PAPER • OPEN ACCESS

Improving the resilience of neural network solution of inverse problems in Raman spectroscopy of multi-component solutions of inorganic compounds to the distortions caused by frequency shift of the spectral channels

View the article online for updates and enhancements.

Improving the resilience of neural network solution of inverse problems in Raman spectroscopy of multi-component solutions of inorganic compounds to the distortions caused by frequency shift of the spectral channels

I V Isaev[1], S A Burikov[1], T A Dolenko[1], K A Laptinskiy[1] and S A Dolenko[1]

[1]M.V. Lomonosov Moscow State University, Leninskie gory 1/2, Moscow, Russia, 119991

e-mail: isaev_igor@mail.ru

Abstract. In this study, we considered the problem of determining the concentrations of ions dissolved in water by the spectra of Raman scattering of light. At the moment, there are no adequate mathematical models describing the studied object, so in fact the only way to solve this problem is use of machine learning methods based on experimental data. As any data resulting from experimental measurements contains noise, there is a need to develop specific approaches to improving the resilience of the solution to noise in the data. Regarding the studied problem, experimental data may contain distortions of three types: variations in the concentrations of ions, error in the determination of the intensity in the channels of a spectrum, and frequency shift of the channels of a spectrum. This study is devoted to the development of approaches to improve the resilience of the neural network solution to the distortions caused by the shift of the spectral channels.

1. Introduction

In various fields of industry and ecology there is a need for diagnostics of individual components in multicomponent water mixtures, for example, a need to identify and determine the concentration of each dissolved ion in technological and waste waters, when testing the composition of drinks, in monitoring of natural waters [1]. For successful solution of such tasks it is necessary to develop new methods that should be: a) express - provide information in real time; b) non-contact – allow obtaining information remotely, without "interference" with the medium under test; c) have high selectivity in relation to each component of a complex mixture. Traditional chemical (analytical) methods provide a high enough accuracy of determining the concentration of ions – from a few hundredths to units of μg/l [2, 3]. However, these methods are contact, they are individual for each ion, their implementation requires a long time, good sample preparation and consumption of expensive reagents [2, 3].

In [4-7], it was proposed to use the method of Raman scattering spectroscopy to meet the requirements for the method to be express and non-contact. The shape of Raman spectrum of a solution is highly sensitive to changes in its ionic composition and to the concentrations of ions present in the solution [5, 7-10]. However, at the moment there is no adequate mathematical model describing these changes in the shape of Raman spectra. Therefore, to solve the problem of determining the concentrations of ions by the shape of Raman spectra, artificial neural networks

(ANN) were applied. ANN have the ability to learn by example without requiring any *a priori* information about the form of the sought-for dependence, so are widely used for pattern recognition [11-13]. The efficiency of their use for the solution of inverse problems in Raman spectroscopy was shown in [14-18].

The described problem belongs to the class of inverse problems (IP), which have some peculiarities complicating their solution. The considered IP, like most other IPs, is incorrect or at least ill-conditioned, which makes the solution sensitive to the presence of noise in the input data. Despite the fact that neural networks by themselves have the ability to work with noisy data, this is often insufficient when solving inverse problems, because the incorrectness of the problem "outweighs" the ability of the network. Development of the approaches to increase the resilience of methods for solution of IPs to various types of noise is an actual problem.

The method of noise addition to the inputs of the neural network during its training was used as a method of increasing the resilience of the neural network solution to noise in the data. It can be based on the studies [19-20], where it was demonstrated that use of this method can improve the generalizing capabilities of the neural network. In [21] it was shown that this method gives similar results in comparison with other ways to improve the generalizing abilities of the network: error regularization, sigmoid gain scaling, target smoothing. In [22] it was shown that the addition of noise during neural network training is equivalent to Tikhonov regularization. In [23], various methods of adding noise during backpropagation training of multilayer feedforward neural networks were compared: adding noise to the inputs, outputs, weight connections, and weight changes. It has been shown that for a classification problem, the improvement of the generalizing abilities of the network for the method of adding noise to the inputs was the same as for the method of adding noise to the weights, and for a regression problem it was superior. For the other methods of adding noise, no improvement in the generalization capabilities of the network was found. In [24], the methods of adding noise to the input were compared: adding noise individually for each pattern at each iteration of training and expanding the training set by including noisy patterns before training. Both methods were shown to improve the result for the classification problem. For the ever-changing noise, the result was slightly better than the one added to the training sample.

In addition to the direct improvement of the generalizing abilities of a neural network, the method of adding noise to the inputs of a neural network allows one to avoid overtraining [25, 26], as well as to accelerate training [27, 28]. Besides that, the addition of noise to the inputs works well in combination with other methods, for example, when training an ensemble of neural networks [29] or when training a neural network by evolutionary methods [30]. Application of the method for training RBF neural networks [31] and for deep neural networks [32] also showed its effectiveness.

In the previous studies of the authors [28, 33] it has been demonstrated that application of this method allowed increasing resilience of neural network solutions of various types of inverse problems to noise in data for different types, statistics and noise levels. At the same time, the best quality of the solution is observed when the noise level in the training and in the test set coincides. However, in these studies, the noise for each input feature was independent from the others. In the present study, we consider a special type of distortion inherent in the inverse problems of spectroscopy, which affects all input features at once – the shift of the channels of a spectrum. The aim of this work was to test the effectiveness of the method of adding noise during training to this type of distortion.

2. Problem Statement

The problem considered in this study was to identify and determine the concentrations of 10 ions (Cl^-, F^-, HCO_3^-, K^+, Li^+, Mg^{2+}, Na^+, NH_4^+, NO_3^-, SO_4^{2-}) contained in a multi-component solution of 10 salts ($MgSO_4$, $Mg(NO_3)_2$, $LiCl$, $LiNO_3$, NH_4F, $(NH_4)_2SO_4$, KF, $KHCO_3$, $NaHCO_3$, $NaCl$) by their Raman spectra (Fig.1). The investigated solutions contained 1 to 5 of the salts in the concentration range 0-1.5 M (mole/l) with an increment of 0.15-0.25 M. Spectra excitation was performed by an argon laser with a wavelength of 488 nm. Registration of the spectra was carried out by a multi-channel detector based on a CCD matrix. For each solution, the spectrum was recorded in 1824 channels in the

range of Raman frequencies of 565...4000 cm⁻¹. The original dataset, which was used to carry out the study, contained 4445 patterns.
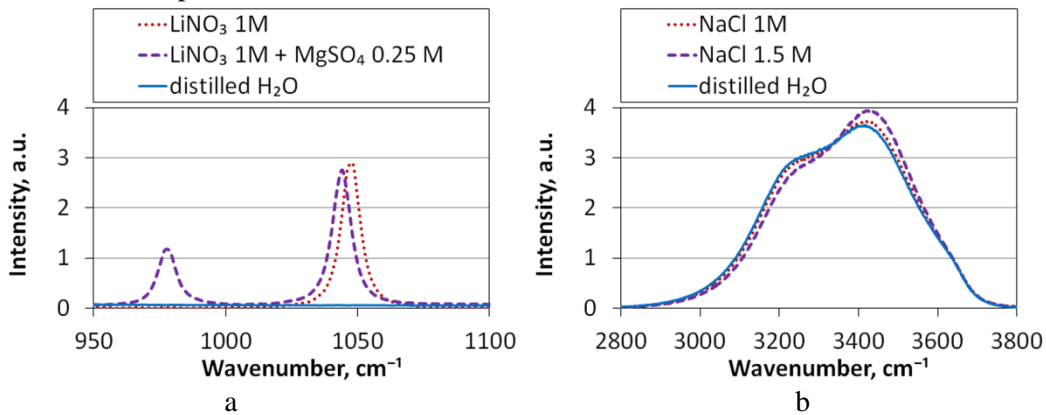

Figure 1. Examples of spectra of ions in the low-frequency (a) and high-frequency (b) ranges.

The possibility of using Raman spectra to diagnose ion composition of water media is due to the high sensitivity of quantitative characteristics of spectral bands to the type and concentration of substances dissolved in water. Many complex ions (sulfides, sulfates, nitrates, phosphates etc.) have their own Raman bands in the region of 300-2000 cm⁻¹ (Figure1) [4-5]. The position of these lines is strictly in accordance with the frequency of oscillation of molecular groups of these ions, the line intensity depends on their concentration in water (Figure 2). For solutions of the same salt, the dependence of the intensity of the line on concentration is almost linear (Figure 2a), however, in presence of other salts, this dependence is disturbed (Figure 2b).
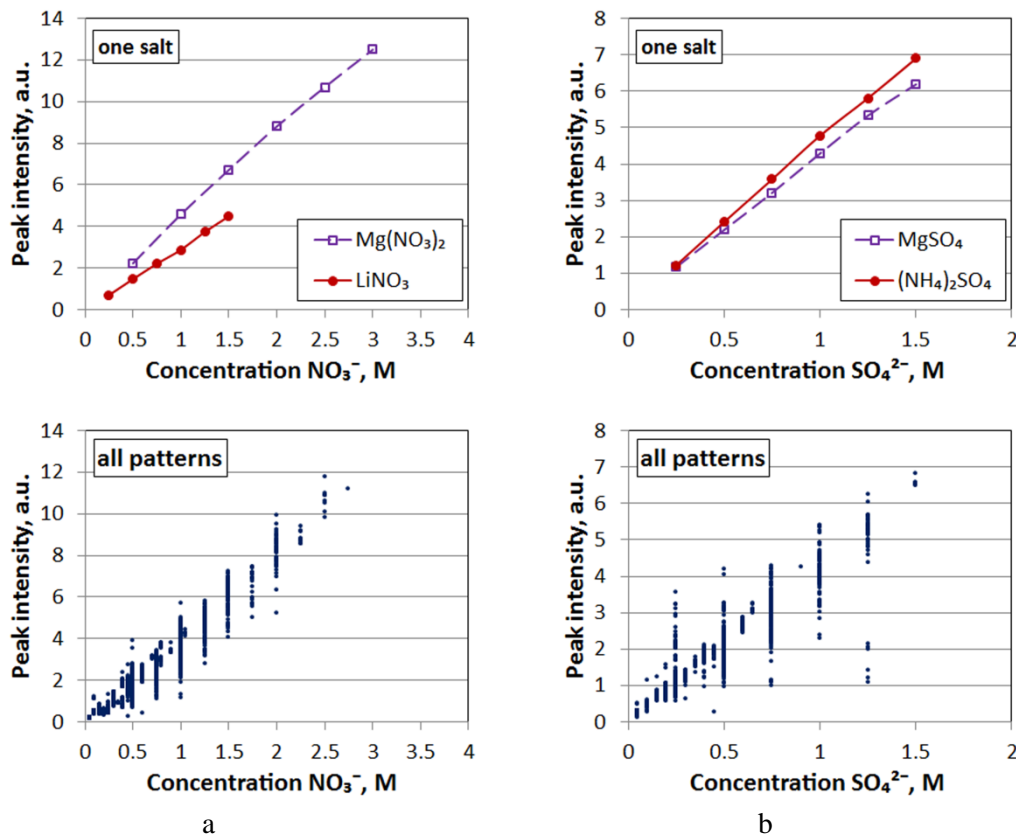

Figure 2. The dependence of the intensity of the peak corresponding to the nitrate ion NO_3^- (a) and to the sulfate ion SO_4^{2-} (b) on the concentration of the ion: top – for solutions containing one salt, bottom – for the whole dataset.

Thus, by Raman spectra of water media in this range of frequencies, one can determine which complex ions are present in water. However, due to the fact that the band intensity is nonlinearly distorted in the presence of other ions, quantitative analysis with traditional methods is difficult.

Monatomic (simple) ions (e.g., Na^+, Cl^-, K^+, Br^-, I^-, Ca^{2+}, Al^{3+} etc.) have no proper Raman lines; however, they have an impact on the Raman valence band of water (Figure 1b). As it has been established in many studies [6-10], the position and shape of the Raman valence band of water (2700–4000 cm^{-1}) depend to a large extent on the type of ion and its concentration in the solution (Figure 1b). When the concentration of ions is increased, the Raman valence band of water is shifted towards higher frequencies, its FWHM decreases, and the intensity of the high-frequency shoulder increases. Different ions cause different change of the position and shape of the Raman valence band of water. However, such joint effect of dissolved ions on the behavior of the Raman valence band of water greatly hampers the solution of inverse problems to determine water parameters with traditional methods.

At the moment, no adequate mathematical models describing such interactions is known, so practically the only way to solve the problem is use of machine learning methods based on experimental data. As any data resulting from experimental measurements contains noise, there is a need to develop specific approaches to improving the resilience of the solution to noise in the data.

3. Description of the Distortions in Data

In the considered problem, experimental data may contain distortions of three types:

a) Variations in the concentrations of ions caused by inaccuracies in the preparation of solutions (the "true" concentrations of ions for each spectrum are not measured by an alternative method, but they are set in the process of preparation of each of the studied solutions).

b) Error in determination of the intensity of the spectra (the dark noise and the error in determining the signal amplitudes by the CCD-detector, influence of cuvette wall and the thickness of measured sample [34]).

c) Channel shifting of the spectrum, which may be due to uncontrolled change in the alignment of the experimental setup when changing the sample (Figure 1a, Figure 3).

From Figure 3 it can be seen that in our dataset, the channel shifting of the spectrum can reach ±2 channels. In addition, the absence of any explicit dependence of the peak position of the characteristic line on the type and quantity of salts and on their concentrations may indirectly indicate that the observed effect is due to the measurement errors and not to the physics of the process.

Thus, the aim of this study was to develop approaches to increase the resilience of the neural network solution to the distortion of the spectroscopic data caused by the channel shifting of the spectrum.

4. Solving the Problem

4.1. Selection of Input Features

To reduce the output dimensionality of the problem, the so-called autonomous determination of parameters was used, where a separate single-output neural network is trained for each determined parameter. Reducing the input dimensionality of the problem was performed on the basis of *a priori* knowledge about the object – the input of the neural network was fed with the features representing the intensity of the spectrum in the channels lying in between 960-1143, 1312-1690, 3014-3601 cm^{-1}, which correspond to the most informative part of the spectrum: to the valence band of water and to the characteristic lines of complex ions. The total number of input features used was 664.

4.2. Use of Neural Networks

We used neural networks containing three hidden layers having 64, 32, and 16 neurons in the 1st, 2nd and 3rd hidden layers, respectively. The activation function was logistic in the hidden layers, and linear in the output layer. Each neural network was trained 5 times with various initial weights values. The statistic indexes of the results of application of the 5 networks were averaged.
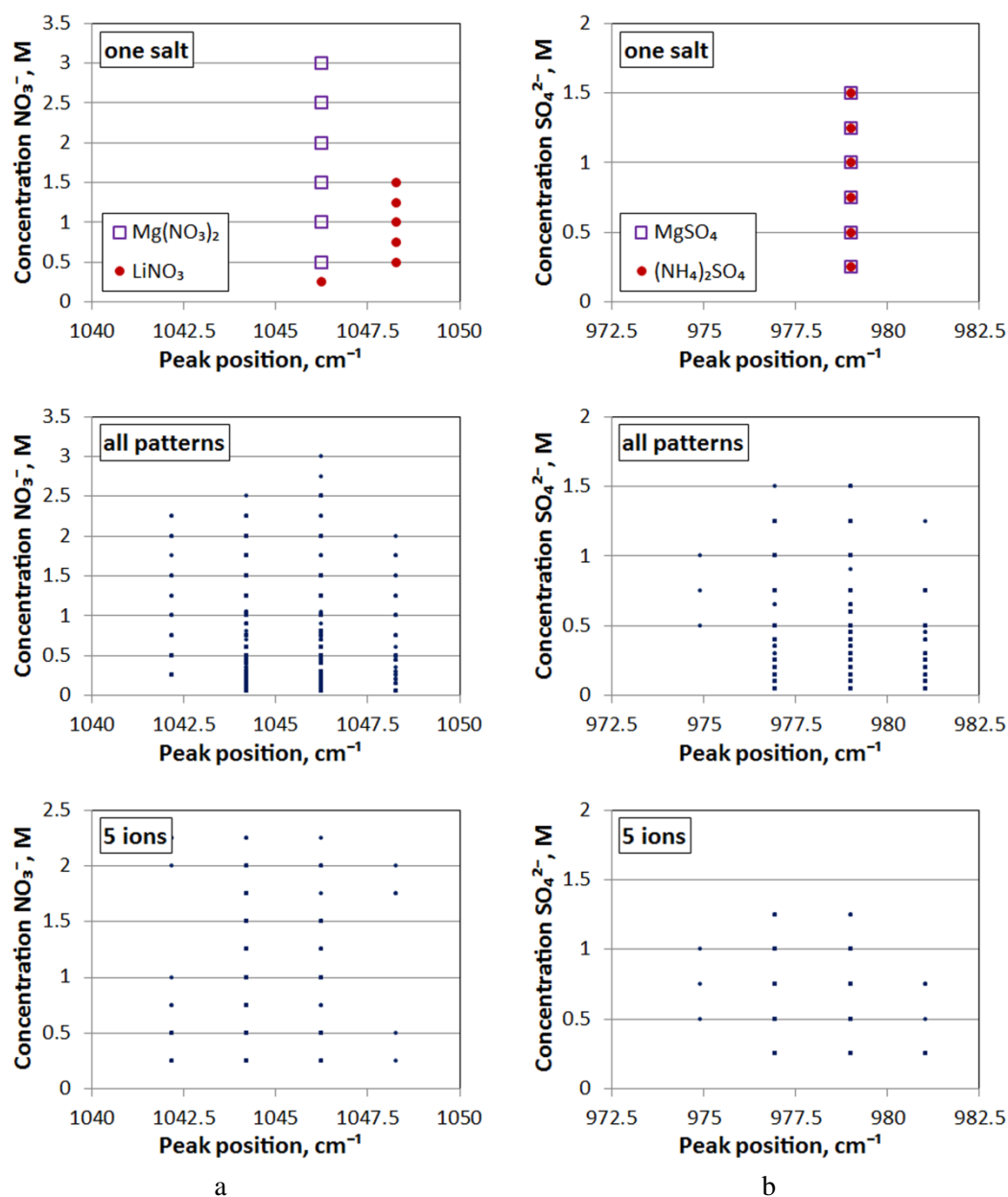
Figure 3. Dependence of the position of the characteristic line of NO_3^- ion (a) and of SO_4^{2-} ion (b) on its concentration in the solution: top – for solutions containing one salt, middle – for the whole dataset, bottom – for solutions containing 5 different ions.

To prevent overtraining of neural networks, the standard technique called early stopping of training was used. The original dataset was randomly split into training, validation and test sets: training was performed on the training dataset, and training was stopped by the validation set (after 1000 epochs with no improvement on the validation set). An independent evaluation of the results was performed on the test set. The number of patterns in training, validation and test sets was 70, 20, 10% of the total number of patterns. Thus, the training dataset contained 3112 patterns, the validation set 889 patterns, and the test set 444 patterns.

4.3. Method of Training with channel shifting
We used neural networks trained on the original dataset and neural networks trained on the dataset containing data with channel shifting. Both types of networks were applied to the test data sets with

and without channel shifting. In the case of data with shifting, each pattern was presented in five options: no shift, shift for 1 and 2 channels to the right, shift for 1 and 2 channels to the left. Thus, for data with shifting, the training dataset contained 15560 patterns, the validation set 4445 patterns, and the test set 2220 patterns.

Thus, we have total of 4 combinations:

No shift to no shift – training on data without shifting, applying to data without shifting

No shift to shift – training on data without shifting, applying to data with shifting

Shift to no shift – training on data with shifting, applying to data without shifting

Shift to shift – training on data with shifting, applying to data with shifting

5. Results

The results of the computational experiment are presented in Fig. 4. One can see that when applied to the test data without channel shifting, both types of neural networks, trained on data with shifting and without shifting, show actually the same results. When neural networks trained on data with no shifting are applied to data with channel shifting, there is a reduction in the quality of solution. This suggests that the ability of neural networks to work with noisy and contradictory data is insufficient to account for distortions of such kind, thus leading to the necessity of using special approaches. In addition, this can serve as indirect evidence of the fact that the representativity of both training and test sets was equal, and in this condition the network was resilient enough to account for the shifting in the test set. Also this may mean that the spectra were taken accurately and the number of spectra with shifted peak in the source dataset was small.
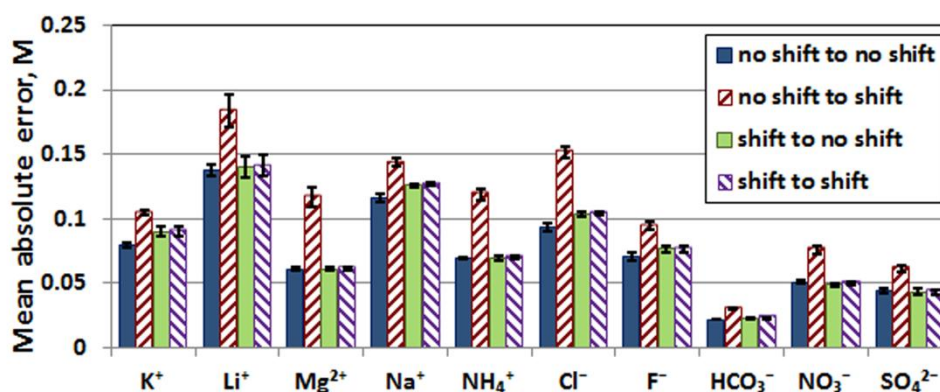

Figure 4. The results of the application of neural networks trained on data with and without shift, with shift and without.

When neural networks trained on data with channel shifting are applied to data also with channel shifting, the obtained results are similar to those obtained when applying the same networks to the data without shifting, thus indicating the efficiency of the approach – adding noise during training to increase noise resilience, also for the channel shifting type of noise.

In addition, one can see that for simple ions and for NH_4^+ ion, whose characteristic lines lie outside the considered area, the drop in problem solution quality when the networks trained on data without channel shifting are applied to data with channel shifting, is higher than for complex ions.

6. Conclusions

As the result of this study, one can draw the following conclusions:

- The approach based on the training of neural networks on data with channel shifting allows one to increase the resilience of neural network solutions to this type of errors in the data.
- This increase in the resilience is more pronounced for simple ions.
- The minimal error on experimental data (containing some patterns with channel shifting) is observed when the neural networks are trained on the same type of experimental data, also containing patterns with channel shifting.

- This confirms the fact that within the experiment-based approach, when experimental data are used to train the neural networks, in the case of sufficient representativity of data in all data sets, the networks turn out to be able to take into account all types of noise present in experimental data, without paying any special attention to the type and intensity of noise.
- If the test data is characterized with increased noise in respect to the training data, the resilience of the solution of an inverse problem may be increased by introducing additional noise of the same type and amplitude into the training set.
- In this study, these conclusions were confirmed for the special type of noise inherent for spectroscopic studies – spectral channel shifting.

7. References

[1] Ahuja S 2013 Monitoring Water Quality *Pollution Assessment, Analysis, and Remediation* (Elseiver) p 400
[2] Crompton T 2002 *Determination of anions in natural and treated waters* (Taylor&Francis) p 828
[3] Michalski R, Jablonska M, Szopa S and Łyko A 2011 Application of Ion Chromatography with ICP-MS or MS Detection to the Determination of Selected Halides and Metal / Metalloids Species *Critical Reviews in Analytical Chemistry* **41** 133-150
[4] Baldwin S and Brown C 1972 Detection of Ionic Water Pollutants by Laser Excited Raman Spectroscopy *Water Research* **6** 1601-1604
[5] Rudolph W and Irmer G 2007 Raman and Infrared Spectroscopic Investigation on Aqueous Alkali Metal Phosphate Solutions and Density Functional Theory Calculations of Phosphate-Water Clusters *Applied Spectroscopy* **61(12)** 274A-292A
[6] Furic K, Ciglenecki I and Cosovic B 2000 Raman Spectroscopic Study of Sodium Chloride Water Solutions *J. Molecular Structure* **6** 225-234
[7] Dolenko T, Churina I, Fadeev V and Glushkov S 2000 Valence Band of Liquid Water Raman Scattering: Some Peculiarities and Applications in the Diagnostics of Water Media *J. of Raman Spectroscopy* **31(8-9)** 863-870
[8] Rull F and De Saja J A 1986 Effect of electrolyte concentration on the Raman spectra of water in aqueous solutions *J. Raman Spectroscopy* **17(2)** 167-172
[9] Gogolinskaia T, Patsaeva S and Fadeev V 1986 The regularities of change of the 3100-3700 cm-1 band of water Raman scattering in salt aqueous solutions *Doklady Akademii Nauk SSSR* **290(5)** 1099-1103
[10] Terpstra P, Combes D and Zwick A 1990 Effect of salts on dynamics of water: A Raman spectroscopy study *J. Chem. Phys* **92(1)** 65-70
[11] Spitsyn V, Bolotova Y, Phan N H and Bui T T T 2016 Using a haar wavelet transform, principal component analysis and neural networks for OCR in the presence of impulse noise *Computer Optics* **40(2)** 249-257 DOI: 10.18287/2412-6179-2016-40-2-249-257
[12] Nikitin M, Konushin V and Konushin A 2017 Neural network model for video-based face recognition with frames quality assessment *Computer Optics* **41(5)** 732-742 DOI: 10.18287/2412-6179-2017-41-5-732-742
[13] Ivanov A, Lozhnikov P and Sulavko A 2017 Evaluation of signature verification reliability based on artificial neural networks, Bayesian multivariate functional and quadratic forms *Computer Optics* **41(5)** 765-774 DOI: 10.18287/2412-6179-2017-41-5-765-774
[14] Gerdova I, Dolenko S, Dolenko T, Churina I and Fadeev V 2002 New opportunity solutions to inverse problems in laser spectroscopy involving artificial neural networks *Izvestiya Akademii Nauk Seriya Fizicheskaya* **66(8)** 1116-1124
[15] Dolenko S, Burikov S, Dolenko T, Efitorov A, Gushchin K and Persiantsev I 2014 Neural Network Approaches to Solution of the Inverse Problem of Identification and Determination of Partial Concentrations of Salts in Multi-Component Water Solutions *ICANN Lecture Notes in Computer Science* **8681** 805-812
[16] Efitorov A, Burikov S, Dolenko T, Persiantsev I and Dolenko S 2015 Comparison of the Quality of Solving the Inverse Problems of Spectroscopy of Multi-Component Solutions with Neural

Network Methods and with the Method of Projection to Latent Structures *Optical Memory and Neural Networks (Information Optics)* **24(2)** 93-101

[17] Efitorov A, Burikov S, Dolenko T, Laptinskiy K and Dolenko S 2015 Significant Feature Selection in Neural Network Solution of an Inverse Problem in Spectroscopy *Procedia Computer Science* **66** 93-102

[18] Dolenko S, Efitorov A, Burikov S, Dolenko T, Laptinskiy K and Persiantsev I 2015 Neural Network Approaches to Solution of the Inverse Problem of Identification and Determination of the Ionic Composition of Multi-Component Water Solutions *Proc. of EANN Communications in Computer and Information Science (CCIS)* **517** 109-118

[19] Holmstrom L and Koistinen P 1992 Using Additive Noise in Back-Propagation Training *IEEE Trans. on Neural Networks* **3(1)** 24-38

[20] Matsuoka K 1992 Noise Injection into Inputs in Back-Propagation Learning *IEEE Trans. Systems, Man, and Cybernetics* **22(3)** 436-440

[21] Reed R, Marks R and Oh S 1995 Similarities of error regularization, sigmoid gain scaling, target smoothing, and training with jitter *IEEE Transactions on Neural Networks* **6(3)** 529-538

[22] Bishop C 1995 Training with noise is equivalent to Tikhonov regularization *Neural Computation* **7(1)** 108-116

[23] An G 1996 The effects of adding noise during backpropagation training on a generalization performance *Neural Computation* **8(3)** 643-674

[24] Zur R, Jiang Y and Metz C E 2004 Comparison of two methods of adding jitter to artificial neural network training *International Congress Series* **1268** 886-889

[25] Zur R, Jiang Y, Pesce L and Drukker K 2009 Noise injection for training artificial neural networks: A comparison with weight decay and early stopping *Medical Physics* **36(10)** 4810-4818

[26] Piotrowski A and Napiorkowski J 2013 A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modeling *Journal of Hydrology* **476** 97-111

[27] Wang C and Principe J 1999 Training neural networks with additive noise in the desired signal *IEEE Transactions on Neural Networks* **10(6)** 1511-1517

[28] Isaev I and Dolenko S 2016 Training with noise as a method to increase noise resilience of neural network solution of inverse problems *Optical Memory and Neural Networks (Information Optics)* **25(3)** 142-148

[29] Peter Zhang G 2007 A neural network ensemble method with jittered training data for time series forecasting *Information Sciences* **177(23)** 5329-5346

[30] Piotrowski A, Rowinski P and Napiorkowski J 2012 Comparison of evolutionary computation techniques for noise injected neural network training to estimate longitudinal dispersion coefficients in rivers *Expert Systems with Applications* **39(1)** 1354-1361

[31] Ho K, Leung C and Sum J 2010 Convergence and objective functions of some fault/noise-injection-based online learning algorithms for RBF networks *IEEE Transactions on Neural Networks* **21(6)** 938-947

[32] Yin S, Liu C, Zhang Z, Lin Y, Wang D, Tejedor J, Zheng T and Li Y 2015 Noisy training for deep neural networks in speech recognition *EURASIP Journal on Audio, Speech, and Music Processing* **2015(2)** 1-14

[33] Isaev I and Dolenko S 2018 Adding Noise During Training as a Method to Increase Resilience of Neural Network Solution of Inverse Problems: Test on the Data of Magnetotelluric Sounding Problem *Studies in Computational Intelligence* **736** 9-16

[34] Nelubina N, Pidgirny V, Bulgakova O, Zvekov A and Kalenskii A 2016 Peculiarities of spectral measurements of colored suspensions in thick-walled cuvettes *Computer Optics* **40(4)** 508-515 DOI: 10.18287/2412-6179-2016-40-4-508-515

Acknowledgements