use spiking neural networks in the systems with little number of neurons.

Long-short term memory (LSTM) is an architecture of a recurrent ANN developed and introduced in 1997 [62]. LSTM network is very good in the tasks of classification, processing and forecasting of the time series. LSTM networks are not receptive to the duration of the time series gaps that gives them a favor in comparison to usual recurrent ANNs. The main distinctive feature of this type of networks is presence of an additional LSTM-block, which is aimed to remember values both for short and long periods of time. These blocks have three or four "gates", the main aim of them is to prevent input of the signals (input gate), control use of the output values (output gate), and overloading of the memory (forget gate). LSTM-blocks could be trained. Backpropagation of the error in time learning process is usually used to train LSTM-blocks. In mathematical form traditional LSTM with a forget gate could be represented as follows (○ operator means the Hadamard product [64], formulas 7-11):

$$f_t = s_g(W_f x_t + U_f h_{t-1} + b_f) \tag{7}$$

$$i_t = s_g(W_i x_t + U_i h_{t-1} + b_i) \tag{8}$$

$$o_t = s_g(W_o x_t + U_o h_{t-1} + b_o) \tag{9}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ s_c(W_c x_t + U_c h_{t-1} + b_c) \tag{10}$$

$$h_t = o_t \circ s_h(c_t) \tag{11}$$

where x_t is an input vector; h_t is an output vector; c_t is a vector of state; W, U, and b – the matrix of parameters and vector; f_t, i_t, o_t are vectors of gates (forget, input and output, respectively); activation functions of the neurons s_g (on the basis of sigmoid function), s_c and s_h (on the basis of hyperboloid tangent).

LSTM with peepholes could be represented as follows (formulas 12-16) [47, 48]:

$$f_t = s_g(W_f x_t + U_f c_{t-1} + b_f) \tag{12}$$

$$i_t = s_g(W_i x_t + U_i c_{t-1} + b_i) \quad (13)$$
$$o_t = s_g(W_o x_t + U_o c_{t-1} + b_o) \quad (14)$$
$$c_t = f_t \circ c_{t-1} + i_t \circ s_c(W_c x_t + b_c) \quad (15)$$
$$h_t = o_t \circ s_h(c_t) \quad (16)$$

Convolutional LSTM is represented by the following set of mathematical functions (formulas 17-21) where * operator means the operation of convolution:

$$f_t = s_g(Wf{*}xt + Uf{*}ht_{-1} + b_f + V_f \circ c_{t-1}) \quad (17)$$
$$i_t = s_g(Wi{*}xt + Ui{*}ht_{-1} + b_i + V_f \circ c_{t-1}) \quad (18)$$
$$o_t = s_g(Wo{*}xt + Uo{*}ht_{-1} + b_o + V_f \circ c_{t-1}) \quad (19)$$
$$c_t = f_t \circ c_{t-1} + i_t \circ s_c(Wc{*}xt + Uc{*}ht_{-1} + b_c) \quad (20)$$
$$h_t = o_t \circ s_h(c_t) \quad (21)$$

The main learning algorithms for this type of ANNs are gradient descent combined with backpropagation of errors. Besides, LSTM could be trained by using the combination of evolution algorithm for the weights in the hidden layers and method of the reference matrix in the output layers. Reinforcement learning is also used in LSTM networks [115].

Recurrent networks of the second level use the weights of the higher levels instead of common weights. Input parameters and parameters of state could be obtained as products.

Gated recurrent units is a mechanism of gates introduced in 2014. It was proved that its efficiency is as good in solving some types of modeling tasks as LSTM [25]. This neural network has no output gate. The architecture of the gated neural network could be presented as follows in mathematical form (formulas 22-24):

$$z_t = s_g(W_z x_t + U_z h_{t-1} + b_z) \quad (22)$$
$$r_t = s_g(W_r x_t + U_r h_{t-1} + b_r) \quad (23)$$
$$h_i = z_i \circ h_{t-1} + (1 - z_t) \circ s_h(W_h x_t + U_h(r_t \circ h_{t-1}) + b_h) \quad (24)$$

where x_t is an input vector; h_t is an output vector; z_t is a vector of renew gate; r_t is a vector of dump gate; W, U, and b are the matrix and vector. Activation functions are on the basis of sigmoid (s_g) or hyperboloid tangent (s_h).

Hopfield neural network is a special case of recurrent ANN with the symmetric matrix of connections. Hopfield network is used as an auto-associated memory, a filter, or for solving the tasks of optimization. The main distinctive thing about this type of ANN is that it works until the balance in the system is achieved, and not until the answer is obtained after the certain number of cycles [139].

The Hopfield network uses a set of McCulloch-Pitts artificial neurons. Each artificial neuron in the Hopfield net can be in two states (it is bipolar): 1; -1. Sometimes the neurons of the Hopfield net are called "spins". The interactions between the neurons in the ANN are described by the function 25:

$$E = \frac{1}{2}\sum_{ij=1}^{N} w_{ij} x_i x_j \quad (25),$$

where w_{ij} is an element of the matrix W. The matrix is symmetric ($w_{ij}=w_{ji}$).

The process of learning of the network is based on remembering of the number of sample vectors, which form the "memory" or "experience" of the network that is further used during its functioning. The learning process of the Hopfield network is supervised. In the case of this ANN the symmetry condition is necessary, but insufficient. The regime of the network functioning is also very important, and it has to be not synchronistic.

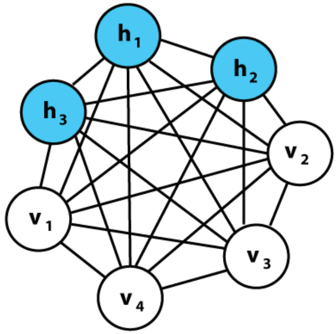


Figure 8. The structure of Boltzmann machine with 3 hidden ($h_{1\text{-}3}$) and 4 visible ($v_{1\text{-}4}$) artificial neurons (the image was taken from en.wikipedia.org)

The main drawbacks of the Hopfield neural network are relatively small volume of the system memory bank, which can be estimated using the formula 26:

$$M \quad \frac{N}{2\ln N} \quad (26)$$

The neural network fails to recognize more images if you try to write them into the ANN memory bank. As you can see, the more neurons you use in the net, the more is its memory bank. However, exceeding number of neurons make the network overloaded that results in bad functioning and issues related to the ANN performance and speed. Storkey in 1997 proposed a new learning rule for the Hopfield network, which increased its capacity in comparison to the commonly used Hebbian's rule [125].

Another disadvantage is that the network cannot guarantee the correct output answer for the task by achieving the balance in the system.

Boltzmann machine is another case of stochastic recurrent neural network [1]. This network can be represented as a generic variant of the Hopfield network. Boltzmann machine is also known as a specific case of Markov random field. The network architecture in a simplified manner is presented in the Fig. 8. The ANN uses simulated annealing learning algorithm, and it is an unsupervised system. Potentially, Boltzmann machine can deal with the most complex and difficult combinatorial tasks. However, the technique did not find wide implementation in practice. The chain of the Boltzmann machines with the limited connections between the neurons (so-called restricted Boltzmann machines) are used in development of deep belief networks. Restricted Boltzmann machines has the neurons, which are not interconnected within the borders of one class. Such ANN architecture was called Harmonium [121] was known since 1986, and became popular after the discovery of fast learning algorithms by Hinton in 2000th. The contrastive divergence training algorithm is used for estimation of the optimal weights of the matrix [60]. This algorithm uses Gibbs sampling for the gradient descent procedure.

Boltzmann machine is a neural network with the determined “energy”. The calculation of the global energy in the system is performed by the formula 27:

$$E=-\sum_{i<j} w_{ij} s_i s_j - \sum_i \theta_i s_i \ (27), \qquad (27)$$

where w_{ij} is a strength of connection between two neurons; s_i describes the status of the neuron; θ_i is the neuron’s threshold.

1.2.4. Convolutional Artificial Neural Networks

Convolutional artificial neural network is a specific type of ANN architecture, which was developed and introduced by LeCun in 1988 [79]. The design of convolutional network is inspired by the vision processes, which take place in living organisms. This ANN type is aimed to perform image recognition and it is based on the principles of deep learning. The idea of convolutional ANN is in interchanges of convolution layers and pooling layers. The structure of the ANN is multi-layer, but has no feedback ties. Convolutional ANN can use various learning algorithms; however, backpropagation is mostly used one. Activation functions of the neurons vary greatly.

The key difference of convolutional ANN is that it uses a convolutional core instead of traditional system, where every neuron has its own weight coefficient. A convolutional core is a matrix of weights, which is moving along the processed layer to form the activation signal for the neurons after every shift. So, the same matrix is used for different neurons of the output layer. The next layer obtained as a result of the above-mentioned process shows the coordinates of the desired feature in the processed layer, and then this information is used for creation of the so-called “feature map”. It should be mentioned that convolutional cores are not set by the investigator, they are created by the neural network during the process of learning.

The subsampling or pooling operation performs reducing of the feature map’s dimensions. The network chooses from a number of neighboring neurons the one with the maximum weight and reduces others to this one. The compaction of the feature map dimensions is performed in such way. Besides, this reduction makes the network more flexible to the scope of the inputs, and makes computations faster.

So, there a lot of layers in the network. First of all, the input signals come to the series of convolution layers, where they are processed alternately by the convolution and pooling (subsampling). The interchange of the layers provides the possibility of feature maps creation, which are further gradually reduced in their dimensions. On the outputs of the convolutional layers several layers of the perceptron are added. The perceptron operates with the final versions of the developed feature maps.

The main layer of the network is a convolutional one. It contains the filters for each input channel, and convolutional core. The weight coefficients are unknown in the beginning. They are determined in the process of learning. The feature of the layer is

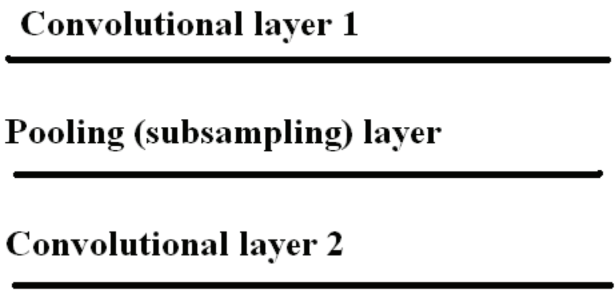


Figure 9. Consequence of the layers in the convolutional neural network

a comparatively little quantity of parameters set in the learning.

The scalar result of every convolution is fed to the activation function. Usually, a non-linear function is used in such type of networks, for example, $f(x)=tanh(x)$, or $f(x)=(1+e^{-x})^{-1}$. The activation layer containing the function is woven into the convolutional layer. Nowadays another function is used more often. The function is called rectified linear unit (ReLU), and could be expressed in the simplified form in the equation $f(x)=max(0,x)$. This function is more preferable because it provides faster training of the ANN [76, 77].

The pooling (or subsampling) layer is used for the non-linear compaction of the feature maps. The mostly used function in this layer is the function of maximum. The pooling layer is inserted after the convolutional layer before the next convolutional layer of the network (Fig. 9).

Sometimes, the functions of mean value and L2-normalization are also used in the pooling layer. Some convolutional ANNs have no pooling layers at all [126].

After the processes of feature maps creation and compaction, the simplified abstract signals are fed directly to the following perceptron (often multi-layer) with its own structure for further processing.

Previously we have mentioned that convolutional networks often use backpropagation learning algorithm. It is a standard one for this type of ANNs. However, some other variants of learning are also possible. For example, there is a learning algorithm called patch-based learning (with autonomic training of the convolutional filters using the auto-association or k-means methodology). There are also several deep learning algorithms used in convolutional ANN's based on the probabilistic mathematical apparatus, sparse coding, etc. [82, 150]. For enhancement of the ANN's performance and stability a number of empirical, viz., drop-out [99], drop-connect, stochastic pooling and artificial data methods, and explicit, viz., early stopping, project gradient descent methods of training are also implemented in convolutional networks.

Advantages of convolutional ANNs can be summarized as follows:

– One of the best algorithms for image classification and recognition, because they do not suffer from the curse of dimensionality (traditional multi-layer perceptrons are also good in this type of tasks, but they are not suitable for dealing with the images of high resolution).

– Better performance in comparison with perceptron in its pure from.

– Convenient paralleling of computations makes possible realization of the algorithm on graphical chipsets.

– Tolerance to shifts and curves of the image.

– Classical (backpropagation) and more advanced methods of learning are possible.

All the above-mentioned features of this type of network make them highly demanded in the spheres of image recognition, especially, face recognition [124]. Convolutional neural networks are prospective components of the ANNs for video recognition and action recognition and detection [5]. These networks are also interesting for dealing with the tasks of natural language processing. Convolutional ANNs are helpful in the medicine. They are used for modeling of possible interactions between the chemical drug compounds and biological substances of the human body. The combination of the convolutional neural network with the Cox-Gompertz proportional hazards model was used in the study of aging and death prediction depending on the different reasons with connection to the biological markers of aging [107]. Maybe less useful but not less interesting implementation of the convolutional neural networks we can find in the development of ANNs that can play checkers and Go games even better than the best human players in the world.

However, convolutional networks are not free from drawbacks and weak points. For example, there are a lot of variables in the ANN, and it is very difficult to set up the network. It is realy hard to develop the network configuration for the certain task.

The ANNs of hybrid structure, for example, convolutional deep belief networks are very prospective for solving various image and signal processing tasks.

1.3. Artificial Neural Networks: Directions of Use

The ANN technology is widely implemented in many branches of modern science and practice. The advantages of this computation method mentioned in the Chapter 1.1 make this technique irreplaceable in a number of various tasks, especially, dealing with natural processes and phenomena, which can hardly be explained, predicted and simulated by the means of conventional mathematics. The range of tasks handled by the modern ANN-based systems is considerable, however, it is possible to descent it to several main directions of the ANN technology usage:

1. The first type of tasks are those dealing with classification. Due to their ability of learning and changing their algorithms by getting new experience ANNs are the best tools for solving of the complicated problems connected with division of the huge massive of scattered data or other objects (including visual and audio ones) into the consanguineous groups depending on some features or peculiarities, which are set for this classification task. The most common task of classification is a recognition of hand-written text, which is generally conducted within the Feed-forward ANN. The ANN is provided with a number of the handwriting examples, which are associated with the concrete number or letter. And on the basis of these examples, it learns to recognize the hand-written text and transform it into the typed variant. This direction of usage is very popular among the agricultural scientists, for example, in conduction of the breeding work, creation of the soil maps, natural resources' status analysis, etc. It is also widely implemented in medicine for the purposes of helping the physicians in precise diagnosis of their patients.

2. The second scope of the ANN technology implementation is connected with predictions. ANNs provide us with a wide range of possible instruments and give us a wide field for choosing from various types of simulation, modeling, and function approximation (like in regression analysis) to obtain the most precise understanding of how the modeled process work and what should we expect in the future. In this case, a particular ANN is trained to provide users with possible variations of the future happenings related to the certain natural or artificial processes on the basis of the previously entered information on these processes in the past. Feed-forward ANNs are mostly used for this type of tasks. Forecasting ability of ANNs is one of the most usable features of the technology among the scientists in almost all branches of modern science and technology. The highest interest in accurate predictions is shown by the economists. They use ANNs in financial predictions, stock market predictions, trading predictions, currency predictions, futures predictions, bond ratings, economic recessions and failures predictions, etc. We use the method in prediction of crop yields under the different natural conditions and cultivation technologies, natural resources' status changes under the anthropogenic activity and agricultural land-use, weather forecasts, programming of the rational use of natural and artificial resources in agri-

cultural production, etc. This set of tasks is considered to be of the most importance for agricultural science and practice, and it is the most developed one at the time.

3. Cluster analysis of the complex and large data sets is another valuable way of the ANN technology use. There are three major types of ANNs for solving the clustering tasks: simple competitive networks, adaptive resonance theory networks, Kohonen self-organized maps (these are just a variation of the simple competitive network). Cluster analysis is also used in agriculture, for example, in breeding science and in the tasks of optimization (logistic, production, etc.). Besides, ANNs are successfully used to conduct the comparatively easy tasks of data filtering within huge and often unorganized arrays of information or other objects.

4. The tasks of association are very useful in restoring damaged or noisy data sets. An ANN is previously trained to "remember" the correct data sets, and when the corrupted ones are inserted in the ANN it returns to the user the last actual version of the unharmed data, which are the closest to the entered data. ANNs are not widely used for these purposes in agricultural science, however, this feature of the networks is very useful and helpful. IT-specialists use Hopfield ANN for the tasks of association and image compression.

5. At the last time the ANN technology is implemented in robotics. The technology is a prospective one in the creation of the machines with artificial intelligence (AI). And if the AI creation sounds like a subject of sci-fi stories so far, the ANN approaches are introduced in development of modern prostheses. One way or another, nowadays ANN-based methods are also used in control engineering[5]. And we think that nobody would be astonished with a self-driving harvester in the field or an idea of a "smart house", which were also created not without ANN usage.

If you are going to use the technology in your scientific and research work you should know key moments on how to get the best outlet from the implementation of the ANN approach and not get into trouble of obtaining incorrect results. So, first of all, you must clearly realize what goal you are pursuing through the application of the ANN technology. You must make your mind and then determine the group of tasks under which your purposes fall. And then you can make a selection of the appropriate type of the ANN for dealing with this kind of task. Besides, you must be careful preparing the data set for your model because ill-conceived and badly arranged inputs and training design lead to mistakes in the model outputs, resulting in failure of all the system. You can try several different learning algorithms and find the best one before the final application of it to your ANN model.

5 Control engineering or control systems engineering is a branch of engineering science, which deals with the questions of application of the automation in an artificial and natural environment to provide sustainable and human-free monitoring and management of the particular systems [131].

1.4. Advantages and Drawbacks of Artificial Neural Networks

Artificial neural networks are getting the increasing popularity in scientific community throughout the world for their suitability and good performance in solving of a great variety of theoretical and practical tasks. They are so loved among the scientists because of some crucial advantages over the other statistical computation techniques. The main advantages of ANNs are the following:

1.First of all, ANNs provide an opportunity of working with extremely big datasets, including enormous amounts of numeric values for other computational technologies. The most interesting thing about it is that ANN's deep learning performance just increases with increased data amounts, that is not fair for other learning algorithms (Fig. 10). Great datasets give ANNs a chance to show their potential, and often ANN is the only one suitable technique to perform the task with incomplete data.

2. Ability to learn and handle large unorganized data sets. ANNs will not fall down even if the information provided for the analysis or modeling is stochastic, noisy or incomplete, while the most used traditional methods, for example, regression and correlation analyses, will not be able to go on with such datasets.

3. Non-linear modeling advantage over the classical statistical methods both in fitting and prediction capacities makes an ANN-based approach in modeling of natural processes and time series predictions the most precise and reliable one [18, 32, 129]. Engagement of several different learning algorithms within one network provides a considerable increase in the ANN performance.

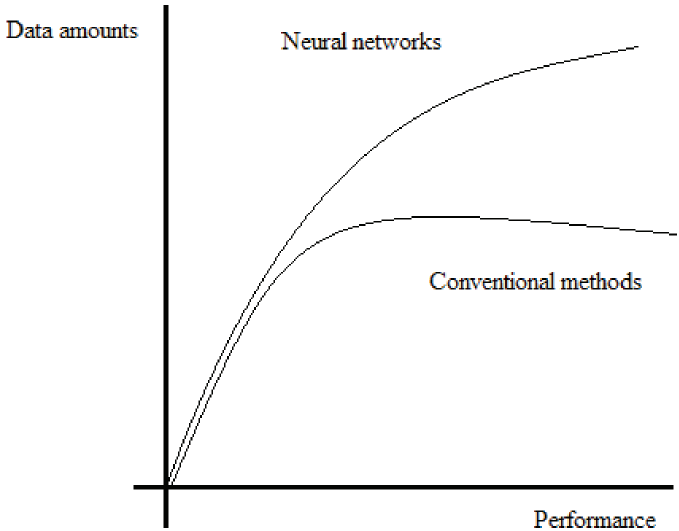


Figure 10. Performance of conventional data processing methods in comparison to artificial neural networks in dependence on the amounts of data

3. Self-organization is another important advantage over conventional statistical methods. This feature makes the life of explorer easier and provides time saving. Besides, self-organized computational system is more insured against the errors, which may occur for the human factor.

4. Artificial neural networks are very good in the tasks of generalization, classification and clusterization. That is why they are widely used in the systems of recognition. simulations are performed.

5. Is it possible to operate fully with a non-numeric information within the traditional computation techniques of mathematical statistics? You know that the answer is "No". But in case of an ANN approach to such tasks there are no hindrances: networks of the specific type can operate with the non-numeric information (See Paragraphs 1.2 and 1.3 of the book for more details).

6. Modern techniques used for the realization of artificial neural networks provide comparatively high tolerance to failures caused by the probable corruptions of the ANN. This guarantees you that your data, in-between computations and simulations will not be lost in case of emergency.

7. Ability to handle parallel processing is another important distinctive feature of ANNs, which make them better than a number of other data processing technologies.

8. Scientists have some prospects concerning the understanding of how the human brain works by deep studying of the ANN technology. Furthermore, the technology is considered to be the most prospective one for the development of artificial intelligence.

Notwithstanding the fact that artificial neural networks are considered to become the most widely used instrument of mathematical statistics in the nearest future, they have a number of significant drawbacks yet. Here we shall place some of them.

1. Powerful hardware is required to make ANNs functioning properly and fast. And this restriction is not so easily eliminated as it seems to be at the first sight. Of course, our computers become more and more powerful each year, however, if we are speaking about solving difficult and complex tasks of modern science, very often we are facing the problem of lack of the computer power to handle the developed artificial neural network properly. The computations consume much time, become less accurate because of the need to simplify the network to provide its functioning. Besides, purchase of the newest and most powerful machines is not always possible for our scientists. Therefore, the issue of high requirements of the ANNs upon the hardware is very sensitive and may become one of the major factors of further inhibition of artificial networks spreading.

2. "Black box" nature of the networks functioning. Very often you will never know how your ANN got its conclusion. And this is also very important disadvantage. Neural networks should be more open and accessible to provide better understanding on how they are operating the data and reaching their output decisions.

3. Sometimes it is also very difficult to understand why this or that architecture of

the ANN would be better for solving the concrete task. So, you can waste a lot of time selecting from a great number of possible architectures the only one, which will provide the best performance.

4. The ANN technique seems to be more difficult than conventional mathematical statistics for most people [129]. This problem could be solved by the development and introduction in colleges and universities of comprehensive educational programs and training courses for people who are desirous to learn how to use the technology for their purposes.

5. Empirical nature of the ANN-based models is a relative disadvantage because almost all traditional methods of modeling are also of empirical nature [129].

6. Sometimes the problem of overfitting may occur. However, it can be solved by the careful development and adjustment of the network [129].

All in all, we believe that the ANN technology is one of the most prospective mathematical methods for handling huge and incomplete stochastic data sets, which require non-linear processing and analysis. The technology should be studied and developed furthermore. The main advances and improvements related to ANN are connected with:

Integration of fuzzy logic into the learning processes.

Further development of the pulsed (spiking) ANNs as one of the most prospective.

Hardware development. Besides, software for the realization of artificial neural networks in computer environment should become as optimized and user-friendly as possible.

Widening of the ANN algorithms application in science and practice, seeking for better learning algorithms and new fields of the technology implementation in real life.

CHAPTER 2

THE OVERVIEW OF COMPUTER SOFTWARE FOR ARTIFICIAL NEURAL NETWORKS REALIZATION

2.1. Software Overview

There are a lot of software applications for creation, training and comprehensive usage of the ANN computation and decision-making technology. These programs provide a vast diversity of possible computer realization of the technology by using various types of ANNs with versatile learning mechanisms suitable for solving a number of theoretical and practical tasks. They differ in the form of the interface, computation facilities (for example, some applications are able only to deal with some kind of information and data, or they can provide a limited specter of services, excluding some valuable ones), system requirements and requirements to the skills and education of user (for example, some programs require good knowledge and experience in programming and mathematics, while others do not require any additional skills or

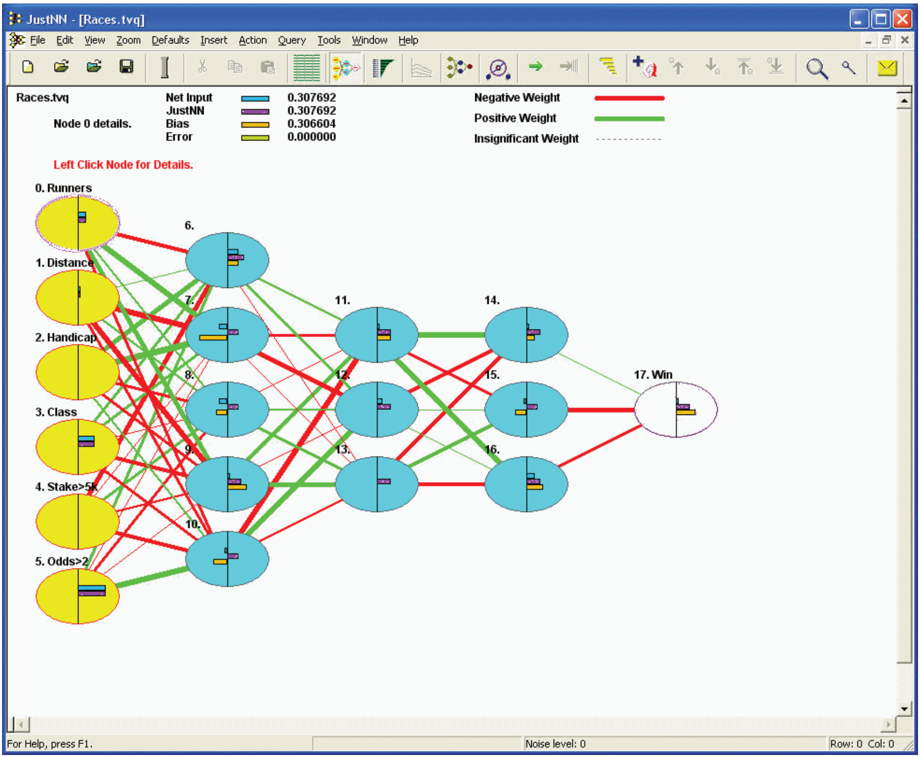


Figure 11. The interface of JustNN software application

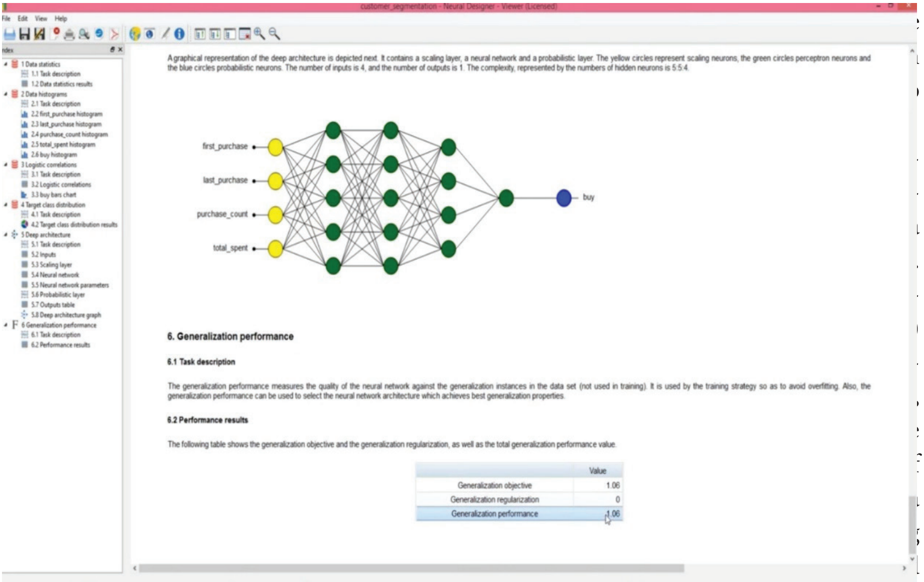


Figure 12. The interface of Neural Designer software application

prediction. JustNN uses backpropagation learning algorithm with a sigmoid function within the feed-forward ANN architecture.

Neural Designer is a computer application designed for usage of the ANN approach in data analysis (Fig. 12). It was developed by Artelnics as a cross-platform program available for Windows, Mac OS X, and most Linux distributions (including such popular as Ubuntu and Mint). Neural Designer is a commercial software, and requires your purchasing the license for a month (250 EURO) or a year (2500 EURO). However, the app has a trial period, and you are able to test it before paying the money. Neural Designer is one of the simplest ANN-based analysis tools on the market. The interface of the program is very clear and user-friendly, and it is easy to work within the application. Neural Designer provides a huge and powerful package of the ANN-based tools, such as creation of large ANNs with a complicated deep learning algorithm to perform tasks of descriptive statistics, diagnosis of data sets, prediction, and prescription. You feed the program with your input data, set the task and simply obtain the output, which is presented in the form of the predictive model. Neural Designer nowadays is the gauge software developed for practical implementation of the ANN-based approaches in data analysis and modeling.

NeuroXL Predictor is an Excel add-in compatible with Microsoft Office 2000-2016 Fig. 13). It was developed to simplify and make easy for everyone use of the ANN approach for predictions and forecast models. NeuroXL Predictor is one of the easiest programs for implementation of the ANN forecast technique to the concrete practical needs. It is really helpful for most people that the add-in does not require any special

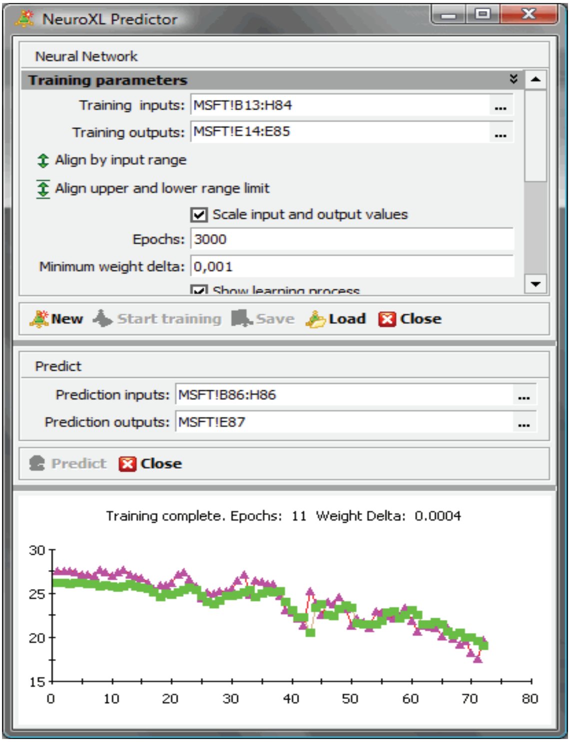


Figure 13. The interface of NeuroXL Predictor add-in for Microsoft Excel 2000-2019

skills or knowledge about the ANN technology to be successfully used. The main disadvantage of the product for some number of potential users is the need of purchasing the license for 100 USD. However, if you just have a look through the price list of the nearest analogs you will find that this price is not as high as it seems to be at the first sight. Besides, you can use NeuroXL Predictor together with **NeuroXL Clusterizer** – another powerful tool, which uses the ANN-based approach for deep classification and clusterization of the data (Fig. 14). Both products use feed-forward ANN structure, have several learning options including a possibility of choice between different types of activation function. If you need an adjustable and flexible tool with a great number of possible learning options and giving more comprehensive and wide feedback you should not use NeuroXL products, because they lack customization options and everything is doing without user's interference. The instruments of NeuroXL are not suitable for studying the ANN functioning too because of above-mentioned reasons: the program is working like a "black box" and you will never know exactly how the results were really obtained during the data processing. However, the application is a valuable staff for those people who need a reliable forecasting program and do not want bother themselves with mathematical complications.

Neural Lab is a free ANN simulator developed for creation, training and testing of various neural networks within the visual computer environment (Fig. 15). The application is integrated with Microsoft Visual Studio. It can be used for prediction, mapping, association, classification and simulation. Neural Lab found its adherents predominantly among the scientists and researchers engaged in engineering, information technologies, economics, etc. The program primarily uses two main activation functions of neurons: sigmoid and (y) = tanh(ay). Neural Lab provides a possibility of the

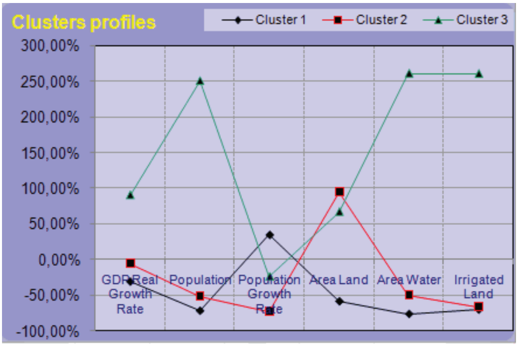


Figure 14. The interface of NeuroXL Clusterizer add-in for Microsoft Excel 2000-2019

creation of feed-forward multi-layer perceptron, probabilistic, or Kohonen network with various learning algorithms including conjugate gradient method, the Levenberg-Marquardt minimization method [95], method of simulated annealing [80], the method of genetic algorithm which is in core not a single algorithm, but a complex of various algorithms and computation techniques [67]. The diversity of training algorithms makes Neural Lab a great and flexible tool for dealing with a vast majority of tasks and calculations in different branches of modern science [81]. However, Neural Lab is not an easy software to use, and requires some experience, programming skills and knowledge of neural networks core to make this application work properly and fully open its computational potential. That is why Neural Lab cannot be recommended for every man wanting to implement artificial neural networks approach in his work.

STATISTICA Neural Networks (together with an extension pack Automated Neural Networks) is one of the most powerful tools of computer realization of the ANN approach to solving the complicated tasks (Fig. 16). The program provides possibilities of working with the newest ANN architectures. It is very useful that a user has an opportunity to adjust the learning process and ANN structure in all details, and control the process of the ANN training and testing. The complex and comprehensive ANN-based approaches are united within the on program with comparatively easy understandable interface, which makes STATISTICA package a big favor. A user is provided with a great diversity of the instruments for editing and analysis of the data sets. One of the useful features of the application is a possibility of being connected to the external database for operating with the data in the “cloud” regime. Some users consider STATISTICA Neural Networks package to be a complicated tool because of a great number of settings available for thorough adjustment of the ANN built. For these types of users, there is an additional package named Automated Neural Networks, which deals with questions of the ANN architecture design and presets “by itself” so that a user would not be puzzled over the tuning and choosing appropriate network parameters. STATISTICA Neural Networks package can solve a great variety of practical and theoretical tasks (regression, classification, cluster analysis, time series prediction, etc.) by using such types of the ANN models: feed-forward multi-layer perceptron, radial basis networks, self-organized Kohonen maps. Besides, you can organize several ANNs in the Network Ensemble to increase the performance of the program

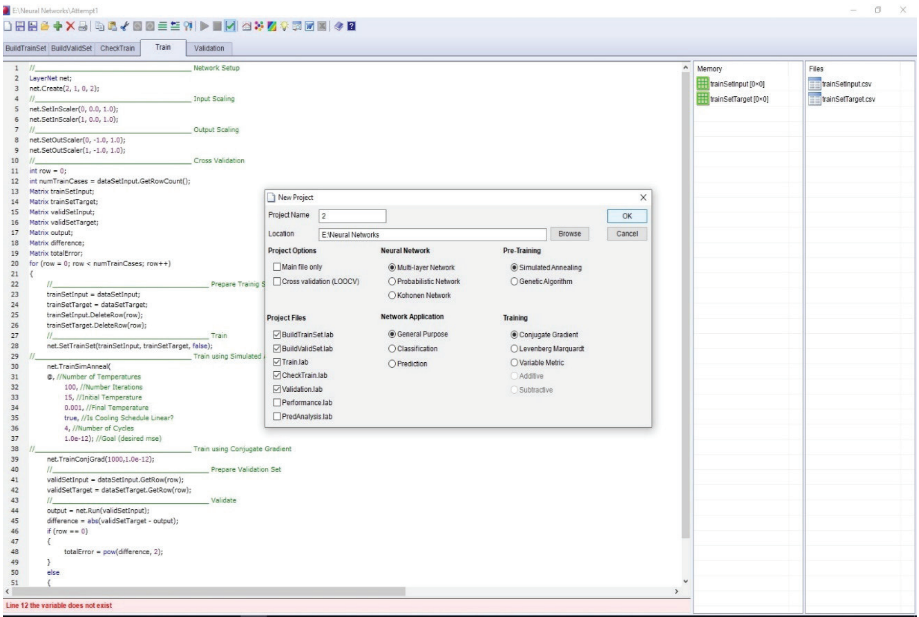


Figure 15. Neural Lab working interface

data analysis. The latter variant is a helpful treatment for noisy data and small data sets. Besides, you can get an "expert advice" from the integrated in the package ANN-analyzing system about your network: is it good enough for the solving task or maybe it needs some improvement or simplifying. A great "plus" of STATISTICA Neural Networks package is the presence of the powerful graphical tools for making the results of the ANN work more understandable and easier for analysis. Users are also provided with the possibility of the ANN productivity improvement and choosing one of the vast diversities of activation functions for the neurons in the ANN, for example, identical, exponential, hyperbolic, logistic (sigmoid) or sinusoidal. It is believed that STATISTICA Neural Networks is equipped with two best learning algorithms, viz. conjugate gradients method and an algorithm of non-linear optimization or so-called BFGS-algorithm. However, you can use some less perfect algorithm, for example, gradient descent, method of k-means, Kohonen's algorithm. If you want, you can stop learning of the ANN whenever you need. After the learning process has been finished you can move on to the testing of the ANN. STATISTICA Neural Networks has a very powerful testing environment. Such indexes as mean square error, confusion matrix and correlation coefficients are calculated automatically. Kohonen network has the window of topological map where user can visually observe the processes of the network activation. Finally, we should admit a useful feature of inserting your data in the previously built network. This option efficiently saves your time. You also can export your results to some external software environment, for example, Microsoft Excel. The source sys-

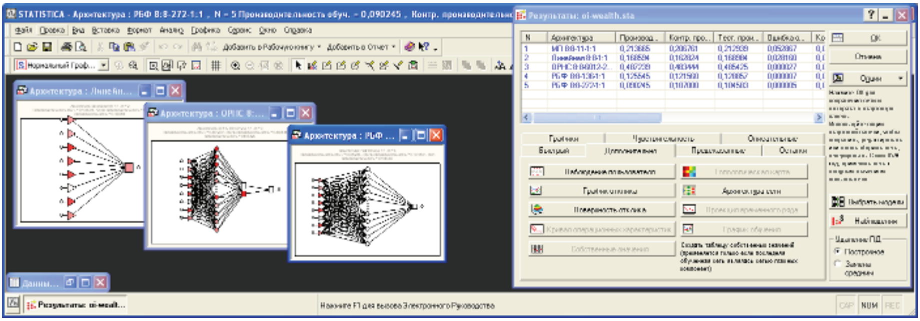


Figure 16. The interface of STATISTICA Automated Neural Networks

tem code of the ANN models can be generated and expressed on the languages C, C++, C#, Java, PMML (Predictive Model Markup Language). STATISTICA Neural Networks can boast for its good service support. Another positive thing about the program is its multi-lingual interface option. It is a great "plus" for people who have troubles with English interface.

NeuroPro v.0.25 is a user-friendly Russian product distributed for free (Fig. 17). The program is a manager of ANNs, which can provide the user with such features and operations:

– creation of ANN;

– connection of the database file (.dtb or.db extension) to the created network;

– correction of the database;

– addition of new layers to the ANN (can handle up to 10 layers with 100 neurons in each layer);

– training of the ANN for solving the tasks of prediction or classification (or both tasks simultaneously);

– testing of the ANN and giving the feedback about the accuracy of the task solving;

– calculation of the input weights and importance;

– simplifying of the ANN;

– generation of the verbal description of the ANN;

– choice of the learning algorithm from the possible options.

NeuroPro is a good tool for studying the feed-forward ANN approach. However, the application is very limited, has imperfections in learning algorithm, and sometimes gives inaccurate predictions. But the fact of distribution on "free of charge" basis makes this application a favor in comparison to the expensive analogs.

NeuroSolutions 7 Pro is another one Microsoft Excel add-in that enhances the spreadsheets efficiency and makes it possible to use the ANN approach in the well-known software environment (Fig. 18). The program is quite an expensive one with a single-license price of 1500 USD. However, the NeuroSolutions 7 Pro package performs a huge number of operations and can be used for cluster analysis, forecasting,

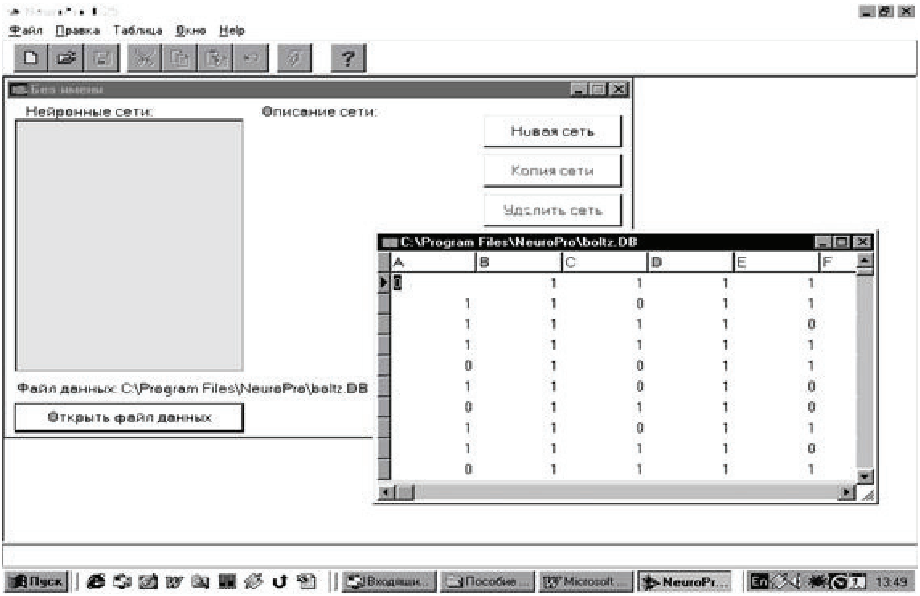


Figure 17. The interface of NeuroPro software application

classification, association. The tool is good enough both for novice in the ANN and advanced users. The number of possible options is amazing. User is able to create and implement to his purpose's linear regression, multi-layer perceptron, different types of feed-forward ANNs, probabilistic ANN, self-organization Kohonen maps, principal component analysis, general regression ANN, radial basis function, neuro-fuzzy ANNs, Hopfield ANN, recurrent ANN, and some other innovative methods of statistical and mathematical data processing. NeuroSolutions 7 Pro provide you an opportunity to use a vast diversity of learning algorithms: backpropagation, recurrent backpropagation, backpropagation through time, Hebbian learning, Ojas learning, Sangers learning, competitive learning, and Kohonen learning. So, the application is a perfect instrument for comprehensive study and practical use of the ANN-based approaches in data analysis and processing.

Neuronica is a free extension for LibreOffice Calc (Fig. 19). It is compatible with LibreOffice version 4.0 and newer (sometimes problems with running this extension occur with the latest versions of LibreOffice Calc on Linux; some bugs and computational errors occurred while running the extension on Windows with LibreOffice version higher than 5). Neuronica is a comparatively easy tool with an understandable interface. It is a little bit similar to the previously reviewed NeuroXL Excel add-in. The extension provides a user with a possibility of creating a feed-forward multi-layer perceptron and uses one of the following learning algorithms: classic backpropagation, resilient propagation, scaled conjugate gradient. A user has a privilege of setting ap-

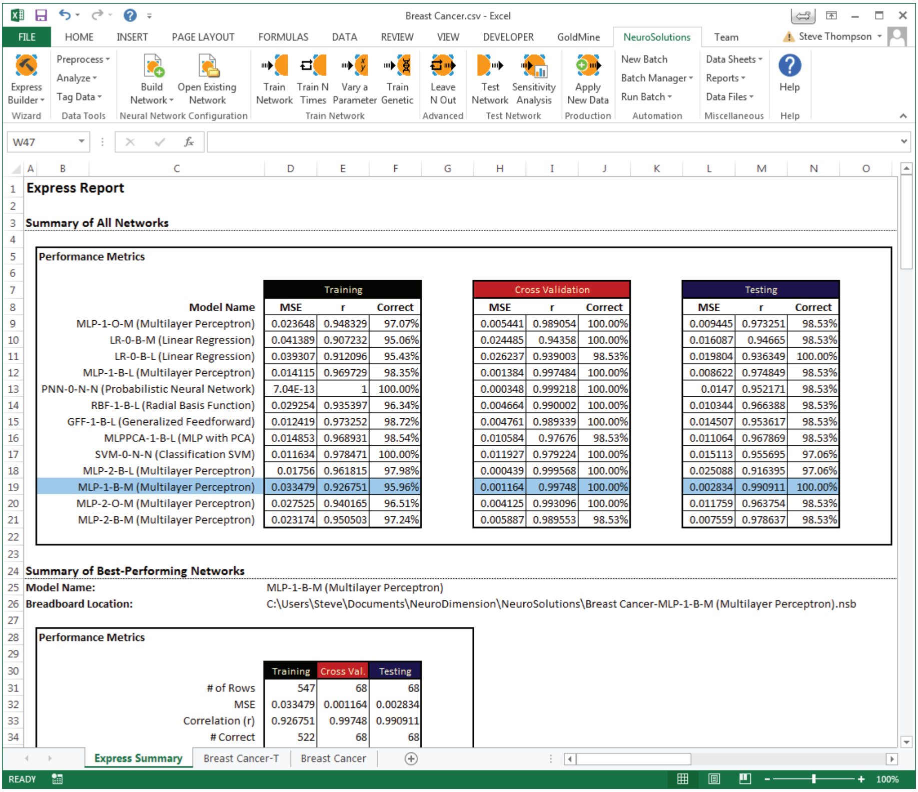


Figure 18. The interface of NeuroSolutions 7 Pro software application

propriate parameters of machine learning (momentum and learning rate), which is quite helpful in many cases because provides an opportunity to obtain the most precise prediction and regulate the learning speed. The trained ANN could be tested within the application afterwards to estimate the errors of the designed model. Notwithstanding the above-mentioned options, Neuronica remains quite a limited instrument with lots of bugs occurred while functioning, so it does not suit for conduction of serious computations and handle complicated projects with a huge number of data. Neuronica is not a tool of choice if you need a comprehensive and powerful instrument for development and work with serious data analysis by using the ANN-based approach. However, it is a good and simple instrument for those modest people who want to have a look and try how the ANN technique would work on solving their tasks.

GMDH Shell is a commercial software for dealing with the tasks of time series prediction, pattern recognition and data processing by using the ANN approach (Fig. 20). The positive thing about the program for an average user is its comparative easiness in usage, because the complicated neural algorithms and structures are hidden

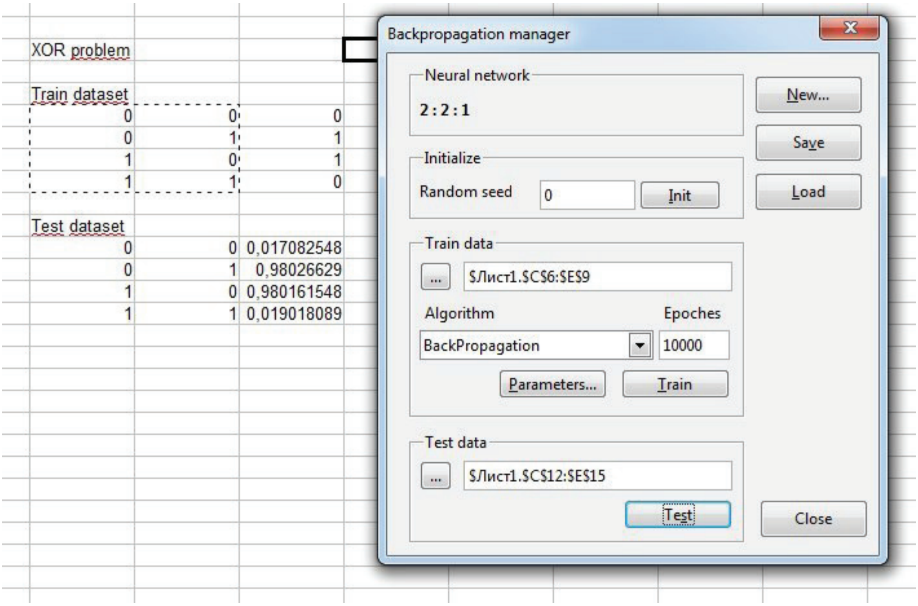


Figure 19. The interface of Neuronica software application

under the user-friendly cover of the software interface. However, this fact makes the Shell to perform the tasks as a "black box", because you never know what the application is really doing with your data at the moment. So, GMDH Shell's suitability for educational purposes is limited. We should mention that the program carefully uses hardware resources of your computer so that provide you with the fastest performance without losing the quality even on not powerful machines. In addition, we can say that GMDH Shell determines the best structure of model automatically, provides the best inter-relation between the performance and working speed, creates precise models and gives accurate predictions even with small data sets, combines ANN-based instruments for prediction, classification, clusterization and regression analysis within one software environment. However, the price of GMDH Shell for personal usage is quite high – 1999 USD. But the developers provide scientific institutions and researchers with the opportunity of getting the copy of the program for free within the cooperation framework.

SAS Software, which is quite popular in the USA, can also be used for realization of the ANN approach. Naturally, SAS Software is not a neural network tool in the proper sense. It is a powerful instrument for statistical data processing and analysis. However, SAS Enterprise Miner 13 could be used for creation and training of ANNs [113]. SAS Enterprise Miner contains two nodes: Neural Network and AutoNeural. The first node is used to configure user's own ANN, while the second one is used to create the ANN in the automated regime by using the support tools built in the pro-

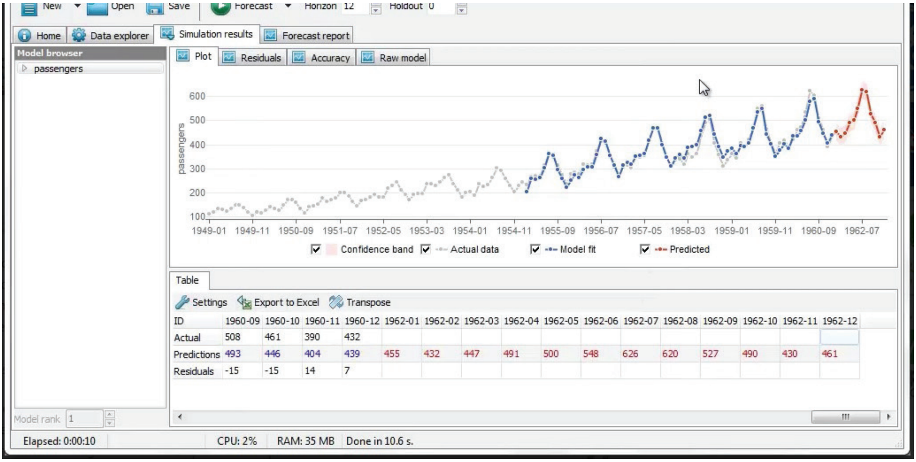


Figure 20. The interface of GMDH Shell software application

gram. SAS provides a possibility to implement the ANN approach to data processing and analysis by using the common multi-layer perceptron. The process of the network creation, training and testing can be a little bit sophisticated to most of the users, so we just stopped on the SAS Software to point out the possibility of its usage in this field. Besides, the program is an expensive one, and it is not widely implemented for solving the tasks by the means of the ANN.

ProForecaster is another Microsoft Excel add-in, which was firstly created to help people who have got no special mathematical education in their business predictions. ProForecaster provides them with an opportunity of getting reliable time series predictions and regression analysis within convenient and familiar Excel environment. It should be mentioned that the application performs predictions on the basis of hybrid algorithm, so we deal with a combined instrument (not pure artificial neural network) while working within proForecaster. Users can perform their predictions in an Automatic Mode. That means that the application finds itself the best algorithm for model creation and forecast among the wide range of available options, such as Smoothing Models, Growth Functions, Neural Networks (including different architectures and structures). Besides, user can create several different models of the same phenomena and compare them within the application by using statistical methods (which are additionally accompanied with charts and figures). People who are familiar with mathematical modeling and statistics can refuse help of the Automatic Mode and take all controls in their hands performing analysis and predictions in the Manual Expert Mode, which provides an opportunity of flexible adjustment and setting the parameters of models. ProForecaster is offered in two editions: free and professional. Free edition is completely functional but has some serious limitations in choosing forecasting methods (only 8 available). Professional edition of the add-in has more forecasting

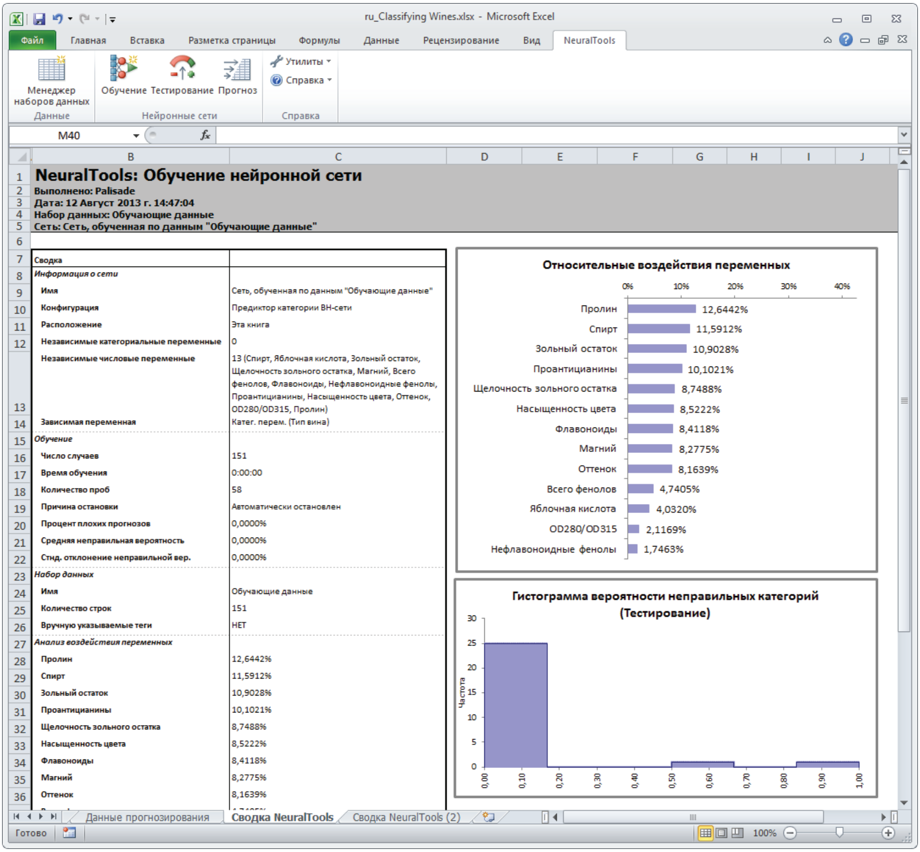


Figure 21. Palisade NeuralTools Microsoft Excel add-in working interface

options (15 models), provides users with a powerful artificial neural network engine, can handle several forecasting models simultaneously, provides an option of an Expert ranking of the models (this option is very helpful in finding the best model over the created).

Palisade NeuralTools is a well-known time series forecasting tool working in the MS Excel environment (Fig. 21). It is widely used in economic risks predictions, diseases diagnostic, agricultural production forecasting, environmental monitoring, marketing, archaeology and other branches of modern science and practice where an accurate prediction is often needed to prevent unexpected failures and risks. NeuralTools are the multilingual application, so you do not need to be good in English for comfortable work with the program. However, the add-in is quite expensive and requires to pay at least 435 GBP. NeuralTools does not require any special knowledge about ANN approaches and mathematical statistics. All predictions and models are run in automatic mode, and if you want, the program will find itself the best fitting algorithm for

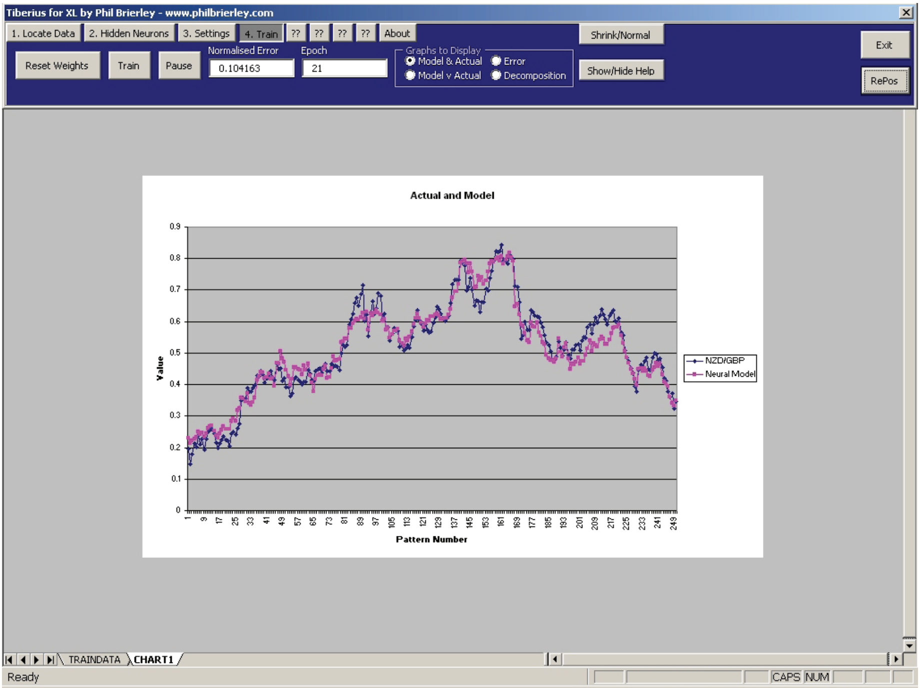


Figure 22. The interface of TiberiusXL add-in for Microsoft Excel

your concrete case to provide you with precise results. As a result of forecasting you can get the ranking of inputs effects on the target values to see which of them seems to be the most important in the certain case. Previously created ANNs can be saved and used in further studies. Another advantage of the application is that it is suitable for working both with categorical and numerical data. Also, user can get the resulting information not only in figures and tables but in graphical form to make it easier to analyze and interpret the results and have forecast trends on sight.

Tiberius XL is a free independently developed by Phil Brierley MS Excel add-in [19, 20], which was created to provide all interested in artificial neural networks approach people with a possibility of studying and designing their own networks on the basis of the VBA code (Fig. 22). Tiberius XL is a limited application, which can "as it is" handle no more than 5 neurons in the hidden layer, and 1000 epochs of training process. The add-in uses a feed-forward multilayer perceptron approach with the backpropagation learning algorithm realized within the user-friendly environment of MS Excel. Tiberius XL is compatible with all modern versions of the Office software, which makes it possible to be run on the computers with installed MS Office 2000-2019. The extension is completely integrated into the MS Office environment. However, the application cannot be used for the wide range of purposes. For example, you

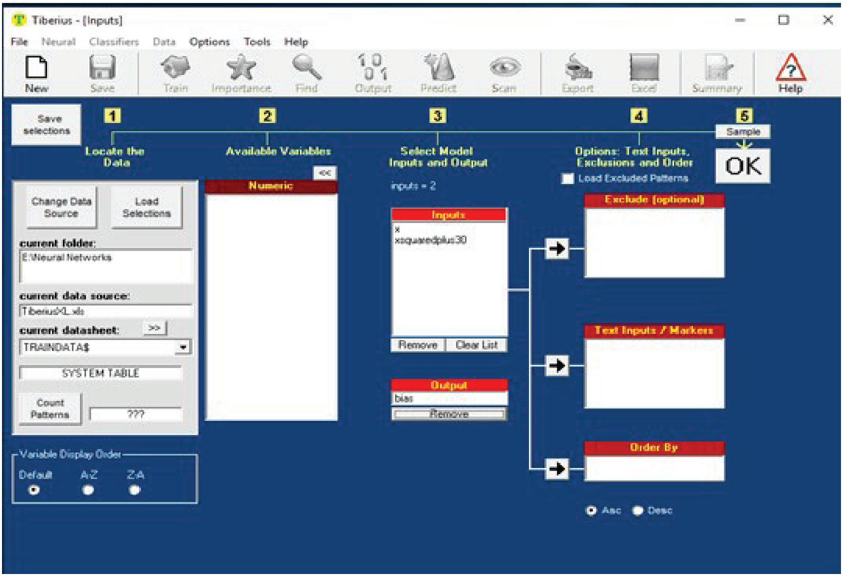


Figure 23. Tiberius Data Mining Predictive Modeling Software 7 working interface

cannot use Tiberius XL for the predictions, classification or clusterization tasks. If you want you may purchase the source code of the add-in and create your own extension on the basis of Tiberius XL with the extended functionality. Of course, there is another option to obtain a version of Tiberius software with the additional functionality including predictive modeling. You can download Tiberius Data Mining Predictive Modeling Software 7 for free (if you provide the proofs that you are an academic user, not a commercial one). This application gives a possibility of application neural networks approaches to data analysis for getting precise and reliable forecasts (Fig. 23). Tiberius 7 provides such powerful flexible tools as decision tree, non-linear regression analysis, advanced visualization tools including 3D graphs and models, deep model analysis toolkit, etc. Tiberius Predictive Modeling Software was designed for the purposes both of researchers and ordinary customers, which are interested in getting fast and precise predictions in the friendly environment (the application is good integrated with most of popular statistical software as MS Excel, SAS, SPPS, etc.).

2.2. A Comparison of Different Software for Realization of Artificial Neural Networks-based Modeling in Agriculture

Taking into account vast variety of available software for the development and implementation of the ANN-based approach to modeling and forecasting, we think that it is necessary not only to overview their functionality capacities but test their real performance in solving some actual tasks, for example, yield prediction. We performed the testing by creation of a predictive model for sweet corn yields by using the yielding data of sweet corn ears without husks, which were obtained within the field trials conducted in 2014-2016 within the scientific framework of Kherson State Agricultural

Table 1.
Sweet corn yield predictions by the artificial neural network models developed in different software environment

Inputs of the predictive models (the studied factors)			Sweet corn yields, t ha-1 true values obtained in the trials)	Outputs of the predictive models		
Factor A (Moldboard plowing depth, cm)	Factor B (Fertilizers application rates, kg ha-1)	Factor C (Plants density, plants ha-1)		NeuroXL Predictor	JustNN	TiberiusXL
20	0	35000	2.67	3.19	2.90	2.72
20	0	50000	2.85	2.70	2.95	2.38
20	0	65000	3.01	2.88	3.02	3.19
20	0	80000	2.96	3.44	3.10	3.48
20	60	35000	5.56	5.57	5.92	5.76
20	60	50000	6.31	6.20	7.30	5.98
20	60	65000	7.67	7.27	8.10	7.55
20	60	80000	6.80	6.71	8.35	7.45
20	120	35000	7.53	7.57	7.10	8.53
20	120	50000	8.81	9.16	8.33	9.22
20	120	65000	10.93	10.17	8.91	10.60
20	120	80000	9.58	10.19	9.04	10.14
28	0	35000	3.00	3.03	3.18	1.69
28	0	50000	3.34	3.44	3.33	1.22
28	0	65000	3.57	3.51	3.47	2.61
28	0	80000	3.37	2.96	3.58	3.29
28	60	35000	4.89	4.34	4.92	4.38
28	60	50000	5.55	5.95	5.77	4.28
28	60	65000	6.25	6.48	6.26	5.85
28	60	80000	5.64	5.76	6.44	6.06
28	120	35000	6.23	6.28	6.21	8.34
28	120	50000	7.56	7.74	7.11	8.63
28	120	65000	8.59	7.98	7.44	9.44
28	120	80000	7.56	7.78	7.40	9.10

University [91]. The average crop yield was used as the output for the model while technological parameters were the inputs. The model of yield prediction was developed by using both commercial and free software, namely: NeuroXL Predictor, JustNN and TiberiusXL. We set the model parameters in every program to provide the maximum available accuracy of the predictive model. So, there were three neural networks developed:

1) In NeuroXL Predictor: ANN structure 3-10-1, activation function – zero-based log-sigmoid.

2) In JustNN: ANN structure 3-6-1, simple sigmoid activation function.

3) In TiberiusXL: ANN structure 3-5-1, simple sigmoid activation function.

As a result of our modeling, we obtained such predictions (Table 1) and estimated their accuracy by the calculation of the RSQ (coefficient of determination) using the following formula 28:

$$\mathrm{RSQ} = 1 - \frac{V(y|x)}{V(y)} \tag{28}$$

where V/x is the dispersion of the dependent argument.

As a result of the calculation of RSQ for the studied predictive models we obtained such values of the coefficient of determination, namely: 0.978 for NeuroXL Predictor model, 0.922 for JustNN model, and 0.913 for TiberiusXL model, correspondingly. So, it was proved that accuracy of all the developed predictive networks was satisfactory, and laid in the interval of 0.9-1.0 above 90%). However, notwithstanding the fact that each of the created predictive networks can be successfully used for sweet corn yield prediction, their performance is not equal. The best predictive performance was granted by the commercial software NeuroXL Predictor, free software is considerably less precise in the predictions (by 5.73-6.65%).

So, it is not all the same what software to use for the yield predictions. If you need to obtain the most accurate and reliable model you have to test some different options and then decide which one suits best for solving the concrete task. Besides, commercial proposals are considered to be significantly better than free distributed software.

CHAPTER 3

ARTIFICIAL NEURAL NETWORKS USE IN DIFFERENT PRACTICAL AND THEORETICAL TASKS OF MODERN AGRICULTURAL SCIENCE

3.1. An Empirical Model of Chickpea Productivity Prediction in Dependence on Cultivation Technology Treatments in the Conditions of the South of Ukraine

The chickpea productivity prediction model was created on the basis of the results of the three-year field experiments devoted to study of the crop response to different cultivation technology treatments [78]. The field experiment was carried out during the period from 2012 to 2014 on the irrigated lands of agricultural cooperative farm "Radianska Zemlia" (Bilozerskyi district of Kherson region, Ukraine).

The soil of the experimental field was represented by dark-chestnut solonets soil. The soil has a clear differentiation of the profile on humus-eluvial and humus-illuvial horizons. The total depth of the humus horizon is 50-55 cm, the color of the horizon is dark gray with a brown shade. Hydrochloric acid boils at the depth of 60-70 cm. The content of humus in the soil fluctuates within the rates of 2.5-3.0%. Groundwater layer is situated at the depth of 3-5 m, and the water does not take part in the processes of soil creation.

The field experiments were devoted to study of such agrotechnological factors as:

Factor A – moldboard plowing depth (20-22 and 28-30 cm);

Factor B – mineral fertilizers application doses (no fertilizers, $N_{45}P_{45}$, $N_{90}P_{90}$);

Factor C – sowing rate (0.5, 1.0, 1.5 M per ha);

Factor D – humidification conditions (with or without artificial humidification by the means of overhead sprinkler irrigation machine).

The experiments were conducted in four replications by using the split plot design method. All the observations, measurements and calculations were carried out with accordance to the common rules of experimental work in irrigated agriculture [96].

Chickpea cultivation technology based on the general recommendations for the crop cultivation in the conditions of the Dry Steppe zone. Rosanna chickpea cultivar was used in the experiments. After the previous crop (winter wheat) harvesting, double disking of the stubble was conducted on the depth of 6-8 and 10-12 cm. Primary tillage was carried out with accordance to the previously stated experimental design by the means of standard moldboard plow. Mineral fertilizers were applied in pre-plowing period in the doses pointed out in the study design scheme by the means of a seed drill. Two weeks later additional cultivation on the depth of 12-14 cm was carried out to level the soil surface and destroy weeds after the plowing. Harrowing was conducted in the early

Table 2.
Average chickpea grain yields depending on moldboard plowing depth, mineral fertilizers doses, sowing rate and humidification conditions for the period of 2012-2014, t ha-1

Moldboard plowing depth, cm (Factor A)	Mineral fertilizers application dose, kg per ha (Factor B)	Sowing rate, in millions per ha (Factor C)		
		0.5	1.0	1.5
No irrigation water applied (Factor D)				
20-22	No fertilizers	1.26	1.48	1.55
	N45P45	1.41	1.68	1.77
	N90P90	1.52	1.80	1.90
28-30	No fertilizers	1.28	1.50	1.60
	N45P45	1.44	1.72	1.83
	N90P90	1.56	1.85	1.98
Irrigation water applied (Factor D)				
20-22	No fertilizers	2.18	2.48	2.70
	N45P45	2.66	3.02	3.31
	N90P90	2.83	3.24	3.53
28-30	No fertilizers	2.22	2.53	2.74
	N45P45	2.71	3.10	3.38
	N90P90	2.89	3.33	3.60

ANOVA results (the analysis was performed at the probability level of 95%): the least significant differences (hereinafter – LSD) for the studied factors and their interactions, t ha^{-1}: A, D – 0.035-0.048; B, C – 0.043-0.059; AD – 0.050-0.068; BD, CD, AB, AC – 0.061-0.083; BC – 0.075-0.102; ABD, ACD – 0.086-0.118; BCD, ABC – 0.106-0.144; ABCD – 0.150-0.204.

spring. Pre-sowing cultivation was conducted on the depth of sowing (5-7 cm). Chickpea seeds were sown by the means of a seed drill John Deere 740A with accordance to the trials design. The seeds were treated by the active symbiotic bacteria (with preparations Rizophyt, Phosphoenterin, Biopolicyde). The crops were rolled instantly after sowing. Soil herbicide Hezaguard 500 FW in the dose of 3.0 L per ha was used before sprouting of the crop. Insecticide Nurel D in the dose of 1.0 L per ha was used at the stage of "budding-flowering". The soil moisture was maintained at the level of 75-80% of the field water holding capacity on the irrigated experimental variants by the means of irrigation machine "Kuban". Chickpea grain harvesting was conducted by using the self-propelled combine harvester at the stage of full ripeness of beans.

The method of generalized regression ANN (GRNN) was applied for chickpea grain productivity prediction in dependence on the studied technological treatments. Cross-check of the prediction model was performed by using the statistical criteria, viz., mean error, mean absolute error, standard deviation, mean relative error, coefficient of correlation, etc. [78, 103]. The ANN approach was realized within Windows OS in the application STATISTICA Advanced, STATISTICA Automated Neural Networks v.10. The yield data were previously generalized and analyzed by the means of multi-factor ANOVA to prove significance of the impact of all the studied factors on the crop productivity. The results of yield estimation and statistical analysis are repre-

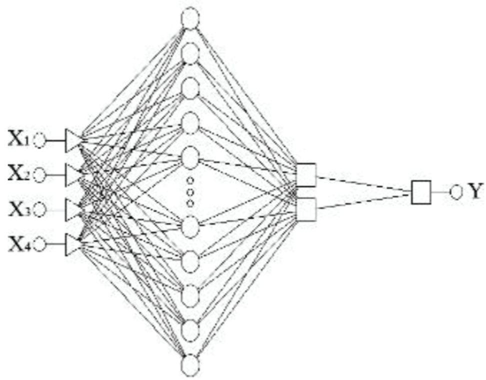


Figure 24. The structural scheme of the developed GRNN (4-54-2-1)

sented in the Table 2.

The GRNN with 2 hidden layers of the structure 4-54-2-1 was developed (Fig. 24).

The results of its training are as follows: productivity of training – 0.22; control productivity – 0.37; testing productivity – 0.36; training error – 0.29; control error – 0.45; testing error – 0.47. Multiple nonlinear correlation between the studied factors and chickpea yields in the model was 0.96, coefficient of determination – 0.92. This proves high accuracy of the developed ANN model for yield prediction. The errors of the GRNN model are represented in the Table 3.

Besides, by the means of the GRNN model the ranking of the studied factors by their effect on the crop yields was performed. So, the highest influence on chickpea grain yields had irrigation (with a relative share expressed in points of 3.47), then followed mineral fertilizers doses (1.35 points), sowing rate (1.25), and moldboard plowing depth (0.97). The share of the studied factors may also be represented in more usual and convenient way – in the percentage. So, the share of irrigation in the crop productivity formation is 49.29%; the share of mineral fertilizers – 19.18%; the share of the sowing rate – 17.75%; the share of the primary tillage depth – 13.78%

The results of the study proved high reliability and accuracy (of 92%) of the GRNN modeling approach for prediction and estimation of chickpea yields depending on the cultivation technology in the above-mentioned environmental conditions.

3.2. Sweet Corn Productivity Prediction by Using Several Algorithms of Artificial Neural Network Realization within Different Software. A Comparison between Artificial Neural Network and Multiple Linear Regression Models for the Crop Yields Prediction

Artificial neural network approach was implemented to predict sweet corn yields in dependence on the crop cultivation technology. The ANN-based model was developed on the basis of the results of three-year field experiments devoted to the improvement of the crop cultivation technology in the irrigated conditions of the South of Ukraine [90, 91].

The field experiments were carried out by using the split-plot design method in four replications during 2014-2016 in the basic Agricultural Cooperative Farm “Radianska Zemlia” (Bilozerskyi district of the Kherson region, Ukraine; latitude 46°43′42″N, longitude 32°17′38″E, altitude 42 m).

Table 3.
Errors of the GRNN model for chickpea grain yield prediction

Criteria	Error of the model
Mean error, t per ha	-0.0182
Mean absolute error, t per ha	0.1699
Standard deviation, t per ha	0.2204
Correlation	0.96
Mean relative error, %	7.92

The soil of the experimental plots was represented by the dark-chestnut solonets soil. The humus content in the 0-50 cm soil layer was 2.5%. The bulk density of the 0-100 cm soil layer was 1.35 t m^{-3}. The lightly-hydrolized Nitrogen content (determined by the methodology of Kornfield) was 35 mg kg^{-1}, the mobile Phosphorus content (determined by the methodology of Machygin) was 32 mg kg^{-1}, the exchangeable Potassium content (determined by the methodology of Machygin) was 430 mg kg^{-1} in the arable soil layer.

The program of investigations covered studying of the following factors: Factor A – tillage (moldboard plowing at the depth of 20-22 and 28-30 cm); Factor B – mineral fertilizers application doses (no fertilizers applied; $N_{60}P_{60}$; $N_{120}P_{120}$ of active substance applied); Factor C – plants density (35000, 50000, 65000, 80000 plants ha^{-1}). Sweet corn yields (assessed in the ears without husks) were determined by the hand-harvesting of fruits in their technical ripeness from the entire area of every experimental plot with further weighing of the picked ears on the electronic analytical weighs. Sweet corn cultivar used in the field experiments was Brusnytsia (standard sweet – *su*), originated by the Skvyrska Research Station of the Institute of Vegetable and Melon Growing (Instytut Ovochivnytstva I Bashtannytstva), National Academy of Agrarian Sciences of Ukraine. Sweet corn cultivation technology in the field experiments was based on the general recommendations for growing the crop under the irrigated conditions in the South of Ukraine. Mineral fertilizers (ammonium nitrate and superphosphate) were applied with accordance to the experimental design in the pre-plowing period by the means of a seed drill. The previous crop was winter wheat. Disk tillage at the depth of 10-12 cm, which was followed by the moldboard plowing, was conducted after the previous crop harvesting. Two cultivator tillage at the depth of 8-10 and at the depth of 5-6 cm were conducted during the spring. Sweet corn was sown at the depth of 5-6 cm with the standard inter-row spacing of 70 cm. The terms of sowing were: 1st of May in 2014, 22nd of May in 2015 and 21st of May in 2016, respectively. Herbicide Harnes (*Acetochlor*, 900 g l^{-1} of the active substance) was applied in pre-sowing period in the 2.0 l ha^{-1} dose. Karate Zeon insecticide (*Lambda-cyhalothrin,* 50 g l^{-1} of the active substance) was used at the 3-5 leaves crop stage in the 0.2 l ha^{-1} dose. Master Power herbicide (*Foramsulfuron,* 31.5 g l^{-1}, *Iodosulfuron*, 1.0 g l^{-1}, *Tienecarbazon-methyl*, 10 g l^{-1}, *Cyprosulfamide* (antidote), 15 g l^{-1} of the active substances) was applied at the

7-8 leaves crop stage in the 1.25 l ha^{-1} dose. Koragen insecticide (*Chlorantraniliprole*, 200 g l^{-1} of the active substance) was used at the beginning of the panicle earing crop stage in the 0.1 l ha^{-1} dose. Soil humidity during the sweet corn vegetation was maintained at the 80% of the field water-holding capacity level by the means of drip irrigation. The drip tape was placed in every row of the sweet corn crops. We used *Eurodrip* 5 mil drip-tape with spacing of drippers of 20 cm and discharge rate of 1.2 L per hour. Water application doses were: in 2014 – 10 times at the rate of 5 mm until the stage of 7-8 leaves of crop and 12 times at the rate of 10 mm in the rest of the period; in 2015 – 6 times at the rate of 5 mm until the stage of 7-8 leaves of crop and 9 times at the rate of 10 mm in the rest of the period; in 2016 – 8 times at the rate of 5 mm until the stage of 7-8 leaves of crop and 12 times at the rate of 10 mm in the rest of the period. The total average content of the irrigation water applied was 150 mm during the crop vegetation period in average for three years of the study. The total water consumption of the crops was determined by the balance method of Kostiakov [96, 136, 137].

Two different software applications were used to assess and forecast sweet corn yields depending on the cultivation technology treatments. First of all, we must pay attention to the realization of ANN model within the free software JustNN. An easy ANN with one hidden layer of neurons was developed using the built-in tools of the program. The means of JustNN are limited, so the feed-forward ANN had very simple structure 3-6-1. A sigmoid function is likely to be used in the software as an activation function of the artificial neuron, although, this information is not stressed anywhere. As a result of the training process the values of training efficiency were obtained, which were as follows: training error – 0.0098 with the maximum of 0.0788); mean model

Table 4.
Sweet corn yields in ears without husks depending on the agrotechnological treatments, t ha-1 average for the period of 2014-2016)

Factor A (Moldboard plowing depths, cm)	Factor C (Plants densities, plants ha-1)	Factor B (Fertilizers application doses, kg ha-1)			Mean values by the Factor A
		No fertilizers	N60P60	N120P120	
20-22	35000	2.67 (±0.30)	5.56 (±0.57)	7.53 (±0.88)	6.22
	50000	2.85 (±0.28)	6.31 (±1.03)	8.81 (±1.31)	
	65000	3.01 (±0.34)	7.67 (±0.75)	10.93 (±1.32)	
	80000	2.96 (±0.35)	6.80 (±1.15)	9.58 (±1.03)	
28-30	35000	3.00 (±0.33)	4.89 (±0.55)	6.23 (±0.86)	5.45
	50000	3.34 (±0.38)	5.55 (±0.54)	7.36 (±0.87)	
	65000	3.57 (±0.43)	6.25 (±0.69)	8.59 (±1.02)	
	80000	3.37 (±0.39)	5.64 (±0.60)	7.56 (±0.92)	
Mean values by the Factor B		3.10	6.08	8.32	
Mean values by the Factor C		4.98	5.70	6.67	5.99

ANOVA results (the analysis was performed at the probability level of 95%): the LSD for the studied factors and their interactions, t ha^{-1}: A – 0.10; B – 0.07; C – 0.12; ABC – 0.32.

Table 5. Sweet corn yields data grouped for the development of mathematical models

Inputs			Output
Factor A (Moldboard plowing depths, cm)	**Factor B (Fertilizers application doses, kg ha-1)**	**Factor C (Plants densities, plants ha-1)**	**Sweet corn yields, t ha-1**
20	0	35000	2.67
20	0	50000	2.85
20	0	65000	3.01
20	0	80000	2.96
20	60	35000	5.56
20	60	50000	6.31
20	60	65000	7.67
20	60	80000	6.80
20	120	35000	7.53
20	120	50000	8.81
20	120	65000	10.93
20	120	80000	9.58
28	0	35000	3.00
28	0	50000	3.34
28	0	65000	3.57
28	0	80000	3.37
28	60	35000	4.89
28	60	50000	5.55
28	60	65000	6.25
28	60	80000	5.64
28	120	35000	6.23
28	120	50000	7.56
28	120	65000	8.59
28	120	80000	7.56

Table 6.
Results of multiple linear regression analysis of sweet corn yields

Treatments	Coefficients of correlation (R)	Coefficients of determination (RSQ)	Coefficients of regression (b)	Student criterion	Student criterion at $p<0.05$
$X_1X_2X_3$	0.947	0.897	4.0270	3.388	2.069
X_1	-0.166	0.028	-0.0972	-2.319	
X_2	0.913	0.833	0.0436	12.729	
X_3	0.190	0.036	0.0265	2.655	

Notes. X_1 is the moldboard plowing depth expressed in cm; X^2 is the mineral fertilizers application rates expressed in kg ha^{-1} of the active substance; X_3 is the plants density expressed in plants ha^{-1}.

error – 0.0383 t per ha. Additionally, JustNN provided the share of inputs (cultivation technology parameters expressed in the numeric form) effect on the final output value (crop yields). So, the share of moldboard plowing depth was 3.32 points or 8.94%, the

Table 7.
Comparison of true and predicted sweet corn yields by using the ANN-based and multiple linear regression models, t ha-1

True yields values	ANN predicted yields	Linear regression predicted yields	Residuals for the ANN prediction	Residuals for the linear regression prediction
2.67	3.19	3.01	-0.52	-0.34
2.85	2.70	3.41	0.15	-0.56
3.01	2.88	3.81	0.13	-0.80
2.96	3.44	4.20	-0.48	-1.24
5.56	5.57	5.63	-0.01	-0.07
6.31	6.20	6.02	0.11	0.29
7.67	7.27	6.42	0.40	1.25
6.80	6.71	6.82	0.09	-0.02
7.53	7.57	8.24	-0.04	-0.71
8.81	9.16	8.64	-0.35	0.17
10.93	10.17	9.04	0.76	1.89
9.58	10.19	9.44	-0.61	0.14
3.00	3.03	2.23	-0.03	0.77
3.34	3.44	2.63	-0.10	0.71
3.57	3.51	3.03	0.06	0.54
3.37	2.96	3.43	0.41	-0.06
4.89	4.34	4.85	0.55	0.04
5.55	5.95	5.25	-0.40	0.30
6.25	6.48	5.64	-0.23	0.61
5.64	5.76	6.04	-0.12	-0.40
6.23	6.28	7.46	-0.05	-1.23
7.56	7.74	7.86	-0.18	-0.30
8.59	7.98	8.26	0.61	0.33
7.56	7.78	8.66	-0.22	-1.10

share of mineral fertilizers dose was 25.52 points or 68.66%, and the share of plants density was 8.33 or 22.40%.

Besides, another ANN-based model was created to compare the accuracy of ANN feed-forward networks with backpropagation learning algorithm with the multiple linear regression model of the crop productivity. The ANN was developed within NeuroXL Predictor add-in in MS Excel 2010 environment [102]. We used default settings of the program to train the network. The activation function type used was a zero-based log-sigmoid. Amount of the artificial neurons in the hidden layer was 10. RSQ (or coefficient of determination) value was used to compare the accuracy of regression

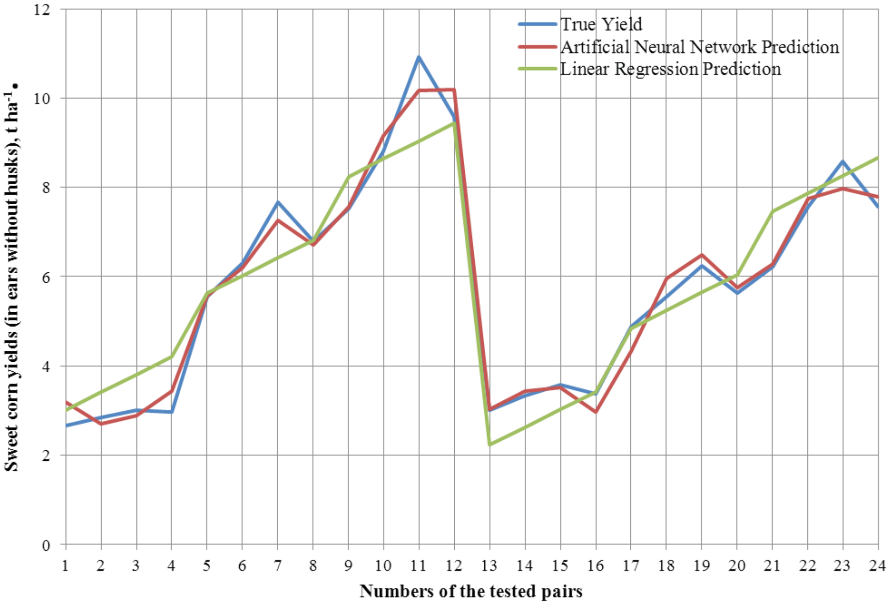


Figure 25. Graphical comparison of the ANN-based and multiple linear regression model of sweet corn yields

and ANN-based prediction of sweet corn yields. The RSQ value was determined by using the formula 28 [30].

Average sweet corn yields were used as the output of the developed models, and agrotechnological factors expressed in the numeric form were used as the inputs. The yields of sweet corn ears without husks are represented in the Table 4. Significance of the established differences between the studied variants of the crop cultivation was proved by the results of multi-factorial ANOVA, which was conducted within AgroStat add-in for Microsoft Excel [96].

To use the yields data in the mathematical modeling they were previously grouped in the certain consequence as it is represented in the Table 5.

Then the multiple regression analysis was performed by using the built-in package of analysis of MS Excel 2010 software. The results of the analysis are represented in the Table 6.

By the results of multiple linear regression analysis, the model of sweet corn yields depending on the cultivation technology factors was developed:

$$Y = 4.0270\text{-}0.0972X_1+0.0436X_2+0.0265X_3,$$

where: Y is the sweet corn yields; X_1 is the moldboard plowing depth expressed in cm; X_2 is the mineral fertilizers application rates expressed in kg ha^{-1} of the active substance; X_3 is the plants density expressed in plants ha^{-1}.

This model was run with the set of the inputs used for its development to test its reliability and accuracy. The same thing was performed with the ANN-based model, however, NeuroXL conducted all computations in the "closed" mode as most of the ANN software. NeuroXL does not provide developers with the mathematical expression of the prediction model. The results of computations performed for the regression and ANN-based models are represented in the Table 7.

As could be seen from the Table 7, the amplitude of the residuals was 2.44 times less of the ANN-based model (1.28 t ha^{-1}) than of the regression model (3.13). Besides, RSQ value of the linear regression model was 0.897 that is sufficiently less than of the ANN-based model with RSQ of 0.978. So, the advantage in accuracy and reliability of the ANN-based approach to the crop yields modeling is quite obvious. The graphical expression of the prediction models is provided in the Fig. 25.

3.3. Prediction of Sweet Corn Total Water Consumption by Using an Artificial Neural Network

Artificial neural network approach could be successfully used not only for modeling crops' yields. It is also an effective tool for simulations in other fields of agriculture, for example, getting predictions of the total evapotranspiration or water consumption of the irrigated and non-irrigated crops. Below you will find an example of implementation of the technology for prediction of sweet corn total water consumption depending on the cultivation technology by using the ANN-based model developed within JustNN software application.

The data used for creation of the ANN-based model were taken from the results of the previously described trials with sweet corn (the methodology and conditions of the experiment could be found in the Paragraph 3.2) conducted during 2014-2016 [137]. We created the ANN with the structure of 3-6-1-3-1 with 3 hidden layers of neurons. Accuracy of the model was also assessed by calculation of the RSQ value. The results of the training process of the network is represented in the Table 8.

We also obtained the results of the inputs weight in the output value. They are represented in the Table 9.

Therefore, we can see that the highest effect on the crop water-use is caused by the factor of plants density (41.25%), and the least effect on the water consumption had tillage depth (only 22.89%). The RSQ value of the ANN-based model of sweet corn water consumption was 0.937. This value is high enough to make us conclude about sufficient reliability and accuracy of the ANN-based model. Besides, we represent the results of cross-testing of the modeled and true water consumption values in the Table 10.

The total amplitude of the residuals reached the maximum of 5.4 mm. This is an amount of one irrigation, or even less (if we are talking of the later stages of the crop

Table 8.
Training errors of the ANN-based model for sweet corn total water consumption

Type of the training error	Average value of the error, mm	Standard deviation, mm
The minimum	0.000011	0.000057
Average	0.009913	0.000080
The maximum	0.044620	0.001982

Table 9.
The inputs' weight in sweet corn total water consumption

The factor	Absolute significance, points	Standard deviation, points	Percentage
Plants density	7.28	0.23	41.25
Doses of mineral fertilizers	6.33	0.43	35.86
Moldboard plowing depth	4.04	0.80	22.89

Table 10.
Comparison of the true and ANN-predicted values of sweet corn total water consumption

Doses of mineral fertilizers	**Plants density**	**Sweet corn water consumption, mm ± standard deviation**		
		True values	Predicted values	Residuals
Moldboard plowing at the depth of 20-22 cm				
No fertilizers	35000	258.3±3.42	261.5±3.88	-3.2±1.13
	50000	261.6±3.40	262.4±4.33	-0.8±0.95
	65000	265.7±4.71	264.7±5.75	+1.0±1.07
	80000	266.8±5.17	269.1±8.29	-2.3±3.21
N60P60	35000	262.4±4.48	262.9±4.65	-0.5±1.50
	50000	268.1±6.81	266.0±6.66	+2.1±0.17
	65000	271.4±7.34	270.9±8.78	+0.5±1.50
	80000	274.0±1.096	274.6±8.51	-0.6±2.58
N120P120	35000	267.1±6.45	267.9±7.76	-0.8±1.42
	50000	270.7±6.31	272.8±8.96	-2.1±2.65
	65000	277.1±9.22	275.7±8.35	+1.4±0.92
	80000	277.6±9.00	276.7±7.97	+0.9±1.25
Moldboard plowing at the depth of 28-30 cm				
No fertilizers	35000	259.5±3.20	262.1±3.96	-2.6±0.85
	50000	262.3±3.84	263.5±4.11	-1.2±0.75
	65000	266.2±5.14	266.6±4.31	-0.4±1.85
	80000	267.5±5.36	270.8±5.06	-3.3±0.55
N60P60	35000	263.3±3.64	264.1±4.27	-0.8±0.86
	50000	268.6±4.26	267.7±4.47	+0.9±1.47
	65000	271.7±4.26	272.0±5.40	-0.3±1.33
	80000	274.8±7.48	274.9±6.70	-0.1±1.59
N120P120	35000	267.6±5.60	269.1±4.74	-1.5±1.68
	50000	271.4±5.42	273.3±5.95	-1.9±0.78
	65000	277.5±7.48	275.6±7.09	+1.9±0.40
	80000	278.6±7.19	276.5±7.55	+2.1±0.36

growth and development). So, the proposed model can be used for operative planning and sweet corn irrigation management under the conditions of trials conduction.

3.4. Weather Forecast on the Basis of an Artificial Neural Network and Holt-Winters Triple Exponential Smoothing Method: Performance Comparison

Weather forecasts is another important field of implementation of artificial neural networks. Precise prophesying of meteorological situation in general and weather phenomena in particular is in a great demand not only for emergency services and public information services but for agrometeorological forecasts, too. Agrometeorology is a branch of modern agricultural science, which deals with questions of usage of weather and climate information to provide further development of agricultural economy, especially, concerning an introduction and expand of crops and trees, and enhancement of their productivity under the current meteorological conditions in the certain region. Agrometeorology is an interdisciplinary branch of studies involving not only traditional meteorology and agriculture but hydrology, soil science, horticulture, forestry, physics, chemistry, mathematical statistics and informational technologies to provide sustainable development of economy. One of the most important subdisciplines of agrometeorology is usage of predictive informational technologies to obtain both short and long-term forecasts of the complicated complex meteorological phenomena and deep understanding of the conditions of their occurrence [2].

Table 11.
Air temperature prediction by using Holt-Winters exponential smoothing approach with handling the seasonal effects

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Average
True air temperature values, °C													
2010	-4.0	2.0	3.4	10.7	17.6	22.5	24.7	26.1	17.7	7.8	10.5	1.6	11.7
2011	-2.7	-3.8	2.5	9.7	16.9	21.3	24.7	22.3	18.4	9.5	2.2	3.8	10.4
2012	-1.7	-7.3	2.6	13.2	20.8	23.3	26.5	23.7	19.1	14.6	6.6	-0.9	11.7
2013	-0.3	2.2	3.1	12.0	20.7	23.0	23.1	24.2	15.1	9.3	7.4	0.4	11.7
2014	-1.5	0.0	6.9	11.5	18.0	20.8	25.0	24.5	18.3	9.2	3.2	-0.2	11.3
2015	-0.2	0.8	5.2	9.3	17.0	20.9	23.4	24.1	20.9	9.4	7.3	2.2	11.7
2016	-3.6	4.0	6.3	12.7	16.2	22.1	24.4	24.6	18.0	8.4	4.0	-1.2	11.3
2017	-4.7	-0.7	7.0	9.3	16.3	22.0	23.4	25.4	19.9	11.3	5.4	5.9	11.7
Predicted air temperature values, °C													
2018	-1.8	2.8	10.4	10.8	15.9	19.7	24.3	25.8	20.0	8.5	1.5	0.3	11.5
2019	-3.4	1.7	9.4	8.3	12.7	19.2	23.3	24.3	21.8	8.1	-1.2	2.7	10.6
2020	-2.4	-1.8	9.5	11.8	16.6	21.2	25.1	25.8	22.5	13.2	3.2	-0.7	12.0
2021	-0.9	7.7	10.0	10.6	16.5	20.9	21.7	26.5	18.5	7.9	4.0	1.3	12.0
2022	-2.2	5.5	13.8	10.1	13.8	18.7	23.6	26.9	21.7	7.8	-0.2	0.7	11.7
2023	-3.7	4.5	12.9	7.6	10.5	18.1	22.6	25.4	23.5	7.3	-2.9	3.1	10.7
2024	-2.7	1.0	13.0	11.1	14.4	20.1	24.4	26.9	24.2	12.4	1.5	-0.3	12.2
2025	-1.3	10.5	13.5	9.9	14.3	19.8	21.0	27.6	20.2	7.1	2.3	1.8	12.2

Table 12.
Air temperature prediction by using the artificial neural network

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Average
True air temperature values, °C													
2010	-4.0	2.0	3.4	10.7	17.6	22.5	24.7	26.1	17.7	7.8	10.5	1.6	11.7
2011	-2.7	-3.8	2.5	9.7	16.9	21.3	24.7	22.3	18.4	9.5	2.2	3.8	10.4
2012	-1.7	-7.3	2.6	13.2	20.8	23.3	26.5	23.7	19.1	14.6	6.6	-0.9	11.7
2013	-0.3	2.2	3.1	12.0	20.7	23.0	23.1	24.2	15.1	9.3	7.4	0.4	11.7
2014	-1.5	0.0	6.9	11.5	18.0	20.8	25.0	24.5	18.3	9.2	3.2	-0.2	11.3
2015	-0.2	0.8	5.2	9.3	17.0	20.9	23.4	24.1	20.9	9.4	7.3	2.2	11.7
2016	-3.6	4.0	6.3	12.7	16.2	22.1	24.4	24.6	18.0	8.4	4.0	-1.2	11.3
2017	-4.7	-0.7	7.0	9.3	16.3	22.0	23.4	25.4	19.9	11.3	5.4	5.9	11.7
Predicted air temperature values, °C													
2018	-5.0	1.8	7.2	10.7	17.3	21.5	23.6	25.0	19.6	10.1	4.4	2.7	11.6
2019	-5.0	2.2	7.3	10.6	18.1	21.4	23.5	25.0	19.8	10.1	4.2	3.0	11.7
2020	-5.0	2.5	7.4	10.6	18.9	21.4	23.4	25.0	20.0	10.1	4.0	3.2	11.8
2021	-5.0	2.7	7.5	10.5	19.6	21.3	23.3	25.0	20.1	10.1	3.8	3.4	11.9
2022	-5.0	2.9	7.5	10.5	20.1	21.2	23.2	25.0	20.3	10.2	3.6	3.6	11.9
2023	-5.0	3.1	7.5	10.4	20.4	21.2	23.1	25.0	20.4	10.2	3.5	3.7	12.0
2024	-5.0	3.3	7.5	10.4	20.7	21.1	23.1	24.9	20.5	10.2	3.3	3.8	12.0
2025	-5.0	3.4	7.5	10.3	20.9	21.1	23.1	24.9	20.6	10.3	3.2	3.9	12.0

We used the ANN-based approach to forecast air temperature in Kherson region of Ukraine on the basis of the 8-year temperature data obtained at Kherson regional hydro-meteorological station. Besides, to compare the forecasts, we used another statistical method for the temperature prediction – Holt-Winters triple exponential smoothing, which is well-known as a method of high accuracy for the time series forecast with the effect of seasonality [16, 28, 40, 45, 46, 83].

The time line with accordance to the method is presented as: yi, …, yt, $y \in R$. The task of the time line seasonal forecasting is programed by using the following number of formulae 29:

$$y_{t+d} = a_t (T_t)^d \omega_t + (dmods) - s \tag{29}$$

where s is seasonality, ω_t $i \in 0, \ldots, s\text{-}1$ describes a season profile, T_t is a trend parameter, and a_t is a forecast parameter without influence of the trend and seasonality, respectively.

The results of the forecasts performed by using different methodological approaches are represented in the Tables 11 and 12.

The ANN-based air temperature predictive model was developed within NeuroXL Predictor add-in for Microsoft Excel (for more details go to Paragraph 2.1). The neural network with 50 neurons in hidden layer, momentum of 0.60, initial weights of 0.30, learning rate of 0.30 was created. We used a zero-based log-sigmoid activation function to perform training within 3000 epochs.

Table 13.
The comparison of the artificial neural network and Holt-Winters exponential smoothing based average air temperature forecast by the values of the discrepancy, oC

Year	Forecasting method		Discrepancy
	Holt-Winters	Neural network	
2018	11.5	11.6	+0.1
2019	10.6	11.7	+1.1
2020	12.0	11.8	-0.2
2021	12.0	11.9	-0.1
2022	11.7	11.9	+0.2
2023	10.7	12.0	+1.3
2024	12.2	12.0	-0.2
2025	12.2	12.0	-0.2

To compare the predictions of air temperature made by the means of Holt-Winters approach and ANN, the data of ANN-based forecast is provided in the Table 12. It should be mentioned that the ANN model for air temperature had the RSQ value of 0.02-0.92 that shows high fluctuations in the forecast in dependence on the month and allows us to conclude that the ANN-provided forecast is not quite a reliable one. Besides, it is obvious that the ANN-based forecast did not manage to handle the effect of seasonality that is a very important disadvantage. Inability to recognize and handle the effect of seasonality resulted in the linear nature of the predictions even notwithstanding the fact of usage of non-linear logistic sigmoid activation function.

As you can see in the Table 13, the comparison did not show any advantage of the ANN-based forecast over the traditional mathematical computations by using the exponential smoothing. The discrepancies are not very high, although they are quite distinctive. The reason of the ANN failure could be put on the little amount of the training data because we know that neural networks performance in predictions is better with larger number of the learning samples. Besides, the computation environment did not allow us to engage the factor of seasonality in the air temperature changes into the predictive model.

Therefore, traditional forecasting computations have to be admitted to be more reliable in the tasks of time series predictions of the environmental phenomenon with complicated seasonal affections. One thing that we should mention here is that often it is possible to get some increase (and in particular cases this increase is really sufficient) of the ANN-based predictive model performance by adjustment of the neural network architecture, choice of the most appropriate activation function and learning algorithm parameters, especially, momentum, number of epochs, etc. Of course, modern commercial software makes the forecasting easier (for example, such software as NeuroXL Predictor or Tiberius Data Mining Predictive Modeling Software 7) than in case of performing traditional mathematical calculations in the half-manual regime. But the neural network approach, as you could see in the example, needs to be implemented carefully and thoroughly tested to avoid dramatic mistakes in the forecasts and not make erroneous conclusions.

3.5. Tomato Productivity in the Irrigated Conditions of the South of Ukraine Depending on Technological and Biological Factors

The study devoted to the estimation of tomato productivity in the irrigated conditions of the South of Ukraine was carried out on the experimental fields of the Institute of Irrigated Agriculture of the National Academy of Agrarian Sciences of Ukraine with accordance to the modern standards of research and scientific work in agronomy [15, 96].

By the results of the crop productivity study, an empirical neural model was created by the means of STATISTICA Neural Networks software (v. 6.1) to determine the impact of each investigated factor on the yield of tomato. The feed-forward network had the structure of 5-11-1, and used a combination of different activation functions for the neurons: radial basis function and linear function (Fig. 26). The first input was the depth of primary tillage expressed in cm; the second input was the sum of the effective temperatures above 10°C; the third input was represented by the longevity of sun radiation in hours; the fourth input was the total water consumption of the crop expressed in m^3 per hectare; the fifth input were the doses of mineral fertilizers applied to the experimental field expressed in the kg of the active substance per hectare. The output was the yield of tomato fruit expressed in t per ha.

The results of the ANN-based estimation of the inputs' weights allowed us to determine that the highest effect on the crop productivity was provided by the factors of the effective temperatures sum and sunshine period duration. However, the other studied factors and their interactions had also significant effect on the crop yield formation.

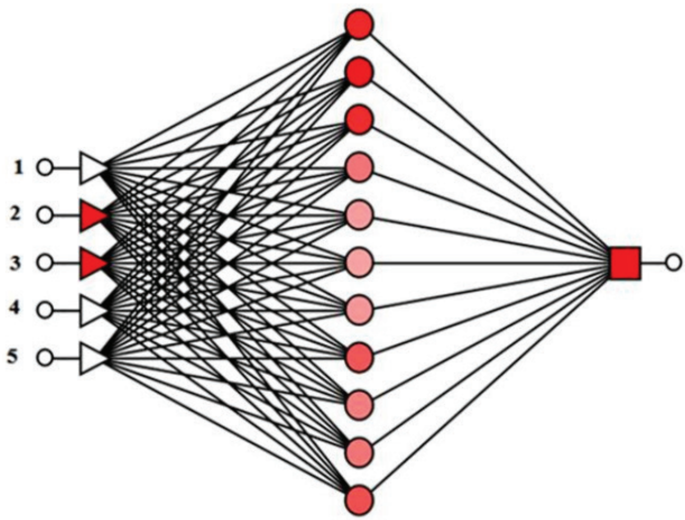


Figure 26. The structure of the artificial neural network for tomato fruit yield modeling

Table 14.
The main indexes of the artificial neural network for tomato fruit yield modeling, t ha-1

Architecture	Training productivity	Control productivity	Testing productivity	Training error	Control error	Testing error
RBF 5:5-11-1:1	0.1647	0.1990	0.1409	0.5148	0.9774	0.8592
Linear 4:4-1:1	0.2822	0.3555	0.3169	0.0732	0.0790	0.0997
MLP 4:4-6-1:1	0.1714	0.2865	0.2094	0.0446	0.0675	0.0655
Linear 5:5-1:1	0.2095	0.2272	0.2403	0.0543	0.0576	0.0753
MLP 5:5-6-1:1	0.2734	0.3404	0.2277	0.0232	0.0462	0.0479

The results of the ANN training results, errors and productivity are given in the Table 14.

The results of the above-mentioned investigations are the intermediate stage of the development of complex synergistic predictive models for the crop yield forecasting.

3.6. Predictive Models for the Irrigation Water Quality of the Ingulets Irrigation System: Artificial Neural Network Approach versus Conventional Exponential Smoothing Forecast

The deficiency and low quality of water for irrigation requires appropriate solutions [104, 116]. One of them is use of ameliorated water from different contaminated sources. To supply fields of Kherson and Mykolaiv regions of Ukraine with water the Ingulets irrigation system is used. The water of the system is contaminated by the effluent disposals and wastes of the metallurgic factories, so it needs significant amelioration to be safe for plants and soils [85, 117]. The new way of water quality improvement, which is based on the mixture of the Ingulets water with fresh water from the Karachuniv reservoir, was introducted in 2010. The goal of the study was to determine water quality of the Ingulets irrigation system by the agronomical criteria due to the new amelioration technique functioning with accordance to the international FAO requirements and Ukrainian standards. Also, we tried to make short-term forecast of the water quality by using the Holt-Winters multiplicative exponential smoothing algorithm [9, 118].

The study was conducted each year during the period from 2007 to 2017. Water samples from the Ingulets irrigation system main channel (latitude 47°0'55"N and longitude 32°47'20"E) were taken each month within the period from April to September. The collected water samples were analyzed in the laboratory of the Mykolaiv Regional Office of Water Management by the generally accepted procedures [3, 140].

Sodium adsorption ratio (*SAR*) was calculated by using the formula 30 [4]:

$$SAR = \frac{Na}{\sqrt{\frac{Ca+Mg}{2}}} \tag{30}$$

where *SAR* is the sodium adsorption ratio, meq L^{-1}; *Na, Ca, Mg* – ions content, expressed in meq L^{-1}.

Water toxicity might be expressed in the equivalents of Chloride ions (eCl^-) that was calculated by using the formula 31 [140]:

$$eCl^- = Cl^- + 0{,}2SO_4^{2-} + 0{,}4HCO_3^- + 10CO_3^{2-} \tag{31}$$

where Cl^-, SO_4^{2-}, HCO_3^-, CO_3^{2-} – ions content, expressed in meq L^{-1}.

Magnesium adsorption ratio (*MAR*) is an indicator of the inter-relations between the ions of Ca^{2+} and Mg^{2+} in water. The optimal value has to be less than 1.0. The index was calculated by using the formula 32:

$$MAR = \frac{Mg}{Ca} \tag{32}$$

where *MAR* is the magnesium adsorption ratio, units; *Ca, Mg* – ions content, expressed in meq L^{-1}.

Sodium percentage is an important indicator of alkalization hazard (*SP*). The index was calculated by using the formula 32 [128]:

$$SP = \left(\frac{Na}{Na+K+Mg+Ca}\right) \times 100\% \tag{33}$$

where *SP* is sodium percentage, expressed in per cents (%); *Na, K, Mg, Ca* – ions content, expressed in meq L^{-1}.

Kelly's ratio (*KR*) is used to categorize irrigation water by its potential dangerousness for soils and plants. The water is considered to be favorable for irrigation without any limitations when the index is <1.0; if it is in the range of 1.0-2.0 the water is marginal; and if the ratio scores over 2.0 points the water is considered to be harmful and unsuitable for irrigation purposes [70]. The index was calculated by the formula 34:

$$KR = \frac{Na}{Ca+Mg} \tag{34}$$

where *KR* is the Kelly's ratio, units; *Na, Ca, Mg* – ions content, expressed in meq L^{-1}.

Permeability index was proposed by Doneen to classify irrigation water affectation on soil permeability by using the ions content in the water [85]. The index was calculated by the formula 35:

$$PI = \frac{\sqrt{HCO_3}+(Na+K)}{Ca+Mg+Na+K} \times 100 \tag{35}$$

where *PI* is the permeability index, units; *Na, K, Ca, Mg,* HCO_3 – ions content, expressed in meq L^{-1}.

Standard deviation (*SD*) of the water quality criteria was calculated by using the formula 36 [42, 87]:

$$SD = \frac{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2}}{N-1} \tag{36}$$

where *SD* is the standard deviation; x_1, …, x_n are the observed values of the water quality criteria; *N* is the number of observations.

The coefficient of variation of the water quality criteria was calculated by using the formula 37 [34]:

$$CV = \frac{SD}{\bar{x}} \tag{37}$$

where *C*V is the coefficient of variation; *SD* is the standard deviation; *x* is the mean value of the water quality criterion.

Water quality in the Ingulets irrigation system was forecasted by using the triple exponential smoothing with handling of the seasonal effects [16, 28, 40, 45, 46, 83].

Table 15.
The results of the Ingulets irrigation water analysis with accordance to DSTU 2730-94 and FAO standards

Years	TDS, mg L-1	pH, units	SAR, meq L-1	eCl-, meq L-1	MAR, units	SP, per cents	KR, units	PI, units
2007	2180	7.94	7.94	21.63	1.25	58.9	1.44	1.06
2008	2008	7.31	5.50	19.05	1.18	48.8	0.95	1.03
2009	2186	7.84	5.37	25.89	1.50	46.4	0.87	0.87
2010	2280	8.72	7.94	20.54	1.47	57.8	1.37	0.98
2011	1673	8.48	5.78	12.07	1.38	52.8	1.12	1.19
2012	1600	8.32	5.00	11.70	1.03	48.2	0.93	1.20
2013	1471	8.24	4.33	11.30	1.13	47.2	0.88	1.30
2014	1578	8.33	5.78	12.22	1.35	53.9	1.17	1.29
2015	1458	8.35	4.79	10.96	1.68	47.8	0.92	1.18
2016	1448	8.23	4.79	10.42	1.39	49.2	0.97	1.29
2017	1489	8.30	5.11	10.49	1.24	50.0	1.00	1.21
SD	332	0.37	1.21	5.55	0.18	4.28	0.20	0.14
x	1761	8.19	5.67	15.11	1.33	51.0	1.05	1.14
CV,%	18.85	4.57	21.29	36.72	13.88	8.39	19.05	12.28

Multiplicative method of the Holt-Winters algorithm was used [40]. The essence of the applicated method is in solving the task of the time line forecasting. The time line is presented as: $yi, \ldots, yt, yt \in R$. The task of the time line forecasting looks as follows (formulas 38-41):

$$y_t + d = a_t(\tau_t)^d \Omega_t + (d mod s) - s \quad (38)$$

$$a_t = a_1 y_i / \Omega_{i-s} + (1 - a_1) a_{t-1} \tau_{t-1} \quad (39)$$

$$\tau_t = a_3 a_i / a_{i-1} + (1 - a_3) \tau_{t-1} \quad (40)$$

$$\Omega_t = a_2 y_i / a_i + (1 - a_2) \Omega_{t-s} \quad (41)$$

where s – seasonality, $\Omega_i, i \in 0,\ldots,s-1$ – season profile, τ_t – trend parameter, a_t – forecast parameter without influence of the trend and seasonality.

Forecasting by Holt-Winters methodology was performed within LibreOffice 5.4 and Microsoft Excel 2019 software.

Forecasting by using the ANN-based approach by the means of multi-layer perceptron of structure 1-12-1 for each qualitative index of the Ingulets water was performed within NeuroXL Predictor add-in for Microsoft Excel 2019. The inputs were represented by time series of the years of the water quality study. The perceptron had 12 neurons in hidden layer. We used training process within 3000 epochs with the initial weight of 0.30, learning rate of 0.30, momentum of 0.40. A hyperbolic tangent function was used to activate the neurons.

It was established that water quality in the Ingulets irrigation system is still poor, but it has been significantly improved since 2010 by the new amelioration technique that resulted in lower values of the main quality criteria (Table 15). The water belongs

Table 16.
The results of the Ingulets irrigation water quality forecasting by Holt-Winters exponential smoothing

Years	TDS, mg L-1	pH, units	SAR, meq L-1	eCl-, meq L-1	MAR, units	SP, per cents	KR, units	PI, units
2018	1450	8.21	4.82	10.00	1.13	49.3	0.84	1.24
2019	1366	8.23	4.56	9.47	1.29	48.3	0.90	1.41
2020	1397	8.15	4.73	9.05	1.07	49.5	0.96	1.46
2021	1315	8.17	4.48	8.52	1.22	48.6	0.79	1.33
2022	1343	8.09	4.64	8.11	1.00	49.7	0.85	1.50
2023	1263	8.11	4.40	7.56	1.14	48.9	0.90	1.55
2024	1290	8.03	4.55	7.16	0.93	50.0	0.73	1.42
2025	1212	8.05	4.31	6.61	1.05	49.3	0.79	1.59

to the second class "Limited suitable for irrigation" according to the DSTU 2730-94 requirements. FAO standards also define the Ingulets irrigation system water as water with restrictions for use in irrigation.

The results of the water quality analysis have shown that the main problems in the Ingulets irrigation system are high total dissoluble salts (TDS), toxic ions (eCl^-) and sodium content: 1489-2280 mg/L, 10.49-21.63 me/L and 11.83-21.97 me/L respectively. Irrigation with such water leads to deterioration of the physical, chemical and biological properties of soils, decrease in crops growth, productivity and yield quality [4, 7, 37, 70, 72, 88, 93, 100, 146]. The statistical data processing has also shown that the most variable quality criteria in the Ingulets irrigation system water are chloride and sodium ions content (CV was 42.76% and 25.84% respectively), and the most stable one was power of hydrogen (pH) with CV at 8.39%. The results of the triple exponential smoothing conducted by using the Holt-Winters multiplicative algorithm have shown probability of significant improvement of the Ingulets irrigation system water quality till 2025. If current water amelioration system function properly, significant decrease of TDS and toxic ions content will be achieved: to 1212 mg/L and 6.61 me/L respectively.

The forecast by Holt-Winters exponential smoothing method has also shown that the SP and SAR values will probably leave on a higher, then optimal for use of the water for irrigation without any restrictions, level: 48.3-50.0% and 4.31-4.82 me/L respectively (Table 16). The designed forecast model should be useful for management and control of the Ingulets irrigation water quality. Application of the Holt-Winters multiplicative algorithm to forecasting water quality in the Ingulets irrigation system keeps on the modern trend of the mathematical modeling use in environmental management [86, 101, 108, 120].

The results of the ANN-based prediction of the water quality indexes are quite different (Table 17).

Table 17.
The results of the Ingulets irrigation water quality forecasting by the artificial neural network

Years	TDS, mg L-1	pH, units	SAR, meq L-1	eCl-, meq L-1	MAR, units	SP, per cents	KR, units	PI, units
2018	1458	8.30	4.83	10.20	1.37	49.6	0.98	1.23
2019	1454	8.31	4.81	10.10	1.37	49.5	0.98	1.22
2020	1451	8.31	4.80	10.03	1.38	49.5	0.98	1.22
2021	1449	8.32	4.79	9.98	1.38	49.4	0.97	1.21
2022	1447	8.32	4.78	9.94	1.38	49.4	0.97	1.20
2023	1446	8.33	4.78	9.91	1.39	49.3	0.97	1.19
2024	1445	8.33	4.77	9.89	1.39	49.3	0.97	1.19
2025	1444	8.33	4.77	9.87	1.40	49.3	0.96	1.18

Table 18.
The values of the coefficient of determination for the artificial neural network predictive model of the Ingulets water quality

Index	RSQ
TDS, mg L-1	0.9882
pH, units	0.9876
SAR, meq L-1	0.4062
eCl-, meq L-1	0.9856
SP, %	0.3602
MAR, units	0.0185
KR, units	0.4065
PI, units	0.9206

The ANN-based prediction is more pessimistic than Holt-Winters based one. We can see no dramatical improvement of the water quality until 2025 by any qualitative parameter, excluding only water toxicity in eCl^-, which is expected to decrease to appropriate levels of <10 meq L^{-1}, and Kelly's ratio, which is augured to stabilize on the levels of <1.0 point. The forecast provided by the multi-layer perceptron is more restrain, however, the RSQ values of the predictive model conceive us to believe in this prediction – the coefficient of determination values is very high for most predicted qualitative parameter, excluding sodium percentage, magnesium adsorption ratio, sodium adsorption ratio and Kelly's ratio, which are the derivative indexes of the artificial manual calculations (Table 18).

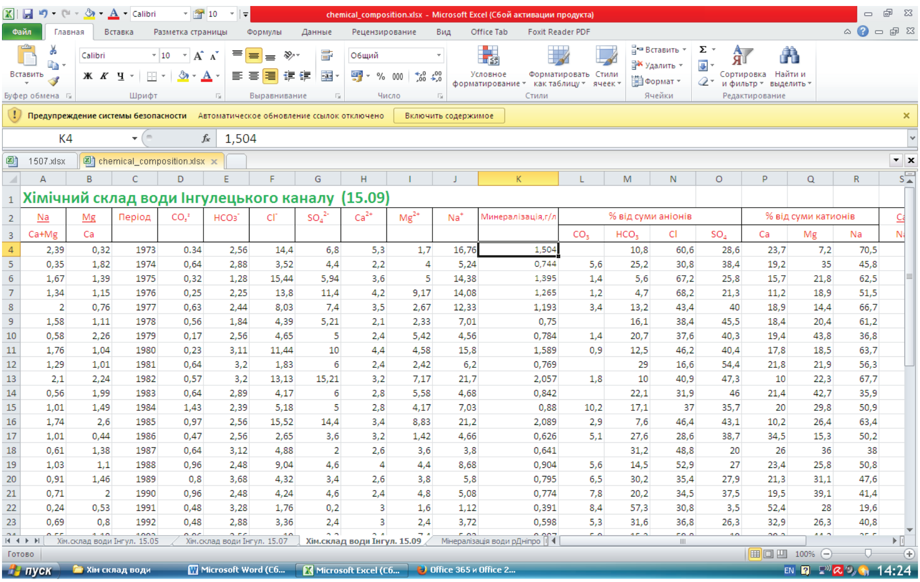
As it can be observed, the predictive model provided by the ANN-based methodology seems to be more natural and trustworthy than the model created by the means of traditional mathematical statistics of Holt-Winters. Therefore, we think that artificial neural networks should be preferred in dealing with similar practical and theoretical tasks.

3.7. Practical Aspects of Cluster Analysis of the Ingulets Irrigation Water and Water of the River Dnipro

Productivity of irrigation greatly depends on the quality of water because artificial humidification with water of low quality (2nd class or 3rd class water be agronomical criteria) leads to considerable deterioration of soil properties (soil crusting, salinization, alkalization, decrease of permeability, etc.) and results in yield losses. Therefore, low-quality water should be applied to the fields with a great care and only has to be accompanied by the respective amelioration measures.

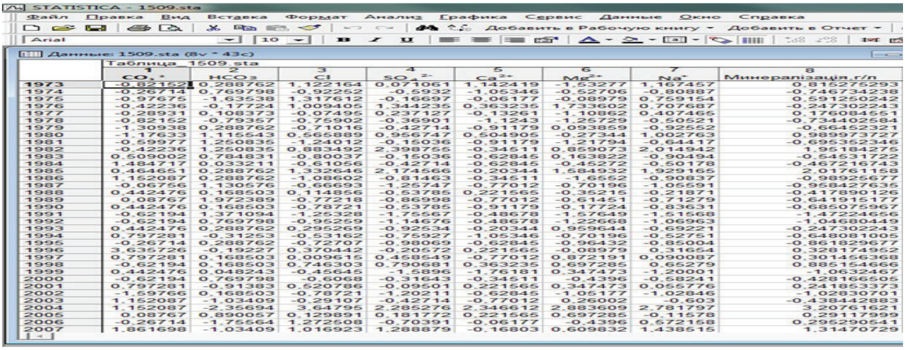
Irrigation water of the Ingulets irrigation system belongs to the 2nd class water with accordance to the State Standard of Ukraine DSTU 2730-94 [140]. Uncontrolled irrigation with this water is dangerous and increases salinity hazard, let alone the risks of soil fertility losses due to the malfunction of favorable soil biota [4]. Scrupulous monitoring of the water quality by agronomical and toxicological parameters has to be continuously conducted to prevent unexpected unfavorable consequences of irrigation. The study devoted to deep classification by the means of the ANN-based cluster analysis was performed to assess a dynamic of the Ingulets irrigation water quality in the long-term period of 1973-2015.

Cluster analysis of the water quality was carried out within the software complex of STATISTICA 6.1. The data related to the water quality parameters was previously generalized and represented in the MS Excel spreadsheet format before the export of them into STATISTICA environment (Fig. 27, 28).


Хімічний склад води Інгулецького каналу (15.09)

Na / Ca+Mg	Mg / Ca	Період	CO_3^{2-}	HCO_3^-	Cl^-	SO_4^{2-}	Ca^{2+}	Mg^{2+}	Na^+	Минералізація,г/л	% від суми аніонів: CO_3	HCO_3	Cl	SO_4	% від суми катионів: Ca	Mg	Na
2,39	0,32	1973	0.34	2,56	14,4	6,8	5,3	1,7	16,76	1,504		10,8	60,6	28,6	23,7	7,2	70,5
0,35	1,82	1974	0,64	2,88	3,52	4,4	2,2	4	5,24	0,744	5,6	25,2	30,8	38,4	19,2	35	45,8
1,67	1,39	1975	0,32	1,28	15,44	5,94	3,6	5	14,38	1,395	1,4	5,6	67,2	25,8	15,7	21,8	62,5
1,34	1,15	1976	0,25	2,25	13,8	11,4	4,2	9,17	14,08	1,265	1,2	4,7	68,2	21,3	11,2	18,9	51,5
2	0,76	1977	0,63	2,44	8,03	7,4	3,5	2,67	12,33	1,193	3,4	13,2	43,4	40	18,9	14,4	66,7
1,58	1,11	1978	0,56	1,84	4,39	5,21	2,1	2,33	7,01	0,75		16,1	38,4	45,5	18,4	20,4	61,2
0,58	2,26	1979	0,17	2,56	4,65	5	2,4	5,42	4,56	0,784	1,4	20,7	37,6	40,3	19,4	43,8	36,8
1,76	1,04	1980	0,23	3,11	11,44	10	4,4	4,58	15,8	1,589	0,9	12,5	46,2	40,4	17,8	18,5	63,7
1,29	1,01	1981	0,64	3,2	1,83	6	2,4	2,42	6,2	0,769		29	16,6	54,4	21,8	21,9	56,3
2,1	2,24	1982	0,57	3,2	13,13	15,21	3,2	7,17	21,7	2,057	1,8	10	40,9	47,3	10	22,3	67,7
0,56	1,99	1983	0,64	2,89	4,17	6	2,8	5,58	4,68	0,842		22,1	31,9	46	21,4	42,7	35,9
1,01	1,49	1984	1,43	2,39	5,18	5	2,8	4,17	7,03	0,88	10,2	17,1	37	35,7	20	29,8	50,9
1,74	2,6	1985	0,97	2,56	15,52	14,4	3,4	8,83	21,2	2,089	2,9	7,6	46,4	43,1	10,2	26,4	63,4
1,01	0,44	1986	0,47	2,56	2,65	3,6	3,2	1,42	4,66	0,626	5,1	27,6	28,6	38,7	34,5	15,3	50,2
0,61	1,38	1987	0,64	3,12	4,88	2	2,6	3,6	3,8	0,641		31,2	48,8	20	26	36	38
1,03	1,1	1988	0,96	2,48	9,04	4,6	4	4,4	8,68	0,904	5,6	14,5	52,9	27	23,4	25,8	50,8
0,91	1,46	1989	0,8	3,68	4,32	3,4	2,6	3,8	5,8	0,795	6,5	30,2	35,4	27,9	21,3	31,1	47,6
0,71	2	1990	0,96	2,48	4,24	4,6	2,4	4,8	5,08	0,774	7,8	20,2	34,5	37,5	19,5	39,1	41,4
0,24	0,53	1991	0,48	3,28	1,76	0,2	3	1,6	1,12	0,391	8,4	57,3	30,8	3,5	52,4	28	19,6
0,69	0,8	1992	0,48	2,88	3,36	2,4	3	2,4	3,72	0,598	5,3	31,6	36,8	26,3	32,9	26,3	40,8

Figure 27. Screenshot of the MS Excel spreadsheet with the information regarding ions content in the water of the Ingulets irrigation system for the period of 1973-2015



STATISTICA - 1509.sta

Данные: 1509.sta (8v * 43c)

Таблица_1509.sta

	1 CO_3^{2-}	2 HCO_3	3 Cl	4 SO_4^{2-}	5 Ca^{2+}	6 Mg^{2+}	7 Na^+	8 Минерапізація,г/л
1973	-0,82152	0,288762	1,122164	0,071061	1,142419	-1,53277	1,167457	0,815275293
1974	-0,26714	0,769798	-0,92252	-0,5932	-1,05346	-0,52706	-0,80887	-0,746734238
1975	-0,97675	-1,63538	1,317612	-0,16697	-0,06177	-0,08979	0,759154	0,591250242
1976	-0,42236	-0,17724	1,009405	1,344235	0,363235	1,733602	0,707687	-0,247302243
1977	-0,28931	0,108373	-0,07495	0,237127	-0,13261	-1,10862	0,407465	0,176084551
1978	-0,82152	-0,79357	-0,75902	-0,36901	-1,1243	-1,257729	-0,50521	-0,734402584
1979	-1,30938	0,288762	-0,71016	-0,42714	-0,91179	0,093859	-0,92552	-0,66452321
1980	-1,17633	1,115543	0,565889	0,956747	0,504905	-0,27344	1,002763	0,989973727
1981	-0,59977	1,250835	-1,24012	-0,15036	-0,91179	-1,21794	-0,64417	-0,695352346
1982	-0,42236	1,250835	0,883492	2,398755	-0,34511	0,859073	2,014942	1,95184275
1983	0,509002	0,784831	-0,80037	-0,15036	-0,62845	0,163822	-0,90494	-0,54531722
1984	1,484717	0,033211	-0,61056	-0,42714	-0,62845	-0,45272	-0,50178	-0,467216743
1985	0,464651	0,288762	1,332646	2,174566	-0,20344	1,584932	1,929165	2,01761158
1986	1,152087	0,288762	-1,08602	-0,81463	-0,34511	-1,6552	-0,90837	-0,98925677
1987	-0,06756	1,130576	-0,66693	-1,25747	-0,77012	-0,70196	-1,055591	-0,958427635
1988	0,442476	0,168503	0,114856	-0,53785	0,221565	-0,352215	-0,21871	-0,417890126
1989	0,08767	1,972389	-0,77218	-0,86998	-0,77012	-0,61451	-0,71279	-0,641915177
1990	0,442476	0,168503	-0,78721	-0,53785	-0,91179	-0,17724	-0,83631	-0,685075967
1991	-0,62194	1,371094	-1,25328	-1,75567	-0,48678	-1,57649	-1,51568	-1,47224656
1992	-0,62194	0,769798	-0,95259	-1,14676	-0,48678	-1,22668	-1,06963	-1,04680449
1993	0,442476	0,288762	0,295269	-0,92534	-0,20344	0,959644	-0,69221	-0,247302243
1994	0,797281	-0,31253	-0,53162	-0,75927	-1,05346	-0,70196	-0,52751	-0,648081005
1995	-0,26714	0,288762	-0,72707	-0,98069	-0,62845	-0,96432	-0,850004	-0,861829677
1996	3,635726	-0,19227	0,370442	-0,20572	0,221565	-0,08979	0,31654	0,328174952
1997	0,797281	0,168503	0,0009615	0,458549	-0,77012	0,872191	0,090087	0,301456368
1998	-0,62194	0,168503	0,746303	0,790681	0,363235	0,697285	0,65279	0,885154666
1999	0,442476	0,048243	-0,45645	-1,5896	-1,76181	0,347473	-1,20001	-1,0632467
2000	-0,62194	0,769798	-0,6068	-0,31643	-0,34511	-0,4396	-0,58241	-0,428166505
2001	0,797281	-0,91383	0,520786	-0,09501	0,221565	0,347473	0,055776	0,241853373
2002	-1,59766	0,168503	-0,78721	-1,20211	-0,62845	-1,05177	-1,02846	-1,02830701
2003	1,152087	-1,03409	-0,29107	-0,42714	-0,77012	0,26002	-0,603	-0,438442883
2004	1,152087	-2,35694	3,64795	2,285276	2,346612	2,883609	2,781797	3,20761621
2005	0,08767	0,890057	0,129891	0,181772	0,221565	0,697285	-0,11578	0,29117999
2006	-0,26714	-1,75564	1,272508	-0,70391	-0,06177	-0,4396	0,572158	0,295290541
2007	1,861698	-1,03409	1,016923	1,288879	-0,16803	0,609832	1,438515	1,31470729

Figure 28. Standardized data of the tables containing mineralization of the water dated 15th of September, 1973-2015

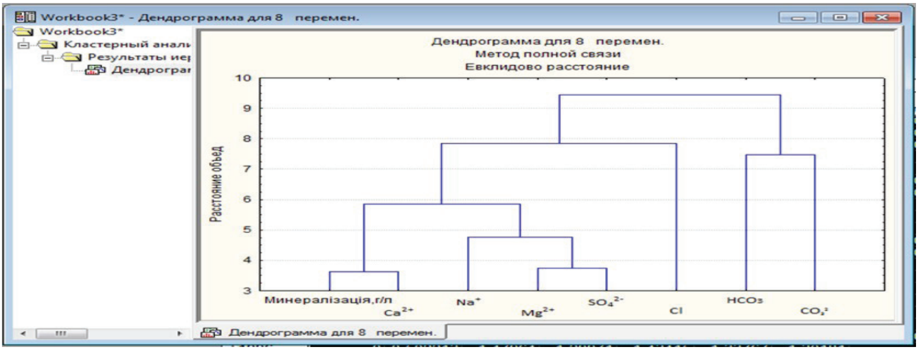


Figure 29. The hierarchic model of the cluster analysis of the water from the Ingulets irrigation system for the period of 1973-2015

After the operations of the data standardization we continued to the "Hierarchic classification" tool of the "Cluster analysis" module of the program. To form a hierarchic diagram, we choose the method of full connection, which determines the distance between clusters as the maximum distance between two objects in different clusters (the most distant positions). As a result of this operation we obtained the hierarchic tree presented in the Fig. 29.

On the basis of the data obtained we formed clusters, which are created for the vertical diagram by using the "top-down" model. The indexes, which are the most similar, are embraced and united into one cluster. Every node of the diagram is a union of two or more clusters, and situation of the nodes on the vertical axis determines the distance on which certain clusters were connected.

The level of is calculated by the following formula 42:

$$d(x,y) = \sqrt{\sum_{i-1}^{n}(x_i - y_i)^2} \qquad (42)$$

Analyzing the results of classification, we conjectured that the studied characteristics of the ion and salt content of the water are formed in four natural clusters (1st

Table 19.
The results of ANOVA related to the cluster analysis of the Ingulets irrigation water by k-means method (for the water samples collected on 15th of September)

Studied parameter	Inter	cc	Intro	cc	F	Significance
CO_3^{2-}	18.22074	4	24.77926	49	9.96119	0.000052
HCO^{3-}	11.14581	4	40.86419	49	4.69041	0.006844
Cl^-	25.41665	4	16.58445	49	19.92459	0.000000
SO_4^{2-}	40.89749	4	11.10251	49	46.17805	0.000000
Ca^{2+}	19.44948	4	22.65052	49	11.10541	0.000021
Mg^{2+}	19.27717	4	22.72284	49	11.02870	0.000022
Na^+	44.78601	4	7.21499	49	62.68642	0.000000
Mineralization	42.10275	4	9.89725	49	42.16684	0.000000

Table 20.
The results of ANOVA related to the cluster analysis of the Ingulets irrigation water by k-means method (for the water samples collected on 15th of May)

Studied parameter	Inter	cc	Intro	cc	F	Significance
CO_3^{2-}	26.57292	4	15.42708	49	22.49241	0.000000
HCO_3^-	9.17411	4	42.82589	49	4.64421	0.020996
Cl^-	16.74214	4	25.25786	49	8.61704	0.000164
SO_4^{2-}	22.09274	4	19.90727	49	14.42716	0.000002
Ca^{2+}	15.55882	4	26.44118	49	7.64961	0.000487
Mg^{2+}	19.75545	4	22.24455	49	11.54544	0.000015
Na^+	20.44408	4	21.65692	49	12.21144	0.000009
Mineralization	24.68272	4	17.41728	49	18.52921	0.000000

cluster – mineralization, Ca^{2+}; 2nd cluster – Na^+, Mg^{2+}, SO_4^{2-}; 3rd cluster – Cl^-; 4th cluster – CO_3^{2-}, HCO_3^-).

The above-mentioned hypothesis was checked by using the method of k-means. To perform this operation, we choose "Clusterization by k-means" in the dialog window of "Cluster analysis" module of STATISTICA and set the necessary parameters for further classification. The initial number of clusters was set as 4. As a result of the k-means classification we obtained the intergroup variances by the studied parameters of the chemical composition of the water, which are further compared to intragroup variances for making the decision whether means for particular arguments in different groups are statistically significant. The results of ANOVA are represented by STATISTICA in the form of Table 19.

The ANOVA results showed that on the assumption of the amplitude and significance levels of F-statistics, arguments Na^+, mineralization, SO_4^{2-} and Cl^- are the main factors in the question of distribution of the objects by the clusters.

However, ANOVA results for the water samples collected on 15th of May certified that the leading role in taking the decision of placement of the objects in concrete

Table 21.
The results of ANOVA related to the cluster analysis of the Ingulets irrigation water by k-means method (for the water samples collected on 15th of July)

Studied parameter	Inter	cc	Intro	cc	F	Significance
CO_3^{2-}	21.90194	4	20.09806	49	14.16680	0.000002
HCO_3^-	11.22676	4	40.77424	49	4.74269	0.006484
Cl^-	12.89094	4	29.10907	49	5.75704	0.002426
SO_4^{2-}	41.04442	4	10.96658	49	46.78762	0.000000
Ca^{2+}	42.70105	4	9.29895	49	45.71640	0.000000
Mg^{2+}	24.64500	4	17.45500	49	18.46068	0.000000
Na^+	27.65270	4	14.44740	49	25.05594	0.000000
Mineralization	29.86498	4	12.14602	49	41.99004	0.000000

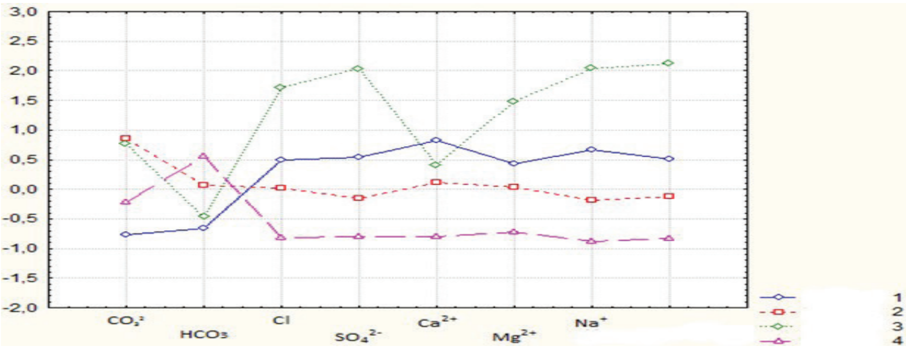


Figure 30. The graph of the means of the ions content of the Ingulets irrigation water for the period of 1973-2015 water samples collected on 15th of September)

cluster were arguments CO_3^{2-}, mineralization and SO_4^{2-} (Table 20), and for the water samples collected on 15th of July – Ca^{2+}, SO_4^{2-}, mineralization (Table 21).

Another way of determination of the nature of clusters is a check of the means for each cluster, which is presented as a graph of the means. Usually this graph gives the best understanding of the results (Fig. 30).

The fourth cluster is characterized by the higher values of mineralization, SO_4^{2-}, Cl^-, Na^+, Mg^{2+} comparatively to the clusters 1, 2, 3. For example, the components of the 2nd cluster are characterized by the higher content of CO_3^{2-}, the elements of the 1st cluster are characterized by the lowest indexes of CO_3^{2-}.

Looking at the graphs of the means for the samples of water collected on 15th of May and 15th of July it is possible to observe that in different vegetation periods the indexes of ions and salt content have different behavior.

For example, the dynamics of the graph for the period of 15th of May is characterized with sharp rapid fluctuations between the objects (Fig. 31).

The graph of the means of the 1st cluster differs from the other by the highest value of CO_3^{2-}, and the lowest value of Mg^{2+}; the years from the 2nd cluster have high CO_3^{2-}

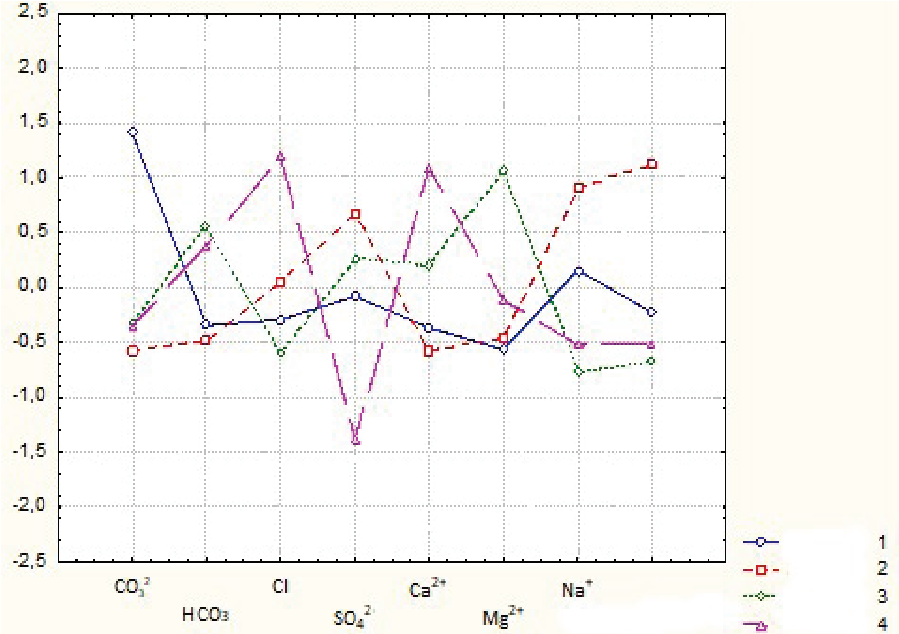


Figure 31. The graph of the means of the ions content of the Ingulets irrigation water for the period of 1973-2015 water samples collected on 15th of May)

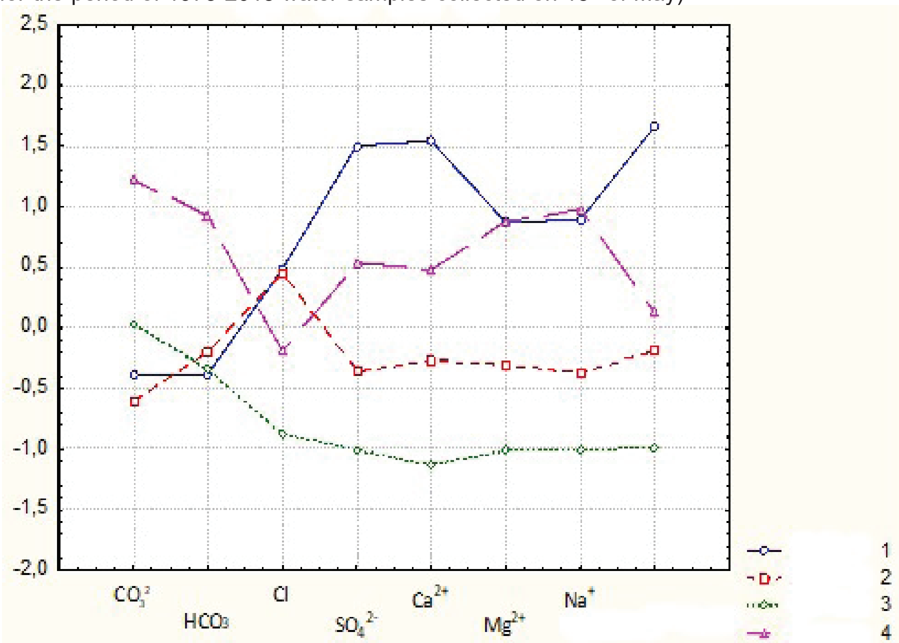


Figure 32. The graph of the means of the ions content of the Ingulets irrigation water for the period of 1973-2015 water samples collected on 15th of May)

Table 22.
The indexes of ions content and mineralization of the Ingulets irrigation water for the samples collected on 15th of May (1973-2015) with the results of classification by clusters

Period	CO_4^{2}	HCO_4	Cl	SO_4^{2-}	Ca^{2+}	Mg^{2+}	Na^{+}	Mineralization, g L-1	Cluster
1973	1.12	2.16	9.60	4.90	4.60	4.20	10.98	1.82	2
1973	0.95	1.85	11.45	2.74	4.47	4.67	9.82	1.65	1
1975	1.25	1.91	14.60	2.25	4.80	4.80	11.54	1.18	4
1976	0.80	1.75	10.78	9.27	2.64	5.00	24.49	2.47	2
1977	1.44	1.89	9.21	7.60	4.80	4.84	11.50	2.11	1
1978	1.20	2.16	7.14	15.00	4.20	4.42	16.87	0.75	2
1979	0.40	2.49	10.49	18.10	4.0	6.84	20.44	2.04	2
1980	1.14	4.11	8.08	11.79	4.05	7.00	14.06	1.98	1
1981	1.28	4.56	16.41	44.21	8.80	14.42	40.86	4.50	2
1982	1.29	4.00	22.11	11.21	7.85	8.00	21.47	2.28	4
1984	0.44	4.04	6.96	6.60	4.60	4.00	10.42	1.78	2
1984	0.17	2.44	8.96	12.21	4.80	4.00	15.87	1.55	2
1985	0.80	4.04	12.00	16.60	4.60	9.00	18.84	1.15	2
1986	0.72	4.20	6.65	7.40	2.60	5.42	9.24	1.12	1
1987	0.80	2.72	18.72	14.00	4.60	12.00	19.64	0.68	2
1988	0.42	2.80	4.60	4.20	4.20	4.60	4.12	0.70	4
1989	0.65	4.04	5.84	4.00	4.40	4.60	4.88	0.75	4
1990	0.80	2.08	6.64	4.60	4.40	4.20	7.52	0.89	1
1991	0.96	2.88	16.88	9.00	6.60	8.80	14.42	0.84	4
1992	0.16	2.56	1.28	2.80	2.40	1.80	2.60	0.47	4
1994	0.64	2.40	4.48	6.00	2.60	4.20	6.72	0.87	1
1994	0.48	2.96	9.12	6.80	5.00	5.60	8.76	1.21	4
1995	0.96	2.24	4.44	6.80	4.20	2.00	8.24	0.89	1
1996	0.48	4.12	9.84	9.60	4.20	9.80	9.04	1.42	4
1997	1.60	4.04	10.16	19.00	5.00	10.80	18.00	0.54	1
1998	0.48	4.68	10.42	15.40	4.20	12.00	14.68	0.85	4
1999	0.64	4.28	9.44	5.40	5.60	6.60	6.56	1.16	4
2000	0.64	4.46	9.28	12.40	4.60	8.60	12.48	1.64	3
2001	0.80	4.20	9.12	14.60	5.20	8.40	14.12	0.99	4
2002	0.42	4.64	7.52	11.00	4.20	10.80	10.08	1.54	4
2004	0.64	4.12	7.60	4.40	4.80	5.20	6.96	0.98	4
2004	0.42	4.00	7.04	8.20	5.00	8.40	6.16	1.24	4
2005	0.96	4.00	7.04	12.00	2.40	10.40	11.20	1.54	3
2006	0.96	4.68	8.00	12.00	5.00	8.20	11.47	1.57	4
2007	0.64	4.20	7.04	15.60	2.60	2.40	11.08	1.44	2
2008	0.48	4.92	10.08	15.80	4.24	7.40	17.48	1.97	2
2009	0.16	4.40	11.20	11.40	4.94	8.48	14.16	1.54	4
2010	0.91	4.10	8.28	14.28	5.29	9.74	15.48	1.77	2
2011	0.87	4.86	7.29	9.48	4.99	8.29	11.98	1.61	2
2012	0.79	4.48	9.48	11.02	4.02	7.88	14.55	2.21	1
2014	0.88	4.41	6.98	10.44	4.20	8.41	7.88	2.12	1
2014	0.71	4.29	7.11	10.98	4.40	9.01	12.74	1.98	4
2015	0.69	4.17	8.18	11.49	4.80	8.04	10.67	2.10	3

values, the maximum of SO_4^{2-}, mineralization and Na^+; the 3rd cluster unites the years with the highest values of Mg^{2+} and HCO_3^- with the lowest mineralization, Cl^- and Na^+ content; the 4th cluster embraces the years with the maximum Cl^- and Ca^{2+}.

The graph of the water samples, which were collected on 15th of July is more stable in its fluctuations (Fig. 32).

Table 23.
The indexes of ions content and mineralization of the Ingulets irrigation water for the samples collected on 15th of July (1973-2015) with the results of classification by clusters

Period	CO_4^{2}	HCO_4	Cl	SO_4^{2-}	Ca^{2+}	Mg^{2+}	Na^{+}	Mineralization, g L-1	Cluster
1973	0.74	2.96	6.92	6.10	2.00	4.00	9.98	2.04	1
1973	0.24	2.84	5.68	5.90	4.10	4.40	7.26	0.95	1
1975	0.48	2.96	4.98	4.59	4.66	5.02	6.49	0.88	1
1976	0.98	4.14	4.72	4.21	2.80	1.84	5.44	0.64	4
1977	0.57	4.00	6.94	7.50	4.10	4.17	10.74	2.25	1
1978	0.80	4.04	7.77	8.50	4.10	4.67	11.44	1.25	1
1979	0.57	2.48	5.75	5.60	2.20	4.58	8.62	0.94	1
1980	0.40	4.44	8.41	12.21	4.80	5.50	15.06	2.15	2
1981	0.80	2.97	11.24	14.21	4.60	6.58	17.00	1.78	2
1982	0.48	2.92	10.48	9.40	4.40	5.84	12.50	1.45	2
1984	0.80	4.20	4.96	4.80	4.20	4.58	6.98	0.89	1
1984	0.47	2.95	14.27	9.00	4.00	5.58	15.60	1.59	2
1985	0.64	2.48	11.84	14.00	4.80	6.84	18.40	1.76	2
1986	0.44	4.04	5.69	7.00	2.80	4.84	9.40	1.05	1
1987	0.57	4.04	4.96	2.60	2.20	4.20	4.20	0.82	4
1988	0.48	4.04	2.00	2.60	2.20	2.00	4.42	1.54	4
1989	0.58	4.52	1.84	4.20	2.60	4.20	2.76	0.65	3
1990	0.64	2.64	1.60	2.60	2.40	2.40	2.60	0.50	4
1991	0.42	2.64	4.56	2.60	2.80	4.20	4.12	1.44	4
1992	0.48	2.88	1.84	1.60	2.20	1.60	4.00	0.46	4
1994	0.64	2.56	2.16	1.20	2.40	2.00	2.16	0.44	4
1994	0.16	2.64	1.52	1.20	2.00	2.20	1.42	0.47	4
1995	0.64	2.56	4.20	4.00	2.40	2.80	4.20	0.61	4
1996	0.80	2.42	1.84	1.00	2.00	2.00	1.96	0.41	4
1997	1.28	2.88	11.28	7.40	4.80	6.40	11.64	0.72	4
1998	0.42	4.76	5.60	5.60	4.20	2.60	8.48	1.68	1
1999	0.42	4.04	5.92	6.80	4.60	4.40	9.08	1.05	1
2000	0.64	2.88	6.08	4.00	2.80	5.20	4.60	0.78	3
2001	0.64	4.46	4.80	4.40	4.20	7.40	1.60	1.84	1
2002	0.42	2.28	5.04	1.20	2.80	4.20	4.80	0.64	4
2004	0.80	2.24	5.60	4.80	2.40	4.60	6.44	1.55	1
2004	0.10	4.04	16.40	14.00	5.40	10.20	17.02	2.24	2
2005	1.44	2.72	7.28	8.00	4.00	6.20	10.21	0.87	4
2006	0.64	4.02	4.08	4.20	4.60	4.60	5.60	1.84	4
2007	1.44	2.48	7.52	5.20	4.80	4.00	7.01	1.05	4
2008	0.80	4.84	14.92	9.40	6.00	7.40	14.10	1.15	2
2009	0.54	2.45	9.60	7.29	4.28	8.80	11.00	1.75	2
2010	0.42	2.78	10.04	5.44	5.40	9.45	12.45	1.76	3
2011	0.85	4.26	12.92	10.28	6.40	7.45	4.50	1.56	2
2012	1.44	2.49	8.49	9.88	4.87	6.40	2.45	1.94	4
2014	0.47	2.65	5.29	9.27	4.67	7.70	10.45	2.28	1
2014	0.94	4.22	4.95	8.98	4.45	5.45	6.49	2.61	1
2015	0.77	4.45	9.41	6.97	5.48	8.57	9.40	1.98	3

For example, here the 1st cluster unites the years with the highest SO_4^{2-}, Ca^{2+} and mineralization, whilst the members of the 2nd cluster have the lowest content of CO_3^{2-}.

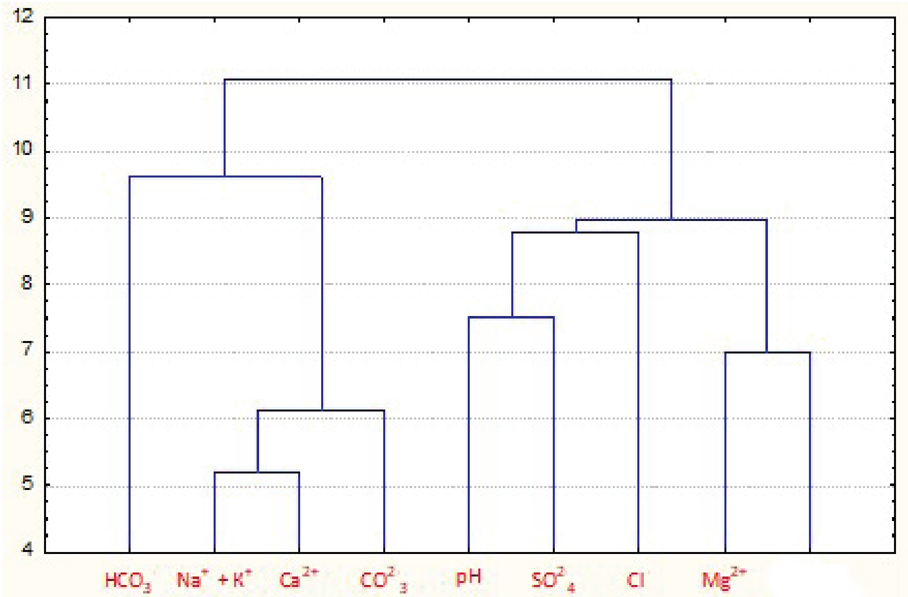


Figure 33. The hierarchic model of the cluster analysis of the Dnipro water for the period of 1973-2015

The 3rd cluster is characterized by the record values of Cl^-, SO_4^{2-}, Ca^{2+}, Mg^{2+}, Na^+ and mineralization. And the 4th cluster has the highest indexes of CO_3^{2-} and Na^+.

The results of the generalized cluster analysis performed within STATISTICA 6.1 software are represented in the Tables 22, 23.

The similar work was performed for the Dnipro water classification. The water samples were collected during the same period of time (1973-2015).

The hierarchic tree allowed concluding that the studied parameters of the Dnipro water could be united in five clusters by the principle:

1 – HCO_3^-;

2 – Na^+, K^+, Ca^{2+}, CO_3^{2-};

3 – pH, SO_4^{2-};

4 – Cl^-;

5 – mineralization and Mg^{2+}.

Such a distribution tells about the highest level of interrelation, close connection between the indexes of HCO_3^-, Mg^{2+} and mineralization (Fig. 33).

By the results of k-means classification and ANOVA we conclude that the most significant indexes in distribution of the water by the clusters are HCO_3^- and Mg^{2+} (Table 24).

Table 24.
The results of ANOVA related to the cluster analysis of the Dnipro water by k-means method

Studied parameter	Inter	cc	Intro	cc	F	Significance
CO_3^{2-}	22.76751	4	19.24249	48	11.24615	0.000004
HCO_3^-	18.72754	4	24.27247	48	7.64472	0.000127
Cl^-	29.72869	4	12.27141	48	24.01485	0.000000
SO_4^{2-}	5.11775	4	46.88224	48	1.41821	0.280747
Ca^{2+}	20.69974	4	21.40027	48	9.24216	0.000026
Mg^{2+}	18.99244	4	24.00766	48	7.84205	0.000104
Na^+	29.48208	4	12.51792	48	22.47441	0.000000
Mineralization	22.87284	4	19.12717	48	11.46048	0.000004

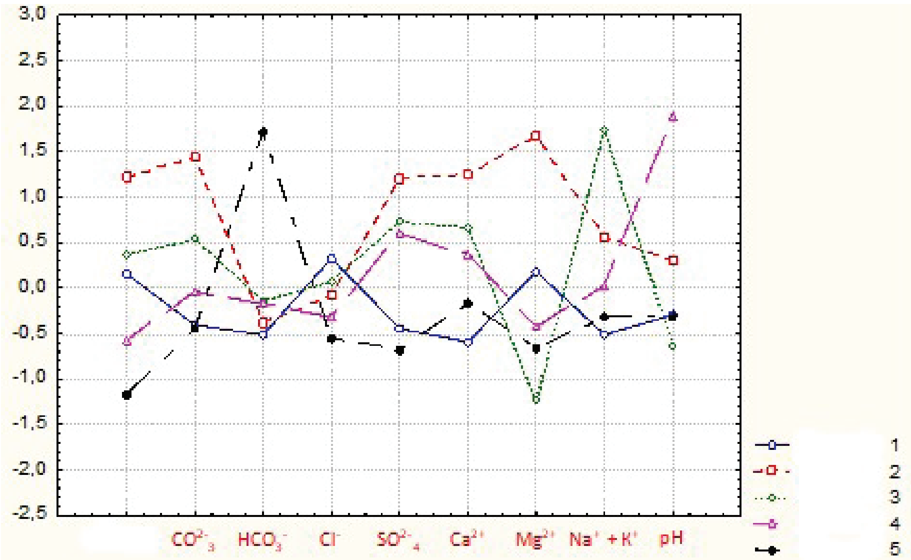


Figure 34. The graph of the means of the ions content of the Dnipro water for the period of 1973-2015

Analysis of the clusterization graph allowed the observation of the following tendencies: 1st class is characterized by the years with the highest Cl^-; the years of the 2nd cluster have the maximum levels of CO_3^{2-}, Mg^{2+}, SO_4^{2-}, Ca^{2+} and mineralization; the years of the 3rd cluster unites the components with the lowest indexes of Mg^{2+} and pH while the values of Na^+, K^+ are the highest; the 4th cluster embraces the waters with the maximum pH value; the 5th cluster is characterized by the lowest indexes of mineralization, CO_3^{2-}, Cl^-, SO_4^{2-} and the highest HCO_3^- values (Fig. 34).

Table 25.
The indexes of ions content and mineralization of the Dnipro water for the samples collected on 15th of May (1973-2015) with the results of classification by clusters

Years	Miner-alization	Anions, meq L-1				Cations, meq L-1			pH	Cluster
		CO^{2}_{-3}	HCO_4^-	Cl^-	SO^{2-}_4	Ca^{2+}	Mg^{2+}	$Na^+ + K^+$		
1973	0.29	0.21	2.12	0.72	1.15	2.64	1.45	1.41	7.48	1
1973	0.42	0.26	2.15	0.49	1.42	2.12	1.77	1.06	8.12	5
1975	0.45	0.04	4.05	0.84	1.24	2.41	0.72	1.40	7.26	1
1976	0.48	0.28	2.84	1.15	1.81	1.85	1.24	1.46	7.49	5
1977	0.42	0.18	4.11	1.15	1.84	1.91	1.46	1.24	7.97	5
1978	0.29	0.11	289	0.97	1.11	2.45	1.48	1.45	8.09	5
1979	0.42	0.1	2.99	0.99	1.01	2.61	0.99	1.44	8.12	1
1980	0.45	0.14	2.6	0.94	1.4	2.15	1.41	1.52	7.99	2
1981	0.48	0.1	2.81	1.6	1.22	2.7	1.92	1.12	7.49	2
1982	0.42	0.12	2.94	1.95	1	2.17	1.89	2.44	7.97	1
1984	0.49	0.1	4.55	1.02	1.95	2.45	1.77	2.5	8.1	1
1984	0.47	0.14	2.71	1.45	1.5	2.01	2.61	1.08	8.02	4
1985	0.46	0.14	2.45	1.25	2.46	2.42	1.88	1.86	8.08	2
1987	0.49	0.22	2.51	0.86	1.28	2.4	1.22	1.44	8.45	2
1988	0.44	0.07	2.64	0.85	1	2.44	1.04	1.18	8.24	5
1989	0.47	0.14	2.95	0.91	1.47	2.46	1.41	1.6	8.49	2
1990	0.49	0.29	2.82	0.95	1.74	2.26	1.76	1.78	8.68	1
1991	0.41	0.06	2.91	0.99	1.21	2.45	1.44	1.29	8.48	5
1992	0.45	0.08	4.06	0.96	0.94	2.68	0.92	1.44	8.4	2
1994	0.44	0.06	2.8	1.12	1.2	2.4	0.2	1.02	8.4	1
1994	0.49	0.11	2.92	1.18	1.75	1.75	4.1	1.01	7.51	2
1995	0.41	0.75	2.8	1.68	2.6	1.4	4.2	1.48	7.4	4
1996	0.47	0.12	2.6	1.44	2.5	2.9	4.1	1.4	8.45	4
1997	0.47	0.04	4.1	0.95	0.95	2.67	0.94	1.44	8.45	4
1998	0.45	0.07	2.15	1.2	1.45	1.7	1.08	1.99	8.21	1
1999	0.47	0.07	2.75	0.95	1.2	2.4	1.2	1.47	8.49	2
2000	0.74	0.05	4.15	0.9	1.41	2.49	1.44	1.68	8.19	2
2001	0.51	0.07	2.7	0.84	0.99	2.4	1.05	1.24	8.44	1
2002	0.48	0.1	4.1	0.94	1.74	2.44	1.76	1.77	8.1	2
2004	0.44	0.11	2.92	0.9	1.15	2.41	1.44	1.24	8.28	1
2004	0.45	0.06	4.1	0.94	1.76	2.41	1.77	1.78	8.49	2
2005	0.40	0.1	4.14	0.9	0.9	2.67	0.94	1.44	8.44	1
2006	0.48	0.14	2.61	1.01	1.21	2.15	1.45	1.46	8.21	1
2007	0.42	0.08	2.7	1.45	1.54	2.1	2.52	1.08	8.09	2
2008	0.46	0.11	2.81	1.04	1.56	2.4	1.75	1.45	8.27	4
2009	0.42	0.09	2.91	0.8	1.62	1.8	1.28	1.66	8.47	2
2010	0.48	0.08	4.08	0.94	0.94	1.41	1.77	1.47	8.48	4
2011	0.45	0.11	2.51	1.41	1.14	2.49	0.88	1.05	8.17	2
2012	0.48	0.15	2.64	1.25	1.47	2.26	1.47	1.55	7.94	4
2014	0.49	0.14	2.78	1.07	1.27	2.25	1.16	1.82	7.29	4
2014	0.59	0.05	4.17	1.2	1.45	2.48	1.46	1.46	8.15	4
2015	0.46	0.09	2.44	0.92	1.06	2.21	2.24	1.48	8.44	4

The dynamics of mineralization and ions content in the Dnipro water is represented in the form of the Table 25.

The results of the study are useful for deep analysis and monitoring of the water quality dynamics in two sources of water for irrigation and show possible way of the neural networks approach to modeling of environment.

3.8. Development of the Artificial Neural Network for Oil-seed Flax Production Processes Management

Usually, while modeling of agroecosystems' productivity parameters of such classes is highlighted: soil; hydrothermal conditions; data about the studied crop. However, the results of numerous experiments revealed the following drawbacks of such models: structure of the network and its training has to be carried out for each crop; the task has a high computation power requirements because of a great number of the network parameters; the difficulty of solving the reversed task; adding new parameters into the model requires its retraining. To eliminate the above-mentioned drawbacks, it is rational to implement a modular structure of a neural network.

The modularity of a neural network allows to make a hierarchic decomposition of a complicated task into a number of the modest sub-tasks, and the corresponding structure of a network might be optimized for the concrete task. To outline many parameters of agroecosystem it is necessary to highlight three input elements (modules): soil fertility, weather conditions, agrotechnological operations.

The formed modules have the following general characteristics:

– module «Soil fertility» allows reflecting the qualitative and quantitative indexes of fertility. Relying on the modern opinions in this question, the input parameters of the module should include the content of humus, nutritive elements; physical and mechanical soil properties, its biological activity, etc. The output of the module should be characterized as the index of potential soil fertility. This module includes mineral nutrition backgrounds with accordance with mineral fertilizers application norm and humidification conditions;

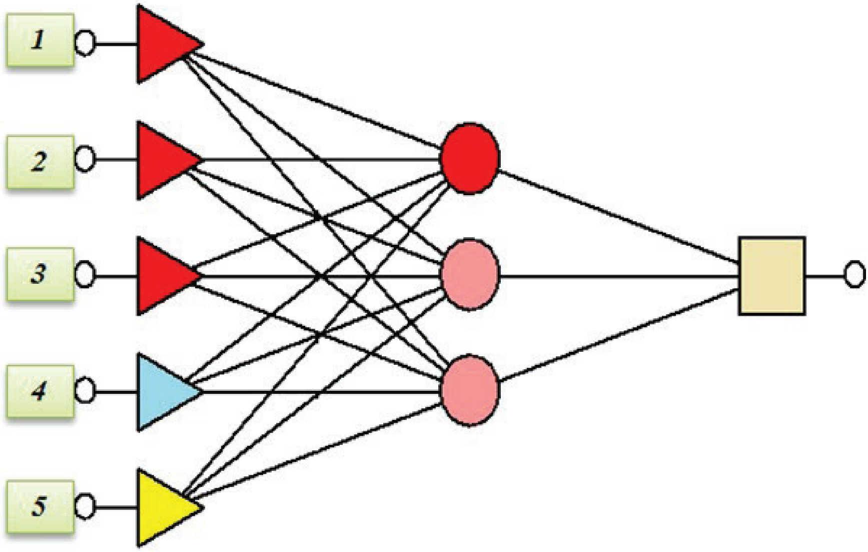
– module «Weather conditions» embraces main meteorological parameters that directly affect plants' productivity: solar insolation, precipitation amounts, indexes of temperature regime, relative air humidity, calculative bioclimatic indexes. The module includes indexes of the sum of effective air temperatures and precipitation amounts during the period of crop vegetation;

– module «Agrotechnological operations» consists of numerous elements that are foreseen by the crop cultivation technology and are directed to the improvement of growth conditions, regulation of harmful organisms quantity.

Implementation of this scheme to the oil-seed flax productivity model allows using the modules «Soil fertility» and «Weather conditions» for further modeling of the other crops' productivity within the crop rotation without any complications and reconstruction of the whole network.

The module «Agrotechnological operations» is developed exclusively for oil-seed flax taking into account peculiarities of its growth and development, in particular, under the cultivation in irrigated conditions.

If necessary, the model might be supplemented with the additional modules without complete reconstruction.



Notes: ***1*** – mineral nutrition background (kg ha^{-1}); ***2*** – inter-row spacing (cm); ***3*** – sowing norm (million ha^{-1}); ***4*** – precipitation amounts during the vegetative period (mm); ***5*** – sum of the effective air temperatures (^{o}C).

Figure. 35. Architecture of the neural network for oil-seed flax yield under the conditions of natural humidification in dependence on the effect of agrotechnological and natural factors (architecture 5:5-3-1:1; training productivity – 0.1540, control productivity – 45.5418, testing productivity – 22.3901)

Using the program complex STATISTICA we performed modeling of oil-seed flax productivity for the rain-fed and irrigated lands. It was proved that the studied factors (mineral nutrition background, inter-row spacing and sowing norm) have an important value from the point of view of oil-seed flax productivity formation, therefore in the architecture of the modeled neural network they are colored in red (Fig. 35).

It should be mentioned that the effect of non-regulated natural factors (precipitation amounts and sum of the effective air temperatures) included in the neural network also had an influence, moreover, the increase of precipitation amounts had an additional effect, and the increase of the sum of effective air temperatures had the contrary effect.

The triad of factors' interaction reflects the maximum efficiency of the system of mineral fertilizers application, inter-row spacing and sowing norm, which provide obtaining of high seed yield of the studied crop. The interaction of the second and third elements of the triad has pink color that testified about the decrease of neural connection among the input factors of influence. We also should note high indexes of the control and testing productivity of the ANN – 45.5 and 22.4, respectively.

By the results of training of the created ANN of oil-seed flax productivity under the cultivation on the non-irrigated lands of the South of Ukraine, it was determined that

Table 26.
Constructive elements of the neural network of agroecological productivity of oil-seed flax in the non-irrigated conditions of the South of Ukraine in dependence on natural and agrotechnological factors

No.	Architecture	Training productivity	Control productivity	Testing productivity	Training error	Control error	Testing error
1.	Linear 5:5-1:1	0.0012	1010.9	880.1	0.0002	1.2554	1.936
2.	MP 3:3-4-1:1	0.0311	8.327	5.158	0.0083	0.0095	0.0115
3.	MP 2:2-4-1:1	0.2834	7.966	5.222	0.0075	0.0090	0.0118
4.	RBF 5:5-3-1:1	0.1540	45.542	22.390	0.0054	0.0089	0.0077
5.	RBF 5:5-1-1:1	0.2873	48.018	42.31	0.0099	0.0071	0.0141

Notes: * – gradation of the factors in the Fig. 35; MP – multi-layer perceptron; RBF – radial basis function.

the architecture of the neurons in the first element (mineral fertilizers background) was linear – 5:5-1:1 Table 26).

In the second (inter-row spacing) and in the third (sowing norm) elements of the network there were fixed multi-layer perceptrons – MP 3:3-4-1:1 and MP 2:2-4-1:1, respectively. In the fourth (precipitation amounts) and fifth (sum of the effective air temperatures) input elements of the ANN there were fixed radial basis functions – RBF 5:5-3-1:1 and RBF 5:5-1-1:1, respectively.

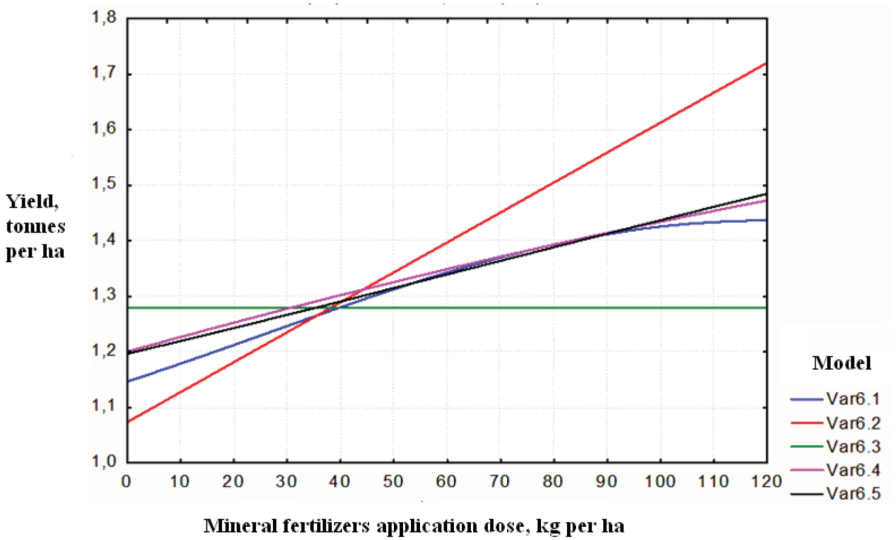
The minimum training productivity of 0.0012 was established in the first element of the model (mineral fertilization background) at the linear structure of the ANN architecture but in this variant control and testing productivity had the highest values – 1011 and 880, respectively, that is more than of the other elements of the model 21.1-170.6 times.

The errors of the developed ANN – training, control, testing ones fluctuated within 0.0002-0.0115, excluding control and testing errors of the first element of the model where it raised up to 1.16-1.94.

Interesting results were obtained by the results of the formation of graphs of response between the modeled indexes of the mineral nutrition background and oil-seed flax yield in interaction with the input indexes of the ANN (Fig. 36).

It was determined that the maximum positive modeled interaction was between the mineral nutrition background and inter-row spacing (Var6.2) when the level of theoretical oil-seed flax yield over scores 1.7 t ha^{-1}.

The non-regulated elements of the neural network – precipitation amounts during the crop vegetation (Var6.4) and sum of the effective air temperatures (Var6.5) provided stable interaction with the increase of mineral nutrition background with the positive dynamic of theoretical oil-seed flax yield from the level of 1.20-1.21 t ha^{-1} at zero fertilizers applied to 1.46-1.48 t ha^{-1} at nitrogen application dose of 120 kg ha^{-1}.



Notes: Var6.1 – mineral nutrition background (kg ha^{-1}); **Var6.2** – inter-row spacing (cm); **Var6.3** – sowing norm (million ha^{-1}); **Var6.4** – precipitation amounts during the crop vegetation; **Var6.5** – sum of the effective air temperatures (^{o}C).

Figure 36. Graph of the response between the modeled indexes of mineral fertilizers application norms (kg ha^{-1}) and oil-seed flax yield (t ha^{-1}) at the conditions of natural humidification in interaction with the input indexes of the neural network

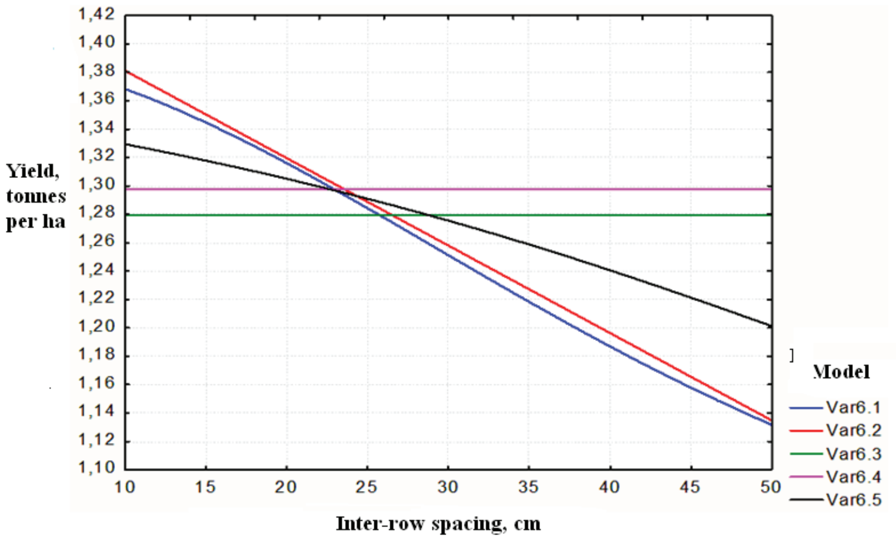
In contrary, the third element of the model – sowing norm (Var6.3) did not interact with the mineral nutrition background with the stable dynamic of theoretical yield at the level of 1.28 t ha-1. This is explained by the proportional redistribution of the nutritive area at the different distance between the plants in the row and the increased need in mineral nutrition at the higher sowing norm.

Quite different tendencies of inter-relation were observed in the graph of the response of the second element of the ANN Var6.2 – «inter-row spacing» (Fig. 37).

A complete absence of interconnection and stability of the indexes of the theoretical seed yield of the crop at the level of 1.28-1.30 t ha^{-1} were fixed in the Var6.3 sowing norm) and Var6.4 precipitation amounts during the vegetation) elements of the developed ANN.

The sum of the effective air temperatures (Var6.5) reflects slow negative dynamic with the decrease of theoretical seed yield from 1.33 t ha^{-1} at the modeled inter-row spacing of 10 cm to 1.20 t ha^{-1} – at the widening of the inter-row spacing to 50 cm.

Significant decrease of the calculative oil-seed flax yield at the widening of the inter-row spacing from 10 to 50 cm was determined in the first element of the model – mineral nutrition background (Var6.1) when this index decreased from 1.36 to 1.13 t ha^{-1} or by 20.4%.



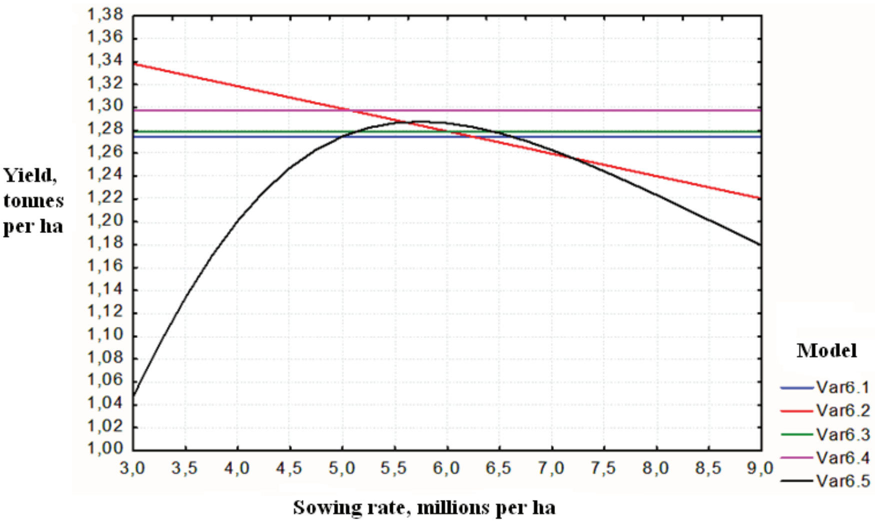
Notes: Var6.1 – mineral nutrition background (kg ha^{-1}); **Var6.2** – inter-row spacing (cm); **Var6.3** – sowing norm (million ha^{-1}); **Var6.4** – precipitation amounts during the crop vegetation; **Var6.5** – sum of the effective air temperatures (^{o}C).

Figure 37. Graph of the response between the modeled indexes of inter-row spacing (cm) and oil-seed flax yield (t ha^{-1}) at the conditions of natural humidification in interaction with the input indexes of the neural network

The high amplitude of fluctuation of the theoretical values of oil-seed flax yield was provided by the comparison of interaction between Var6.5 sum of the effective air temperatures) and Var6.3 sowing norm). In the range of sowing norms from 3.0 to 5.2 million per ha the positive interaction and steep yield increase from 1.05 to 1.27 t ha^{-1} were determined (Fig. 38). At the further increase of the calculative sowing norm from 5.0 to 6.5 million per ha there was fixed stabilization of seed productivity of the plants (Plato stage), and an increase of sowing norm from 6.5 to 9.0 million per ha negative interaction manifested itself with the decrease of the yield from 1.27 to 1.18 t ha^{-1}.

The slow decrease of the theoretical yield from 1.34 to 1.22 t ha^{-1} in regard to the interaction with the inter-row spacing demonstrates the second element of the ANN – inter-row spacing (Var6.2). The other studied elements of the model – mineral nutrition background (Var6.1) and precipitation amounts during the vegetative period (Var6.4) had zero interaction with the sowing norm that is explained by the proportional increase of the expenses of nutrients and moisture at the increase of sowing norm of oil-seed flax.

On the non-irrigated plots precipitation amounts (Var6.4) were in the maximum interaction with mineral nutrition background (Var6.1) regarding the formation of



Notes: Var6.1 – mineral nutrition background (kg ha^{-1}); **Var6.2** – inter-row spacing (cm); **Var6.3** – sowing norm (million ha^{-1}); **Var6.4** – precipitation amounts during the crop vegetation; **Var6.5** – sum of the effective air temperatures (^{o}C).

Figure 38. Graph of the response between the modeled indexes of sowing norm (millions per ha) and oil-seed flax yield (t ha^{-1}) at the conditions of natural humidification in interaction with the input indexes of the neural network

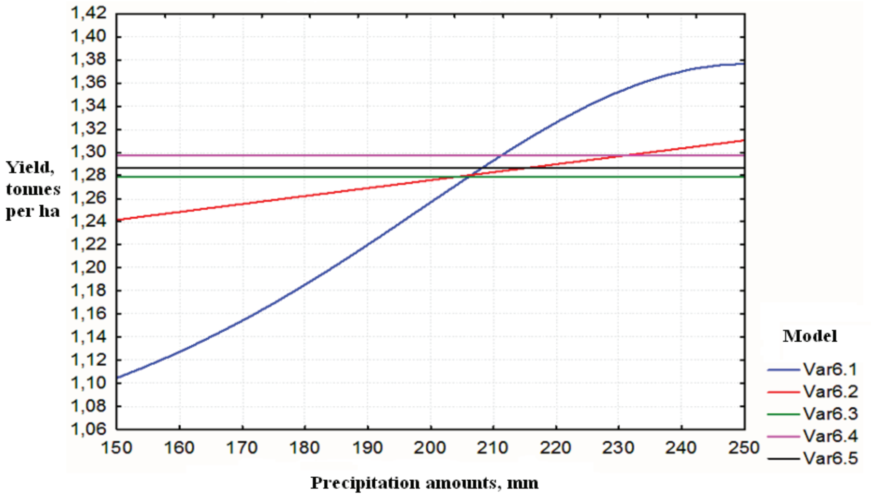
seed productivity (Fig. 39). So, at the amount of 150 mm in the interaction with Var6.1 the theoretical seed yield equaled 1.11 t ha^{-1}, and at the increase of precipitation to 250 mm we observed the raise of the yield to 1.38 t ha^{-1} or by 24.3%.

An insignificant positive interaction was fixed in regard to the second element of the neural network – inter-row spacing (Var6.2) when the raise of the theoretical seed yield of the crop from 1.24 t ha^{-1} by 5.6% was observed.

The other elements of the developed model were in a stable state with the seed yield of oil-seed flax within 1.29-1.30 t ha^{-1} and did not change under the influence of precipitation amounts.

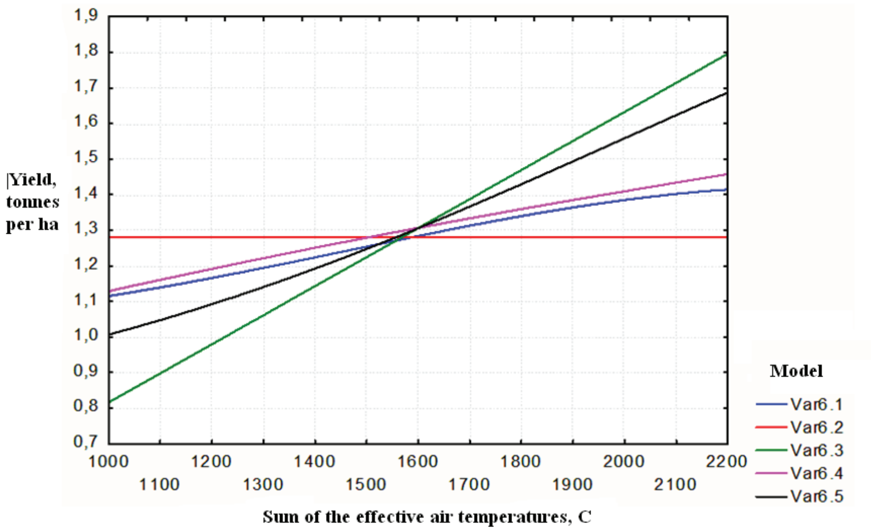
The final element of the developed ANN was the sum of the effective air temperatures (Var6.5), which was characterized by the positive or zero influence on the theoretical yield of oil-seed flax (Fig. 40).

The maximum positive interaction concerning the increase of the seed yield from 0.82 t ha^{-1} 2.2 times in the graph of response was manifested between the sum of the effective temperatures and sowing norm (Var6.3). Mineral nutrition background (Var6.1) and precipitation amounts during the vegetative period (Var6.4) also led to the increase of the theoretical seed yield from 1.12-1.14 to 1.42-1.46 t ha^{-1}, however, it was significantly lower than at the Var6.3 and averaged to 24.6-30.4%.



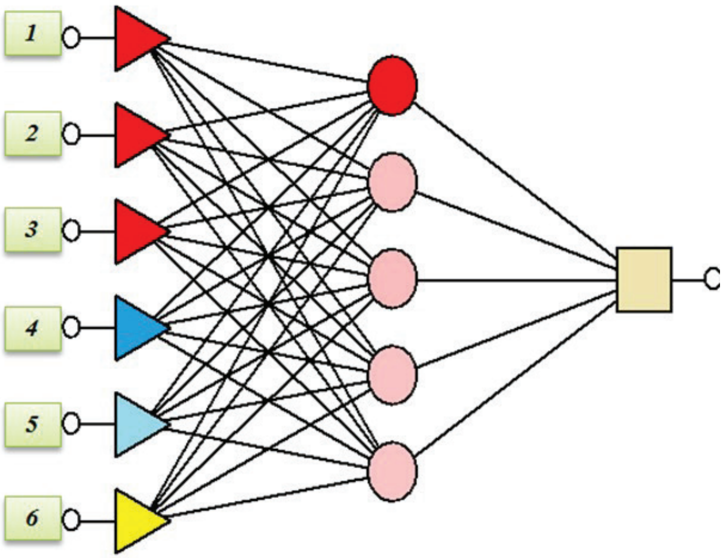
Notes: Var6.1 – mineral nutrition background (kg ha^{-1}); **Var6.2** – inter-row spacing (cm); **Var6.3** – sowing norm (million ha^{-1}); **Var6.4** – precipitation amounts during the crop vegetation; **Var6.5** – sum of the effective air temperatures (^{o}C).

Figure 39. Graph of the response between the modeled indexes of precipitation amounts (mm) and oil-seed flax yield (t ha^{-1}) at the conditions of natural humidification in interaction with the input indexes of the neural network



Notes: Var6.1 – mineral nutrition background (kg ha^{-1}); **Var6.2** – inter-row spacing (cm); **Var6.3** – sowing norm (million ha^{-1}); **Var6.4** – precipitation amounts during the crop vegetation; **Var6.5** – sum of the effective air temperatures (^{o}C).

Figure 40. Graph of the response between the modeled indexes of the sum of the effective air temperatures (^{o}C) and oil-seed flax yield (t ha^{-1}) at the conditions of natural humidification in interaction with the input indexes of the neural network



Notes: Var7.1 – mineral nutrition background (kg ha^{-1}); **Var7.2** – inter-row spacing (cm); **Var7.3** – sowing norm (million ha^{-1}); **Var7.4** – irrigation norm (m^3 ha^{-1}); **Var7.5** – precipitation amounts during the crop vegetation (mm); **Var7.6** – sum of the effective air temperatures (^{o}C).

Figure 41. Architecture of the neural network of oil-seed flax yield in the irrigated conditions in dependence on the influence of natural and agrotechnological factors (architecture: RBF 6:6-5-1:1; training productivity – 0.0358, control productivity – 5.9767, testing productivity – 8.1346)

The inter-row spacing (Var6.2) did not have any interaction with the sum of the effective temperatures, and the level of the theoretical oil-seed flax yield was 1.29 t ha^{-1}.

At the irrigated conditions we created an ANN with the aggregate of the same elements as at the non-irrigated conditions with the addition of the amount of irrigation water applied as a valuable factor of influence on the productivity of oil-seed flax (Fig. 41).

Interaction of the constituent elements of the neural network consists of five blocks, and the first one demonstrates a high level of interaction (red color) and reflects the cumulative effect of Var7.1 mineral nutrition background) and Var7.2 inter-row spacing) elements of the network. The other four blocks of interaction have pink color that testifies about the medium level of interaction of all the elements of the ANN.

The general architecture of the ANN was formed by the radial basis function (RBF 6:6-5-1:1) with less indexes of training, control and testing productivity.

The architecture of the ANN changed in the separate elements as follows: Var7.1 mineral nutrition background) and Var7.2 inter-row spacing) were formed of the radial basis functions – RBF 4:4-1-1:1 and RBF 4:4-3-1:1; Var7.3 sowing norm) and Var6.6 sum of the effective air temperatures) – linear functions – 6:6-1:1 and 2:2-4-1:1;

Table 27.
Constituent elements of the neural network of the agroecological model of oil-seed flax productivity in dependence on natural and agrotechnological factors at irrigation

No.	Architecture	Training productivity	Control productivity	Testing productivity	Training error	Control error	Testing error
1.	RBF 4:4-1-1:1	0,9710	1,003	1,430	4,030	759,5	4,771
2.	RBF 4:4-3-1:1	0,2651	0,9963	0,2334	1,393	753,8	1,529
3.	Linear 6:6-1:1	0,0001	1,111	0,0002	0,0001	57,71	0,0003
4.	MP 2:2-1-1:1	0,6931	1,005	0,7432	0,2003	52,26	0,2179
5.	MP 5:5-6-1:1	0,1771	0,9936	0,1902	0,1336	51,59	0,1461
6.	Linear 2:2-4-1:1	0,1080	0,9906	0,2966	0,0308	51,48	0,0603

Notes: * – gradation of the factors as in the Fig. 41; MP – multi-layer perceptron; RBF – radial basis function.

Var7.4 irrigation norm) and Var7.5 precipitation amounts during the vegetative period) – multi-layer perceptrons – 2:2-1-1:1 and 5:5-6-1:1 Table 27).

Training productivity of the highest level – 0.69-0.97 was in the fourth and first elements, and the least one – 0.0001 in the third element (sowing norm of oil-seed flax).

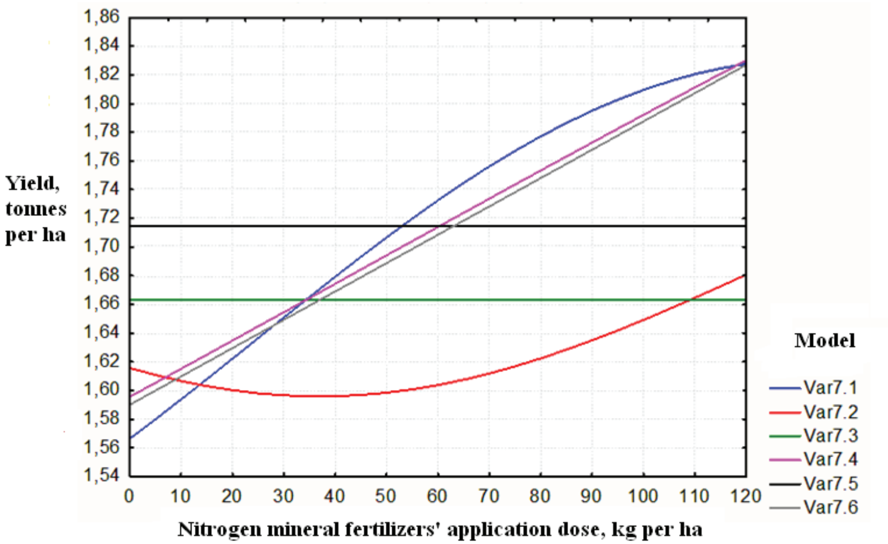
Among the errors of the ANN the highest value had control errors in the first and second elements, which were 760 and 754, respectively, and the lowest level of these indexes within 0.0001-0.0003 was in the third element (sowing norm) on the training and testing errors.

The graphs of the created ANN testify about the difference in interactions between the separate elements of the network both in directions and in quantitative terms. The maximum level of the theoretical yield of oil-seed flax at the level of 1.83 t ha^{-1} is formed under the interaction of Var7.1 mineral nutrition background) with Var7.4 irrigation norm) and Var7.6 sum of the effective air temperatures) elements (Fig. 42).

We should mention that mutual amplification of the influence of Var7.1, Var7.4 and Var7.6 elements of the ANN might be put on the steep increase of the fertilizers' efficiency at the improvement of the water regime of plants on the background of the increased temperatures.

Sowing norm (Var7.2) in close connection with mineral nutrition (Var7.1) firstly had negative direction in the range of yields from 1.62 to 1.59 t ha^{-1} at application of mineral fertilizers (by nitrogen in the range from N_0 to N_{30}), further after coming of the fertilization norm over N_{58} kg ha^{-1} Plato stage was observed, and at the increase of the calculative doses of nitrogen to N_{120} kg ha^{-1} oil-seed flax yield raised to 1.66 t ha^{-1}.

Precipitation (Var7.5) at the absence of interaction with the mineral nutrition background (Var7.1) provides the formation of theoretical yield at the level of 1,72 t ha^{-1}.



Notes: Var7.1 – mineral nutrition background (kg ha^{-1}); **Var7.2** – inter-row spacing (cm); **Var7.3** – sowing norm (million ha^{-1}); **Var7.4** – irrigation norm (m^3 ha^{-1}); **Var7.5** – precipitation amounts during the crop vegetation (mm); **Var7.6** – sum of the effective air temperatures (^{o}C).

Figure 42. Graph of the response between the modeled indexes of mineral fertilizers application norms (kg ha^{-1}) and oil-seed flax yield (t ha^{-1}) at the irrigated conditions in interaction with the input indexes of the neural network

Absence of interaction was also fixed in regard to the differentiation of sowing norms (Var7.2) while the level of calculative yield decreased comparing to the fifth element to 1.66 t ha^{-1}.

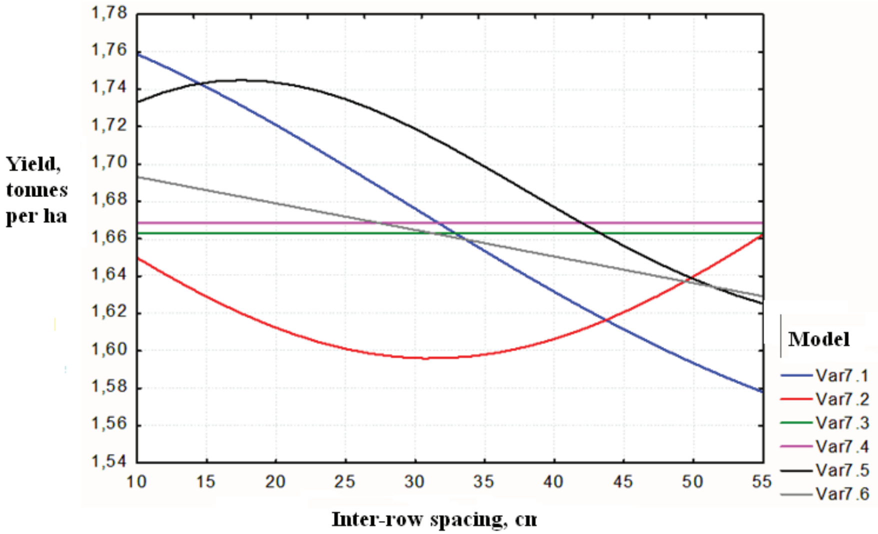
Absolutely different directions and indexes of interaction were observed at the modeling of the graphs of response for Var7.2 – inter-row spacing (Fig. 43).

It was determined that in regard to the mineral nutrition background (Var7.1) decrease of theoretical level of oil-seed flax in the range from 1.76 t ha^{-1} at the inter-row spacing of 10 cm to 1.58 t ha^{-1} at the widening of inter-row spacing to 55 cm.

Sowing norm (Var7.3) and irrigation norm (Var7.4) did not interact with inter-row spacing (Var7.2), herewith the level of theoretical oil-seed flax yield was 1.66-1.67 t ha^{-1}.

Precipitation amounts (Var7.5) interacted in a wavelike way with inter-row spacing (Var7.2) that is explained by the decrease of their efficiency at using artificial humidification. So, in the range of 10-25 cm the level of yield was 1.73-1.74 t ha^{-1}, and at the widening of inter-row spacing to 55 cm a seed productivity decrease was observed by 6.8-7.4%.

Slight interaction was observed in regard to the sowing norm of oil-seed flax (Var7.3) excluding Var7.5 precipitation amounts) that might be explained by the impact of additional humidification and increased productive processes of the plants under the proportional increase of their quantity per the unit of the sown area



Notes: Var7.1 – mineral nutrition background (kg ha^{-1}); **Var7.2** – inter-row spacing (cm); **Var7.3** – sowing norm (million ha^{-1}); **Var7.4** – irrigation norm (m^3 ha^{-1}); **Var7.5** – precipitation amounts during the crop vegetation (mm); **Var7.6** – sum of the effective air temperatures (^{o}C).

Figure 43. Graph of the response between the modeled indexes of inter-row spacing (cm) and oil-seed flax yield (t ha^{-1}) at the irrigated conditions in interaction with the input indexes of the neural network

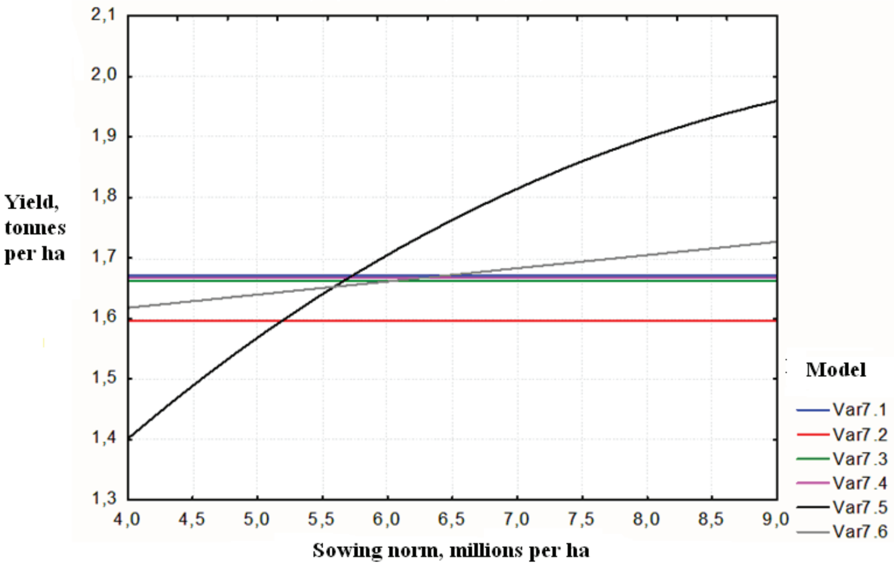
(Fig. 43). It was determined that at the sowing norm of 4-5 million per ha the level of theoretical yield was 1.41-1.55 t ha^{-1}, and at further thickening of the plants from 5.5 to 9.0 million per ha the calculative seed productivity increased by 5.2-39.0%.

Changes in the mineral nutrition background (Var7.1), inter-row spacing (Var7.2) and irrigation norm (Var7.5) were characterized by the absence of mathematical interaction with the sowing norm (Var7.3) of the studied crop. Herewith the level of theoretical oil-seed flax yield was 1.66, 1.60 and 1.67 t ha^{-1}, respectively.

Heat supply of oil-seed flax, which reflects the sum of the effective temperatures (Var7.6), had slight positive interaction with the indexes of sowing norm (Var7.3). So, at the sowing norm in the range of 4.0-4.5 million per ha the level of theoretical yield is in the range of 1.40-1.49 t ha^{-1}, and at the density of 8.5-9.0 million per ha it increases to 1.92-1.97 t ha^{-1}.

The highest level of positive interaction with Var7.4 irrigation norm) had the element Var7.3 sowing norm) (Fig. 45). At the calculative irrigation norm of 890 m^{3}ha^{-1} the level of theoretical seed yield is 1.2 t ha^{-1}, and at the increase to 990 m^3 ha^{-1} we defined it raise up to 2.5 t ha^{-1} or almost twice.

In the second place there was the element of the ANN Var7.5 precipitation amounts during the vegetative period). At the highest and lowest values of irrigation norm



Notes: Var7.1 – mineral nutrition background (kg ha^{-1}); **Var7.2** – inter-row spacing (cm); **Var7.3** – sowing norm (million ha^{-1}); **Var7.4** – irrigation norm (m^3 ha^{-1}); **Var7.5** – precipitation amounts during the crop vegetation (mm); **Var7.6** – sum of the effective air temperatures (^{o}C).
Figure 44. Graph of the response between the modeled indexes of sowing norm (millions per ha) and oil-seed flax yield (t ha^{-1}) at the irrigated conditions in interaction with the input indexes of the neural network

(Var7.4) in the interaction with precipitation (Var7.5) the increase of theoretical seed yield from 1.48 t ha^{-1} by 48% was observed.

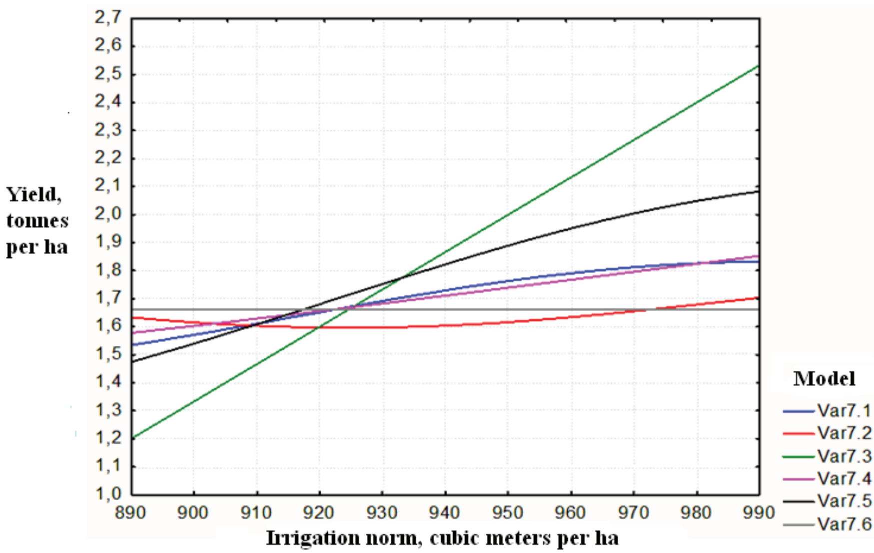
The improvement of mineral nutrition (Var7.1) in the interaction with irrigation (Var7.4) provided the increase of theoretical yield in average from 1.48 to 1.82 t ha^{-1}.

Inter- row spacing (Var7.2) was characterized with a weak curvilinear negative interaction because at the irrigation norm of 920-940 m^3 ha^{-1} a slight decrease of the theoretical seed yield to 1.6 t ha^{-1} or by 1.9-5.2%) was fixed.

The sixth element of the developed ANN (Var7.6) had zero interaction with irrigation norm (Var7.4), however, it provided stable level of oil-seed flax yield of 1.68 t ha^{-1}.

Precipitation amounts (Var7.5) manifested with multi-direction connections with the other elements of the ANN (Fig. 46).

In relation to the first element of the model (Var7.1) direct positive effect on the plants' productivity was established. At the calculative precipitation amount of 165 mm mineral fertilizers' application (Var7.1) provided theoretical seed yield at the level of 1.54 t ha^{-1}, and at the increase of precipitation (Var7.5) to 235 mm – seed productivity increased to 1.79 t ha^{-1}.



Notes: Var7.1 – mineral nutrition background (kg ha^{-1}); **Var7.2** – inter-row spacing (cm); **Var7.3** – sowing norm (million ha^{-1}); **Var7.4** – irrigation norm (m^3 ha^{-1}); **Var7.5** – precipitation amounts during the crop vegetation (mm); **Var7.6** – sum of the effective air temperatures (°C).

Figure 45. Graph of the response between the modeled indexes of irrigation norm (m^3 ha^{-1}) and oil-seed flax yield (t ha^{-1}) at the irrigated conditions in interaction with the input indexes of the neural network

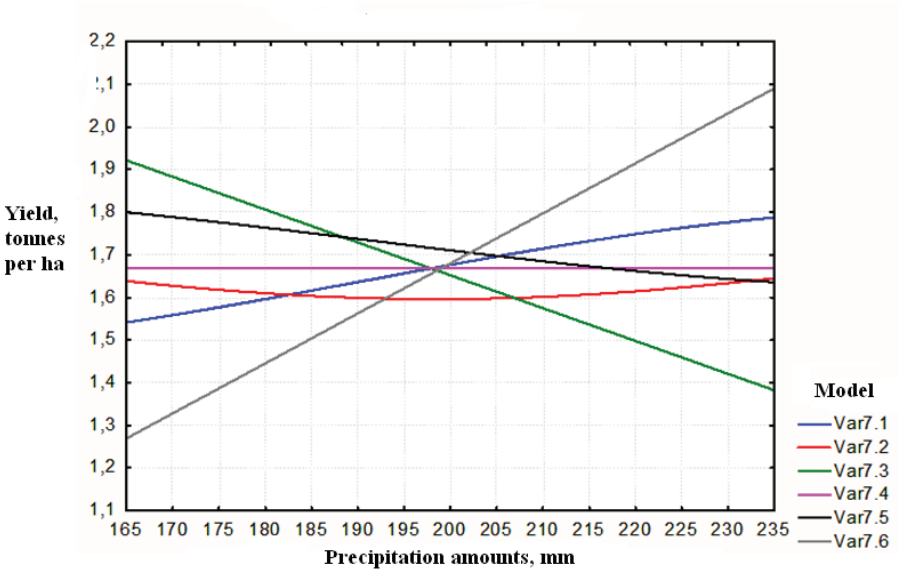
Inter-row spacing (Var7.2) showed poor negative interaction with the decrease of theoretical seed yield of the studied crop to 1.60-1.62 t ha^{-1} at the modeled precipitation income (Var7.5) within the range of 185-215 mm.

Sowing norm (Var7.3) was in significant direct negative interaction with precipitation amounts (Var7.6). At the minimum precipitation amount (165 mm) the calculative seed yield was 1.92 t ha^{-1}, and at the increase of natural humidification up to 235 mm the decrease of the yield by 38.1% to 1.39 t ha^{-1} was determined.

Irrigation norm (Var7.4) did not change theoretical yield of oil-seed flax, which was at the level of 1.67 t ha^{-1}, relatively to the fluctuations of precipitation amounts (Var7.5) within165-235 mm.

The maximum direct positive interaction in regard to the changes of precipitation (Var7.5) in the network of oil-seed flax yield was defined at the comparison of the sum of the effective air temperatures (Var7.6). At the minimum precipitation (165 mm) the level of theoretical seed yield was 1.28 t ha^{-1}, at the increase of precipitation to 200 mm – modeled productivity averaged to 1.67 t ha^{-1}, and at 235 mm – to 2.19 t ha^{-1}, respectively.

The last element of the created ANN of the oil-seed flax yield – sum of the effective air temperatures (Var7.6) – was characterized by the positive interaction with



Notes: Var7.1 – mineral nutrition background (kg ha^{-1}); **Var7.2** – inter-row spacing (cm); **Var7.3** – sowing norm (million ha^{-1}); **Var7.4** – irrigation norm (m^3 ha^{-1}); **Var7.5** – precipitation amounts during the crop vegetation (mm); **Var7.6** – sum of the effective air temperatures (°C).

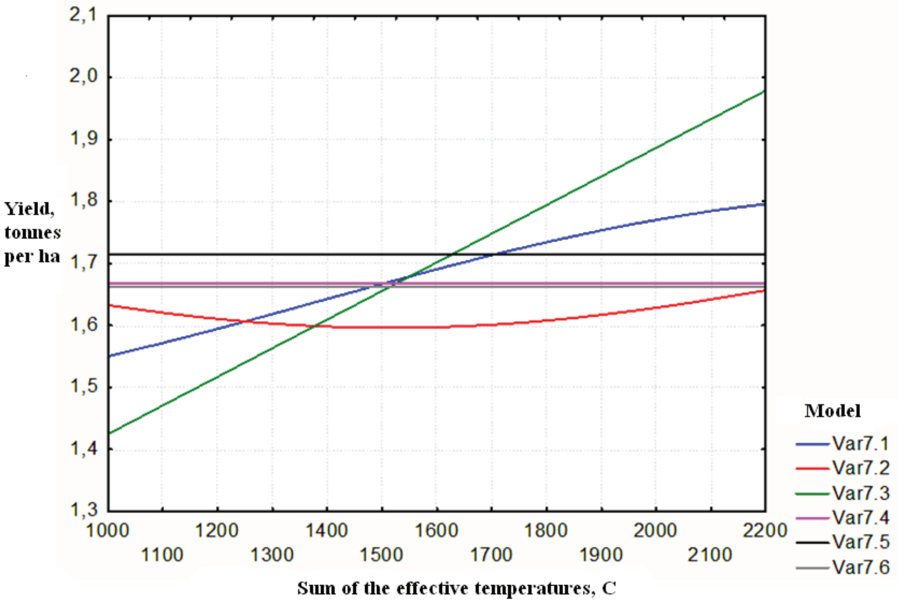
Figure 46. Graph of the response between the modeled indexes of precipitation amounts (mm) and oil-seed flax yield (t ha^{-1}) at the irrigated conditions in interaction with the input indexes of the neural network

Var7.3 and Var7.1 element, and also by the absence of interaction with the other elements of the network (Fig. 47).

Sowing norm had the maximum effect on the value of theoretical oil-seed flax yield at the interaction with temperature regime in the period of vegetation. It should be mentioned that at the modeled sum of the effective temperatures of 1000-1100°C the theoretical yield was 1.43-1.47 t ha^{-1}. At the increase of temperatures up to 2000-2200°C the theoretical yield of the crop increased by 27.9-37.8%.

Mineral fertilizers' application (Var7.1) also had additional influence on the increase of seed productivity at the increase of the sum of effective air temperatures (Var7.6). At the minimum modeled level of heat supply of the crops (1000°C) the seed yield was 1.55 t ha^{-1}, and at the maximum value (2200 °C) the productivity increases to 1.80 t ha^{-1} or by 16,1%).

Other elements of the ANN – irrigation norm (Var7.4), precipitation amounts during the vegetative period (Var7.5) did not have interaction with the sum of effective temperatures, and theoretical seed yield was 1.66-1.71 t ha^{-1}, and inter-ow spacing (Var7.2) had insignificant negative interaction with the decrease of the modeled yield by 2.5-3.8%.



Notes: Var7.1 – mineral nutrition background (kg ha^{-1}); **Var7.2** – inter-row spacing (cm); **Var7.3** – sowing norm (million ha^{-1}); **Var7.4** – irrigation norm (m^3 ha^{-1}); **Var7.5** – precipitation amounts during the crop vegetation (mm); **Var7.6** – sum of the effective air temperatures (°C).

Figure 47. Graph of the response between the sum of the effective air temperatures (°C) and oil-seed flax yield (t ha^{-1}) at the irrigated conditions in interaction with the input indexes of the neural network

Having generalized experimental data the neural network, which describes the effect of the developed constituent elements of cultivation technology on the formation of conditional pure profit in dependence on the oil-seed flax yields determined by the complex of agrotechnological (I) and economic (II) factors, was developed (Fig. 48).

The architecture of the ANN (MP 10:10-8-1:1; N = 10) is based on the ten elements (neurons), which have influence on the intensity of the growth processes. All the input neurons are divided on two blocks:

I. <u>Technological factors:</u>

1. Variety.
2. Fertilizers.
3. Irrigation.
4. Sowing terms.
5. Inter-row spacing.
6. Sowing norm.

II. <u>Economic factors:</u>

7. Grain units outcome per ha.

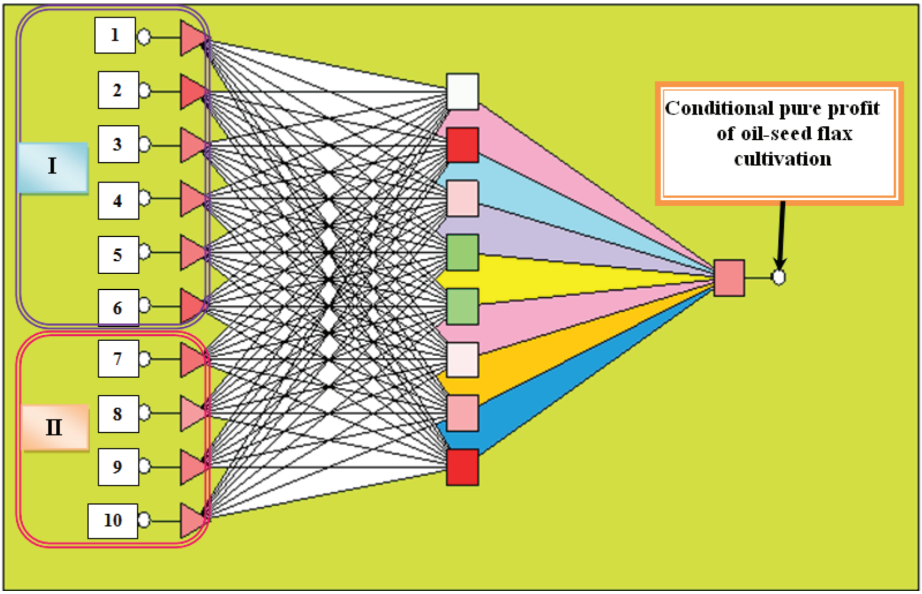


Figure 48. Neural network of oil-seed flax productivity and economic efficiency formation in the conditions of the South of Ukraine, architecture: MP 10:10-8-1:1, training productivity – 0.1059, control productivity – 1.0247, testing productivity – 0.1877

8. Gross value of products, thousands UAH.
9. Productive expenses, thousands UAH.
10. Seed production cost, UAH per 1 tonne.

By the results of the created neural network it was established that mineral fertilizers have the maximum direct influence on the conditional pure profit (marked with the most intensive red color). The interaction of the input factors was observed for mineral fertilizers and irrigation, and also for gross value of the products, productive expenses and, especially, the seed production cost per 1 tonne.

By the results of the neural analysis a high effect of nutrition background, inter-row spacing, sowing norm, precipitation income and sum of the temperatures on the crop productivity formation under the non-irrigated conditions was determined. The highest positive modeled interaction was observed between mineral fertilizers doses and inter-row spacing width. The unregulated elements of the ANN provided stable interaction under the increase of mineral nutrition background with the dynamic of growth of the seed yield at the level of 1.20-1.21 t ha^{-1}. Widening of the inter-row spacing resulted in the decrease of the yield by 20.4%. Interaction between the sum of effective air temperatures and sowing norm demonstrates the increase of theoretical yield from 0.82 to 1.80 t ha^{-1}.

At the irrigated conditions the general architecture of the ANN was formed by the scheme of radial basis function (RBF 6:6-5-1:1). Training productivity of the network of the highest rank (0.69-0.97) was determined for irrigation norm and mineral nutrition background, and the least values were fixed for sowing norm. The maximum level of theoretical yield (1.83 t ha^{-1}) is formed at the interaction of mineral nutrition background with irrigation norm and sum of the effective air temperatures. The improvement of mineral nutrition in combination with irrigation provided the increase of theoretical yield in average by 23%. The share of interaction between the temperatures and sowing norm in the effect on seed yield was 27.9-37.8%. It was also proved that the maximum effect on oil-seed flax production profitability in the South of Ukraine is provided by mineral fertilizers and irrigation.

CONCLUSIONS

Artificial neural networks are very prospective method of mathematical data processing. They find a number of useful implementations in modern agricultural science, which make the processes of taking the right decisions in complex multi-factorial situations easier. ANNs provide science and practice with a huge number of instruments for deep analysis of data, for example, classification and generalization of the large-scale data, creation of linear and non-linear simulation models with different learning algorithms to provide precise and fast predictions, investigation of the processes in their core, efficient control and management, etc. Weather forecasting, economic risks avoidance, crop growth and development simulations, ecological monitoring and environmental management, decision support systems for agricultural producers and other various practical applications of artificial neural networks are the illustrative example of how modern mathematics combined with computer information technologies can improve the life of people nowadays. Development of the ANN-based approach for the comprehensive investigation of the natural processes, which take place in every agro-ecological system, is one of the major ways for achieving the efficient agricultural environmental management, resource-saving, and economically efficient production of high-quality food in the modern conditions. Modern agricultural technologies could not be imagined without the deep integration of informational technologies, which are used not only by scientists, but are widely implemented in such usual things as agricultural machines and agricultural producer support systems. We see artificial neural networks integration in the tractors, syringes and combined harvesters, which are equipped with modern self-studying systems of GPS-tracking and cruise control options. We see neural networks in the systems, which are used by logisticians for solving the tasks of finding the best way of shipping their products. We see neural networks in the bank offices and economic departments of farms where the specialists of the field are computing possible risks, profits and other economic indexes to guarantee their enterprise stable further development. So, we can hardly name the sphere of modern science, technology and production where ANN approach has not been recently introduced and used. And in the nearest future the ANN-based approach will just strengthen its positions in the world scientific community. Of course, artificial neural networks have some essential drawbacks and limitations, which have to be eliminated. We hope that the task of the technology improvement will be solved as soon as possible. Also, it is not doubtful that new ways of the technology realization will be found, and new fields of its implementation will be discovered. Especially, taking into account the growing demand for robotics and artificial intelligence technologies. Therefore, agricultural scientists have to go with the times and adapt the artificial intelligence technologies and methods for the purposes and needs of the science for efficient dealing with modern challenges, which are on the table nowadays. In particular, Ukrainian agricultural sector of national economy is one of the most prospective sectors of getting income to the national budget of the country, so it is very important to develop and raise our agricultural production on the new level with accordance to the modern standards. It is an important task of Ukrainian agrarian science to provide agricultural producers with national developments both in technological and informational fields. We should understand that competitive production of food could be achieved only in the close cooperation between science and practice.

REFERENCES

1. Ackley D.H., Hinton G.E., Sejnowski T.J. A learning algorithm for Boltzmann machines // Cognitive Science. – 1985. – №9(1). – P. 147-169.
2. Agrometeorology / Seemann, Chirkov, Lomas, Primault, Primault B. и др. Под ред. Seemann J., Chirkov Y.I., Lomas J., Primault B.Springer Science & Business Media, 2012.
3. APHA Standard methods for the examination of water and waste water. – Washington DC: American Public Health Association, 1995.
4. Ayers R.S., Westcott D.W. Water quality for agriculture. FAO irrigation and drainage paper 29. – Rev. 1. – Rome: Food and Agriculture Organization of the United Nations, 1985.
5. Baccouche M., Mamalet F., Wolf C., Garcia C., Baskurt A. Sequential deep learning for human action recognition // Human Behavior Unterstanding. Lecture Notes in Computer Science. – Berlin Heidelberg: Springer, 2011. – P. 29-39.
6. Balabin R.M., Safieva R.Z., Lomakina E.I. Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction // Chemometr Intell Lab. – 2007. – №88(2). – P. 183-188.
7. Banjaw D.T., Megersa H.G., Lemma D.T. Effect of water quality and deficit irrigation on tomatoes yield and quality: a review // Advanced Crop Science and Technology. – 2017. – №5. – P. 295.
8. Bartsev S.I., Okhonin V.A. Adaptive networks of information processing. – Krasnoyarsk: Institute of Physics of AS of the USSR, 1986. – 20 pp.
9. Beck M.B. Water quality modeling: a review of the analysis of uncertainty // Water Resources Research. – 1987. – №23(8). – P. 1393-1442.
10. Bekolay T. Learning in large-scale spiking neural networks: Thesis.... Master – University of Waterloo.
11. Ben-Hur A., Horn D., Siegelmann H., Vapnik V.N. Support vector clustering // Journal of Machine Learning Research. – 2001. – №2. – P. 125-137.
12. Bertsekas D.P. Dynamic programming and optimal control: approximate dynamic programming. – Vol. 2. Athena Scientific, 2012.
13. Bertsekas D.P., Tsitsiklis J.N. Neuro-dynamic programming. Athena Scientific, 1996.
14. Bianucci A.M., Micheli A., Sperduti A., Starita A. Application of cascade correlation networks for structures to chemistry // Applied Intelligence. – 2000. – №12(1-2). – P. 117-147.
15. Biliaieva I.M. Theoretical bases and agro-ecological substantiation of the measures for the irrigated lands productivity improvement in conditions of the South of Ukraine: Thesis... Doctor Agricultural Sciences: 06.01.02. – Kherson, 2018. – 422 pp.
16. Billah B., King M.L., Snyder R.D., Koehler A.B. Exponential smoothing model selection for forecasting // International Journal of Forecasting. – 2006. – №22. – P. 239-247.
17. Bongard M.M. Problems of recognition. – Moscow: Fizmatgiz, 1967.
18. Bourquin J., Schmidl, H., van Hoogevest P., Leuenberger H. Advantages of artificial neural networks (ANNs) as alternative modelling technique for data sets showing non-linear re-

lationships using data from a galenical study on a solid dosage form // European Journal of Pharmaceutical Sciences. – 1998. – №7(1). – P. 5-16.

19. Brierley P. Some practical applications of neural networks in the electricity industry: Thesis.... Ph.D. – Cranfield University, UK, 1998.
20. Brierley P., Batty B. Data mining with neural networks – an applied example in understanding electricity consumption patterns // Knowledge Discovery and Data Mining. IEE, 1999. – P. 240-303.
21. Carpenter G.A., Grossberg S. Adaptive resonance theory // The Handbook of Brain Theory and Neural Networks. – Cambridge: MIT Press, 2003. – P. 87-90.
22. Carpenter G.A., Grossberg S. ART 2: Self-organization of stable category recognition codes for analog input patterns // Applied Optics. – 1987. – №26(23). – P. 4919-4930.
23. Carpenter G.A., Grossberg S. ART 3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures // Neural Networks. – 1990. – №3. – P. 129-152.
24. Carpenter G.A., Grossberg S., Rosen D.B. Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system // Neural Networks. – 1991. – №4. – P. 759-771.
25. Chung J., Gulcehre C., Cho K., Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling // CoRR. – 2014.
26. Ciresan D., Meier U., Schmidhuber J. Multi-column deep neural networks for image classification // IEEE Conference on Computer Vision and Pattern Recognition. – 2012. – P. 3642-3649.
27. Cortes C., Vapnik V.N. Support-vector networks // Machine Learning. – 1995. – №20(3). – P. 273-297.
28. De Livera A.M., Hyndman R.J., Snyder R.D. Forecasting time series with complex seasonal patterns using exponential smoothing // Journal of the American Statistical Association. – 2011. – №106(496). – P. 1513-1527.
29. Debok G., Kohonen T. Analysis of financial data by using self-organizing maps. Alpina Publisher, 2001. – 317 pp.
30. Devore J.L. Probability and Statistics for Engineering and the Sciences. – Boston: Cengage learning, 2011.
31. Donald O.H. The organization of behavior. – New York: Wiley, 1949.
32. Dumitru C., Maria V. Advantages and disadvantages of using neural networks for predictions // Ovidius University Annals, Series Economic Sciences. – 2013. – №13(1).
33. Elman J.L. Finding structure in time. // Cognitive Science. - 1990. - P. 179-211.
34. Everitt B., Skrondal A. The Cambridge dictionary of statistics (Vol. 106). – Cambridge: Cambridge University Press, 2002.
35. Exploring the limits of language modeling / Jozefowicz, Vinyals, Schuster, Shazeer, Wu, 2016.
36. Farley B.G., Clark W.A. Simulation of self-organizing systems by digital computer // IRE Transactions on Information Theory. – 1954. – №4(4). – P. 76-84.
37. Feizi M., Hajabbasi M.A., Mostafazadehfard B. Saline irrigation water management strategies for better yield of safflower (Carthamus tinctorius L.) in an arid region // Australian Journal of Crop Science. – 2010. – №4. – P. 408-414.
38. Fernández S., Graves A., Schmidhuber J. An application of recurrent neural networks to discriminative keyword spotting // Proceedings of the 17th International Conference on Artificial Neural Networks. ICANN’07. – 2007. – P. 220-229.

39. FitzHugh R. Mathematical models of threshold phenomena in the nerve membrane // The bulletin of mathematical biophysics. – 1955. – №17(4). – P. 257-278.
40. Forecasting with exponential smoothing: the state space approach / Hyndman R., Koehler A.B., Ord J.K., Snyder R.D. Springer Science & Business Media, 2008.
41. Frasconi P., Gori M., Sperduti A. A general framework for adaptive processing of data structures // IEEE Transactions on Neural Networks. – 1998. – №9(5). – P. 768-786.
42. Furness R.W., Bryant D.M. Effect of wind on field metabolic rates of breeding northern fulmars // Ecology. – 1996. – №77(4). – P. 1181- 1188.
43. Gallicchio C., Micheli A. Tree echo state networks // Neurocomputing. – 2013. – №101. – P. 319-337.
44. Galves A., Löcherbach E. Infinite systems of interacting chains with memory of variable length – a stochastic model for biological neural nets // Journal of Statistical Physics. – 2013. – №151(5). – P. 896-921.
45. Gardner E.S. Exponential smoothing: the state of the art – Part II. // International journal of forecasting. – 2006. – №22(4). – P. 637-666.
46. Gelper S., Fried R., Croux C. Robust forecasting with exponential and Holt-Winters smoothing // Journal of forecasting. – 2010. – №29(3). – P. 285-300.
47. Gers F. A., Schmidhuber J. LSTM recurrent networks learn simple context free and context sensitive languages // IEEE Transactions on Neural Networks. – 2001. – №12(6). – P. 1333-1340.
48. Gers F., Schraudolph N., Schmidhuber J. Learning precise timing with LSTM recurrent networks // Journal of Machine Learning Research. – 2002. – №3. – P. 115-143.
49. Gerstner W., Kistler W.M. Spiking neuron models: Single neurons, populations, plasticity. – Cambridge: Cambridge university press, 2002.
50. Godfrey K. Compartmental models and their application. Academic Press, 1983.
51. Golubev Y.F. Neural networks methods in mechatronics. – Moscow: MSU Publishing house, 2007. – 157 pp.
52. Goodfellow I., Bengio Y., Courville A. Deep learning. MIT Press, 2016. – 196 pp.
53. Graves A., Liwicki M., Fernandez S., Bertolami R., Bunke H., Schmidhuber J. A novel connectionist system for improved unconstrained handwriting recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2009. – №31(5).
54. Graves A., Schmidhuber J. Offline handwriting recognition with multidimensional recurrent neural networks // NIPS'08 Proceedings of the 21st International Conference on Neural Information Processing Systems. – 2009. – P. 545-552.
55. Hammer B., Micheli A., Sperduti A., Strickert M. A general framework for unsupervised processing of structured data // Neurocomputing. – 2004. – №57. – P. 3-35.
56. Hammer B., Micheli A., Sperduti A., Strickert M. Recursive self-organizing network models // Neural Networks. – 2004. – №17. – P. 1061-1085.
57. Han J., Morag C. The influence of the sigmoid function parameters on the speed of backpropagation learning // From Natural to Artificial Neural Computation. 1995. – P. 195-201.
58. Hindmarsh J.L., Rose R.M. A model of neuronal bursting using three coupled first order differential equations // Proc. R. Soc. London, Ser. B. – 1984. – №221. – P. 87-102.
59. Hinton G Deep belief networks // Scholarpedia. – 2009. – №4(5). – P. 5947.
60. Hinton G.E. Training products of experts by minimizing contrastive divergence // Neural Computation. – 2002. – №14(8). – P. 1771-1800.
61. Hinton G.E., Osindero S., Teh Y. A fast learning algorithm for deep belief nets // Neural Computation. – 2006. – №18(7). – P. 1527-1554.

62. Hochreiter S., Schmidhuber J. Long short-term memory // Neural computation. – 1997. – №9(8). – P. 1735-1780.
63. Hodgkin A.L., Huxley A.F. A quantitative description of membrane current and its application to conduction and excitation in nerve // The Journal of physiology. – 1952. – №117(4). – P. 500-544.
64. Horn R. A., Johnson C.R. Matrix analysis. – Cambridge: Cambridge University Press, 2012.
65. Ivakhnenko A.G., Lapa V.G. Cybernetics and forecasting techniques. American Elsevier Pub. Co, 1967.
66. Jaeger H., Haas H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication // Science. – 2004. – №304(5667). – P. 78-80.
67. Jones M.T. Artificial intelligence: a systems approach. – Massachusetts: Infinity Science Press LLC, 2008.
68. Jordan M.I. Serial order: A parallel distributed processing approach. // Institute for Cognitive Science Report 8604. - University of California, 1986.
69. Kaelbling L.P., Littman M.L., Moore A.W. Reinforcement learning: a survey // Journal of Artificial Intelligence Research. – 1996. – №4. – P. 237-285.
70. Kelly W. P. Use of saline irrigation water // Soil Science. – 1963. – №95(4). – P. 355-391.
71. Khakimov B.B. Modeling of correlation dependencies by splines on the example of geology and ecology. – Moscow: MSU Publishing house, 2003. – 144 pp.
72. Kim H., Jeong H., Jeon J., Bae S. Effects of irrigation with saline water on crop growth and yield in greenhouse cultivation // Water. – 2016. – №8(4). – P. 127.
73. Kohonen T. Self-organizing maps. – Vol. 3. – New York: 2001. – 501 pp.
74. Kosko B. Bidirectional associative memories // IEEE Transactions on Systems, Man, and Cybernetics. – 1988. – №18(1).
75. Kosko B. Competitive adaptive bi-directional associative memories // Proceedings of the IEEE First International Conference on Neural Networks. – 1987. – №2. – P. 759-766.
76. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolutional neural networks // Communications of the ACM. – 2017. – №60(6). – P. 84-90.
77. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolutional neural networks // Advances in Neural Information Processing Systems. – 2012. – №1. – P. 1097-1105.
78. Lavrenko S.O., Lavrenko N.N., Pichura V.I. Neural network modeling of chickpea grain yield on ameliorated soils // Nauchnyi zhurnal Rossiiskogo NII problem melioratsii. – 2015. – №2(18). – P. 16-30.
79. LeCun Y., Boser B., Denker J.S., Henderson D., Howard R.E., Hubbard W., Jackel L.D. Backpropagation applied to handwritten zip code recognition // Neural Computation. – 1989. – №1(4). – P. 541-551.
80. Ledesma S., Avina G., Sanchez R. Practical considerations for simulated annealing implementation // Simulated Annealing. I-Tech Publishing, 2008. – P. 401-420.
81. Ledesma S., Ibarra-Manzano M.A., Garcia-Hernandez M.G., Almanza-Ojeda D.L. Neural lab a simulator for artificial neural networks // Computing Conference, 2017. 2017. – P. 716-721.
82. Lee H., Grosse R., Ranganath R., Ng A.Y. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations // ICML'2009. 2009.
83. Lewis C.D. Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. Butterworth-Heinemann, 1982.
84. Li X., Wu X. Constructing Long Short-Term Memory based Deep Recurrent Neural Networks for Large Vocabulary Speech Recognition. – 2014.

85. Likhovid P.V. Analysis of the Ingulets irrigation water quality by agronomical criteria // Success of Modern Science and Education. – 2015. – №5. – P. 10-12.
86. Liu S., Tai H., Ding Q., Li D., Xu L., Wei Y. A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction // Mathematical and Computer Modelling. – 2013. – №58(3). – P. 458-465.
87. Logan M. Biostatistical design and analysis using R: a practical guide. John Wiley & Sons, 2011.
88. Lozovitsii, P.S. Monitoring of the humus status of soils of the Ingulets irrigation system // Eurasian Soil Science. – 2012. – №45(3). – P. 336- 349.
89. Lykhovyd P.V. Global warming inputs in local climate changes of the Kherson region: current state and forecast of the air temperature // Ukrainian Journal of Ecology. - 2018. – №8(2). – P. 39-41.
90. Lykhovyd P.V. Prediction of sweet corn yield depending on cultivation technology parameters by using linear regression and artificial neural network methods // Biosystems Diversity. - 2018. – №26(1). – P. 11-15.
91. Lykhovyd P.V. The productivity of sweet corn depending on soil treatment, fertilizers, plants thickness as a result of drip irrigation: Thesis... Ph.D. Agricultural Sciences: 06.01.02. – Kherson, 2017. – 256 pp.
92. Lykhovyd P.V., Kozlenko Y.V. Assessment and forecast of water quality in the River Ingulets irrigation system // Ukrainian Journal of Ecology. - 2018. – №8(1). – P. 350-355.
93. Lykhovyd P.V., Lavrenko S.O. Influence of tillage and mineral fertilizers on soil biological activity under sweet corn crops // Ukrainian Journal of Ecology. – 2017. – №7(4). – P. 18-24.
94. Maass W. Networks of spiking neurons: the third generation of neural network models // Neural networks. – 1997. – №10(9). – P. 1659-1671.
95. Masters T. Neural, novel & hybrid algorithms for time series prediction. – New York: John Wiley and Sons, 1995.
96. Methodology of field experiment (Irrigated agriculture): Textbook / Ushkarenko, Kokovikhin, Holoborodko, Vozhehova, – Kherson: Hrin DS, 2014. – 443 pp.
97. Minsky M., Papert S. Perceptrons: An introduction to computational geometry. MIT Press, 1969.
98. Multilingual language processing from bytes / Gillick, Brunk, Vinyals, Subramanya, 2015.
99. Nitish S., Hinton C.G., Krizhevsky A., Sutskever I., Salakhutdino R. "Dropout: a simple way to prevent neural networks from overfitting // Journal of Machine Learning Research. – 2014. – №15(1). – P. 1929-1958.
100. Ould Ahmed B.A., Yamamoto T., Inoue M. Response of drip irrigated sorghum varieties growing in dune sand to salinity levels in irrigation water // Journal of Applied Sciences. – 2007. – №7. – P. 1061-1066.
101. Palani S., Liong S.Y., Tkalich P. An ANN application for water quality forecasting // Marine Pollution Bulletin. – 2008. – №56(9). – P. 1586- 1597.
102. Patel P.D., Patel S.P. Prediction of weld strength of metal active gas (MAG) welding using artificial neural network // International Journal of Engineering Research and Applications. - 2011. – №1(1). – P. 36-44.
103. Pavliuk Y.V., Terekhin E.A., Pichura V.I. Use of neural network technologies for modeling of time series processes of rivers water content // Erosion and canal processes and modern methods of their study: materials of the X Seminar of young scientists of the universities connected by the council on erosion, canal and estuary processes. – 2014. – P. 141-148.

104. Pereira L.S., Oweis T., Zairi A. Irrigation management under water scarcity // Agricultural Water Management. – 2002. – №57(3). – P. 175-206.
105. Petrov A.P. About the capacities of perceptron // News of AS of the USSR, Technical cybernetics. - 1964. - №6.
106. Ponulak F., Kasinski A. Introduction to spiking neural networks: Information processing, learning and applications // Acta neurobiologiae experimentalis. – 2011. – №71(4). – P. 409-433.
107. Pyrkov T., Slipensky K., Barg M., Kondrashin A., Zhurov B., Zenin A., Pyatnitskiy M., Menshikov L., Markov S., Fedichev P.O. Extracting biological age from biomedical data via deep learning: too much of a good thing? // Scientific Reports. – 2018. – №8(1). – P. 5210.
108. Reckhow K.H. Water quality prediction and probability network models // Canadian Journal of Fisheries and Aquatic Sciences. – 1999. – №56(7). – P. 1150-1158.
109. Rochester N., Holland J.H, Habit L.H., Duda W.L. Tests on a cell assembly theory of the action of the brain, using a large digital computer // IRE Transactions on Information Theory. – 1956. – №2(3). – P. 80-93.
110. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain // Psychological Review. – 1958. – №65(6). – P. 386-408.
111. Rumelhart D., McClelland J.L., PDP Research Group. Parallel distributed processing. – Vol. 1. – Cambridge: MIT Press, 1986. – 567 pp.
112. Russell S.J., Norvig P. Artificial intelligence: a modern approach. – Third Edition. – Prentice Hall, 2009. – 1152 pp.
113. SAS Getting started with SAS Enterprise Miner 13.1. 2013.
114. Schmidhuber J. Learning complex, extended sequences using the principle of history compression // Neural Computation. – 1992. – №4. – P. 234-242.
115. Schmidhuber J., Wierstra D., Gagliolo M., Gomez F. Training recurrent networks by Evolino // Neural Computation. – 2007. – №19(3). – P. 757-779.
116. Seckler D., Barker R., Amarasinghe U. Water scarcity in the twenty-first century // International Journal of Water Resources Development. – 1999. – №15(1-2). – P. 29-42.
117. Shakhman I. A., Bystriantseva A. N. Assessment of ecological state and ecological reliability of the lower section of the Ingulets river // Hydrobiological Journal. – 2017. – №53(5).
118. Shang X.S., Lin W.D., Tang Y.K. Development and application of a combined water quality prediction model based on exponential smoothing and GM (1, 1). A case study of iron and manganese levels in Yongjiang River // Huanjing Kexue yu Jishu. – 2011. – №34(1). – P. 191-195.
119. Shi X., Chen Z., Wang H., Yeung D., Wong W., Woo W. Convolutional LSTM network: A machine learning approach for precipitation nowcasting // NIPS'15 Proceedings of the 28th International Conference on Neural Information Processing Systems. – 2015. – Vol. 1. – P. 802-810.
120. Singh K.P., Basant A., Malik A., Jain G. Artificial neural network modeling of the river water quality – a case study // Ecological Modelling. – 2009. – №220(6). – P. 888-895.
121. Smolensky P. Information processing in dynamical systems: Foundations of harmony theory // Parallel Distributed Processing: Explorations in the Microstructure of Cognition. – Cambridge: MIT Press, 1986. – P. 194-281.
122. Socher R., Lin C., Ng A.Y., Manning C.D. Parsing natural scenes and natural language with recursive neural networks // The 28th International Conference on Machine Learning. – 2011.
123. Sperduti A., Starita A. Supervised neural networks for the classification of structures // IEEE Transactions on Neural Networks. – 1997. – №8(3). – P. 714-735.

124. Steve L., Giles C.L., Tsoi A.C., Back D. Face recognition: a convolutional neural network approach // Neural Networks. – 1997. – №8(1). – P. 98-113.
125. Storkey A. Increasing the capacity of a Hopfield network without sacrificing functionality // Artificial Neural Networks – ICANN'97. 1997. – P. 451-456.
126. Striving for simplicity: the all convolutional net / Springenberg, Dosovitskiy, Brox, Riedmiller M. 2014.
127. Sutskever L., Vinyals O., Le Q. Sequence to sequence learning with neural networks // Electronic Proceedings of the Neural Information Processing Systems Conference. – 2014. – №27. – P. 5346.
128. Todd D.K. Groundwater hydrology. – New York: John Wiley and Sons, 1980.
129. Tu J. V. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes // Journal of clinical epidemiology. – 1996. – №49(11). – P. 1225-1231.
130. URL: https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/index.html
131. URL: https://engineering.case.edu/eecs/node/213
132. URL: https://towardsdatascience.com/introduction-to-neural-networks-advantages-and-applications-96851bd1a207
133. URL: https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/
134. van Otterlo M., Wiering M. Reinforcement learning and Markov decision processes // Reinforcement Learning. – Berlin Heidelberg: Springer, 2012. – P. 3-42.
135. von Neumann J. The computer and brain. Yale University Press, 1986. – 136 pp.
136. Vozhehova R.A., Lavrynenko Y.O., Kokovikhin S.V., Lykhovyd P.V., Biliaieva I.M., Drobitko A.V., Nesterchuk V.V. Assessment of the CROPWAT 8.0 software reliability for evapotranspiration and crop water requirements calculations // Journal of Water and Land Development. - 2018. – №39. – P. 147-152.
137. Vozhehova R.A., Lykhovyd P.V., Lavrenko S.O., Kokovikhin S.V., Lavrenko N.M., Marchenko T.Yu., Sydyakina O.V., Hlushko T.V., Nesterchuk V.V. Artificial neural network use for sweet corn water consumption prediction depending on cultivation technology peculiarities // Research Journal of Pharmaceutical, Biological and Chemical Sciences. - 2019. – №10(1). – P. 354-358.
138. Warren M., Pitts W. A logical calculus of ideas immanent in nervous activity // Bulletin of Mathematical Biophysics. – 1943. – №5(4). – P. 115-133. Hebb,
139. Wasserman F. Neural computing. Theory and practice. – Moscow: Mir, 1992. – 240 pp.
140. Water quality for irrigation. Agronomical criteria: DSTU 2730-94. – Kyiv: Derzhstandart Ukrajiny, 1994.
141. Weng J. Why have we passed neural networks do not abstract well? // Natural Intelligence: the INNS Magazine. – 2011. – №1(1). – P. 13-22.
142. Weng J., Ahuja N., Huang T.S. Cresceptron: a self-organizing neural network which grows adaptively // Proc. International Joint Conference on Neural Networks. – 1992. – №1. – P. 576-581.
143. Weng J., Ahuja N., Huang T.S. Learning recognition and segmentation of 3-D objects from 2-D images // Proc. 4th International Conf. Computer Vision. – 1993. – P. 121-128.
144. Weng J., Ahuja N., Huang T.S. Learning recognition and segmentation using the Cresceptron // International Journal of Computer Vision. – 1997. – №25(2). – P. 105-139.
145. Wiener N. Cybernetics or control and communication in the animal and the machine. – New York: John Wiley & Sons, 1948.

146. Wilcox L.V. Classification and use of irrigation water. Circular No. 969. – Washington: USDA, 1955. – P. 19.
147. Wilson H.R., Cowan J.D. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue // Kybernetik. – 1973. – №13(2). – P. 55-80.
148. Wu X., Guo G., Weng J. Skull-closed autonomous development: WWN-7 dealing with scales // Proc. International Conference on Brain-Mind. – Michigan: 2013. – P. 1-9.
149. Yadav N., Yadav A., Kumar M. An introduction to neural network methods for differential equations. Springer, 2015. – 100 pp.
150. Zeiler M., Krishnan D., Taylo, G., Fergus R. Deconvolutional networks // CVPR'2010. 2010.
151. Zell A. Simulation neuronaler netze. – Vol. 1. Addison-Wesley, 1994. – 73 pp.

Vozhehova R. A., Lykhovyd P. V., Kokovikhin S. V., Biliaieva I. M., Markovska O. Y., Lavrenko S. O., Rudik O. L.

MONOGRAPH

Vozhehova R. A., Lykhovyd P. V., Kokovikhin S. V., Biliaieva I. M., Markovska O. Y., Lavrenko S. O., Rudik O. L. Artificial neural networks and their implementation in agricultural science and practice/ Vozhehova R. A., Lykhovyd P. V., Kokovikhin S. V., Biliaieva I. M., Markovska O. Y., Lavrenko S. O., Rudik O. L. - Warsaw: "Diamond trading tour", 2019. - 108 p.

ISBN 978-83-66401-06-8

Signed in the press on 01.06.2019.
Format 60x94/16
Offset paper. Offset printing.
Conventional printed sheets 6.1
Circulation 500 copies.
Order No. 255
Wydawca: Sp. z o.o. «Diamond trading tour»
Adres wydawcy I redacji:00-728 Warszawa, ul. S. Kierbedzia, 4 lok.103