Adaptation of the neural network to the study of electrical characteristics of dispersed materials in sorption processes

G V Kalashnikov, S V Makeev and E S Bunin

1 Voronezh State University of Engineering Technologies, 19, Revolutsii Ave., Voronezh, 394036, Russia

E-mail: kagen5@yandex.ru

Abstract. Modern technological processes cannot exist without optimizing their parameters. The current direction remains their improvement and identification of such regimes, in which without losing the quality of finished products there is a reduction of energy expenditures and resource saving. Since some processes have a complex structure, and several parameters have to be varied to optimize them, it is advisable to use mathematical devices and models at the primary stage. The subject matter of artificial neural networks refers to the interdisciplinary field of knowledge related to biocibernetics, electronics, applied mathematics, statistics and automation. Artificial neural networks are versatile for developing new technological solutions and analyzing the detected data with the most effective result. The conclusions obtained from the experiments make it possible to understand how the main parameters of sorption processes affect its intensity, as well as to determine the intervals of change of input factors for statistical analysis. All this makes it possible to ensure stability of technological parameters during all experiments and high reproducibility of experiments. This technology is gaining popularity at present due to the growth of computing capabilities and modernization of equipment. Therefore, the direction of research and improvement of all technological parameters and modes is relevant.

1. Introduction

One promising application of artificial neural networks is industrial production. In this area, there is a tendency to move to production modules with a high level of automation, which requires an increase in the number of intelligent self-regulating and self-adjusting machines. However, manufacturing processes are characterized by a wide variety of dynamically interacting parameters, making it difficult to create adequate analytical models. Modern production is constantly becoming more complicated. This slows down the introduction of new technological solutions. In addition, in some cases successful analytical mathematical models show failure due to lack of computing capacity. As a result, there is increasing interest in alternative approaches to modeling production processes using artificial neural networks, which provide opportunities to create models that operate in real time with low errors and can be further learned during use.

2. Research objective

To prove expediency of application of neural networks to analysis of sorption processes, to consider algorithms of their training with selection of operating modes for optimization of parameters of power consumption and resource saving.
3. Research object
Based on the analysis of the theory and technique of sorption processes, we concluded that it would be useful to apply a neural network to the study of electrical characteristics of dispersed materials. One of the main electrical characteristics of the electric smoking process is the corona discharge current intensity; the intensity of the process is characterized by such a parameter as the degree of deposition of smoking components on the product. The strength of the corona current, or rather its dependence on the voltage applied to the corona electrodes, is essentially a voltagamp characteristic of the electrocoptic plant and determines the load on the high-voltage generator, and therefore characterizes the energy consumption on the electric smoking process. It is natural to assume that these characteristics are dependent on such basic process parameters as relative humidity in the smoking chamber and speed of movement of the smoke-air mixture [1].

4. Materials and methods
The efficiency of the process is characterized by a degree of precipitation $K$, which is determined according to expression (1) as the ratio of the difference in the concentrations of the smoke-air mixture at the inlet $c_1$ and the outlet $c_2$ to the concentration at the inlet of the smoking chamber $c_1$.

$$K = \frac{c_1 - c_2}{c_1}. \quad (1)$$

The scientific significance of the research carried out in the process of precipitation of smoke components lies in finding optimal intervals of technological parameters that would ensure the highest efficiency of the process with minimum energy consumption [2–4]. As input parameters we chose the speed of air-flue mix in the smoking camera $\nu (\text{m/s})$ and relative humidity of the air-flue mix $\varphi (%)$. The ranges of each of the input parameters are shown in table 1.

| Name of parameter | Unit of measure | Lower level | Top level | Change interval |
|-------------------|-----------------|-------------|-----------|-----------------|
| Corona electrode voltage, $U_k$ | kV | $x_1$, 5 | 60 | 2 |
| Speed of air-flue mix $\nu$, m/s | m/s | $x_2$, 0.1 | 1.2 | 0.1 |
| Relative humidity of smoke-air mixture, $\varphi$ | % | $x_3$, 50 | 90 | 6 |

**Table 1. Ranges of input factors variation**

![Figure 1](image-url)  
**Figure 1.** Experimental curves: (a) - experimental dependence of deposition degree on voltage applied to corona electrodes -Δ- at $\nu=0.6$ m/s, $\varphi=65\%$; -Ο- at $\nu=0.4$ m/s, $\varphi=75\%$; -□- at $\nu=0.2$ m/s, $\varphi=85\%$; (b) - Experimental dependence of corona current on voltage -Δ- at $\nu=0.6$ m/s, $\varphi=65\%$; -Ο- at $\nu=0.4$ m/s, $\varphi=75\%$; -□- at $\nu=0.2$ m/s, $\varphi=85\%$. 

1. [1]  
2. [2]  
3. [3]  
4. [4]
Since the experimental plant has a relatively high level of automation and is equipped with a centralized control system for data collection, the removal of such a large enough number of data points did not pose special technical problems. Figure 1 shows a view of some of the resulting curves.

In view of the sufficient complexity of the type of dependencies obtained, their approximation by standard methods is extremely difficult. Therefore, we decided to use a mathematical apparatus based on neural networks to process the obtained experimental data.

5. Discussion of the results

In accordance with the principles of functioning of biological neurons, various mathematical models have been created, which more or less realize the properties of the natural nervous cell. Another important factor is the choice of training strategy. There are two approaches: learning with a teacher (supervised learning) and learning without a teacher (unsupervised learning).

The unipolar function is generally represented by the formula:

\[ f(x) = \frac{1}{1 + e^{-\beta x}}. \]  

(2)

In these formulas the \( \beta \) parameter is selected by the user. Its value affects the form of the activation function. A sigmoid neuron is generally trained with a teacher on the principle of minimizing the objective function, which for a single learning tuple \( <x, d> \) \( i \)-th of the ith neuron is defined as:

\[ E = \frac{1}{2} (y_i - d_i)^2, \]  

(3)

where

\[ y_i = f(u_i) = f \left( \sum_{j=0}^{N} w_{ij} x_j \right). \]  

(4)

The function \( f(u_i) \) is sigmoidal, \( x \) – is the input vector, \( x = [x_0, x_1, ..., x_k]^T \) with \( x_0 = 1 \) in the presence of polarization and \( x_0 = 0 \) in the absence of polarization, and \( d_i \) – is its expected value at the output of the \( i \)-th neuron. The use of a continuous activation function allows the use of gradient methods in training.

The multilayer network consists of neurons arranged at different levels, besides the input and output layers there is at least one more inner, i.e. hidden layer. As already noted in the literature on neural networks, such a neural system is called a multilayer perceptron. The weights of the hidden layer neurons are marked with the upper index (1) and the output layer weight vectors with the upper index (2). The output signals of the hidden layer neurons are indicated by \( u_j (j = 1, 2, ..., K) \), and output layer \( y_j (j = 1, 2, ..., M) \). We assume that the neuronal activation function is given in sigmoidal unipolar. To simplify the description, we will use the extended designation of the input network vector as \( x = [x_0, x_1, ..., x_k]^T \), where \( x_0 = 1 \) corresponds to a unit polarization signal. Two input network vectors are connected to vector \( x \) vector of actual output signals \( y = [y_0, y_1, ..., y_M]^T \) and expected output vector \( d = [d_0, d_1, ..., d_M]^T \). The purpose of training is to select such weights \( w_{ij}(1) \) and \( w_{ij}(2) \) for all layers of the network, so that at the given input vector \( x \) the values of signals \( u_i \) are obtained at the output, which with the required accuracy will coincide with the expected values \( d_i \) for \( i = 1, 2, ..., M \). If we consider the unit polarization signal as one of the components of the input vector \( x \), the polarization weights can be added to the weight vectors of the respective neurons of both layers. With this approach, the output of the \( i \)-th neuron of the hidden layer can be described by the function:

\[ v_j = f \left( \sum_{j=0}^{N} w_{ij} x_j \right), \]  

(5)

Since the experimental plant has a relatively high level of automation and is equipped with a centralized control system for data collection, the removal of such a large enough number of data points did not pose special technical problems. Figure 1 shows a view of some of the resulting curves.

In view of the sufficient complexity of the type of dependencies obtained, their approximation by standard methods is extremely difficult. Therefore, we decided to use a mathematical apparatus based on neural networks to process the obtained experimental data.
wherein index 0 corresponds to a signal and polarization weights \( v_0 = 1, x_0 = 1 \). In the output layer, the k-th neuron produces an output defined as:

\[
y_k = f \left( \sum_{i=0}^{K} w_{ki}^{(2)} v_i \right) = f \left( \sum_{i=0}^{K} w_{ki}^{(2)} f \left( \sum_{j=0}^{N} w_{ij}^{(1)} x_j \right) \right).
\]  

(6)

It follows from the formula that the value of the output signal is influenced by the weights of both layers, whereas the signals generated in the hidden layer are independent of the weights of the output layer. In the case of a single learning sample \((x, d)\), the objective function is defined as:

\[
E(w) = \frac{1}{2} \sum_{k=1}^{M} (y_k - d_k)^2.
\]  

(7)

In more training samples \( j \) \((j = 1, 2, \ldots, p)\) the objective function becomes the sum of all selections:

\[
E(w) = \frac{1}{2} \sum_{j=1}^{p} \sum_{k=1}^{M} (y_k^{(j)} - d_k^{(j)})^2.
\]  

(8)

For the sake of simplicity, it can be considered that the purpose of the training is to determine the values of the neuronal weights of each network layer so that at a given input vector the values of the signals \( y \), coinciding with the required accuracy with the expected values \( d \), at \( i = 1, 2, \ldots, M \).

Network training using the reverse error propagation algorithm is carried out in several stages. After obtaining the output values \( y \), it becomes possible to calculate the actual value of the objective function \( E(w) \). The second step minimizes the value of this function. At that refinement of weights vector (training) is performed by formula:

\[
w(k + 1) = w(k) + \Delta w,
\]  

(9)

where

\[
\Delta w = \eta p(w),
\]  

(10)

\( \eta \) – training coefficient, a \( p(w) \) - direction in multidimensional space \( w \).

Teaching a multi-layer network using gradient techniques requires defining a gradient vector relative to the weights of all layers of the network, which is necessary to properly select the \( p(w) \) direction. This problem has an obvious solution only for the weights of the output layer. For other layers a special strategy has been created, which in the theory of artificial neural networks is called the algorithm of inverse propagation of error (English: error back propagation), identified, as a rule, with the procedure of network training. According to this algorithm, the following steps are identified in each training cycle.

The objective function is defined by the expression:

\[
E = \frac{1}{2} \sum_{k=1}^{M} \left[ f \left( \sum_{i=0}^{K} w_{ki}^{(2)} v_i \right) - d_k \right]^2 = \frac{1}{2} \sum_{k=1}^{M} \left[ f \left( \sum_{i=0}^{K} w_{ki}^{(2)} f \left( \sum_{j=0}^{N} w_{ij}^{(1)} x_j \right) \right) - d_k \right]^2.
\]  

(11)

First of all, the weights of neurons of the output layer are selected. For weekend scales we get:

\[
\frac{\partial E}{\partial w_{ij}^{(2)}} = (y_i - d_i) \frac{df(u_i^{(2)})}{du_i^{(2)}} v_j,
\]  

(12)

where \( u_i^{(2)} = \sum_{j=0}^{K} w_{ij}^{(2)} v_j \). If to enter designation \( \delta_{ij}^{(2)} = (y_i - d_i) \frac{df(u_i^{(2)})}{du_i^{(2)}} \), then the corresponding gradient component with respect to the neuronal weights of the output layer can be represented as:
\[ \frac{\partial E}{\partial w_{ij}^{(2)}} = \delta_i^{(2)} v_j. \] (13)

The gradient components with respect to the neurons of the hidden layer are determined by the same principle; however they are described by a different, more complex dependency following from the existence of a function given in the form:

\[ \frac{\partial E}{\partial w_{ij}^{(1)}} = \sum_{k=1}^{M} (y_k - d_k) \frac{dy_k}{dv_i} \frac{dv_i}{dw_{ij}^{(1)}}. \] (14)

After specifying the individual components of this expression, we obtain:

\[ \frac{\partial E}{\partial w_{ij}^{(1)}} = \sum_{k=1}^{M} (y_k - d_k) \frac{df(u_k^{(2)})}{du_k^{(2)}} w_{ki}^{(2)} \frac{df(u_i^{(1)})}{du_i^{(1)}} x_j. \] (15)

If to enter designation:

\[ \delta_i^{(1)} = \sum_{k=1}^{M} (y_k - d_k) \frac{df(u_k^{(2)})}{du_k^{(2)}} w_{ki}^{(2)} \frac{df(u_i^{(1)})}{du_i^{(1)}}, \] (16)

then we get an expression defining the components of the gradient relative to the weights of the neurons of the hidden layer in the form:

\[ \frac{\partial E}{\partial w_{ij}^{(1)}} = \delta_i^{(1)} x_j. \] (17)

In the classical inverse error propagation algorithm, the \( p(w) \) factor sets the direction of the negative gradient, so:

\[ \Delta w = \eta \nabla E(w). \] (18)

The results of the work performed prove that the neural network has the ability to learn and generalize the accumulated knowledge in relation to the sorption process. The network is able to summarize the information received and show good results on data not used in the training process. A trained approximation network plays the role of a universal multiple-variable function approximation that implements a nonlinear function of the form \( y = f(x) \), where \( x \) – the input vector and \( y \) – the implemented multiple-variable function \[3\].

6. Conclusion

Based on the system approach, complex theoretical and experimental studies have been carried out, as a result of which recommendations on scientific and practical support of sorption processes have been developed. The relevance of the use of neural network for the study of electrical characteristics of dispersed materials has been proved. Analysis of neural network training algorithms allowed identifying optimal parameters of sorption processes in terms of energy consumption reduction and resource saving growth.

References

[1] Drozdetskaya I S and Berezovikova I P 2018 Influence of liquid smoking flavoring on rheological characteristics of fish mince *Proceedings of VSUET* **80** (1) 193-198

[2] Semenova J A, Levshenko M T, Kolovolova A Y, Ilukhina N V and Kurbanov M N 2019 Establishment of a test culture to study the effects of ionizing radiation on opportunistic and pathogenic microflora of food products *Proceedings of VSUET* **81** (2) 245-249

[3] Zotov A P and Makeev S V 2010 Modeling of the process of production of liquid food media
with smoke aroma Proceedings of VSUET 43 (1) 94-98

[4] Mikhailova O V, Belova M V, Korobkov A N and Novikova G V 2019 Development of a plant for exfoliation of rapeseed in an electromagnetic field of ultra-high frequency Proceedings of VSUET 81 (2) 27-34