Application of Artificial Neural Networks to Solve Problems of Identification and Determination of Concentration of Salts in Multi-Component Water Solutions by Raman Spectra[1]

S. A. Burikov[a], S. A. Dolenko[b], T. A. Dolenko[a], and I. G. Persiantsev[b]

[a]*Department of Physics, M.V. Lomonosov Moscow State University, Leninsky Gory 1/2, Moscow, 119991 Russia*
e-mail: burikov@lid.phys.msu.ru, tdolenko@lid.phys.msu.ru
[b]*D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University*
Leninsky Gory 1/2, Moscow, 119991 Russia
e-mail: dolenko@srd.sinp.msu.ru, ipers@srd.sinp.msu.ru
Received March 31, 2010; in final form, April 15, 2010

Abstract—In this paper, the results of elaboration and comparative analysis of approaches concerned with application of neural network algorithms for effective solution of problem of pattern recognition (inverse problem with discrete output) along with inverse problem with continuous output are presented. Consideration is carried out at the example of problem of identification and determination of concentrations of inorganic salts in multi-component water solutions by Raman spectra. The studied approach is concerned with solution of both problems (classification and determination of concentrations) using a single neural network trained on experimental or quasi-model data.

Keywords: Neural networks, inverse problems, identification, multi-component solution, Raman spectroscopy.

INTRODUCTION

It is well known that artificial neural networks (ANN) are a class of mathematical algorithms showing very high efficiency for solution of various problems of approximation, prediction, evaluation, classification, pattern recognition etc. ANN are used also for solution of inverse problems (IP) where their properties, such as training by examples, high noise stability, stability to contradictory data etc., play a special role (see, for example, [1]).

Most often, IP of regression type are under consideration—those with continuous output (for example, determination of temperature of plasma by its proper glow spectrum). However, more complicated situations can take place, when simultaneous solution of problems of classification or pattern recognition (determination of components contributing to observed properties of object, for example, to its spectrum), i.e. IP with discrete output, and traditional IP of regression type for every found component, is required. Call this type of IP complex inverse problem.

In this study, elaboration of method of solution of such problems is carried out at the example of complex IP of identification and determination of partial concentrations of salts in multi-component water solution by Raman spectra. The problem of determination of concentrations of substances dissolved in water is very important for oceanology, ecological monitoring, and control of technical waters. The methods most required are express non-contact methods which can be implemented in remote express mode.

It is very promising to use Raman spectra for express remote determination of concentrations of inorganic substances dissolved in water. In [2, 3] it is suggested to use Raman spectra of complex ions for determination of type and concentrations of salts in water solution (including remote sensing mode). In Raman spectra near 1000 cm^{-1}, narrow bands of valence vibrations of anions NO_3^-, SO_4^{2-}, PO_4^{3-}, CO_3^{2-} are present. The type of anion was determined by position of corresponding band, its concentration—by the band intensity. This method has a disadvantage—it is suitable only for analysis of substances having their

---

[1] The article is published in the original.

own Raman bands, i.e. for salts with complex anions. However, concentration of such salts in natural waters is much lower than, for example, concentration of metal halides.

The authors of [4–11] elaborated a method of determination of concentration of dissolved salts by water Raman valence band. Principle opportunity of use of spectral Raman bands of water for diagnostics of solutions arises from high sensitivity of their characteristics to type and concentration of salts dissolved in water. In [4, 5], concentration of one salt in the solution was determined using water Raman valence band. The authors of [8–11] suggested and elaborated a method of determination of partial concentrations of several salts in multi-component solutions by water Raman valence band. It was possible because of application of ANN for solution of this IP. The authors of this paper are unaware of other studies where concentration of several salts in a solution would be measured by optical methods. Nevertheless, the problem already solved by the authors of the present paper—identification of type of salts and determination of their partial concentrations by water Raman valence band by traditional NN approach in 3-component solution [9–11]—does not satisfy urgent problems of oceanology and ecology. It is necessary to identify a larger number of salts in the solution and at the same time to retain accuracy of determination of their concentrations or better to improve it. Solution of such multi-parametrical inverse problem requires elaboration of optimal methodical approaches to application of NN algorithms.

1. METHODICAL APPROACHES TO SOLUTION OF IP USING NEURAL NETWORKS

Training of ANN for solution of IP requires a representative dataset, i.e. a set of data that would reflect all characteristic aspects of behavior of the object. This dataset can be obtained in different ways. In compliance with this, one can mark out three different methodical approaches to solution of IP using ANN.

(a) "Model-based" approach. If an adequate analytical model of solution of the direct problem is available, it can be used for generation of arrays of data with necessary representativity. The main disadvantage of this approach is that elaboration and realization of an adequate model is often impossible or very difficult. Unfortunately, this is the case with the studied IP of determination of concentrations of components using Raman spectra. Because of extraordinary complexity of the object, there is no adequate model based on physical reasons that would allow obtaining the dependence of water Raman spectrum (i.e. of the intensity in every spectral channel) on concentration of dissolved salts, especially taking into account their non-linear interaction. That is why this approach is, unfortunately, unacceptable in this case.

(b) "Experiment-based" approach. The data used for ANN training are obtained in experiment. (In this paper, the data are Raman spectra of various solutions with different combinations and concentrations of components, obtained using laser Raman spectrometer). This approach does not require a model to be available, and it allows taking non-linear properties of the object into account. However, obtaining a representative data set can be a non-trivial experimental problem. In this study, 8695 experimental spectra for 4268 different solutions were obtained for realization of the "experiment-based" approach.

(c) "Quasi-model" approach. If no adequate physically grounded analytical model of solution of the direct problem is available, one can replace it by a parametrical "quasi-model" based on experimental data. This model formally describes the dependence of the observed data on the sought-for parameters. In the present case, a quasi-model (a set of quasi-models) describes the dependence of intensity in every spectral channel on concentrations of components. The simplest quasi-model is a linear one when it is supposed that the intensity in each spectral channel is a linear combination of partial concentrations of the components. In order to construct more complicated quasi-models, which can describe the desired dependence better, one can use more efficient adaptive methods for construction of models—for example, group method of data handling (GMDH) or some types of ANN.

In the studies of one-, two-, and three-component solutions carried out before, several approaches were used—"experiment-based", "linear quasi-model", and "GMDH quasi-model". It has been demonstrated [14], that the "linear quasi-model" approach is unusable for determination of concentrations of salts in multi-component solutions. This fact confirms the conclusion about non-linear influence of salts on Raman valence band in multi-component solutions. For this reason, the "linear quasi-model" approach was not used in this study.

2. BASIC METHODS OF SOLUTION OF PROBLEM AND CRITERIA OF MODEL EVALUATION

The considered complex IP naturally breaks down into two problems—the problem of identification of components of the solution and the problem of determination of their partial concentrations. The method applied in this paper provides simultaneous solution of these two problems when a single ANN is

used. The values at the outputs of the ANN are treated as estimations of concentrations of components; if the value at some output is lower than a pre-defined threshold, it is considered that the corresponding component is not present in the solution. This method was also used in all the above mentioned earlier experiments with one-, two- and three-component solutions.

The NN architectures used in this study to solve the complex IP were perceptrons with one and three hidden layers, trained by error back-propagation method, General Regression NN (GRNN) [12] and Group Method of Data Handling (GMDH) [13]. For construction of quasi-models, perceptrons and GRNN were tested.

For evaluation of quality of models, four basic criteria were used in this study. In the description of these criteria, the following notation was used:

y—estimation of the value of an output variable made by the model (neural network);

d—the desired value of this output variable;

$\bar{d}$—average value of the desired output over the whole concerned dataset;

N—the number of patterns in the concerned dataset.

Summation is carried out over all the patterns of the dataset for which the criterion is calculated (from 1 to N).

1. Coefficient of multiple determination R^2 is calculated according to the following formula:

$$R^2 = 1 - \frac{\Sigma(d-y)^2}{\Sigma(d-\bar{d})^2}. \tag{1}$$

This criterion compares the error of the constructed model with the error of trivial reference model (the estimate provided by this reference model is the mean value of the estimated variable over all the patterns of the dataset). When the estimate is absolutely accurate, R^2 equals 1. If the accuracy of the estimate is worse than the accuracy of the trivial model, R^2 is negative. For many kinds of problems, R^2 is the most substantial universal criterion of evaluation of model quality. This criterion is dimensionless.

2. Mean squared error (MSE) is calculated according to the following formula:

$$MSE = \sqrt{\frac{\Sigma(d-y)^2}{N}}. \tag{2}$$

This criterion has the same dimension as the estimated variable. In our case (determination of concentration), it is M (molarity—quantity of moles of substance per liter of solution).

3. Mean absolute error (MAE), calculated according to

$$MAE = \frac{\Sigma|d-y|}{N}. \tag{3}$$

This criterion also has the same dimension as the estimated variable.

4. Mean relative error (MRE), calculated according to

$$MRE = \frac{1}{N}\Sigma\frac{|d-y|}{d} \times 100\%. \tag{4}$$

This criterion has no dimension and is usually expressed in percent.

Because of limited size of the article, no comparative analysis of areas of application, advantages and disadvantages of these criteria is presented here.

3. EXPERIMENT

The scheme of the laser Raman spectrometer is presented in Fig. 1.

Excitation of Raman spectra was performed by argon laser (wavelength 488 nm, output power 350 mWt). In order to remove elastic scattering signal, edge-filter (Semrock) was used. It allowed approaching laser line to 200 cm^{-1}. Registration of spectra was performed by monochromator (Acton 2500i, grade 900 l/mm, focal length 500 mm) and CCD-camera (Synapse 1024 × 128 BIUV, Jobin Yvon). Spectra were measured in two regions: 200−2300 cm^{-1} and 2300−4000 cm^{-1} for every sample. Practical resolution of the spectrometer was 2 cm^{-1}, duration of accumulation of one spectrum was 1 s. The temperature of samples during experiment was stabilized at 22.0 ± 0.2°C. Spectra were normalized to laser
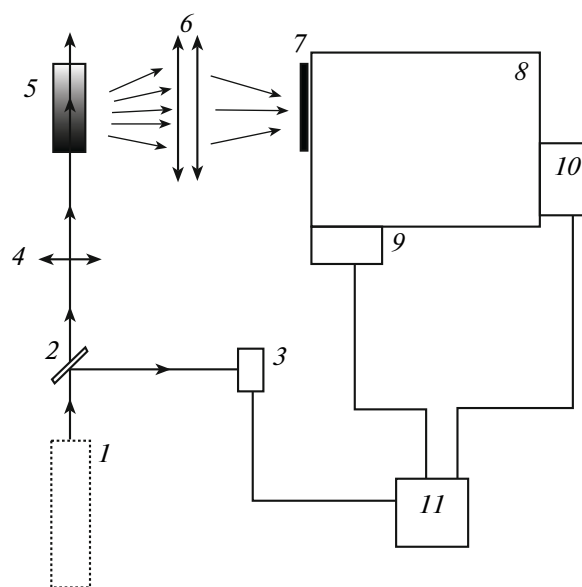
power, duration of registration of the spectrum and spectral sensitivity of the detector.

The objects of research were water solutions of the salts NaCl, NH_4Br, Li_2SO_4, KNO_3, CsI. These salts are present in natural waters at significant concentrations. Concentration of every salt in the solution was changed from 0 to 2.5 M (with increment 0.2–0.25 M). The range of concentrations was chosen according to the following considerations: salinity of seawater 35‰ corresponds to concentration 0.5 M of the most widespread salt NaCl. In mineral waters concentration of certain salts can reach 1 M, in waste waters—up to 1–2 M. For preparation of experimental solutions, bidistilled water and analytically pure reagents were used.
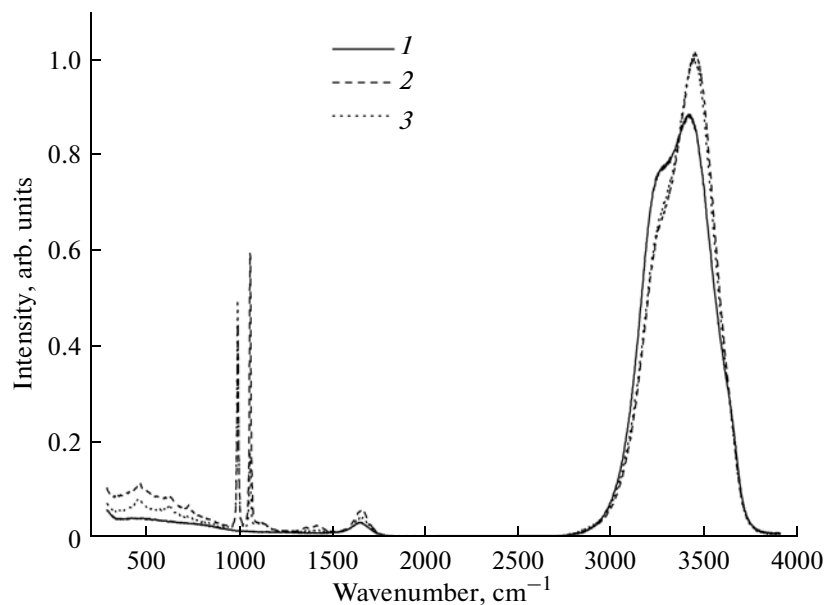

Fig. 1. Experimental setup for Raman spectroscopy of liquid water and water solutions: *1*—argon laser (488 nm), *2*—beam splitter, *3*—indicator of laser power, *4*—lens, *5*—cuvette, *6*—focusing system, *7*—polarizer, *8*—monochromator, *9*—photomultiplier (PMT), *10*—CCD-camera, *11*—computer.

In Fig. 2, some of the obtained Raman spectra of water and water solutions of salts are presented. In the low-frequency region of Raman spectrum (200–1800 cm^{-1}) one can observe bands of valence vibrations of anions SO_4^{2-} (3) and NO_3^- (2). Their intensities depend on the concentrations of the corresponding salts. As experimental results demonstrate, under increase of concentration of salts, water Raman valence band (2600–4000 cm^{-1}) shifts towards high frequencies, its half-width decreases, the intensity of its high-frequency part increases, the intensity of the low-frequency part decreases. The changes in position and shape of water valence band depend on type and concentration of salt. Moreover, anions influence the behavior of water valence band stronger than cations [6].

Fig. 2. Raman spectra: *1*—distilled water; *2*—water solution with concentration of salts: NaCl, NH_4Br, Li_2SO_4—0.4 M, KNO_3 and CsI—0.6 M, *3*—water solution with concentration of salts: 0.6 M for NaCl and Li_2SO_4, 0.2 M for NH_4Br, KNO_3 and CsI.

4. DATA PRE-PROCESSING AND PREPARATION OF DATASETS

During experiment, two bands of Raman spectrum were recorded, in frequency regions 200...2300 cm^{-1} (low-frequency band) and 2300...4000 cm^{-1} (valence band). Preparation of experimental spectra to work with ANN included the following steps:

(1) To work with ANN, narrower ranges of spectrum were used: 766 channels in the range 281...1831 cm^{-1} for low-frequency region and 769 channels in the range 2700−3900 cm^{-1} for the valence band.

(2) For each of the bands, the pedestal caused by non-controlled elastic scattering in cuvette with sample, was subtracted. Then the spectra were normalized to the area of the valence band in the pointed region.

(3) For the data array obtained after this processing (1535 features—intensities in every channel, 8695 patterns), average value and standard deviation were calculated for every feature. Then the low-informative features with standard deviation less than 2.5×10^{-5} and the features corresponding to low-frequency region below 950 cm^{-1} (see Fig. 2) were excluded from further consideration. Variability of the features in the region below 950 cm^{-1} is caused mainly by non-controlled changes of light scattering in the cuvette with sample ("clutter"). Total amount of the remaining features was 704 : 185 channels in the range 950...1673 cm^{-1} for low-frequency region and 519 channels in the range 2870...3682 cm^{-1} for valence band.

(4) It is well known [6, 7] that for single-component solutions of salts, dependence of the intensity of Raman spectrum on concentration of salt in the range 0...2.5 M is linear. That is why for single-component solutions, spectra only for 5 different concentrations of each salt were measured. This was enough to make sure that dependence of intensity on concentration was linear, and to use this dependence for generation of quasi-model linearly-interpolated spectra of single-component solutions for another 45 intermediate concentrations of each salt. As the result, the total amount of patterns in the data array was 9144 : 160 spectra of distilled water (measured in different time simultaneously with measurement of spectra of different salts); 100 spectra (real and interpolated) for single-component solutions of each of 5 salts, 240 spectra for each of 10 two-component and each of 10 three-component combinations of salts in solutions; 420 spectra for each of 5 four-component combinations of salts; 1584 spectra of five-component solutions.

(5) The thus obtained total array of data (704 features, 9144 patterns) was used to calculate of minimal and maximal values of each feature for normalization of input and output variables.

(6) As for each solution with some combination of concentrations two spectra (A and B) were measured, partitioning into sets of data necessary for use of NN was performed not by separate spectra but by combinations of concentrations (excluding spectra of distilled water). In order to provide equal conditions for every combination of salts ("combinational class"), the same proportion of training, test and examination sets should have been observed for all combinations of salts. That is why partitioning into datasets was performed in following way:

(a) The whole data array was divided into 63 sets—two (containing A spectra and B spectra) for each combinational class with solutions and one for the class containing spectra of distilled water.

(b) Each of these 63 sets was randomly divided into training, test and examination sets (with 70 : 20 : 10 ratio). For each set, the same random seed was used for random number generator. This procedure guaranteed placement of spectra with the same concentrations (A and B) to the same set and therefore equal proportion among the three datasets fro all the combinational classes. This provided equality and necessary independence of sets, and this allowed avoiding illusions in evaluation of the results.

(c) By appending corresponding parts A and B to each other, 31 training, test and examination sets were obtained—one for each combinational class. The corresponding sets for the class with distilled water were obtained at the stage b).

(d) To work with the whole data array without separation into combinational classes, training sets for all combinational classes were merged to obtain the common training set. Similar routine was performed for test and examination sets.

Information about number of patterns in different datasets obtained as the result of the described procedure is summarized in Table 1.

5. RESULTS OF SOLUTION OF THE PROBLEM USING THE "EXPERIMENT-BASED" APPROACH

The inputs of ANN were fed with all the 704 features chosen according to the routine described in previous section, without any additional selection, compression or other pre-processing connected with further reduction of input dimensionality of the problem. Each feature separately was normalized into 0...1 range for the whole data array (see step 5 in the preceding section).

Table 1. Number of patterns in different datasets

Number of salts in solution	Number of classes	Total number of patterns	Training patterns	Test patterns	Examination patterns
0	1	160	112	32	16
1	5	100	70	20	10
2	10	240	168	48	24
3	10	240	168	48	24
4	5	420	294	84	42
5	1	1584	1110	316	158
Total	32	9144	6402	1828	914

Table 2. Comparison of the results of determination of salts concentrations by Raman spectra in multi-component solutions (MAE on examination set, M)

Method	NaCl	NH_4Br	Li_2SO_4	KNO_3	CsI
Perceptron	0.029	0.024	0.020	0.019	0.023
GRNN	0.102	0.047	0.064	0.058	0.050
GMDH	0.059	0.046	0.032	0.033	0.046

The ANN had five outputs according to the maximal number of components in a solution. Each output corresponded to one of the considered salts, and its desired value corresponded to concentration of this salt in the solution.

Three models were used to solve the considered problem: perceptrons with three hidden layers, GRNN and GMDH. The results (MAE on examination set, M) are presented in Table 2.

As can be seen from Table 2, the best results were demonstrated by perceptron with three hidden layers. The hidden layers contained 40, 20 and 10 neurons. Linear activation function was used in the output layer, and logistic activation function was used in hidden layers. The following values of training parameters were used: learning rate—0.01; moment—0.5; stop training criterion—1000 epochs after minimum of error on test dataset.

Table 3 presents comparison of the results obtained by the authors of this study before [11] for diagnostics of three-component solutions using only valence band, and the results obtained in this study using both bands of spectrum and valence band only.

Significantly better results obtained in this study compared to the results of [10, 11], in a wider range of concentrations and for five-component (instead of three-component) solutions, can be explained by the fact that in the present study the low-frequency region of spectrum (where bands of complex ions are present) was taken into consideration.

6. RESULTS OF CONSTRUCTION OF QUASI-MODELS

One of the main reasons causing decrease of accuracy of solution of this inverse problem using the "experiment-based approach" is the unfavorable ratio of the number of input variables of the problem (704) and the number of patterns in the training set (8229 in training and sets together). One can overcome

Table 3. Mean absolute error (MAE) for determination of concentrations of salts by Raman spectra in multi-component solutions, M (on examination data set)

Experiment	Band	Range, M	NaCl	NH_4Br	Li_2SO_4	KNO_3	CsI	KI
[10, 11] 3 comp.	Valence	0…0.7	0.07	0.06	—	—	—	0.05
[10, 11] 3 comp.	Valence	0…1	0.07	0.11	—	—	—	0.12
This study, 5 comp.	Valence	0…2.5	0.047	0.029	0.040	0.046	0.032	—
This study, 5 comp.	Both	0…2.5	0.029	0.024	0.020	0.019	0.023	—

Table 4. Statistics for construction of quasi-models based on perceptron and based on GRNN (on examination dataset)

Index, model	min	max	mean	standard. deviation
R^2, perceptron	0.681	0.983	0.933	0.048
R^2, GRNN	0.768	0.974	0.933	0.041
MRE, %, perceptron	0.5	272.9	9.9	27.0
MRE, %, GRNN	0.4	75.0	7.1	11.6

it in one of two ways: by further reduction of the number of input variables (this way was not considered in this study) or by increasing the number of patterns.

In the situation when additional experiments cannot be conducted, one can try to get additional patterns by spectra interpolation using "quasi-models"—parametrical or adaptive model of solution of the direct problem. In our case this is the model of dependence of intensity in each channel of the spectrum on concentrations of the components in the solution.

However, one should understand that if a "quasi-model" is not adequate enough, using such "quasi-model" approach can lead not to improvement but to degradation of the quality of solution of the IP. In particular, for this reason, a linear "quasi-model" is unsuitable for modeling of multi-component mixtures [10, 11].

In this study, two types of quasi-models were considered: based on perceptron and based on GRNN. For determination of applicability of the "quasi-model" approach for this problem it was necessary first to choose the best of the constructed quasi-models (i.e. the quasi-model providing the smallest error of solution of direct problem on the examination set).

In Table 4, statistics calculated over all 704 modeled channels are presented for quasi-models based on perceptron and on GRNN. From this table, one can see that average statistics were close, so at this stage there was no reason to prefer either of these two quasi-models.

7. RESULTS OF USE OF "QUASI-MODEL" APPROACH FOR SOLUTION OF THE INVERSE PROBLEM

The data for use of the "quasi-model" approach were prepared in the following way. To create the "quasi-model" datasets, we used a grid with 0.15 M increment for all salts; only spectra with total concentration of salts less than 2.5 M were considered. There were 53 130 such spectra with different concentrations of salts, i.e. more than 10 times as much as experimental spectra with different concentrations. The obtained array of 53 130 "quasi-model" spectra was randomly divided into training, test, and examination datasets with 70 : 20 : 10 ratio.

After the division of the array of "quasi-model" spectra, the thus obtained training and test sets were supplemented with training and test datasets from the experimental array. The examination datasets were left separate. Thus, the following datasets were used for the "quasi-model" approach: training set (43 593 patterns), test set (12 454 patterns), examination quasi-model set (5313 patterns) and examination experimental set (914 patterns).

Two such complete sets have been obtained: one with generation of "quasi-model" spectra with the "quasi-model" based on perceptron, the other one—with the "quasi-model" based on GRNN.

The thus obtained training sets were used to train identical perceptrons, with architecture and training parameters identical to those used to solve the IP within the "experiment-based" approach (Section 5).

Table 5 displays comparison of the results obtained within the "experiment-based" approach on the examination (experimental examination) dataset, with the results obtained within the "quasi-model" approach on quasi-model examination dataset and on experimental examination dataset, for two kinds of "quasi-model" (based on perceptron and based on GRNN).

The provided results allow making the following conclusions.

(1) The "quasi-model" approach failed to meet expectations compared to the "experiment-based" approach. In all cases, the results on the experimental examination dataset, obtained within the "quasi-model" approach, turned out to be worse or significantly worse than the results obtained within the "experiment-based" approach. This is an evidence of low adequateness of the "quasi-models" used.

(2) The "quasi-model" based on GRNN provided the results on the experimental examination dataset that outperformed the results provided by the "quasi-model" based on perceptron.

Table 5. Some statistics for NN determination of concentrations of salts in multi-component solutions by Raman spectra within the "experiment-based" and "quasi-model" approaches.

Notation for the datasets: EE—experimental examination, QE—quasi-model examination. Notation for the approaches: Exp—"experiment-based" approach, QMP and QMGR—"quasi-model" approach with "quasi-model" based on perceptron and GRNN, respectively

Approach	Dataset	NaCl	NH_4Br	Li_2SO_4	KNO_3	CsI
R^2						
Exp	EE	0.990	0.993	0.996	0.996	0.994
QMP	EE	0.969	0.985	0.983	0.986	0.987
QMGR	EE	0.980	0.986	0.992	0.991	0.989
QMP	QE	0.989	0.995	0.990	0.991	0.996
QMGR	QE	0.990	0.992	0.991	0.991	0.992
MSE, M						
Exp	EE	0.041	0.036	0.029	0.031	0.036
QMP	EE	0.070	0.051	0.057	0.054	0.051
QMGR	EE	0.057	0.050	0.040	0.042	0.047
QMP	QE	0.041	0.027	0.037	0.038	0.026
QMGR	QE	0.039	0.034	0.037	0.037	0.035
MAE, M						
Exp	EE	0.029	0.024	0.020	0.019	0.023
QMP	EE	0.050	0.037	0.041	0.035	0.037
QMGR	EE	0.042	0.034	0.029	0.029	0.031
QMP	QE	0.030	0.021	0.027	0.028	0.019
QMGR	QE	0.029	0.025	0.027	0.027	0.025
MRE, %						
Exp	EE	9.2	5.8	5.6	4.3	5.9
QMP	EE	17.4	9.3	10.4	8.2	9.3
QMGR	EE	12.5	8.8	7.7	6.8	7.5
QMP	QE	10.4	7.2	8.6	8.9	6.2
QMGR	QE	11.3	9.5	9.5	9.5	9.1

(3) The results obtained within the "quasi-model" approach on the "quasi-model" examination dataset, are comparable to those obtained within the "experiment-based" approach on the experimental examination set, sometimes outperforming them. This is the manifestation of higher representativity of datasets in the "quasi-model" approach.

CONCLUSIONS

(1) Unique experimental material has been obtained. This material is an array of Raman spectra (in the frequency range from 200 cm^{-1} to 4000 cm^{-1}) of water solutions of inorganic salts (NaCl, NH_4Br, Li_2SO_4, KNO_3, CsI) in the range of total concentrations from 0 to 2.5 M (mole per liter of solution).

(2) Based on the obtained array of spectra, a special procedure providing optimal representativity of datasets was used to form training, test, and examination datasets for subsequent work.

(3) The complex IP of identification of salts and determination of their partial concentrations in 5-component water solution by Raman spectra was solved within the "experiment-based" approach using both bands of Raman spectrum (low-frequency and valence bands), as well as using only water Raman valence band, and also within the "quasi-model" approach using both bands of the spectrum. When experimental Raman spectra used as input data included, besides the valence Raman band of water (2700–3900 cm^{-1}), also the low-frequency part of Raman spectrum of water and ions of the dissolved salts (280–1830 cm^{-1}), the obtained values of the error of determination of concentration on the examination dataset were low enough: the mean absolute error was from 0.019 to 0.029 M in the concentration range

from 0 to 2.5 M. These results significantly outperform the results obtained within the same experiment by water Raman valence band only, and they outperform the results obtained before for 3-component solutions in narrower concentration range several fold (Table 3).

(4) Quasi-models calculated by experimental spectra were built based on perceptron and on GRNN. The quality of approximation of the sought-for dependence was similar for both quasi-models. Quasi-models based on perceptron and GRNN were used to form training, test, and examination datasets with necessary representativity for the "quasi-model" approach to the IP solution.

(5) The "quasi-model" approach failed to meet expectations compared to the "experiment-based" approach. In all cases, the results on the experimental examination dataset, obtained within the "quasi-model" approach, turned out to be worse or significantly worse than the results obtained within the "experiment-based" approach (Table 5). This is an evidence of the fact that it is necessary to elaborate more complex "quasi-models", adequate to the complexity of the solved direct problem.

ACKNOWLEDGMENTS

REFERENCES

1. Dolenko, S.A., Dolenko, T.A., Persiantsev, I.G., Fadeev, V.V., and Burikov, S.A., Solution of Inverse Problems of Optical Spectroscopy Using Neural Networks, *Neurocomputers: Elaboration, Application*, 2005, nos. 1–2, pp. 89–97 [in Russian].

2. Baldwin, S.F. and Brown, C.W., Detection of Ionic Water Pollutants by Laser Excited Raman Spectroscopy (short communication), *Water Research*, 1972, vol. 6, pp. 1601–1604.

3. Rudolph, W.W. and Irmer, G., Raman and Infrared Spectroscopic Investigation on Aqueous Alkali Metal Phosphate Solutions and Density Functional Theory Calculations of Phosphate-Water Clusters, *Applied Spectroscopy*, 2007, vol. 61, no. 12, pp. 274A–292A.

4. Georgiev, G.M., Kalkanjiev, T.K., Petrov, V.P., and Nickolov, Zh., Determination of Salts in Water Solutions by a Skewing Parameter of the Water Raman Band, *Applied Spectroscopy*, 1984, vol. 38, no. 4, pp. 593–595.

5. Furic, K., Ciglenecki, I., and Cosovic, B., Raman Spectroscopic Study of Sodium Chloride Water Solutions, *J. Molecular Structure*, 2000, vol. 6, pp. 225–234.

6. Dolenko, T.A., Churina, I.V., Fadeev, V.V., and Glushkov, S.M., Valence Band of Liquid Water Raman Scattering: Some Peculiarities and Applications in the Diagnostics of Water Media, *J. Raman Spectroscopy*, 2000, vol. 31, pp. 863–870.

7. Bekkiev, A.Yu., Gogolinskaya, T.A., and Fadeev, V.V., Simultaneous Determination of Temperature and Salinity of Sea Water Using the Laser Raman Spectroscopy Method, *Sov. Phys. Dokl.*, 1983, vol. 271, no. 4, pp. 849–853.

8. Burikov, S.A., Dolenko, T.A., Fadeev, V.V., and Sugonyaev, A.V., New Opportunities in the Determination of Inorganic Compounds in Water by the Method of Laser Raman Spectroscopy, *Laser Physics*, 2005, vol. 15, no. 8, pp. 1–5.

9. Burikov, S.A., Dolenko, T.A., Fadeev, V.V., and Sugonyaev, A.V., Identification of Inorganic Salts and Determination Their Concentrations in Water Solutions Above the Water Raman Valence Band Using Artificial Neural Networks, *Pattern Recognition and Image Analysis*, 2005, vol. 15, no. 2, pp. 520–523.

10. Burikov, S.A., Dolenko, T.A., and Fadeev, V.V., Identification of Inorganic Salts and Determination Their Concentrations in Multi-components Water Solutions Over Water Raman Valence Band Using Artificial Neural Networks, *Neurocomputers: Elaboration, Application*, 2007, no. 5, pp. 62–72 [in Russian].

11. Burikov, S.A., Dolenko, T.A., and Fadeev, V.V., Identification of Inorganic Salts and Determination of Their Concentrations in Water Solutions from the Raman Valence Band Using Artificial Neural Networks, *Pattern Recognition and Image Analysis,* 2007, vol. 17, no. 4, pp. 554–559.

12. Specht, D., A General Regression Neural Network, *IEEE Trans. on Neural Networks*, 1991, vol. 2, no. 6, pp. 568–576.

13. Madala, H.R. and Ivakhnenko, A.G., *Inductive Learning Algorithms for Complex Systems Modeling*, CRC Press, 1994.