Neural network modeling of air pollution in tunnels according to indirect measurements

T Kaverzneva, T Lazovskaya, D Tarkhov and A Vasilyev

Peter the Great St-Petersburg Polytechnic University, 29 Politechnicheskaya Str, 195251 Saint-Petersburg, Russia

E-mail: a.n.vasilyev@gmail.com

Abstract.
The article deals with the problem of providing the necessary parameters of air of the working area in dead-end tunnels in the case of ventilation systems powered off. An ill-posed initial-boundary problem for the diffusion equation is used as a mathematical model for a description and analysis of mass transfer processes in the tunnel. The neural network approach is applied to construct an approximate solution (regularization) of the identification problem in the case of the approximate measurement data and the set of interval parameters of the modeled system. Two types of model measurements included binary data are considered. The direct problem solution and the inverse problem regularization for the offered neural network approach are constructed uniformly.

1. Introduction

Air pollution of the human environment reduces the duration of his life, leading to a variety of serious diseases. In his professional activity, a person is exposed to a greater number of harmful factors that constantly requires the development of additional security measures, which can not be carried out without research and theoretical work.

During the construction of underground structures, tunnels for different purposes (for example, underground roads and railways) one of the most important tasks to ensure the safety of the operations is to control the parameters of the air environment, which may be contaminated sources of different origin. Normalization of air mainly achieved using rational ventilation system. However, during the construction and operation may be necessary to conduct a number of studies, for example, repair, in the absence of ventilation or disable it.

2. Problem statement

The paper goes on to study the problems with a nonstandard statement [1-6]. Here we construct a mathematical model for the dissemination of harmful substances from the source, which is located in the depths of the tunnel. This model allows for the concentration of harmful substances at the outlet to determine the distribution of dangerous concentration zones for such substances through the tunnel.

The problem of identification arises when environmental forecasting the state of air quality in the tunnel; a diffusion equation in a moving medium (and some pieces of heterogeneous information about a concentration of the pollutant) is the basis of a mathematical model:

Let the pollutant concentration $\Phi = \Phi(x,t)$ in the area $I = [0; L] \subset R$ at the time $t \in (0; T)$ satisfies the following initial-boundary-value problem

$$\begin{cases} \frac{\partial \Phi}{\partial t} + u\frac{\partial \Phi}{\partial x} - \sigma\frac{\partial^2 \Phi}{\partial^2 x} + \tau\Phi = P, \\ \Phi(x,0) = \Phi_0(x), \\ \Phi(0,t) = \Phi_1(t), \end{cases} \tag{1}$$

here the function $P(x,t) = \sum_{s=1}^{k} p_s(t)\delta_s(x; x_s)$ characterizes the power of pollution sources, $p_s(t)$ – the power of the s-th contamination source distributed in a neighborhood of x_s with a density $\delta_s(x; x_s)$, u – the speed, σ – the coefficient of turbulent diffusion, $\tau > 0$ – the parameter that determines the intensity of the absorption of the pollutant due to its entrainment, deposition, chemical reactions, etc.

It is expected that there is no initial contamination throughout the tunnel $\Phi_0(x) = 0$, $0 \le x \le L$. No contamination is at the entrance to the tunnel $\Phi_1(t) = 0$, $0 \le t \le T$.

By the replacement of variables $\Phi(x,t) = \exp(ax+bt)U(x,t)$, $P(x,t) = \exp(ax+bt)Q(x,t)$, where $a = u/2\sigma$, $b = -\tau - u^2/4\sigma$, the differential equation can be transformed into the heat equation

$$\frac{\partial U}{\partial t} - \sigma\frac{\partial^2 U}{\partial x^2} = Q \tag{2}$$

with appropriate initial and boundary conditions on the function U

$$U(x,0) = U_0(x) = 0, \ \ U(0,t) = U_1(t) = 0. \tag{3}$$

Note that the assignment of boundary conditions of the form $(\frac{\partial}{\partial x}\Phi - \frac{u}{2\sigma}\Phi)(L,t) = 0$ at the end of the tunnel results in the function U to the homogeneous Neumann condition $(\frac{\partial}{\partial x}U)(L,t) = 0$.

Next, we will consider the transformed task for the function U. The coordinates x, t are normalized so that $L = 1$, $T = 1$. The problems with the proposed statement (as power sources and part of the boundary conditions at the end of the tunnel are not listed) are ill-posed - they are in need of additional data.

Suppose that, instead of the missing data conditions, we know some experimental observations

$$U(\tilde{x}_j, \tilde{t}_j) = \varphi_j, \ j = 1, \ldots, m, \tag{4}$$

obtained with a set of sensors, $\{(\tilde{x}_j, \tilde{t}_j)\}_{j=1}^{m} \in [0;1]\times[0;1]$. The challenge is to find an approximate solution to the problem – the function U and the restoration of the source – the function Q as well. Another option of setting the identification problem is the restoration of the initial conditions.

3. Neural network approach

The approach proposed [1-4] allow us to bring together pieces of heterogeneous information about the system in the neural network model, use the regularizing properties of neural networks for solving inverse and ill-posed problems.

An approximate solution of the identification problem can be found in the form of a system output of two artificial neural networks of a given architecture

$$U(x,t;\sigma) = \sum_{i=1}^{N_1} c_i g_i(x,t;\sigma,\mathbf{a}_i), \tag{5}$$

$$Q(x,t) = \sum_{i=1}^{N_2} d_i h_i(x,t;\mathbf{b}_i), \tag{6}$$

which weights – linearly incoming parameters c_i, d_i and nonlinear input parameters $\mathbf{a}_i$, $\mathbf{b}_i$ – be determined in the process of gradual learning network based on the minimization of the error functional given in a discrete form

$$J = \sum_{j=1}^{M} (\frac{\partial U}{\partial t} - \sigma \frac{\partial^2 U}{\partial x^2} - Q)^2 (x_j, t_j) + \lambda_0 \sum_{j=1}^{M_0} U^2(x_j, 0) + \lambda_1 \sum_{j=1}^{M_1} U^2(0, t_j) + \lambda \sum_{j=1}^{m} (U(\tilde{x}_j, \tilde{t}_j) - \varphi_j)^2. \quad (7)$$

Here $\{(x_j, t_j)\}_{j=1}^{M}$ – periodically regenerated sampling points in $[0;1] \times [0;1]$, $\{(x_j, 0)\}_{j=1}^{M_0}$, $\{(0, t_j)\}_{j=1}^{M_1}$, – sampling points on the border areas; $\lambda_k > 0$ – positive penalty coefficients.

In this problem of modeling air pollution in tunnels, the assessment of the concentration in the main part of the tunnel is more important than identifying the source, so we can find the neural network approximation to the function U minimizing the error functional J without the source Q in the first term. In this case, the first sum is taken over the sampling points of the part of the variable region such that it is known that $Q = 0$.

4. Measurement data modeling

We found some analytical solutions of the problem (for the given $Q(x,t)$ and the boundary conditions added at the end of the tunnel); these solutions were used when setting the "experimental measurements" φ_j. We choose, for example, nonhomogeneous Neumann's condition (linear on t increase in derivative) as an additional one $\partial_x U(1,t) = t$, then the approximate model solution is given by [7]

$$\tilde{U}(x,t) \approx tx - \frac{x}{2} + \frac{x^3}{6} + \sum_{n=0}^{100} \frac{\sin \frac{\pi(2n+1)x}{2}}{(2n+1)^4} (-1)^n \exp \frac{-\pi^2 (2n+1)^2 t}{4}. \quad (8)$$

Using $\tilde{U}$ we generate the data of observations $\tilde{U}(x_*, \tilde{t}_j) = \varphi_j$, $j = 1, \ldots, m$, taken near the exit of the tunnel $x_* = 0.1$. Also, this function is used to check the result of solving the problem.

For modeling are offered two cases. If the special device measures data and captures the accurate model values we use the values of the function $\tilde{U}(x,t)$ at equidistant points $\tilde{t}_i$ of the time interval at fixed x_*

$$G_1(t_i) = \tilde{U}(x_*, \tilde{t}_i), \quad (9)$$

In another case, the binary data is created. We consider a concentration above the critical level as the availability and lower as the absence of harmful emissions in the air. So this data has a certain error. The data generation held by the next rule

$$G_2(t_i) = \begin{cases} \beta, \ \tilde{U}(x_*, t_i) + \varepsilon(2\xi - 1) > \alpha; \\ 0, \ \tilde{U}(x_*, t_i) + \varepsilon(2\xi - 1) \le \alpha, \end{cases} \quad (10)$$

where ε is a measurement error and can take the big values, ξ is a standard uniformly distributed random value, α is a sensitivity threshold, $t_i \in [0,1]$, $\beta = 2\alpha$.

In the second case we considered a condition corresponded to the impermeability of the right end of the tunnel. The allocation of constant intensity is on the right end of the tunnel at the part of $d =$ length.

Then the approximate model solution is given by [7]

$$\hat{U}(x,t) = \frac{16}{\pi^3} \sum_{n=0}^{+\infty} \frac{\sin \frac{\pi(2n+1)x}{2} \sin \frac{\pi(2n+1)d}{2}}{(2n+1)^3} \cdot (-1)^n (1 - \exp \frac{-\pi^2 (2n+1)^2 t}{4}), \quad (11)$$

where $d = 0.1$.

Journal of Physics: Conference Series **772** (2016) 012035 doi:10.1088/1742-6596/772/1/012035

5. Calculations

We construct a neural network model for our task on all available data. This is the equation (2), the boundary conditions (3), and measurement data (9)–(10). The selection of the model parameters (the weights of the neural network) is based on the error functional minimizing. The error functional optimization is conducted by RProp and Particle Swarm algorithms [9]. As basic elements of the neural network g_i, different types and number of functions were applied: Gaussians, sigmoids with activation function $\mathrm{th}(\cdot)$, etc.

5.1. *Exact Data*

Here we present results obtained by the neural network solution in the case of 10 neuro elements and two types of functions for the accurate "measurements" (9). The first type – sigmoid functions

$$u(x,t,x_c,t_c,a,b,c,d) = c(\mathrm{th}(a(x-x_c)+b(t-t_c))+d),$$

where x_c, t_c, a, b, c, d are the parameters (weights) of the neural network. As the result, we have the perceptron network with one hidden layer.

The basis functions of the second type are the solutions of the equation (2)

$$u(x,t,x_c,t_c,a,c) = c(\frac{1}{\sqrt{t-t_c}}\exp\frac{-0.25(x-x_c)^2}{t-t_c}+a). \tag{12}$$

The given error functional J (7) takes the form

$$J_0 = \lambda_0\sum_{j=1}^{M_0}U^2(x_j,0)+\lambda_1\sum_{j=1}^{M_1}U^2(0,t_j)+\lambda\sum_{j=1}^{m}(U(x_*,\tilde{t}_j)-G_1(\tilde{t}_j))^2. \tag{13}$$

The counterplot of the neural network solution and the real concentration function is presented in Figure 1 (the perceptron network) and Figure 2 (the special network) at the level of the harmful substances concentration value 0.05. In this model, this level acts as an example of critical exceeding of which is dangerous to humans.

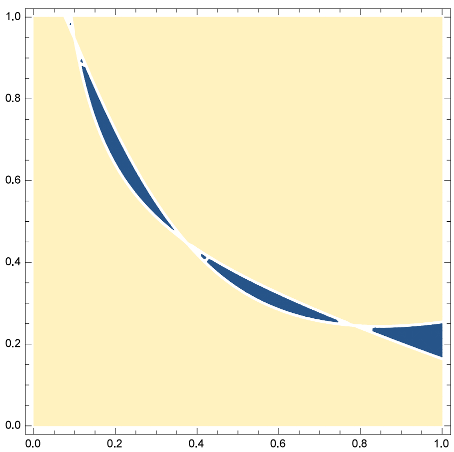

Figure 1. The counterplot of the neural network solution and the model concentration function $\tilde{U}$ of level 0.05. The perceptron network with $n = 10$ neuro elements.

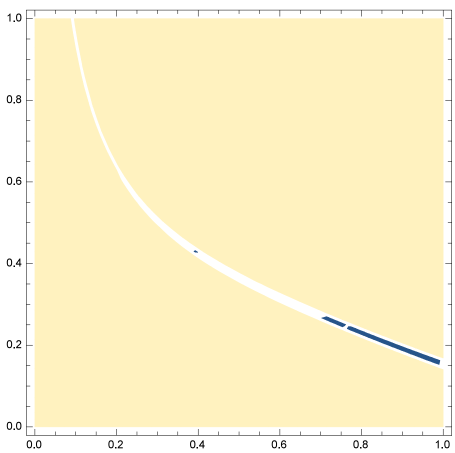

Figure 2. The counterplot of the neural network solution and the model concentration function $\tilde{U}$ of level 0.05. The neural network with $n = 10$ basis neuro elements. The basis functions used are fundamental solutions (12) of a diffusion equation

The plots represent the elevated concentration prevalence through the tunnel from the source ($x = 1$) towards the entrance ($x = 0$) over time ($0 \leq t \leq 1$). As expected, the special features provide a more accurate solution (mismatched area marked by dark color).

Note, that with an increase in the number of neurons up to 20 the universal perceptron network provides a good approximation of a function $\tilde{U}(x, t)$ in an area near the exit of the tunnel. But there is an accumulation of error at the far end of the tunnel.

In the case of the model concentration function $\hat{U}$ (11), we have obtained a good neural network solution using fundamental solutions (12) of a diffusion equation only.

5.2. Binary Data

Neural network solution of considered problem is constructed for the data generated by the formula (10) at the parameters values $\varepsilon = \alpha = 0.03$ and $\beta = 2\alpha$. Note that in evaluating the results it must be taken into account the character of the data and the selected error.

A good result was achieved by the neural network with $n = 5$ special basis functions (12). Figure 3 illustrates the approximation of the concentration function $\tilde{U}(x, t)$ in the measuring point ($x_* = 0.1$).

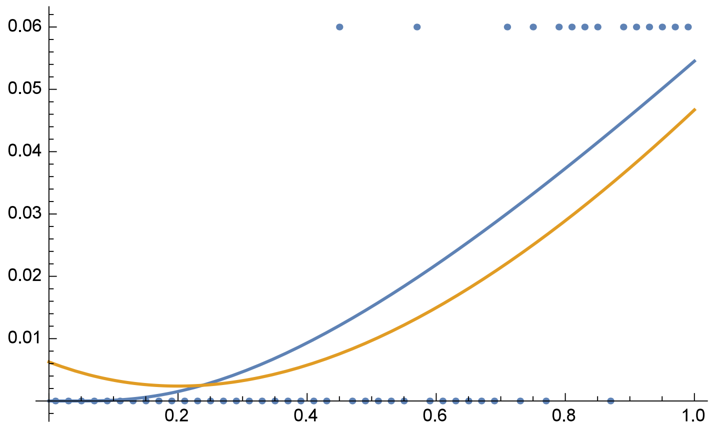

Figure 3. The plot of the neural network solution, the concentration function $\tilde{U}$ at $x_* = 0.1$ and the point data. The special network with $n = 10$ neuro elements.

The approximation of the concentration function $\hat{U}(x, t)$ (11) has a similar appearance.

6. Conclusions

The approximate neural network solutions of the considered problems are constructed. Computing experiment showed only a slight discrepancy between the boundaries of areas in excess of the maximum permissible concentration of harmful substances, obtained by the neural network solution in the case of 10 neuro elements and modeling solution.

As basic elements of the neural network, different types and number of functions were applied, but it is preferable to use the functions that in advance satisfy the diffusion equation, then the first term can be eliminated from the error functional. This approach is less versatile but allows us to get a more accurate solution with a smaller number of basic functions. The insignificant modification of methods described [2, 4] allows considering the case of interval specified parameters.

The statement of problems discussed in our work includes the differential equations. This fact causes the lack of overfitting. Usually, overfitting occurs when it is training a neural network on the finite set of input data in a situation when the capacity of this set does not exceed the number of weights. As a rule, the problem is solved by the division of a data set for training and test sets and the selection of the network structure to minimize the error on the test set. Differential equations are valid on some infinite set of points (in some domain). Therefore, the set of points from which training and test sets are selected is infinite. It makes possible the regeneration of the training and test sets and dimensional changes to eliminate overfitting.

In our problem, the number of times at which the measurements are made, finite and the points, from which we can select a test set, form a square $[0;1] \times [0;1]$, that is, an infinite set.

We are going to consider the possibility of creation forecast of harmful substances spread in a tunnel from any source and definition of zones with dangerous concentrations for protection of the workers who are in a tunnel from the influence of these pollutants.

Acknowledgments

The work is supported by the Russian Foundation for Basic Research (Grant no. 14-01-00660).

References

[1] Kainov N, Tarkhov D and Shemyakina T 2014 Application of Neural Network Modeling to Identification and Prediction Problems in Ecology Data Analysis for Metallurgy and Welding Industry *Nonlinear Phenomena in Complex Systems* **17** 57-63

[2] Lazovskaya T and Tarkhov D 2015 Fresh approaches to the construction of parameterized neural network solutions of a stiff differential equation *St. Petersburg Polytechnical University Journal: Physics and Mathematics* http://dx.doi.org/10.1016/j.spjpm.2015.07.005

[3] Vasilyev A, Tarkhov D and Guschin G 2004 Neural Networks Method in Pressure Gauge Modeling *Proc. of the 10th IMEKO TC7 Int. Symp. on Advances in Measurement Science (Saint-Petersburg, Russia vol 2)* 275-279

[4] Vasilyev A and Tarkhov D 2014 Mathematical Models of Complex Systems on the Basis of Artificial Neural Networks *Nonlinear Phenomena in Complex Systems* **17** 327 35

[5] Tarkhov D and Vasilyev A 2005 New neural network technique to the numerical solution of mathematical physics problems. I: Simple problems *Optical Memory and Neural Networks (Information Optics)* **14** 59-72

[6] Tarkhov D and Vasilyev A 2005 New neural network technique to the numerical solution of mathematical physics problems. II: Complicated and nonstandard problems *Optical Memory and Neural Networks (Information Optics)* **14** 97-122

[7] Polyanin A and Nazaikinskii V 2016 *Handbook of Linear Partial Differential Equations for Engineers and Scientists* (Chapman and Hall/CRC) p 1643

[8] Haykin S 2008 *Neural Networks and Learning Machines* (Prentice Hall) p 936

[9] Riedmiller M and Braun H 1993 *A direct adaptive method for faster backpropagation learning: The Rprop algorithm. Proceedings of the IEEE International Conference on Neural Networks* (IEEE Press) 586-591