NEURAL NETWORK SOLUTION OF AN INVERSE PROBLEM IN RAMAN SPECTROSCOPY OF MULTI-COMPONENT SOLUTIONS OF INORGANIC SALTS: TRAINING WITH NOISE AS A METHOD TO INCREASE NOISE RESILIENCE OF THE SOLUTION

Igor Isaev[1], Ekaterina Vervald[2], Olga Sarmanova[2], and Sergey Dolenko[1]

1. M.V.Lomonosov Moscow State University, D.V. Skobeltsyn Institute of Nuclear Physics, Moscow, Russia
2. M.V.Lomonosov Moscow State University, Physics Department, Moscow, Russia isaev_igor(at)mail.ru, dolenko(at)sinp.msu.ru

In various fields of industry and ecology there is a need for diagnostics of individual components in multi-component aqueous mixtures. For example, it concerns the problems of identification and determination of concentration of each dissolved ion in technological and waste waters, control of the composition of the drinks, monitoring of nature waters. For successful solution of such problems it is necessary to elaborate new methods which should be: a) express - provide information in real time mode; b) non-contact – allow obtaining information remotely, without interference in the studied medium; c) they must have high selectivity in relation to each component of the complex mixture.

Conventional chemical (analytical) methods provide a rather high accuracy of determination of concentration of ions – from units to hundredths of µg/l [1, 2]. However, these methods are contact, they are individual for each ion, they are time-consuming, and they require competent preparation of the samples and consumption of expensive reagents. Therefore, elaboration of express non-contact methods for determination of concentrations of substances dissolved in water is a very important task for ecological monitoring of coastal waters.

These properties can be provided by laser Raman spectroscopy (RS) method. The possibility of using RS for diagnostics of the ion composition of a water medium arises from high sensitivity of the shape and quantitative characteristics of spectral bands to the type and concentration of substances dissolved in water [3,4]. Many complex ions (sulfides, sulfates, nitrates, phosphates, etc.) have their own Raman bands in the area of 300-2000 cm-1. Also, different ions (complex or simple) cause different changes in the position and shape of water Raman valence band. However, such a cooperative effect of various dissolved ions on the behavior of the water Raman valence band greatly hampers the solution of inverse problems of determination of parameters of the solution using traditional methods. Use of adaptive methods of data analysis, primarily, of artificial neural networks (ANN), provides selectivity of the methods and possibility of quantitative determination of the content of each component in the medium in presence of many other components [5].

The problem of determining the concentrations of dissolved substances by spectra belongs to the class of inverse problems that have a number of features, complicating their solution. The considered inverse problem, like most other inverse problems, is usually incorrect or at least ill-conditioned, which makes it sensitive to the presence of noise in the input data. Despite the fact that ANN by themselves have the ability to work with noisy data, this may be not enough for inverse problems, because the incorrectness can "outweigh" the abilities of the network. Development of approaches to increase the resilience of problem solution methods to noise is an urgent task.

In the previous study of the authors [6] it has been demonstrated on a model inverse problem that adding noise to the training set when training neural networks allows increasing the resilience of the ANN solution to noise in the input data. The purpose of the present study was to demonstrate the applicability of this method for the inverse problem of determination of the concentrations of ions in multicomponent salt solutions.

The problem considered in the present study was to identify by Raman scattering spectra and to 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 multicomponent solution of any combinations of the 10 salts (MgSO4, MgNO3, LiCl, $LiNO_3$, NH_4F, $(NH_4)2SO_4$, KF, $KHCO_3$, $NaHCO_3$, NaCl). The investigated solutions contained from 1 to 5 of the salts in the concentration range 0 – 1.5 M (mol/l) with a step of 0.15-0.25 M. Excitation of the spectra was performed by an argon laser with radiation wavelength of 488 nm. For each solution, the spectrum was registered in 1824 channels in the frequency range of 565...4000 cm^{-1}. The whole dataset that was used in this study contained 4445 samples.

Thus, in this study we considered a multi-parameter inverse problem consisting in determination of 10 parameters (ion concentrations) by 1824 input features (channel intensities). To reduce the output dimensionality of the problem, an approach was used in which a separate single-output ANN was trained for each determined ion, the so-called Autonomous determination. To reduce the dimension of the input, selection of significant input features was used.

Two types of noise were considered in this study, additive and multiplicative, and two kinds of noise statistics: uniform noise (uniform distribution) and Gaussian noise (normal distribution).

As a summary of this study, it has been demonstrated that adding noise to the training set when training neural networks allows increasing the resilience of the neural network solution to noise in the input data. This conclusion is valid for all ions, for all types and any statistics of noise. The type and statistics of noise added to the training set should coincide with those of the noise present in the input data when the ANN is applied. There is also an optimal level of the noise to be added during training, which also depends on the expected level of noise contained in the input data when the ANN is applied.

ACKNOWLEDGEMENTS

This study has been performed at the expense of Russian Science Foundation, grant no. 14-11-00579.

The authors would like to thank S.A. Burikov, T.A. Dolenko, and K.A. Laptinskiy for conducting the experiment and for useful discussions.

REFERENCES

1 Crompton T.R. Determination of anions in natural and treated waters. Taylor & Francis, 2002, 828 p.

2 Michalski R., Jablonska M., Szopa S., Łyko A., 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, 2011, V.41, pp.133-150.

3 Rull F., De Saja J.A. Effect of electrolyte concentration on the Raman spectra of water in aqueous solutions. J. Raman Spectroscopy, 1986, V.17, No.2, pp.167-172.

4 Dolenko T.A., I.V. Churina, V.V. Fadeev, S.M. Glushkov. Valence Band of Liquid Water Raman Scattering: Some Peculiarities and Applications in the Diagnostics of Water Media. J. of Raman Spectroscopy, 2000, V.31, No.8-9, pp.863-870.

5 Efitorov A., T. Dolenko, S. Burikov, K. Laptinskiy, S. Dolenko. Solution of an Inverse Problem in Raman Spectroscopy of Multi-component Solutions of Inorganic Salts by Artificial Neural

Networks. A.E.P. Villa et al. (Eds.): ICANN 2016, Part II. Lecture Notes in Computer Science (LNCS), 2016, V.9887, pp.355-362.

6 Isaev, I.V., Dolenko, S.A, Training with noise as a method to increase noise resilience of neural network solution of inverse problems, Optical Memory and Neural Networks, 2016, V.25, No.3, pp.142–148