The Neural Network Method for Solving an Evolutionary Partial Differential Equation in a Bounded Domain with a Piecewise Smooth Surface

Iryna Vergunova

Taras Shevchenko National University of Kiev, 64/13, Volodymyrska Street, Kyiv, 01601, Ukraine

Abstract

The paper proposes the effective method for approximating solutions of evolutionary partial differential equation. This solutions used for dynamic analysis of the process of water erosion in the zone of soil aeration in an artificial agrolandscape as in hydrotechnical system. The method presented in this paper is based on the basis of Galerkin method, Greville algorithm and limited inverse propagation for the evolutionary component. The neural network representation for the solution of the partial differential equation approximately satisfies the boundary conditions at each integration step. The convergence of the approximate solution to the exact one is performed. The advantage of the method is that it is easily applicable to solve the partial differential equation in arbitrary limited areas, provides a reduction in the dimensionality of space before training. In addition, it is more efficient in complexity and accuracy than the finite element method and Deep Galerkin Method. The implementation of this method is shown by the example of the process of propagation of Cs-137, the precipitation of which on the surface occurred instantly at a fixed point in time. The problem is considered for the zone of strips in agricultural use on ramparts-terraces of the hydrotechnical system. The results of test calculations of the dynamics of the pollution propagation for two-dimensional profiles for the layers of the soil for the deep 0-20 cm are presented. These results can be considered as corresponding to the available experimental value of the pollutant concentration.

Keywords [1]

Dynamic process, differential model, contamination, neural networks, pseudoinverse, Galerkin Method, piecewise smooth surface

1. Introduction

Determining the concentration of various anthropogenic pollutants in the soil environment and analysis of their movement is an important modern task. To solve it, various models are used: from the simplest models, which are based on regression equations and do not give the dynamics of spatial distribution of concentration in a selected area, to spatially distributed and evolutionary differential models that require significant computational resources and are often quite difficult to implement. But it is the model in the form of equations in partial derivatives that show us the dynamics of pollution. As examples, we give studies with evolutionary models for one (horizontal or vertical) profile in [1, 2], with models that have exact analytical solutions of differential equations, as in [1-3], with models in which simplification is performed in boundary conditions as in [3, 4] or other simplifications as in [5].

Numerous scientific publications address various issues related to the spread of soil contamination. As is known, the implementation of differential spatially distributed mass transfer models involves the use of numerical methods, as in [6–10]. These are methods using finite-difference approximations (for example, [6, 11–12]). Due to the complexity of their numerical implementation and the problem of increasing measurability, scientific studies consider simplified boundary conditions in regular domains and spatial profiles of the smallest dimension, as well as difference models for which the answer to the stability and accuracy of the obtained approximations is known. The finite element method is popular in the application [10, 13, 14]. This method is common for approximating static models with arbitrary

Information Technology and Implementation (IT&I-2021), December 01–03, 2021, Kyiv, Ukraine
EMAIL: vergunova@hotmail.com
ORCID: 0000-0003-3052-9143

domain shape. The direction of construction and research of neural network approximations of differential equations in partial derivatives is actively developing [15–18]. This opened up the possibility of obtaining approximations of differential models of mass transfer (both static and evolutionary) and does not lead to too large dimensions. Nevertheless, the question of effective solution of many practical problems of mass transfer is still open today. The presented work is an extension of research related to the control of the spread of pollution in the surface layers of the soil and the process of water erosion in the agricultural landscape. Our study is part of a study aimed at solving the problem of water erosion control in agricultural areas because the analysis of water erosion dynamics in this agro-landscape can be reduced to considering the mass transfer of Cs-137 in the surface layers of a given area. The task considered in the paper contains diffusion parameters (the effective diffusion coefficient and the speed of directed diffusion movement), the values of which are determined empirically with the help of special studies. Their reference values are given for certain types of soils under some average conditions. In practice, the values of diffusion parameters change with the change of soil moisture and temperature, and therefore their real values for a particular point in space and time will be slightly different. This point is the cause of additional error in the obtained numerical results.

A neural network was used to solve the problem. It is constructed in a narrowed space and approximately satisfies the boundary conditions at each step of integration. Universal approximation properties of neural networks make it attractive to use them in our research.

2. Problem Statement and its Solution

Given the importance of effectively solving the problem of dynamic control of water erosion and analysis of the spread of surface contamination in agrolandscapes with arbitrary surface shapes, the paper considers the construction of an effective neural network method for solving a model problem based on convective diffusion equation (namely: numerical solution of the problem 137 in a limited area under the condition of instantaneous surface contamination at some initial time). The substance Cs-137 is a low-activity pollutant, the particles of which are mainly in solid form, transported with the movement of moisture, poorly supplied to plants. These properties allow you to use the dynamics of the distribution of this substance to detect the presence and analysis of the process of water erosion. For numerical simulation of the movement of radioactive contaminant Cs-137 in the surface layers of the soil, an artificial agro-landscape with a system of hydraulic rampart-terraces was selected with available experimental data on the distribution of contaminants in layers 0-40 cm deep for 11 years.

The use of difference methods and the finite element method leads to a large number of calculations with matrices of very large sizes, the construction of calculations by these methods involves the accumulation of time error, so neural network methods are becoming increasingly popular.

Therefore, the purpose of our work is to substantiate and develop an effective method for numerical simulation of class of evolution problems by giving a particular example of its application.

The main tasks in the presented work are the construction of an effective method for numerical modeling of the dynamics of pollutant distribution and substantiation of the correctness and convergence of the obtained method. The objects of research are: neural networks and Galerkin-type methods for approximating the dynamics of the pollution process, which is described by the partial differential equation in a limited spatial region with a piecewise smooth boundary under the condition of instantaneous surface contamination.

It is known that Cs-137 is a low-active substance that can be in the soil solution in only one state, is transmitted in the soil layers mainly in unchanged form (in solid form) with the movement of moisture. Due to the fact that the substance almost does not turn into a liquid form, it is poorly supplied to plants. Therefore, you can ignore the influence of vegetation on the process of mass transfer. We have that the migration of Cs-137 occurs convectively with the flow of moisture through the soil profile; the filtration flow is linear; soil saturated with water is isotropic, porous, uncompressed; the process of transfer to the soil solution is subject to Fick's law; the soil layers considered in the study do not intersect with the groundwater layers. Therefore, taking into account the substantiation of the monograph [19], we can use the diffusion approach, namely to use the convective diffusion equation when constructing a model equation in our study. For the practical implementation of the study, a system of hydraulic ramparts-terraces in the artificial agro-landscape of the soil protection system of contour-ameliorative agriculture in the Kyiv region near the village of Khodosiyivka was chosen (Figure 1). In this system, each shaft-terrace has a special construction (as in Figure 2) to prevent water erosion. In our study, only the surface layers of the agricultural strips of the shaft-terrace T3 are considered (Figure 1).

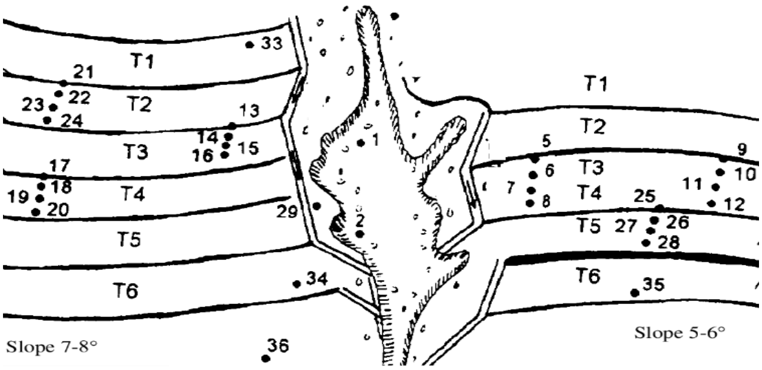

Figure 1: Scheme of agrolandscape (T1-T6 — the rampart-terraces, 1-36 — measuring points of Cs-137 concentration values

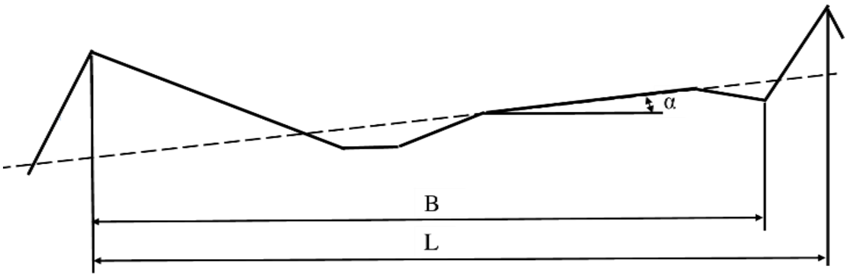

Figure 2: The cross-section of rampart-terrace surface Γ (B — the width of the strip of agricultural use, L — the distance between the ramparts, α — the angle of inclination of the surface of the segment)

In areas of agricultural use during tillage, the soil layers were not turned over (special tillage was used). Therefore, the impact of tillage on the distribution of pollutants did not occur and we can assume that the soil profile is uniform in depth and time. This was also the reason for using the diffusion approach. High difficulty in obtaining values of diffusion parameters in equations based on the convective approach are the reason for using the appropriate values given in reference sources for fixed soil types at average values of humidity, temperature and humus content.

In our model equation, we used reference values of parameters and refined (as a solution of the inverse problem by experimental measurements from study [20]). It is clear that the use of reference values of parameters reduces the accuracy of the calculated results and the best option is to use refined values. Nevertheless, the refined values are calculated from the available experimental measurements, which also contain error. We cannot estimate both types of errors and this is not part of the research objectives. However, their presence affected the obtained numerical results.

To achieve the goal of the study we will consider for the introduced differential model of the problem of controlling the dynamics of surface pollution of the agro-landscape and the available experimental values of pollutant concentration construction of the numerical method as an extension of Galerkin method for space component and Deep Galerkin Method for time component. The solution of the tasks allowed, based on the existing experience of using the Galerkin method, to narrow the space for the construction of neural networks and to obtain a new effective approach that has the properties of dynamism, convergence, adaptability. In our study we considered the surface soil layers of agricultural strips in the rampart-terrace T3 and points 13–16 with deep 0–40 cm in which the pollutant concentration was measured (Figure 1). The measurements of concentrations of contamination Cs-137 were carried out by the Institute of Agriculture of UAAS from May 1986 (immediately after contamination of this territory) to summer 1996 for all 36 points of agrolandscape in soil layers 0-40 cm deep. We selected for numerical simulation only the points located on agricultural strips of rampart-terrace T3 for soil layers 0–20 cm deep. Therefore,Cs-137 is an unaltered substance and mainly transferred by the movement of moisture in unchanged form in the soil and the assumptions set forth above take place. The fall of the pollutant on the agro-landscape surface of occurred instantly at an initial point in time and therefore the mathematical model as in [12] is valid:

$$Zu \equiv \frac{\partial u}{\partial t} + Lu \equiv \frac{\partial u}{\partial t} - \sum_{i=1}^{2} \frac{\partial}{\partial x_i}\left(D(x)\frac{\partial u}{\partial x_i}\right) + \sum_{i=1}^{2} V\frac{\partial u}{\partial x_i} = -\lambda u_0 \varphi(x)\delta(t-0), \quad (1)$$

$$u(x,0)=0,\ x\in\Omega, \tag{2}$$

$$\left.\left(-\sum_{i=1}^{2}D(x)\frac{\partial u}{\partial x_i}+Vu\right)\right|_{x\in\Gamma}=k(\Gamma)c_0q_0(x),\quad t\in[0,T],\quad \left.\left(-\sum_{i=1}^{2}D(x)\frac{\partial u}{\partial x_i}+Vu\right)\right|_{x\in\partial\Omega\setminus\Gamma}=0,\quad t\in[0,T], \tag{3}$$

in bounded area $\overline{Q}=\overline{\Omega}\times[0\le t\le T]$ $(\overline{\Omega}=\Omega\cup\partial\Omega,\Omega\subset R^2)$ with boundary $\partial\Omega$ and piece-smooth surface Γ. In (1)–(3) we denoted: $u(x,t)$ — the concentration of the pollutant in the point $x=(x_1,x_2)\in\overline{\Omega}$ at the moment t; λ — the half-decomposition factor, $\lambda = const > 0$; $\varphi(x_1,x_2)$ — the function describing the surface Γ in the area $\overline{Q}$ (the surface of the rampart-terrace), $\varphi(x_1,x_2)\in L_2(\Omega)$; u_0 — the contamination that has fallen to the surface Γ; $D(x)$ — the effective diffusion coefficient (it's an integral, limited, continuously differentiated function in the spatial region $\overline{Q}$); V — the speed of directed diffusion movement, $V = const > 0$; $k(\Gamma)=k\cdot\cos(\alpha(\Gamma))$, k — the coefficient of surface absorption, $k = const > 0$; $\alpha(\Gamma)$ — the slope of the surface Γ; q_0 — the flow of water coming from atmospheric precipitation with a contamination c_0; $\delta(t-0)$ — Dirac δ-function.

Since the precipitation in this area during the study period did not contain Cs-137 contamination, the boundary conditions (3) took the form:

$$\left.\left(-\sum_{i=1}^{2}D\frac{\partial u}{\partial x_i}+Vu\right)\right|_{x\in\partial\Omega}=0,\ t\in[0,T]. \tag{4}$$

We approximated the function $\varphi(x_1,x_2)$ as $\varphi(x_1,x_2)\cong 1-\left|x_1tg\alpha-x_2\right|/\sqrt{(x_1tg\alpha-x_2)^2+(0{,}01)^2}$ and Dirac δ-function by the hat function [21] (as it converges to the δ-function in the *-weak sense) in the form

$$\delta_\varepsilon(t-0)=\begin{cases}1-|t|/\tau, & |t|\le 2\tau,\\ 0, & |t|>2\tau,\end{cases}$$

where $\varepsilon=2\tau$.

We considered the set of twice continuously differentiated by x in $\overline{\Omega}$ and continuously differentiated by t in $[0,T]$ functions $u(x,t)$ that satisfy the initial-boundary conditions (2)–(3). The functions from this set are defined in space H with the norm

$$\|u\|_H=\left(\int_Q\left[\left(\frac{\partial u}{\partial t}\right)^2+\sum_{i=1}^{2}\left(\frac{\partial u}{\partial x_i}\right)^2\right]dQ\right)^{1/2}.$$

We used the spaces: H^* in which the functions $v(x,t)$ of the conjugate problem are defined; the negative space H^{*-} with the norm

$$\|f\|_{H^{*-}}=\sup_{v\ne 0}\frac{\left|(v,f)_{L_2(Q)}\right|}{\|v\|_{H^*}}.$$

Then the right part of equation (1) is a function from the space H^{*-} since $H_2^-(0,T)\otimes L_2(\Omega)\supseteq H^-(Q)$. Also, we have that the operator Z is not positively defined in the space $L_2(Q)$. But as shown in [12], the following inequalities are true: for any $u\in H$ $\|u\|_{L_2(Q)}\le\hat{c}\|u\|_H$, $\hat{c}=const>0$, $\|Zu\|_{H^{*-}}\le c\|u\|_H$ and $(Zu,u)_{L_2(Q)}\le c\|u\|_H^2$, $c=const>0$. These inequalities guarantee the existence of a unique generalized solution of problem (1)–(3) as function $u\in H$, which is proved in [12]. These facts are important for the construction and correct implementation of the numerical method for solving problems (1)–(3).

3. Method with neural network for the approximation of contamination values

Consider the cross section of the shaft-terrace and select the surface layer with a depth of H. Let us construct an approximate solution $\hat{u}$ with two components of the form:

$$\hat{u}(t,x)=\sum_{k=1}^{\tilde{n}}\tilde{\alpha}_k(t)\tilde{w}_k(x_1,x_2)+\sum_{l=1}^{n}\alpha_l(t)w_l(x_1,x_2),\ t\in[0,T],\ x=(x_1,x_2)\in\overline{\Omega}. \tag{5}$$

In (5) we have two sets of transfer functions defined in $\overline{\Omega}$. The first set $\{\tilde{w}_1(x), \tilde{w}_2(x),..., \tilde{w}_{\tilde{n}}(x)\}$ is used to satisfy the boundary conditions and these functions can be radial-basis functions. The second set $\{w_1(x), w_2(x),..., w_n(x)\}$ – to satisfy the differential equation (1) and these functions can be polynomials. For the first set we can take functions in the form [16]

$$\tilde{w}_k(x)=e^{-\gamma_{1k}(x_1-x_{1k})^2-\gamma_{2k}(x_2-x_{2k})^2}\left\{\frac{e^{[(x-x_k)^T\bar{n}_k]}-1}{e^{[(x-x_k)^T\bar{n}_k]}+1}\right\}. \tag{6}$$

In (6) we have denoted: $\bar{n}_k$ – unit vectors normal to the boundary $\partial\Omega$ at points $x_k=(x_{1k},x_{2k})$, [.] – scalar product of vectors, x_k – the collocation points on the boundary $\partial\Omega$ and the centers of the radial-basis functions, $k=1,...,\tilde{n}$. According to [16] we can take transfer functions of this kind, since these functions are near zero at its center along the boundary $\partial\Omega$ and has a derivative normal to the boundary that is equal to one. Let's write the boundary conditions (3) (or practical condition (4)) in abbreviated form:

$$Z_1u(x,t)\equiv\left(-\sum_{i=1}^{2}D\frac{\partial u}{\partial x_i}+Vu\right)\bigg|_{x\in\partial\Omega}=\phi(x,t)\cdot$$

Then, when evaluating equation (1) at a set of the training or collocation points at the boundary $\partial\Omega$, since the boundary condition is approximately satisfied at points x_k on the boundary $\partial\Omega$ we take into account [16] that $\tilde{\alpha}_k(t)=\tilde{Z}_1^+(\phi-Z_1\alpha_k(t))$, where $\tilde{Z}_1^+$ – right pseudoinverse matrices. More details, we have matrices for points x_k and for transfer functions $\tilde{w}_1(x), \tilde{w}_2(x),..., \tilde{w}_{\tilde{n}}(x)$ and $w_1(x), w_2(x),..., w_n(x)$, that is:

$$\tilde{Z}_1=\begin{pmatrix} Z_1(\tilde{w}_1(x))\big|_{x=x_1} & ... & Z_1(\tilde{w}_{\tilde{n}}(x))\big|_{x=x_1} \\ & ..... & \\ Z_1(\tilde{w}_1(x))\big|_{x=x_N} & ... & Z_1(\tilde{w}_{\tilde{n}}(x))\big|_{x=x_N} \end{pmatrix}, \tag{7}$$

$$Z_1=\begin{pmatrix} Z_1(w_1(x))\big|_{x=x_1} & ... & Z_1(w_n(x))\big|_{x=x_1} \\ & ..... & \\ Z_1(w_1(x))\big|_{x=x_N} & ... & Z_1(w_n(x))\big|_{x=x_N} \end{pmatrix}, \tag{8}$$

that is $\tilde{Z}_1(i,j)=Z_1(\tilde{w}_j(x))\big|_{x=(x_{1i},x_{2i})}$, $Z_1(i,j)=Z_1(w_j(x))\big|_{x=(x_{1i},x_{2i})}$, i =1, 2,…, N.

This fact makes it possible to write (5) as

$$\hat{u}(t,x)=\sum_{k=1}^{n}\left(w^T(x)-\tilde{w}^T(x)\tilde{Z}_1^+Z_1\right)\alpha_k(t)+\tilde{w}^T(x)\tilde{Z}_1^+\phi,\ t\in[0,T],\ x=(x_1,x_2)\in\overline{\Omega}, \tag{9}$$

in which, given (4), the last component is 0.

Then for approximate solution $\hat{u}$ in form (9) [14] takes place:

$$\left(\sum_{q=1}^{n}w_q(x)\frac{d\alpha_q(t)}{dt}-\sum_{i=1}^{2}\sum_{q=1}^{n}\frac{\partial}{\partial x_i}(D\frac{\partial w_q(x)}{\partial x_i})\alpha_q(t)+\sum_{i=1}^{2}\sum_{q=1}^{n}V\frac{\partial w_q(x)}{\partial x_i}\alpha_q(t)-f,w_k(x)\right)_{L_2(\Omega)}=0, \tag{10}$$

k =1, 2,…, n.

From here:

$$\sum_{q=1}^{n}\left(w_q(x),w_k(x)\right)_{L_2(\Omega)}\frac{d\alpha_q(t)}{dt}+ \\ +\sum_{q=1}^{n}\left(\sum_{i=1}^{2}[V\frac{\partial w_q(x)}{\partial x_i}-\frac{\partial}{\partial x_i}(D\frac{\partial w_q(x)}{\partial x_i})],w_k(x)\right)_{L_2(\Omega)}\alpha_q(t)-\left(f_\varepsilon,w_k\right)_{L_2(\Omega)}=0, \tag{11}$$

So, we come to the system of equations (11) on the set of collocation and learning points $x_k = (x_{1k}, x_{2k})$, in which $\alpha_q(t)$ are unknown and for q =1, 2,…, n:

$$\alpha_q(0) = 0 \,. \tag{12}$$

At the collocation points, we now consider the narrowed evolutionary problem (11), (12) to determine the unknown $\alpha_q(t)$. Now all coefficients in system of equations (11) are real numbers.

Denote by $b_{qk} = \left(w_q(x), w_k(x)\right)_{L_2(\Omega)}$, $c_{qk} = \left(\sum_{i=1}^{2}[V \frac{\partial w_q(x)}{\partial x_i} - \frac{\partial}{\partial x_i}(D \frac{\partial w_q(x)}{\partial x_i})], w_k(x)\right)_{L_2(\Omega)}$, $d_k = \left(f_\varepsilon, w_k(x)\right)_{L_2(\Omega)}$ then (11) for k =1, 2,…, n can be written as

$$A_k\alpha \equiv \sum_{q=1}^{n} b_{qk} \frac{d\alpha_q(t)}{dt} + \sum_{q=1}^{n} c_{qk}\alpha_q(t) = d_k \,, \tag{13}$$

To approximate unknown functions $\alpha_q(t)$ using neural networks (we proposed Deep Galerkin Method [20]), we minimize in L_2 the loss functional of the form:

$$\Im(\hat{\alpha}, \bar{\theta}) = \left\| A\hat{\alpha} - d \right\|^2_{L_2(0,T)} + \hat{\alpha}(0)^2 \,. \tag{14}$$

where

$$\hat{\alpha}(t;\theta) = \sum_{i=1}^{H} c_i e^{-\omega_i (t-\tau_i)^2} \,, \tag{15}$$

H – the number of hidden neurons in the layer, $\bar{\theta}$ – the vector of search parameters of neural networks, $\theta = \left(\omega_i\big|_{i=\overline{1,H}}, c_i\big|_{i=\overline{1,H}}, \tau_i\big|_{i=\overline{1,H}}\right) \in R^{2H}$. The loss functional (14) is a measure of how strongly the neural network $\hat{\alpha}$ satisfies the vector of required parameters and initial condition.

We randomly select points in $[0,T]$ as the current mini-batch [15]. Then calculate $\hat{\alpha}(t;\theta)$ from (15), calculate the loss functional (14) and take a descent step [15, 18, 20] by stochastic gradient descent method with the calculation of loss functional gradient. We repeat these steps until $\left\|\theta^h - \theta^{h-1}\right\| \le \varepsilon$.

So, at each step we calculate:

$$c_i^{(h)} = c_i^{(h-1)} - \lambda_{1i}^{(h-1)} \frac{\partial \Im(\hat{\alpha}^{(h-1)}, \theta^{(h-1)})}{\partial c_i^{(h-1)}}, \; i = \overline{1,H}, \tag{16}$$

$$w_i^{(h)} = w_i^{(h-1)} - \lambda_{2i}^{(h-1)} \frac{\partial \Im(\hat{\alpha}^{(h-1)}, \theta^{(h-1)})}{\partial w_i^{(h-1)}}, \; i = \overline{1,H}, \tag{17}$$

$$\tau_i^{(h)} = \tau_i^{(h-1)} - \lambda_{3i}^{(h-1)} \frac{\partial \Im(\hat{\alpha}^{(h-1)}, \theta^{(h-1)})}{\partial \tau_i^{(h-1)}}, \; i = \overline{1,H}. \tag{18}$$

If we write (15) in the form $\hat{\alpha}(t;\theta) = \sum_{i=1}^{H} F(S_i)$, where $F(S_i) = c_i e^{-\omega_i (t-\tau_i)^2}$, then $\frac{\partial \hat{\alpha}}{\partial t} = \sum_{i=1}^{H} 2F(S_i)\omega_i(\tau_i - t)$ and the loss functional (14) takes the following form:

$$\Im(\hat{\alpha}, \theta) = \left\| \sum_{i=1}^{H} \sum_{q=1}^{n} F(S_i)\left[2b_{qk}\omega_i(\tau_i - t) + c_{qk}\right] - d \right\|^2_{L_2(0,T)} + \hat{\alpha}(0)^2 \,. \tag{19}$$

In study [12] we showed that for problem (1)–(3) exist a unique generalized solution as function $u \in H$. The convergence of the approximate solution $\hat{u}$, provided that we already have the time component, can be obtained from [14].

The approximate solution $\hat{u}$ is such that $\hat{u}(x,t;\theta): R^3 \to R$ with $\hat{\alpha}(t;\theta): R^{Hn} \to R^n$ and all activation functions $F(S_i)$ are continuous, bounded and non-constant on $[0,T]$. Let for set of n neural networks we have: $\hat{\alpha}^H = \left\{ \hat{\alpha}: R^{Hn} \to R^n \,\middle|\, \hat{\alpha}(t;\theta) = \sum_{i=1}^{H} F(S_i) \right\}$, $\hat{\alpha}_3 = \bigcup_H \hat{\alpha}^H$. The approximate solution $\hat{\alpha}$ can satisfy the differential operator A on $[0,T]$.

All n functions in $\hat{\alpha}(t;\theta)$ belongs to $C^{\infty}(R)$. Then, given [22] $\hat{\alpha}_3$ is dense in $C([0,T])$. Moreover, using the results in [22] we have that: all activation functions $F(S_i) \in C^{\infty}(R)$ are non-constant, continuous and bounded, then $\hat{\alpha}_3$ is uniformly dense in $C^{\infty}([0,T])$ and dense $C^{1,2}([0,T])$; all activation functions $F(S_i) \in C^{\infty}(R)$ are non-constant, continuous, bounded and all its derivatives are bounded, then $\hat{\alpha}_3$ is uniformly dense in $C^{\infty}([0,T])$ and dense in $L_2(0,T)$ and in $W_2^1(0,\mathrm{T})$.

Then from the neural network approximation theorem for the class of quasilinear parabolic partial differential equations in study [15] and from the above facts it follows that for a class of neural networks with a single hidden layer and n hidden neurons that minimize the loss functional $\Im(\hat{\alpha},\theta)$ there is a sequence of neural networks $\hat{\alpha}^H(t;\theta)$ such that $\Im(\hat{\alpha},\theta) \to 0$, $H \to \infty$, and $\hat{\alpha}^H(t;\theta) \to \hat{\alpha}$, $H \to \infty$, in $L_2(0,T)$ in the approximation sense.

4. Results and Discussion

To study the dynamics of pollution of the surface layers of the landscape rampart-terrace T3 with slope α = 7–8° (experimental points 13–16) was selected (Figure 1). Measurement of Cs-137 content in the surface layers of the soil for rampart-terrace T3 was performed at the Institute of Agriculture UAAS (Ukraine). The use of plowless tillage in this area did not cause soil mixing, which allowed to neglect the process of migration of Cs-137 down the soil profile. Given the aspects of the constructed complex of rampart-terrace for the analysis of the presence and dynamics of water erosion of the terrace shaft T3, we have the necessary data on the spatial distribution of pollution in the selected area. We used the results of experimental measurements of Cs-137 concentration performed in 1986–1998 for the system of rampart-terraces for selected points in a layer 0–20 cm deep on the strip of agricultural use of the terrace. The fact of binding of Cs-137 particles with the upper soil particles and the fact that the adsorbed Cs-137 is transferred with the solid phase particles of the soil from drainage sites to soil deposition during water erosion process allowed comparing the Cs-137 concentration for the stored matrices of spatial values in several moments. These comparisons was the main basis for detecting the process of water erosion in the soil layers. The study considered the cross section (two-dimensional profile) of the rampart-terrace T3, although the proposed technique allows to consider arbitrary ramparti-terraces of the hydraulic system.

To obtain values of Cs-137 concentration, test calculations were performed. The initial approximations are: $w_i^{(0)} \in [-0{,}05; 0{,}05]$, $\lambda_Q^{(0)} \in [0{,}01; 0{,}05]$, $c_i^{(0)} \in [0{,}1; 0{,}5]$. For the condition of the stop we took: $\varepsilon = 1{,}0 \cdot 10^{-10}$. The value of effective diffusion coefficient is $D = 4{,}7 \cdot 10^{-8}$; the value of the speed of directed diffusion movement $V = 1{,}0 \cdot 10^{-6}$. We had the following set of experimental points to generate mini-batches: at the initial moment (t = 0) – 28 points; on $[0,T]$ – 308 points.

The parameters (learning rate) $\lambda_{1i}^{(h)}$, $\lambda_{2i}^{(h)}$, $\lambda_{3i}^{(h)}$, $i = \overline{1,H}$, does not decrease with increasing n. Given the limited size of the set of experimental points, the set of collocation points at the boundary of the region and only 12 points of time in the consideration we had for test calculations: $\tilde{n} < n \leq 21$, $H \leq 7$.

The transfer functions $w_1(x), w_2(x), \ldots, w_n(x)$ defined as: $w_j(x) = x_1^{m_{1j}} x_2^{m_{2j}}$, $m_{1j}, m_{2j} = \overline{0{,}14}$ n = 21. The transfer functions $\tilde{w}_1(x), \tilde{w}_2(x), \ldots, \tilde{w}_{\tilde{n}}(x)$, $\tilde{n}$ = 17, α = 7°, for strip on T3 we have in the form:

$$\tilde{w}_k(x) = e^{-\gamma_{1k}(x_1 - x_{1k})^2 - \gamma_{2k}(x_2 - x_{2k})^2} \left\{ 1 - \frac{2}{e^{[\sin\alpha(x_1 - x_{1k}) - \cos\alpha(x_2 - x_{2k})]}} \right\}.$$

Recurrent formulas of Greville algorithm were used to obtain pseudoinverse matrices in points x_k and for transfer functions $\tilde{w}_1(x), \tilde{w}_2(x), \ldots, \tilde{w}_{\tilde{n}}(x)$. For layers 0–5, 5–10, 10–20 cm for the terrace T3 with slope 7° numerical calculations were performed for the cross-sectional points of the shaft-terrace, which correspond to the experimental measurement points. But due to the limited number of experimental points, there was a strong limitation in our realization. The results of test calculations showed a decrease in the values of pollution for the points of the layer 0–5 cm, a gradual increase in the values of pollution for the points of the lower layers in the first years and a decrease in the following years. This corresponded to the results of experimental measurements and we can conclude about the low level of water erosion in the middle and lower parts of the terrace T3. As a confirmation of this fact, Figure 3 presents graphs of changes in the values of Cs-137 contamination for the middle part of the terrace 3 (point 15) at a depth of 1 cm and 6 cm for the entire study period, and Table 1 shows the corresponding experimental values.

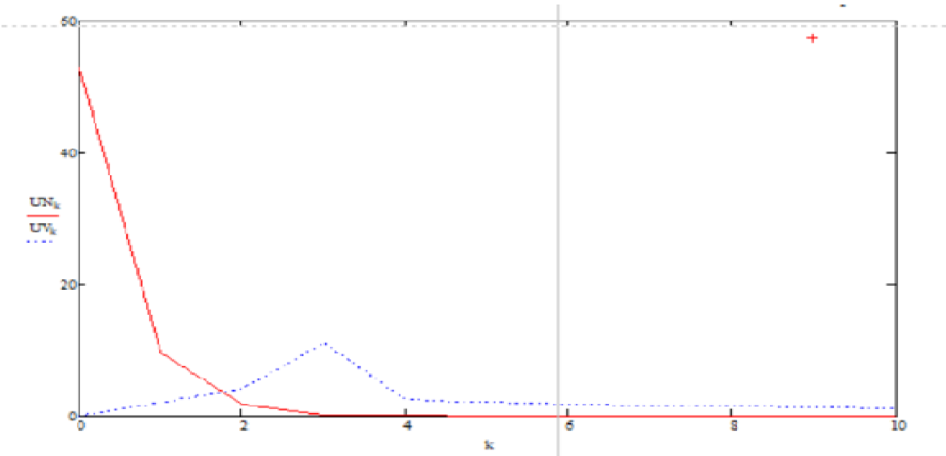

Figure 3: Cs-137 pollution in 1986-1996, terrace T3, depth – 1 cm (UN), 6 cm (UV), Curie/km^2

Table 1
Cs-137 pollution at the middle part of terrace T3, depth – 1-5 cm, Curie/km^2 (experimental data)

Year	Layer, cm	Contamination
1986	0	0,60
1986	5-10	0,00
1989	0-5	0,27
1989	5-10	0,11
1990	5-10	0,04

5. Conclusion

As a result of research, the method for approximation of decisions in the problem of dynamic control of distribution of pollution in a two-dimensional bounded area with a piece-smooth surface is offered. The constructed method is a method that combines the use of Galerkin method and the construction of neural networks. We offered for the introduced differential model for this problem the numerical method as an extension of Galerkin method for space component and Deep Galerkin Method for time component. This approach allowed, based on the existing experience of using the Galerkin method, to narrow the space for the construction of neural networks and to obtain a new effective method with the properties of dynamism, convergence, adaptability. The convergence of the approximate solution obtained by the proposed method to the exact solution of our problem is shown. A test example of the implementation of the developed approach for agrolandscape with a system of hydraulic shafts-terraces, which received surface contamination of Cs-137 at a fixed initial time, is given.

6. Acknowledgements

I express my deep gratitude to National Research Center "Institute of Agriculture of the National Academy of Agrarian Sciences of Ukraine" for the opportunity to participate in research projects: "Control of radioactive contamination of soil and plants"; “Agri-environmental monitoring and modeling of sustainable agro-landscapes and agroecosystems”.

7. References

[1] V. G. Danilov, V. P. Maslov, K. A. Volosov, Mathematical Modelling of Heat and Mass Transfer Processes, Springer Science & Business Media, 2012.

[2] Handbook of Chemical Mass Transport in the Environment Ed. by L. J. Thibodeaux and D. Mackay, CRC Press, 2011.
[3] Heat and Mass Transfer – Modeling and Simulation. Ed. by M. Hossain, InTech, 2011.
[4] A. T. Barabanov, Principles of adaptive-landscape generation and development of soil protection agricultural systems, Geography and Natural Resources **37**(2) (2016) 106–113. doi:10.1134/S1875372816020037.
[5] L. D. Romanchuck, T. P. Fedonyuk, R. G. Fedonyuk, Model of influence of landscape vegetation on mass transfer processes, Biosyst. Divers. **25**(3) (2017) 203–209. doi:10.15421/011731.
[6] S. P. Kundas, I. A. Gishkelyuk, V. I. Kovalenko, O. S. Khilko, Kompyuternoye modelirovaniye migratsii zagryaznyayushchikh veshchestv v prirodnykh dispersnykh sredakh, MGEU im. A.D. Sakharova, 2011. (In Russian).
[7] E. V. Shein, Matematicheskoye modelirovaniye v pochvovedenii, Marakushev, 2016. (In Russian).
[8] G. P. Brovka, I. N. Dorozhok, S. N. Ivanov, Raschetnyye skhemy protsessov konvektivno-diffuzionnogo perenosa vodorastvorimykh soyedineniy, Inzhenerno-fizicheskiy zhurnal **83**(5) (2010) 866–872. (In Russian).
[9] S. B. Patil, H. S. Chore, Contaminant transport through porous media: An overview of experimental and numerical studies, Advances in Environmental Research **3**(1) (2014) 45–69. doi:10.12989/aer.2014.3.1.045.
[10] Y. Zhang, Q. Wang, S. T. Zhang, Numerical simulation of benzene in soil contaminant transport by Finite Difference Method, Advanced Materials Research, 414 (2012) 156–160.
[11] T. V. Boiko, A. O. Abramova, Yu. A. Zaporozhets, Matematychne modeliuvannia mihratsii zabrudniuiuchykh rechovyn u gruntakh, Vostochno-Evropeyskiy zhurnal peredovyih tehnologiy **6**(4) (2013) 14–16. (In Ukrainian).
[12] I. Vergunova, Computational method for the analysis of dissemination of surface pollution in hydrotechnical ramparts, EUREKA: Physics and Engineering, 5 (2018) 38–55. doi:10.21303/2461-4262.2018.00711.
[13] X. Li, S. Cescotto, H. R. Thomas Finite-Element Method for contaminant transport in unsaturated soils, Journal of hydrologic engineering 7 (1999) 265-274. doi: 10.1061/(ASCE)1084-0699(1999)4:3(265).
[14] I. Vergunova, The System Dynamic Control of the Water Erosion in a Hydrotechnical Rampart-Terrace, in: Proceedings of 2019 IEEE International Conference on Advanced Trends in Information Theory (IEEE ATIT 2019), Kyiv, 2019, pp. 344–348. doi:10.1109/ATIT49449.2019.9030434.
[15] J. Sirignano, K. Spiliopoulos, DGM: A deep learning algorithm for solving partial differential equations, Journal of Computational Physics 375 (2018) 1339–1364. doi:10.1016/j.jcp.2018.08.029.
[16] K. Rudd, S. Ferrari, A constrained integration (CINT) approach to solving partial differential equations using artificial neural networks, Neurocomputing 155 (2015) 277–285. doi:10.1016/j.neucom.2014.11.058.
[17] A. N. Vasil'ev, F. V. Porubaev, D. A. Tarhov, Nejrosetevoj podhod k resheniyu nekorrektnyh zadach teploperenosa, Nauchno-tekhn. vedom. SPGPU. Informatika, telekommunikacii i upravlenie 2 (2011) 133–142. (In Russian).
[18] J. Chen, R. Du, K. Wu, A Comparison Study of Deep Galerkin Method and Deep Ritz Method for Elliptic Problems with Different Boundary Conditions, Commun. Math. Res. **36**(3) (2020) 354–376. doi:10.4208/cmr.2020-0051.
[19] V. M. Prokhorov, Migracziya radioaktivnykh zagryaznenij v pochvakh. Fiziko-khimicheskie mekhanizmy i modelirovanie, Energoizdat, 1981. (In Russian).
[20] I. Vergunova, Vergunov V., Rosemann I. Solving the Coefficient Inverse Problem by the Deep Galerkin Method, in: Proceedings of 2021 IEEE Int. Conf. on Advanced Computer Information Technologies (IEEE ACIT 2021), Deggendorf, Germany, 2021, pp. 65–70. doi:10.1109/ACIT52158.2021.9548633.
[21] B. Hosseinia, N. Nigama, J. M. Stockiea, On regularizations of the Dirac delta distribution, Journal of Computational Physics 305 (2016) 423–447. doi:10.1016/j.jcp.2015.10.054.
[22] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neural Networks 4 (1991) 251–257.