

**Использование проекционных методов  
многомерного регрессионного анализа  
для оценки концентраций химических веществ  
в многокомпонентных водных растворах**

*А.Ю. Суковатова*

Алтайский государственный университет (Барнаул, Россия)

**On Application of Multivariate Regression Analysis  
Projective Methods for Estimating Chemical Substances  
Concentrations in Multicomponent Water Solutions**

*A.Yu. Sukovatova*

Altai State University (Barnaul, Russia)

Приведены результаты исследования диэлектрических свойств снеговой воды с малой концентрацией растворенных веществ и воды, отобранной из высокоминерализованных озер равнинной части Алтайского края. С использованием проекционных методов многомерного регрессионного анализа на основе экспериментальных данных построены регрессионные зависимости диэлектрических свойств многокомпонентных растворов от массовых концентраций химических веществ и частоты, на которой измерены диэлектрические свойства.

Построены регрессионные зависимости действительной и мнимой частей комплексной диэлектрической проницаемости от концентраций химических веществ, входящих в состав растворов. Исследовали следующие химические вещества:  $\text{CO}_3^{2-}$ ,  $\text{HCO}_3^-$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{NO}_2^-$ ,  $\text{NO}_3^-$ ,  $\text{PO}_4^{3-}$ ,  $\text{F}^-$  и др. Полученные в работе зависимости могут быть использованы для расчета численных значений комплексной диэлектрической проницаемости по известным значениям массовых концентраций. Построение подобных зависимостей позволит контролировать изменение состояния исследуемых объектов и осуществлять мониторинг экологического состояния территории.

**Ключевые слова:** многомерный анализ, диэлектрические характеристики, многокомпонентные водно-солевые растворы.

DOI 10.14258/izvasu(2014)1.2-20

**Введение.** При исследовании природных объектов важно располагать оперативными и достоверными данными об их состоянии. Таковую информацию можно получить, в частности, с помощью радиофизических систем космического базирова-

This paper presents results of a study of dielectric properties of snow water with low solutes concentrations and of water from strongly mineralized flat land lakes of Altai region. Experimental data are processed with multivariate regression analysis projective methods, and regression dependence of dielectric properties on mass concentrations of chemical substances and measured frequency for multicomponent solutions is elaborated.

In the paper, regression dependencies of real and imaginary parts of the complex permittivity on mass concentrations of chemicals in solutions are presented. The study is conducted for the following chemicals:  $\text{CO}_3^{2-}$ ,  $\text{HCO}_3^-$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{NO}_2^-$ ,  $\text{NO}_3^-$ ,  $\text{PO}_4^{3-}$ ,  $\text{F}^-$ , etc. Numerical values of the complex dielectric constant can be calculated with the proposed dependencies for the known values of mass concentration. Thus, the dependencies can be used in monitoring of environmental situation and state changes of natural objects.

**Key words:** multivariate analysis, dielectric properties, multicomponent water saline solutions.

ния. Изучение радиоизлучательных свойств водных растворов является основой для разработки дистанционных методов мониторинга природных объектов. При дистанционной диагностике природной среды возникает необходимость учета химических

параметров исследуемого объекта, оказывающих влияние на его физические свойства.

Для оценки уровня загрязненности территории и пространственного распределения атмосферных загрязнителей требуется отбирать значительное количество проб с их последующим химическим анализом. Снижение трудоемкости работ возможно путем разработки и внедрения оперативных электрофизических методов диагностики природных сред.

Большинство природных систем являются многопараметрическими. При их изучении необходимо учитывать одновременное изменение множества параметров, что является сложной задачей. Часто в данной ситуации исследователи поступают следующим образом: фиксируют один из параметров, изменяя все остальные, затем данная операция повторяется для остальных параметров. Подобные действия ведут к возникновению больших объемов данных, изучение которых затруднительно. Чтобы избежать возникновения данной проблемы, автор предлагает использовать проекционные методы многомерного анализа данных для моделирования состояния природных объектов [1].

Приведены результаты исследования диэлектрических свойств снеговой воды с малой концентрацией растворенных веществ и воды, отобранной из высокоминерализованных озер равнинной части Алтайского края. На основе экспериментальных данных построены регрессионные зависимости диэлектрических свойств воды от массовых концентраций химических веществ и частоты, на которой измерены диэлектрические свойства.

**Методика эксперимента.** Диэлектрические характеристики воды измеряли в дециметровом диапазоне, в котором наблюдаются заметные различия в диэлектрических свойствах разных водных фракций, образующихся в результате взаимодействия с ионами растворенных солей. Использовали метод мостовых схем, обладающий высокой точностью определения диэлектрических параметров дисперсных смесей и жидкостей, характеризующихся высокой поглощательной способностью. Лабораторная установка, выполненная с использованием измерительного и опорного каналов промышленного фазометра ФК2-18, позволяла измерять диэлектрические параметры жидкостей в диапазоне частот от 300 МГц до 5,0 ГГц и имела следующие технические характеристики: диапазон однозначных измерений фазы  $\pm 180^\circ$ ; диапазон измерения затуханий от 0 до 60 дБ; погрешность измерений ослабления 0,5 дБ, фазы —  $2^\circ$ .

В качестве величины, характеризующей содержание химических веществ в воде, использовали массовую долю  $S$  растворенного вещества, которую определяли по формуле  $S = M_c / M$ , где  $M$ ,  $M_c$  — массы образца и сухого остатка, измеренные путем взвешивания на аналитических весах с точностью до 0,001 г

до и после выпаривания воды из образца в термостате при температуре  $215^\circ\text{C}$ . Измеряли на разных частотах ослабление  $A$  и сдвиг фазы  $\varphi$  электромагнитной волны в исследуемом образце, используемые для расчета действительной ( $\varepsilon'$ ) и мнимой ( $\varepsilon''$ ) частей комплексной диэлектрической проницаемости (КДП)  $\varepsilon = \varepsilon' + i\varepsilon''$ , а также показателей преломления  $n$  и поглощения  $\kappa$ , связанных с  $\varepsilon$  следующей формулой:  $\sqrt{\varepsilon} = n + i\kappa$ . Измерения диэлектрических параметров воды проводили при температуре  $25 \pm 1^\circ\text{C}$  [2, 3].

В данной работе для построения регрессионных зависимостей использованы проекционные методы многомерного анализа данных: проекция на латентные структуры (ПЛС) и регрессия на главные компоненты (РГК).

**Результаты эксперимента и обсуждение.** Построены регрессионные зависимости действительной и мнимой частей КДП от массовых долей химических веществ ( $S_i$ ) и частоты сигнала  $f$  (0,6, 1,11, 1,42 ГГц) следующего вида:

$$\varepsilon' = \sum_{i=1}^n \alpha_i S_i + \gamma f,$$

$$\varepsilon'' = \sum_{i=1}^n \beta_i S_i + \lambda f.$$

Проверку полученных моделей проводили на тестовых наборах данных, не участвовавших в построении модели. Коэффициенты построенных моделей приведены в таблицах. В приведенных таблицах  $S_i$  — концентрации химических веществ, растворенных в воде.

Химические вещества в таблицах:  $S_1$  —  $\text{Cl}^-$ ,  $S_2$  —  $\text{SO}_4^{2-}$ ,  $S_3$  —  $\text{NO}_3^-$ ,  $S_4$  —  $\text{F}^-$ ,  $S_5$  —  $\text{PO}_4^{3-}$ ,  $S_6$  —  $\text{NO}_2^-$ . Исследованные величины изменялись в следующих пределах (мг/л):  $S_1 = 0.6 \div 1.5$ ,  $S_2 = 1.6 \div 3.7$ ,  $S_3 = 1.8 \div 3.1$ ,  $S_4 = 0.02 \div 0.07$ ,  $S_5 < 0.03$ ,  $S_6 < 0.2$ ,  $f = 0.6 \div 1.42$  ГГц.

Химические вещества в таблицах 4–6:  $S_1$  —  $\text{CO}_3^{2-}$ ,  $S_2$  —  $\text{HCO}_3^-$ ,  $S_3$  —  $\text{Cl}^-$ ,  $S_4$  —  $\text{SO}_4^{2-}$ ,  $S_5$  —  $\text{Ca}^{2+}$ ,  $S_6$  —  $\text{Mg}^{2+}$ ,  $S_7$  —  $\text{NO}_2^-$ ,  $S_8$  —  $\text{N-NO}_2$ ,  $S_9$  —  $\text{NO}_3^-$ ,  $S_{10}$  —  $\text{N-NO}_3$ ,  $S_{11}$  —  $\text{N-NH}_4$ ,  $S_{12}$  —  $\text{P}$ .

Исследованные величины изменялись в следующих пределах:  $S_7 = 0 \div 888$  мг/л,  $S_2 = 0.001 \div 3568.5$  мг/л,  $S_3 = 0.001 \div 194597.2$  мг/л,  $S_4 = 0.001 \div 36022.5$  мг/л,  $S_5 = 0.001 \div 4000$  мг/л,  $S_6 = 0.001 \div 63360$  мг/л,  $S_7 = 0.001 \div 9.03$  мг/л,  $S_8 = 0.001 \div 2.708$  мг/л,  $S_9 = 0.001 \div 0.08$  мг/л,  $S_{10} = 0.001 \div 0.02$  мг/л,  $S_{11} = 0.001 \div 516.41$  мг/л,  $S_{12} = 0.001 \div 366$  мг/л,  $S_{13} = 0.6 \div 1.42$  ГГц.

При анализе полученных результатов выяснили, что модели, построенные методом ПЛС, имеют меньшую размерность и значение ошибки прогноза [1].

Химический анализ исследуемых образцов выполнен в химико-аналитическом центре ИВЭП СО РАН.

Полученные зависимости могут быть использованы для расчета численных значений КДП

Таблица 1

Регрессионные коэффициенты модели показателя преломления снеговой воды, метод ПЛС

b0	f	S <sub>1</sub> (Cl <sup>-</sup> )	S <sub>2</sub> (SO <sub>4</sub> <sup>2-</sup> )	S <sub>3</sub> (NO <sub>3</sub> <sup>-</sup> )	S <sub>4</sub> (F <sup>-</sup> )	S <sub>5</sub> (PO <sub>4</sub> <sup>-3</sup> )	S <sub>6</sub> (NO <sub>2</sub> <sup>-</sup> )
8,86	0,57	-0,0001	-0,014	0,06	0,016	-0,13	-0,05

Значение ошибки прогноза: 0,06.

Таблица 2

Регрессионные коэффициенты модели показателя поглощения снеговой воды, метод ПЛС

b0	f	S <sub>1</sub> (Cl <sup>-</sup> )	S <sub>2</sub> (SO <sub>4</sub> <sup>2-</sup> )	S <sub>3</sub> (NO <sub>3</sub> <sup>-</sup> )	S <sub>4</sub> (F <sup>-</sup> )	S <sub>5</sub> (PO <sub>4</sub> <sup>-3</sup> )	S <sub>6</sub> (NO <sub>2</sub> <sup>-</sup> )
0,56	0,0003	0,0021	0,0002	-0,0000	0,0000	-0,0001	-0,0000

Значение ошибки прогноза: 1,09.

Таблица 3

Регрессионные коэффициенты модели показателя преломления воды из высокоминерализованных озер, метод ПЛС

b0	f	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	S <sub>4</sub>	S <sub>5</sub>
9,08	-0,29	0,00032	-0,000062	-0,00001	-0,000015	0,00034

Продолжение таблицы 3

S <sub>6</sub>	S <sub>7</sub>	S <sub>8</sub>	S <sub>9</sub>	S <sub>10</sub>	S <sub>11</sub>	S <sub>12</sub>
-0,000004	0,0068	0,023	5,24	21,81	0,002	0,23

Значение ошибки прогноза: 2,11.

Таблица 4

Регрессионные коэффициенты модели показателя поглощения воды из высокоминерализованных озер, метод ПЛС

b0	f	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	S <sub>4</sub>	S <sub>5</sub>
3,39	-0,05	-0,0012	-0,00028	0,000004	0,000018	0,00026

Продолжение таблицы 4

S <sub>6</sub>	S <sub>7</sub>	S <sub>8</sub>	S <sub>9</sub>	S <sub>10</sub>	S <sub>11</sub>	S <sub>12</sub>
0,000011	0,069	0,23	-1,61	-6,7	0,0012	0,0069

Значение ошибки прогноза: 4,82.

по известным значениям массовых концентраций. Решение обратной задачи (определение концентраций по известным значениям КДП, измеренных на разных частотах) позволит контролировать изменение

состояния исследуемого объекта, в частности, изменение минерализации водоема и уровень загрязнения снежного покрова, и таким образом, осуществлять диагностику экологического состояния территории.

## Библиографический список

1. Эсбенсен К., Анализ многомерных данных. Избр. гл. — Черноголовка, 2005.
2. Романов А.Н., Суковатова А.Ю., Оскорбин Н.М. Моделирование диэлектрических свойств воды из природных минерализованных водоемов с использованием регрессионного анализа // Оптика атмосферы и океана. — 2011. — Т. 24, №11.
3. Романов А.Н., Суковатова А.Ю., Рапута В.Ф. Моделирование диэлектрических свойств снеговой воды с малой концентрацией растворенных веществ // Оптика атмосферы и океана. — 2013. — Т. 26, № 7.