

Геоинформатика. 2022. № 1. С. 32–39.  
*Geoinformatika*. 2022;(1):32–39.

### Моделирование геообъектов и геопроцессов

Научная статья

УДК 504.064.2.001.18

<https://doi.org/10.47148/1609-364X-2022-1-32-39>

## Метод встречного прогнозирования пространственного ряда на примере содержания пыли в снеговом покрове

© 2022 г. — А.С. Буторова<sup>1, 2, а)</sup>, А.П. Сергеев<sup>1, б)</sup>, А.В. Шичкин<sup>1, в)</sup>, А.Г. Бувич<sup>1, д)</sup>, Е.М. Баглаева<sup>1, е)</sup>, М.В. Сергеева<sup>1, ф)</sup>

<sup>1</sup>Федеральное государственное бюджетное учреждение науки Институт промышленной экологии Уральского отделения Российской академии наук (ИПЭ УрО РАН); Россия, Екатеринбург

<sup>2</sup>Федеральное государственное автономное образовательное учреждение высшего образования «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина» (ФГАОУ ВО «УрФУ им. Б.Н. Ельцина»); Россия, Екатеринбург

<sup>а)</sup>автор-корреспондент: [a.s.butorova@urfu.ru](mailto:a.s.butorova@urfu.ru), <sup>б)</sup>[sergeev@ecko.uran.ru](mailto:sergeev@ecko.uran.ru), <sup>в)</sup>[and@ecko.uran.ru](mailto:and@ecko.uran.ru), <sup>д)</sup>[bag@ecko.uran.ru](mailto:bag@ecko.uran.ru),

<sup>е)</sup>[e.m.baglaeva@urfu.ru](mailto:e.m.baglaeva@urfu.ru), <sup>ф)</sup>[marin@ecko.uran.ru](mailto:marin@ecko.uran.ru)

**Аннотация:** В работе предложен оригинальный подход для прогнозирования значений пространственно-координатного ряда. Такой подход может использоваться, в частности, для восстановления пропущенных данных. Метод встречного прогнозирования опробован на модели искусственной нейронной сети (ИНС), которая последовательно обучается на значениях, предшествующих прогнозируемому отрезку ряда слева и справа. Окончательным прогнозом модели являются взвешенно-усредненные результаты этих двух наборов. Мы проверили работу метода на примере прогнозирования содержания пыли в снеговом покрове. 256 проб снега были отобраны с шагом 0,2 м по линии в районе отвалов действующего карьера по добыче медной руды. Для проверки точности моделей на основе полученных данных были созданы два пространственно-координатных ряда: ряд измеренных значений (измеренные значения как есть) и перемешанный ряд (случайным образом перемешанные значения ряда измеренных значений). Прогноз с наименьшими ошибками и наибольшим коэффициентом корреляции был получен для ряда измеренных значений. Наименее точный прогноз был получен для перемешанного ряда. RMSE для ряда измеренных значений оказалось на 58% меньше, чем для перемешанного ряда, среднее значение коэффициента корреляции составило 0,3 для ряда измеренных значений и –0,06 для перемешанного ряда.

**Ключевые слова:** метод встречного прогнозирования, восстановление пропущенных данных, искусственные нейронные сети, снеговой покров, пыль

Для цитирования: Буторова А.С., Сергеев А.П., Шичкин А.В., Бувич А.Г., Баглаева Е.М., Сергеева М.В. Метод встречного прогнозирования пространственного ряда на примере содержания пыли в снеговом покрове // Геоинформатика. — 2022. — № 1. — С. 32–39. <https://doi.org/10.47148/1609-364X-2022-1-32-39>.

### Modeling of geological objects and geoprocesses

Original article

## Counter-prediction method of the spatial series on the example of the dust content in the snow cover

© 2022 — A.S. Butorova<sup>1, 2, а)</sup>, A.P. Sergeev<sup>1, б)</sup>, A.V. Shichkin<sup>1, в)</sup>, A.G. Buevich<sup>1, д)</sup>, E.M. Baglaeva<sup>1, е)</sup>, M.V. Sergeeva<sup>1, ф)</sup>

<sup>1</sup>Institute of Industrial Ecology UB RAS; Ekaterinburg, Russia

<sup>2</sup>Ural Federal University named after B.N. Yeltsin; Ekaterinburg, Russia

<sup>а)</sup>corresponding author: [a.s.butorova@urfu.ru](mailto:a.s.butorova@urfu.ru), <sup>б)</sup>[sergeev@ecko.uran.ru](mailto:sergeev@ecko.uran.ru), <sup>в)</sup>[and@ecko.uran.ru](mailto:and@ecko.uran.ru), <sup>д)</sup>[bag@ecko.uran.ru](mailto:bag@ecko.uran.ru), <sup>е)</sup>[e.m.baglaeva@urfu.ru](mailto:e.m.baglaeva@urfu.ru), <sup>ф)</sup>[marin@ecko.uran.ru](mailto:marin@ecko.uran.ru)

**Abstract:** The paper proposes an original approach for predicting the values of the spatial series. This approach can be used, in particular, to recover missing data. The counter-prediction method was tested on a model of an artificial neural network (ANN), which is sequentially trained on the values preceding the predicted segment of the series on the left and right. The final prediction of the model is the weighted average of the results of these two sets. We have tested the work of the method using the example of predicting the dust content in the snow cover. 256 snow samples were taken with a step of 0.2 m along the line in the area of the dumps of the existing open pit for the extraction of copper ore. To check the accuracy of the models, based on the data obtained, two spatial series were created: a series of measured values (measured values as they are) and a mixed series (randomly mixed values of a series of measured values). The forecast with the minimum errors and the maximum correlation coefficient was obtained for a number of measured values. The least accurate forecast was obtained for a mixed series. RMSE for a series of measured values was 58% less than RMSE for a mixed series, an average value of the correlation coefficient was 0.3 for a series of measured values and –0.06 for a mixed series.

**Key words:** counter-prediction method, missing data recovering, artificial neural networks, snow cover, dust

For citation: Butorova A.S., Sergeev A.P., Shichkin A.V., Buevich A.G., Baglaeva E.M., Sergeeva M.V. Counter-prediction method of the spatial series on the example of the dust content in the snow cover. *Geoinformatika*. 2022;(1):32–39. <https://doi.org/10.47148/1609-364X-2022-1-32-39>. In Russ.

**Введение**

Прогнозы пространственно-координатного распределения признака, основанные на моделях машинного обучения в том числе на моделях ИНС, в последние годы все чаще используют исследователи в разных областях естествознания [7, 15, 23, 27]. Такие модели часто превосходят в точности методы, которые основаны на детерминистском подходе и на геостатистике [5, 13, 25]. Особенно это относится к прогнозированию сложных динамических процессов, в том числе в области экологии, метеорологии и т.п. [3, 4, 8-10, 12, 14, 19-21, 28, 30]. Сюда же можно отнести прогнозирование динамики временных рядов [6, 14, 22].

Часто исследователи сталкиваются с задачами по изучению влияния какого-либо объекта на наблюдаемую среду. В процессе сбора данных для анализа воздействия не всегда можно увидеть полную картину пространственно-координатного распределения того или иного изучаемого признака [2, 11, 24, 29]. Пропущенные области данных заполняются при помощи прогностических моделей, обучаемых на данных слева и справа от пропущенной области. Для таких задач могут подойти ИНС, используемые для прогнозирования временных рядов: например, ИНС типа NARX [1, 16-18, 26]. Но в отличие от временного ряда, где предсказание – всегда экстраполяция, для таких задач модель будет работать как интерполятор, взвешенно усредняя экстраполяции слева и справа.

Разумный прогноз может быть получен для пространственно-координатного ряда, в котором

присутствуют определенные зависимости. Извлекая эту информацию на стадии обучения, модель будет способна выдавать разумный прогноз. Мы предполагаем, что модель на основе ИНС может дать разумный прогноз, поскольку пространственно-координатный ряд, основанный на измеренных значениях (например, содержании примесей в осаждающей среде), должен обладать этим свойством, так как распределения его значений сложились под действием определенных природных и техногенных факторов.

Целью исследования было проверить предложенный подход на примере прогнозирования отрезков пространственно-координатных рядов, полученных на основе измеренных данных о содержании пыли в снеговом покрове рядом с действующим карьером по добыче медной руды.

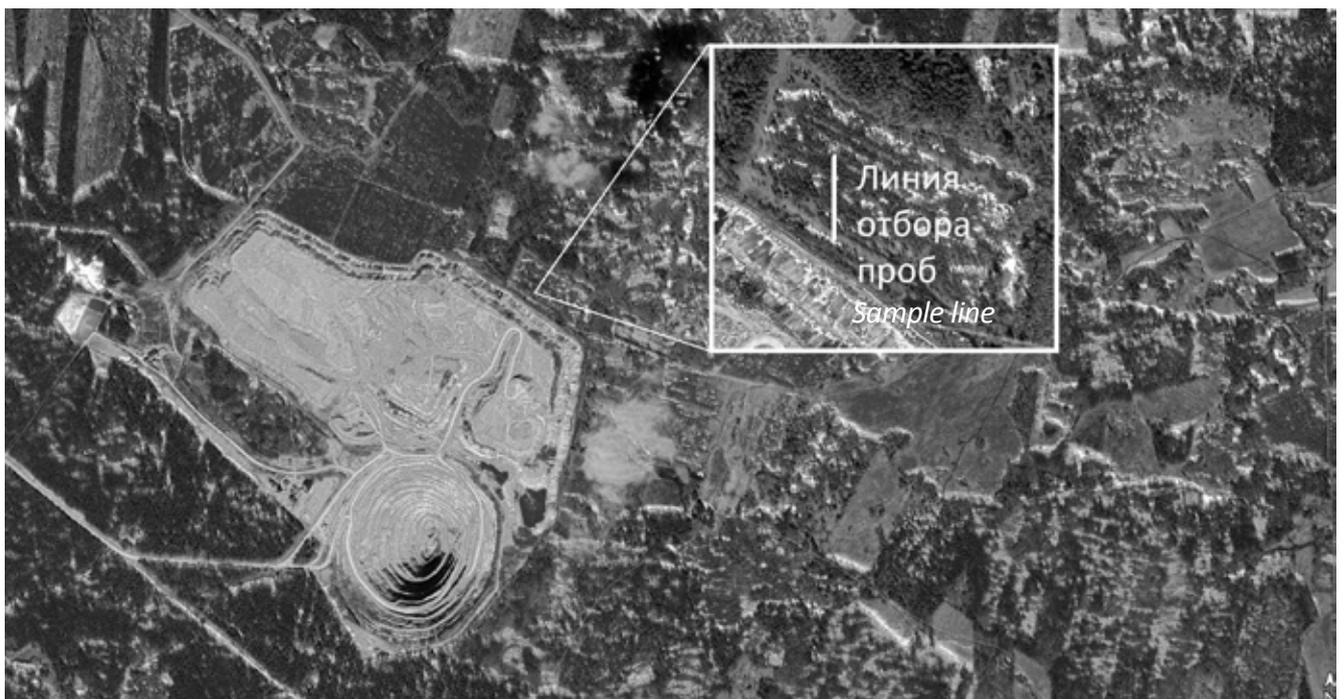
**Материалы и методы**

Для исследования были использованы результаты анализа проб снегового покрова. Пробы были отобраны по профилю в направлении юг-север в 100 метрах от отвалов в северной части действующего карьера по добыче медной руды (рис. 1). Карьер расположен рядом с городом Реж Свердловской области.

Всего было отобрано 256 проб снега. Для моделирования использовались данные о содержании взвешенных веществ (пыли) в снеговой пробе. Пробы отбирались специальным пробоотборником – снегомером ВС-43 диаметром 83 мм по линии, которая была ориентирована с юга на север. Первый

**Рис. 1.** Место проведения исследования (Google Earth)

*Fig. 1. Research location (Google Earth)*



кern снега был взят на расстоянии 70 метров от нижней границы отвала рудника. Проба представляла собой один kern со снегом на полную глубину снегового покрова. Глубина снегового покрова зависела от рельефа подстилающей поверхности и менялась от 0,35 м до 0,61 м. Для исключения попадания частиц почвы в пробу последние 2 см снега отбрасывались. Пробы отбирались максимально близко друг к другу, но так, чтобы не было обрушения стенки снегового покрова между соседними kernами. Опытным путем был определен шаг между соседними kernами, который составил 0,2 м. Общая длина линии отбора проб составила 51 м.

Каждая проба упаковывалась в двойной полиэтиленовый пакет и маркировалась уникальным номером. Отобранные пробы были проанализированы в аккредитованной лаборатории. Измерялось содержание пыли в сухом остатке.

Таким образом был получен пространственно-координатный ряд, который состоял из 256 значений содержания пыли в снеговом покрове.

Для прогнозирования использовалась нелинейная авторегрессионная нейронная сеть (NAR). Такая сеть создается и обучается в разомкнутом цикле, используя в качестве целевых значений измеренные содержания пыли в снеговом покрове. После обучения цикл замыкается, предсказанные значения используются для подачи в сеть новых входных сигналов с некоторой временной задержкой. Для обучения использовался алгоритм Левенберга – Марквардта. Искусственная нейронная сеть NAR, обученная на начальном отрезке длиной 240 значений, предсказывала последние 16 значений исходного ряда. Количество нейронов в скрытом слое изменялось от 20 до 40, временная задержка изменялась от 2 до 6. Для каждой такой структуры было построено 500 сетей; таким образом, всего было построено 52500 сетей. Для дальнейшей работы была выбрана структура сети с минимальной ошибкой RMSE, которая имела один скрытый слой с 35 нейронами и временной задержкой равной 6.

Предложенный нами подход заключается в том, что прогноз для интересующего отрезка пространственно-координатного ряда строится на основе результатов двух предсказаний. Первый прогноз получается при обучении на значениях ряда, которые расположены слева от прогнозируемого отрезка. Второй прогноз получается при обучении на значениях ряда, расположенных справа от прогнозируемого отрезка. Затем оба соответствующих полученных значения («левый» и «правый») взвешенно усредняются. В результате мы получаем про-

гнозируемый отрезок ряда из встречных усредненных значений.

Для оценки точности предложенного подхода мы провели сравнение результатов моделирования для двух координатных рядов, созданных из измеренных значений. Первый – ряд измеренных значений, который состоял из измеренных значений содержания пыли в снеговом покрове по профилю. Второй – перемешанный ряд, который состоит из случайным образом перемешанных значений ряда измеренных значений.

Для оценки точности прогноза модели NAR использовались RMSE (1) и коэффициент корреляции Спирмена (2).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (p - o)^2}{n}}, \quad (1)$$

где  $p$  – прогнозируемая концентрация,  $o$  – измеренная концентрация,  $n$  – количество точек.

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}, \quad (2)$$

где  $d_i$  – разность между рангами сопряженных значений признаков,  $n$  – количество точек.

### Результаты и обсуждение

Для прогнозирования всегда использовались 16 значений ряда. Обучающие наборы для первого прогнозируемого отрезка (16 значений с 9-го по 24-й) состояли из 8 первых значений ряда для прогноза слева и 232 последних для прогноза справа. Предсказываемый отрезок с каждым шагом сдвигался еще на 8 значений ряда. На последнем шаге обучающие наборы состояли из 232 первых значений и 8 последних значений. Таким образом, было получено два встречных прогноза для каждого отрезка ряда длиной 16 значений, один прогноз справа и один слева. Затем соответствующие предсказанные значения (первое значение слева с последним значением справа и т.д.) суммировались и делились пополам. В результате мы получили 16 усредненных предсказанных значений. Затем еще 16 значений со сдвигом на 8 и так далее. Всего было предсказано 29 отрезков. Этот алгоритм был повторен для каждого из рядов.

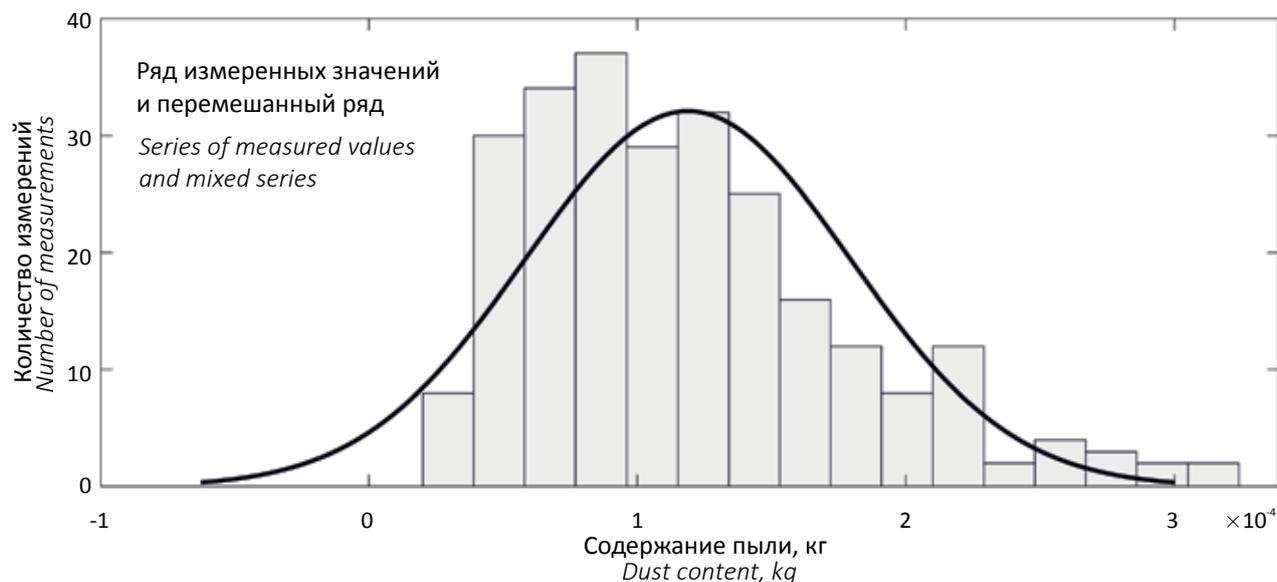
В таблице 1 представлены дескриптивные статистики для исследуемых пространственно-координатных рядов. Дескриптивные статистики для исходного и перемешанного рядов одинаковы, так

Табл. 1. Дескриптивные статистики значений пространственно-координатных рядов

Tab. 1. Descriptive statistics of spatial series values

Ряды	min, кг	max, кг	mean, кг	median, кг	SD	Skewness	Kurtosis
Исходный	0,000025	0,00030	0,00012	0,00011	0,00006	0,897	3,44
Перемешанный	0,000025	0,00030	0,00012	0,00011	0,00006	0,897	3,44

**Рис. 2.** Гистограмма распределения значений содержания пыли для моделируемых рядов

*Fig. 2. Histogram of the distribution of dust content values for the modelled series*


как значения ряда не менялись, но были случайным образом перемешаны.

В таблице 2 показаны средние значения RMSE и коэффициента корреляции для всех прогнозируемых отрезков обоих моделируемых рядов.

Таблица 2 показывает, что лучший результат прогноза с минимальными ошибками был получен для ряда измеренных значений (жирный шрифт). RMSE для этого ряда было на 58% меньше, чем для перемешанного ряда. Разница между средним RMSE

**Табл. 2.** Оценка точности прогноза для пространственно-координатных рядов

*Tab. 2. Estimation of forecast accuracy for spatial series*

Ряды	RMSE, кг	Коэффициент корреляции
Ряд измеренных значений (прогноз слева)	0,000060	0,247
Ряд измеренных значений (прогноз справа)	0,000053	0,243
Ряд измеренных значений (общий прогноз)	<b>0,000045</b>	<b>0,303</b>
Перемешанный ряд (прогноз слева)	0,000130	-0,052
Перемешанный ряд (прогноз справа)	0,000114	-0,059
Перемешанный ряд (общий прогноз)	0,000106	-0,059

для прогнозов слева и справа составляла 11,7% для ряда измеренных значений и 12,5% для перемешанного ряда.

На рисунке 3 и 4 проиллюстрированы результаты, полученные при предсказании содержания пыли для каждого прогнозируемого отрезка. Приведены усредненные по 16 значениям коэффициенты корреляции и RMSE для каждого предсказанного отрезка. Индексы приведены отдельно для предсказаний слева и справа, а также для окончательного прогноза. Индексы для первого отрезка на рисунках являются усредненными для модели, обученной на первых восьми измеренных значениях (для прямого направления), для модели, обученной на первых 232 значениях (в обратном направлении) и для усредненного общего прогноза (сумм соответствующих двух встречных значений для этого отрезка, деленных пополам).

Самый высокий коэффициент корреляции был получен для ряда измеренных значений и в среднем составляет 0,3. Для перемешанного ряда значения коэффициентов корреляции колеблются около 0 (в среднем  $-0,06$  для ряда).

Величина и размах изменения ошибки были наименьшими для ряда измеренных значений и наибольшими для перемешанного ряда. На рисунках хорошо заметно, что при большой разнице в длинах обучающих выборок ошибка встречного прогноза больше (левый и правый края рисунков). Это можно объяснить большой величиной ошибки для предсказания, полученного при обучении модели на малом количестве значений. Когда количество значений для обучающей выборки становится достаточным (в нашем случае это примерно 88 для ряда измеренных значений), а длины обучающих

Рис. 3. Усредненные коэффициенты корреляции для пространственных рядов

Fig. 3. Average correlation coefficients for spatial series

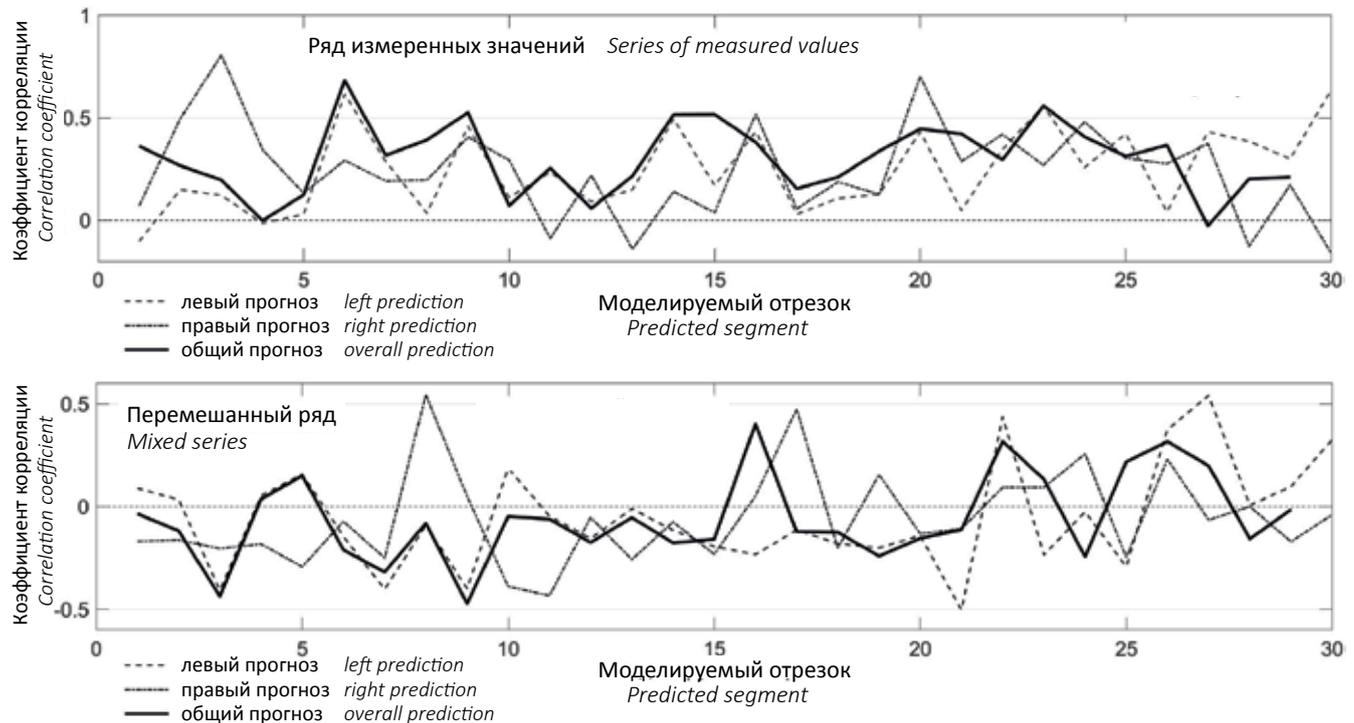
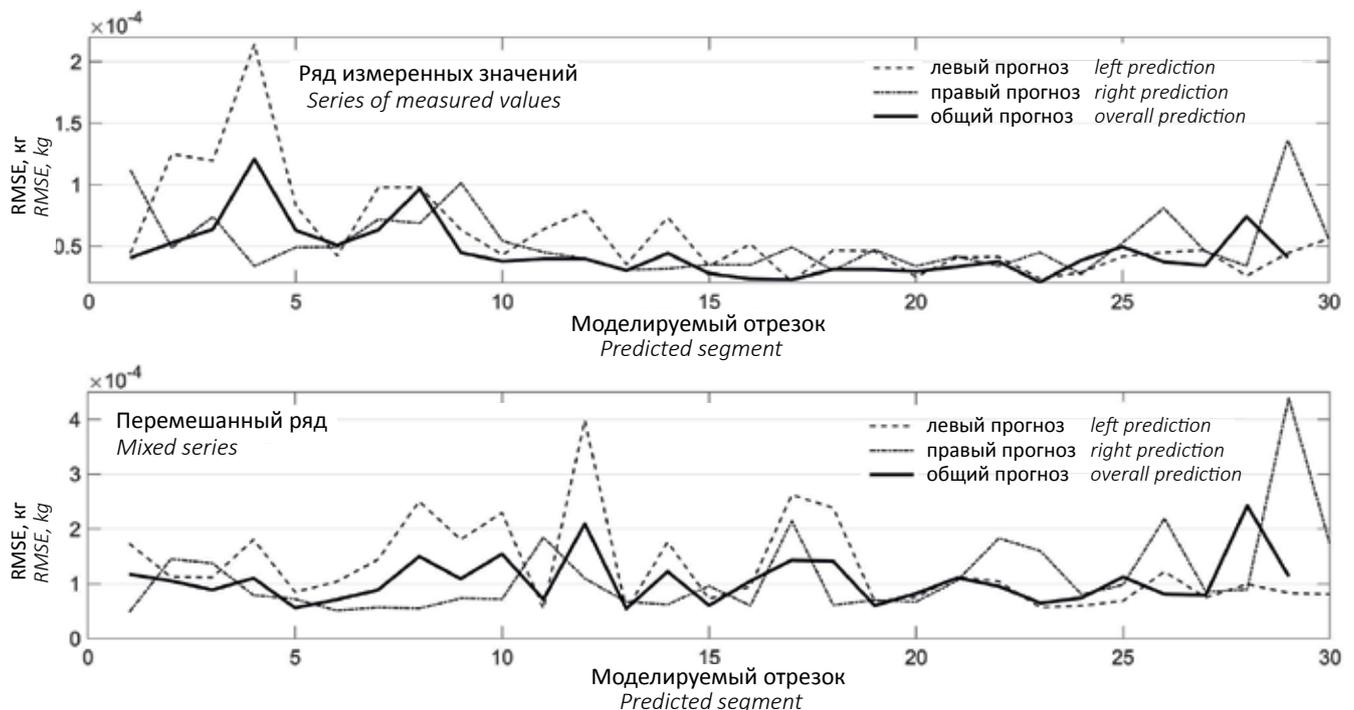


Рис. 4. Усредненные RMSE для пространственных рядов

Fig. 4. Average RMSE for spatial series



выборки с обеих сторон становятся сопоставимы, величины ошибок и колебания их значений становятся минимальными.

У поведения ошибки для перемешанного ряда никаких закономерностей не наблюдается. Вели-

чина ошибки в несколько раз больше, чем для ряда измеренных значений. Она хаотично изменяется и не зависит от длин обучающих выборок. Кроме того, размах колебания ее значений также самый большой.

**Заключение**

В работе предложен подход для прогнозирования отрезков пространственно-координатного ряда. Метод основан на идее взвешенного усреднения двух встречных прогнозов. Каждый прогноз генерируется отдельной моделью, обученной с одной стороны исследуемой области пространственно-координатного ряда. Модель на основе ИНС NAR выявила паттерны в ряду измеренных значений содержания пыли в снеговом покрове. Модель, которая использовала для обучения значения этого ряда, дала точный прогноз для моделируемого отрезка.

Та же модель, обученная на значениях перемешанного ряда (созданного из перемешанных в случайном порядке значений ряда измеренных значений), дала прогноз с большими ошибками и почти нулевым коэффициентом корреляции. Перемешав значения, мы убрали паттерны, таким образом лишив модель «опоры» для предсказания.

Авторы полагают, что предложенный подход будет востребован для восстановления пропущенных данных, например, при оценке воздействия природных и техногенных объектов на исследуемую среду.

**Список источников**

1. *Ardalani-Farsa M., Zolfaghari S.* Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks // *Neurocomputing*. – 2010. – Vol. 73. – Iss. 13–15. – P. 2540–2553. DOI: <https://doi.org/10.1016/j.neucom.2010.06.004>.
2. *Baglaeva E.M., Sergeev A.P., Shichkin A.V., Buevich A.G.* The effect of splitting of raw data into training and test subsets on the accuracy of predicting spatial distribution by a multilayer perceptron // *Mathematical Geosciences*. – 2020. – Vol. 52. – Iss. 1. – P. 111–121. DOI:10.1007/s11004-019-09813-9.
3. *Bayma L.O., Arruda Pereira M.* Identifying finest machine learning algorithm for climate data imputation in the state of Minas Gerais, Brazil // *Journal of Information and Data Management*. – 2018. – Vol. 9. – № 3. – P. 259–259.
4. *Buevich A., Sergeev A., Shichkin A., Baglaeva E.* A two-step combined algorithm based on NARX neural network and the subsequent prediction of the residues improves prediction accuracy of the greenhouse gases concentrations // *Neural Computing and Applications*. – 2021. – Vol. 33. – Iss. 5. – P. 1547–1557. DOI:10.1007/s00521-020-04995-4.
5. *Dai F., Zhou Q., Lv Z., Wang X., Liu G.* Spatial prediction of soil organic matter content integrating artificial neural network and ordinary kriging in Tibetan Plateau // *Ecological Indicators*. – 2014. – Vol. 45. – P. 184–194. DOI: <https://doi.org/10.1016/j.ecolind.2014.04.003>
6. *Dmitriev A.V., Kotin V.V.* Time series prediction of morbidity using artificial neural networks // *Biomedical Engineering*. – 2013. – Vol. 47. – Iss. 1. – P. 43–45. DOI: <https://doi.org/10.1007/s10527-013-9331-z>.
7. *Fernandez Jaramillo J.M., Mayer R.* Sample selection via angular distance in the space of the arguments of an artificial neural network // *Computers & Geosciences*. – 2018. – Vol. 114. – P. 98–106. DOI:10.1016/j.cageo.2018.02.003.
8. *Fernando H.J.S., Mammarella M.C., Grandoni G., Fedele P., Di Marco R., Dimitrova R., Hyde P.* Forecasting PM10 in metropolitan areas: efficacy of neural networks // *Environmental Pollution*. – 2012. – Vol. 163. – P. 62–67. DOI: <https://doi.org/10.1016/j.envpol.2011.12.018>.
9. *Guo G., Wu F., Xie F., Zhang R.* Spatial distribution and pollution assessment of heavy metals in urban soils from southwest China // *Journal of Environmental Sciences*. – 2012. – Vol. 24. – Iss. 3. – P. 410–418. DOI: [https://doi.org/10.1016/S1001-0742\(11\)60762-6](https://doi.org/10.1016/S1001-0742(11)60762-6).
10. *Hussain S., AlAlili A.* A hybrid solar radiation modeling approach using wavelet multiresolution analysis and artificial neural networks // *Applied Energy*. – 2017. – Vol. 208. – P. 540–550. DOI:10.1016/j.apenergy.2017.09.100.
11. *Kadow C., Hall D.M., Ulbrich U.* Artificial intelligence reconstructs missing climate information // *Nature Geoscience*. – 2020. – Vol. 13. – P. 408–413. DOI: <https://doi.org/10.1038/s41561-020-0582-5>.
12. *Klemmer K., Koshiyama A., Flennerhag S.* Augmenting correlation structures in spatial data using deep generative models // *arXiv:1905.09796 [cs.LG]*. – 2019.
13. *Kühnlein M., Appelhans T., Thies B., Nauss T.* Improving the accuracy of rainfall rates from optical satellite sensors with machine learning – A random forests-based approach applied to MSG SEVIRI // *Remote Sensing of Environment*. – 2014. – Vol. 141. – P. 129–143. DOI: <https://doi.org/10.1016/j.rse.2013.10.026>.
14. *Liu Z., Jiang P., Zhang L., Niu X.* A combined forecasting model for time series: Application to short-term wind speed forecasting // *Applied Energy*. – 2020. – Vol. 259. – P. 114137. DOI:10.1016/j.apenergy.2019.114137.
15. *McDermott P.L., Wikle C.K.* Bayesian recurrent neural network models for forecasting and quantifying uncertainty in spatial-temporal data // *Entropy*. – 2019. – Vol. 21. – Iss. 2. – P. 184. DOI: <https://doi.org/10.3390/e21020184>.
16. *Menezes Jr. J.M.P., Barreto G.A.* Long-term time series prediction with the NARX network: An empirical evaluation // *Neurocomputing*. – 2008. – Vol. 71. – Iss. 16–18. – P. 3335–3343. DOI: <https://doi.org/10.1016/j.neucom.2008.01.030>.
17. *Pisoni E., Farina M., Carnevale C., Piroddi L.* Forecasting peak air pollution levels using NARX models // *Engineering Applications of Artificial Intelligence*. – 2009. – Vol. 22. – Iss. 4–5. – P. 593–602. DOI: <https://doi.org/10.1016/j.engappai.2009.04.002>.
18. *Rakhmatova A., Sergeev A., Shichkin A., Buevich A., Baglaeva E.* Three-day forecasting of greenhouse gas CH4 in the atmosphere of the Arctic Belyy Island using discrete wavelet transform and artificial neural networks // *Neural Computing and Applications*. – 2021. – Vol. 33. – Iss. 16. – P. 10311–10322. DOI:10.1007/s00521-021-05792-3.
19. *Reichstein M., Camps-Valls G., Stevens B., Jung M., Denzler J., Carvalhais N., Prabhat.* Deep learning and process understanding for data-driven earth system science // *Nature*. – 2019. – Vol. 566. – P. 195–204. DOI:10.1038/s41586-019-0912-1.
20. *Russo A., Raischel F., Lind P.G.* Air quality prediction using optimal neural networks with stochastic variables // *Atmospheric Environment*. – 2013. – Vol. 79. – P. 822–830. DOI: <https://doi.org/10.1016/j.atmosenv.2013.07.072>.
21. *Sakizadeh M., Mirzaei R., Ghorbani H.* Support vector machine and artificial neural network to model soil pollution: a case study in Semnan Province, Iran // *Neural Computing and Applications*. – 2017. – Vol. 28. – P. 3229–3238. DOI: <https://doi.org/10.1007/s00521-016-2231-x>.
22. *Sergeev A., Shichkin A., Buevich A.* Time series forecasting of methane concentrations in the surface layer of atmospheric air in Arctic region // *AIP Conference Proceedings*. – 2018. – Vol. 2048. – Iss. 1. – P. 060005. DOI:10.1063/1.5082120.

23. Sun Y., Wang X., Chen Y., Liu Z. A modified whale optimization algorithm for large-scale global optimization problems // *Expert Systems with Applications*. – 2018. – Vol. 114. – P. 563–577. DOI:10.1016/j.eswa.2018.08.027.
24. Wang S., Cao J., Yu P. Deep learning for spatio-temporal data mining: A survey // *arXiv :1906.04928 [cs.LG]*. – 2019.
25. Wieland R., Mirschel W., Zbell B., Groth K., Pechenick A., Fukuda K. A new library to combine artificial neural networks and support vector machines with statistics and a database engine for application in environmental modelling // *Environmental Modelling & Software*. – 2012. – Vol. 25. – Iss. 4. – P. 412–420. DOI: <https://doi.org/10.1016/j.envsoft.2009.11.006>.
26. Zemouri R., Gouriveau R., Zerhouni N. Defining and applying prediction performance metrics on a recurrent NARX time series model // *Neurocomputing*. – 2010. – Vol. 73. – Iss. 13–15. – P. 2506–2521. DOI: <https://doi.org/10.1016/j.neucom.2010.06.005>.
27. Ziggah Y.Y., Youjian H., Tierra A.R., Laari P.B. Coordinate transformation between global and local data based on artificial neural network with K-fold cross-validation in Ghana // *Earth Sciences Research Journal*. – 2019. – Vol. 23. – № 1. – P. 67–77. DOI:10.15446/esrj.v23n1.63860.
28. Zhao X., Wang C., Su J., Wang J. Research and application based on the swarm intelligence algorithm and artificial intelligence for wind farm decision system // *Renewable Energy*. – 2019. – Vol. 134. – P. 681–697. DOI:10.1016/j.renene.2018.11.061.
29. Zhang Y.-F., Thorburn P.J., Xiang W., Fitch P. SSIM—a deep learning approach for recovering missing time series sensor data // *IEEE Internet of Things Journal*. – 2019. – Vol. 6. – Iss. 4. – P. 6618–6628. DOI:10.1109/JIOT.2019.2909038.
30. Zhou Q., Jiang H., Wang J., Zhou J. A hybrid model for PM2.5 forecasting based on ensemble empirical mode decomposition and a general regression neural network // *Science of the Total Environment*. – 2014. Vol. 496. – P. 264–274. DOI: <https://doi.org/10.1016/j.scitotenv.2014.07.051>.

## References:

1. Ardalani-Farsa M., Zolfaghari S. Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks. *Neurocomputing*. 2010;73(13–15):2540–2553. DOI: <https://doi.org/10.1016/j.neucom.2010.06.004>.
2. Baglaeva E.M., Sergeev A.P., Shichkin A.V., Buevich A.G. The effect of splitting of raw data into training and test subsets on the accuracy of predicting spatial distribution by a multilayer perceptron. *Mathematical Geosciences*. 2020;52(1):111–121. DOI:10.1007/s11004-019-09813-9.
3. Bayma L.O., Arruda Pereira M. Identifying finest machine learning algorithm for climate data imputation in the state of Minas Gerais, Brazil. *Journal of Information and Data Management*. 2018;9(3):259–259.
4. Buevich A., Sergeev A., Shichkin A., Baglaeva E. A two-step combined algorithm based on NARX neural network and the subsequent prediction of the residues improves prediction accuracy of the greenhouse gases concentrations. *Neural Computing and Applications*. 2021;33(5):1547–1557. DOI:10.1007/s00521-020-04995-4.
5. Dai F., Zhou Q., Lv Z., Wang X., Liu G. Spatial prediction of soil organic matter content integrating artificial neural network and ordinary kriging in Tibetan Plateau. *Ecological Indicators*. 2014;45:184–194. DOI: <https://doi.org/10.1016/j.ecolind.2014.04.003>.
6. Dmitriev A.V., Kotin V.V. Time series prediction of morbidity using artificial neural networks. *Biomedical Engineering*. 2013;47(1):43–45. DOI: <https://doi.org/10.1007/s10527-013-9331-z>.
7. Fernandez Jaramillo J.M., Mayerle R. Sample selection via angular distance in the space of the arguments of an artificial neural network. *Computers & Geosciences*. 2018;114:98–106. DOI:10.1016/j.cageo.2018.02.003.
8. Fernando H.J.S., Mammarella M.C., Grandoni G., Fedele P., Di Marco R., Dimitrova R., Hyde P. Forecasting PM10 in metropolitan areas: efficacy of neural networks. *Environmental Pollution*. 2012;163:62–67. DOI: <https://doi.org/10.1016/j.envpol.2011.12.018>.
9. Guo G., Wu F., Xie F., Zhang R. Spatial distribution and pollution assessment of heavy metals in urban soils from southwest China. *Journal of Environmental Sciences*. 2012;24(3):410–418. DOI: [https://doi.org/10.1016/S1001-0742\(11\)60762-6](https://doi.org/10.1016/S1001-0742(11)60762-6).
10. Hussain S., AlAlili A. A hybrid solar radiation modeling approach using wavelet multiresolution analysis and artificial neural networks. *Applied Energy*. 2017;208:540–550. DOI:10.1016/j.apenergy.2017.09.100.
11. Kadow C., Hall D.M., Ulbrich U. Artificial intelligence reconstructs missing climate information. *Nature Geoscience*. 2020;13:408–413. DOI: <https://doi.org/10.1038/s41561-020-0582-5>.
12. Klemmer K., Koshiyama A., Flennerhag S. Augmenting correlation structures in spatial data using deep generative models. *arXiv :1905.09796 [cs.LG]*. 2019.
13. Kühnlein M., Appelhans T., Thies B., Nauss T. Improving the accuracy of rainfall rates from optical satellite sensors with machine learning – A random forests-based approach applied to MSG SEVIRI. *Remote Sensing of Environment*. 2014;141:129–143. DOI: <https://doi.org/10.1016/j.rse.2013.10.026>.
14. Liu Z., Jiang P., Zhang L., Niu X. A combined forecasting model for time series: Application to short-term wind speed forecasting. *Applied Energy*. 2020;259:114137. DOI:10.1016/j.apenergy.2019.1141.
15. McDermott P.L., Wikle C.K. Bayesian recurrent neural network models for forecasting and quantifying uncertainty in spatial-temporal data. *Entropy*. 2019;21(2):184. DOI: <https://doi.org/10.3390/e21020184>.
16. Menezes Jr. J.M.P., Barreto G.A. Long-term time series prediction with the NARX network: An empirical evaluation. *Neurocomputing*. 2008;71(16–18):3335–3343. DOI: <https://doi.org/10.1016/j.neucom.2008.01.030>.
17. Pisoni E., Farina M., Carnevale C., Piroddi L. Forecasting peak air pollution levels using NARX models. *Engineering Applications of Artificial Intelligence*. 2009;22(4–5):593–602. DOI: <https://doi.org/10.1016/j.engappai.2009.04.002>.
18. Rakhmatova A., Sergeev A., Shichkin A., Buevich A., Baglaeva E. Three-day forecasting of greenhouse gas CH4 in the atmosphere of the Arctic Belyy Island using discrete wavelet transform and artificial neural networks. *Neural Computing and Applications*. 2021;33(16):10311–10322. DOI:10.1007/s00521-021-05792-3.
19. Reichstein M., Camps-Valls G., Stevens B., Jung M., Denzler J., Carvalhais N., Prabhat. Deep learning and process understanding for data-driven earth system science. *Nature*. 2019;566:195–204. DOI:10.1038/s41586-019-0912-1.
20. Russo A., Raischel F., Lind P.G. Air quality prediction using optimal neural networks with stochastic variables. *Atmospheric Environment*. 2013;79:822–830. DOI: <https://doi.org/10.1016/j.atmosenv.2013.07.072>.
21. Sakizadeh M., Mirzaei R., Ghorbani H. Support vector machine and artificial neural network to model soil pollution: a case study in Semnan Province, Iran. *Neural Computing and Applications*. 2017;28:3229–3238. DOI: <https://doi.org/10.1007/s00521-016-2231-x>.

22. Sergeev A., Shichkin A., Buevich A. Time series forecasting of methane concentrations in the surface layer of atmospheric air in Arctic region. *AIP Conference Proceedings*. 2018;2048(1):060005. DOI:10.1063/1.5082120.
23. Sun Y., Wang X., Chen Y., Liu Z. A modified whale optimization algorithm for large-scale global optimization problems. *Expert Systems with Applications*. 2018;114:563–577. DOI:10.1016/j.eswa.2018.08.027.
24. Wang S., Cao J., Yu P. Deep learning for spatio-temporal data mining: A survey. *arXiv* :1906.04928 [cs.LG]. 2019.
25. Wieland R., Mirschel W., Zbell B., Groth K., Pechenick A., Fukuda K. A new library to combine artificial neural networks and support vector machines with statistics and a database engine for application in environmental modelling. *Environmental Modelling & Software*. 2012;25(4):412–420. DOI: <https://doi.org/10.1016/j.envsoft.2009.11.006>.
26. Zemouri R., Gouriveau R., Zerhouni N. Defining and applying prediction performance metrics on a recurrent NARX time series model. *Neurocomputing*. 2010;73(13–15):2506–2521. DOI: <https://doi.org/10.1016/j.neucom.2010.06.005>.
27. Ziggah Y.Y., Youjian H., Tierra A.R., Laari P.B. Coordinate transformation between global and local data based on artificial neural network with K-fold cross-validation in Ghana. *Earth Sciences Research Journal*. 2019;23(1):67–77. DOI:10.15446/esrj.v23n1.63860.
28. Zhao X., Wang C., Su J., Wang J. Research and application based on the swarm intelligence algorithm and artificial intelligence for wind farm decision system. *Renewable Energy*. 2019;134:681–697. DOI:10.1016/j.renene.2018.11.061.
29. Zhang Y.-F., Thorburn P.J., Xiang W., Fitch P. SSIM—a deep learning approach for recovering missing time series sensor data. *IEEE Internet of Things Journal*. 2019;6(4):6618–6628. DOI:10.1109/JIOT.2019.2909038.
30. Zhou Q., Jiang H., Wang J., Zhou J. A hybrid model for PM<sub>2.5</sub> forecasting based on ensemble empirical mode decomposition and a general regression neural network. *Science of the Total Environment*. 2014;496:264–274. DOI: <https://doi.org/10.1016/j.scitotenv.2014.07.051>.

Статья поступила в редакцию 18.10.2021, одобрена после рецензирования 20.11.2021, принята к публикации 06.12.2021.  
The article was submitted 18.10.2021; approved after reviewing 20.11.2021; accepted for publication 06.12.2021.

### Информация об авторах

#### Буторова Анастасия Сергеевна

Инженер-исследователь лаборатории физики и экологии  
Института промышленной экологии УрО РАН;  
620990, Екатеринбург, ул. С. Ковалевской, д. 20  
Аспирант 1 курса Института радиоэлектроники и  
информационных технологий – РТФ Уральского федерального  
университета им. Б.Н. Ельцина;  
620002, Екатеринбург, ул. Мира, д. 19  
e-mail: a.s.butorova@urfu.ru

#### Сергеев Александр Петрович

Кандидат физико-математических наук, ведущий научный  
сотрудник, и.о. заведующего лабораторией физики и экологии  
Института промышленной экологии УрО РАН;  
620990, Екатеринбург, ул. С. Ковалевской, д. 20  
e-mail: sergeev@ecko.uran.ru

#### Шичкин Андрей Васильевич

Научный сотрудник лаборатории физики и экологии Института  
промышленной экологии УрО РАН;  
620990 Екатеринбург, ул. С. Ковалевской, д. 20  
e-mail: and@ecko.uran.ru

#### Бувич Александр Геннадьевич

Научный сотрудник лаборатории физики и экологии Института  
промышленной экологии УрО РАН;  
620990, Екатеринбург, ул. С. Ковалевской, д. 20  
e-mail: bag@ecko.uran.ru

#### Баглаева Елена Михайловна

Кандидат физико-математических наук, старший научный  
сотрудник лаборатории физики и экологии Института  
промышленной экологии УрО РАН;  
620990, Екатеринбург, ул. С. Ковалевской, д. 20  
e-mail: e.m.baglaeva@urfu.ru

#### Сергеева Марина Викторовна

Научный сотрудник лаборатории физики и экологии Института  
промышленной экологии УрО РАН;  
620990, Екатеринбург, ул. С. Ковалевской, д. 20  
e-mail: marin@ecko.uran.ru

### Information about authors

#### Anastasia S. Butorova

Research Engineer of the Laboratory of Physics and Ecology of the  
Institute of Industrial Ecology, UB RAS;  
20, S. Kovalevskoy str., Ekaterinburg, 620990, Russia  
1st year Postgraduate Student of the Institute of Radio Electronics  
and Information Technologies- RtF of the Ural Federal University  
named after B.N. Yeltsin;  
19, Mira str., Ekaterinburg, 620002, Russia  
e-mail: a.s.butorova@urfu.ru

#### Alexander P. Sergeev

Candidate of Physical and Mathematical Sciences, Leading  
Researcher, Acting Head of the Laboratory of Physics and Ecology  
of the Institute of Industrial Ecology, UB RAS,  
20, S. Kovalevskoy str., Ekaterinburg, 620990, Russia  
e-mail: sergeev@ecko.uran.ru

#### Andrey V. Shichkin

Researcher of the Laboratory of Physics and Ecology  
of the Institute of Industrial Ecology, UB RAS,  
20, S. Kovalevskoy str., Ekaterinburg, 620990, Russia  
e-mail: and@ecko.uran.ru

#### Alexander G. Buevich

Researcher of the Laboratory of Physics and Ecology  
of the Institute of Industrial Ecology, UB RAS,  
20, S. Kovalevskoy str., Ekaterinburg, 620990, Russia  
e-mail: bag@ecko.uran.ru

#### Elena M. Baglaeva

Candidate of Physical and Mathematical Sciences,  
Senior Researcher of the Laboratory of Physics and Ecology  
of the Institute of Industrial Ecology, UB RAS,  
20, S. Kovalevskoy str., Ekaterinburg, 620990, Russia  
e-mail: e.m.baglaeva@urfu.ru

#### Marina V. Sergeeva

Researcher of the Laboratory of Physics and Ecology  
of the Institute of Industrial Ecology, UB RAS,  
20, S. Kovalevskoy str., Ekaterinburg, 620990, Russia  
e-mail: marin@ecko.uran.ru