

Геоинформатика. 2022. № 2. С. 42–53.
Geoinformatika. 2022;(2):42–53.

Моделирование геообъектов и геопроцессов

Научная статья

УДК 504.064.2.001.18

<https://doi.org/10.47148/1609-364X-2022-2-42-53>

Применение перестановочного метода к оценке прогностической способности моделей пространственного распределения концентраций меди и железа в верхнем слое почвы

© 2022 г. — А.П. Сергеев^{1, а)}, А.С. Буторова^{1, 2, б)}, А.В. Шичкин^{1, с)}, А.Г. Бувевич^{1, д)}, Е.М. Баглаева^{1, е)}, И.Е. Субботина^{1, ф)}

¹Федеральное государственное бюджетное учреждение науки Институт промышленной экологии Уральского отделения Российской академии наук (ИПЭ УрО РАН), Россия, г. Екатеринбург;

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

^{а)}sergeev@ecko.uran.ru, ^{б)}автор-корреспондент: a.s.butorova@urfu.ru, ^{с)}and@ecko.uran.ru, ^{д)}bag@ecko.uran.ru,

^{е)}e.m.baglaeva@urfu.ru, ^{ф)}iesub@mail.ru

Аннотация: В статье предложено применение перестановочного метода к оценке прогностической способности моделей на основе искусственных нейронных сетей. Для апробации этого метода были реализованы три модели: многослойный перцептрон, сеть радиально-базисных функций, нейронная сеть с обобщенной регрессией. Для моделирования были использованы данные о пространственном распределении меди и железа в верхнем слое почвы (глубина 0,05 м) на территории субарктического города Ноябрьск, Ямало-Ненецкий автономный округ, Россия. Всего было отобрано 237 проб почвы. Для моделирования данные о концентрации меди и железа были разбиты на два подмножества: тренировочное и тестовое. Моделируемые пространственные наборы данных сравнивались с наблюдаемыми значениями тестового подмножества. Для оценки работоспособности построенных моделей использовались три подхода: 1) расчет коэффициентов корреляции, индексов ошибок или согласия, 2) графический подход (диаграмма Тейлора), 3) рандомизационная оценка вероятности получить расхождение между наблюдаемым и смоделированным наборами данных в предположении, что оба этих набора данных извлечены из одной и той же популяции. Для рандомизационного алгоритма были использованы две статистики: разность средних и коэффициент корреляции. Перестановочный метод показал свою продуктивность, так как позволил оценить значимость расхождения между наблюдаемыми и предсказанными наборами данных.

Ключевые слова: *перестановочный метод, рандомизация, прогнозируемые значения, наблюдаемые значения, пространственное распределение, оценка прогностической способности, искусственные нейронные сети*

Для цитирования: Сергеев А.П., Буторова А.С., Шичкин А.В., Бувевич А.Г., Баглаева Е.М., Субботина И.Е. Применение перестановочного метода к оценке прогностической способности моделей пространственного распределения концентраций меди и железа в верхнем слое почвы // *Геоинформатика*. — 2022. — № 2. — С. 42–53. <https://doi.org/10.47148/1609-364X-2022-2-42-53>.

Modeling of geological objects and geoprocesses

Original article

Application of the permutation method to the assessment of predictive ability of the models of spatial distribution of copper and iron concentrations in the topsoil

© 2022 — A.P. Sergeev^{1, а)}, A.S. Butorova^{1, 2, б)}, A.V. Shichkin^{1, с)}, A.G. Buevich^{1, д)}, E.M. Baglaeva^{1, е)}, I.E. Subbotina^{1, ф)}

¹Institute of Industrial Ecology UB RAS, Ekaterinburg, Russia;

²Ural Federal University named after B.N. Yeltsin., Ekaterinburg, Russia

^{а)}sergeev@ecko.uran.ru, ^{б)}corresponding author: amoskalyova11@yandex.ru, ^{с)}and@ecko.uran.ru, ^{д)}bag@ecko.uran.ru,

^{е)}e.m.baglaeva@urfu.ru, ^{ф)}iesub@mail.ru

Abstract: The article proposes the use of the permutation method for assessment of the predictive ability of models based on artificial neural networks. To test this method, three models based on artificial neural networks were implemented: a multilayer perceptron, a radial basis function network, and a generalized regression neural network. For modeling, data on the spatial distribution of copper and iron in the topsoil (depth 0.05 m) on the territory of the subarctic city of Noyabrsk, Yamalo-Nenets Autonomous Okrug, Russia, were used. A total of 237 soil samples were collected. For modelling, the copper and iron concentration data were divided into two subsets: training and test. The modelled spatial datasets were compared with the observed values of the test subset. To assess the performance of the constructed models, three approaches were used: 1) calculation of correlation coefficients, error or agreement indexes, 2) graphical approach (Taylor diagram), 3) randomization assessment of the probability of obtaining a divergence between the observed and modelled datasets, assuming that both of these datasets taken from the same population. For the randomization algorithm, two statistics were used: difference in means and correlation coefficient. The permutation method proved its productivity, as it allowed to assess the significance of the divergence between the observed and predicted datasets.

Key words: *permutation method, randomization, predicted values, observed values, spatial distribution, predictive ability assessment, artificial neural networks*

For citation: Sergeev A.P., Butorova A.S., Shichkin A.V., Buevich A.G., Baglaeva E.M., Subbotina I.E. Application of the permutation method to the assessment of predictive ability of the models of spatial distribution of copper and iron concentrations in the topsoil. *Geoinformatika*. 2022;(2):42–53. <https://doi.org/10.47148/1609-364X-2022-2-42-53>. In Russ.

Введение

Современные научные исследования и разработки, касающиеся воздействия человека на окружающую среду, неотъемлемо связаны с компьютерным моделированием [2, 13, 20, 25]. Моделирование представляет собой вызов для многофакторных динамических систем, демонстрирующих сложное поведение. Примером такой системы может служить биосфера Земли [12]. Компьютерное моделирование использует различные подходы, но в последние годы наибольшую эффективность демонстрируют модели на основе искусственных нейронных сетей (ИНС) [9, 22, 28, 29, 32]. При этом одной из задач в экологическом моделировании является векторно-значная оценка производительности построенных моделей [1, 3, 10].

Сравнение предсказанных и наблюдаемых значений является основным подходом к оценке производительности моделей. Для этого традиционно используются показатели ошибок и показатели точности (например, средняя абсолютная ошибка, среднеквадратическая ошибка, коэффициент корреляции и т.д.). Полученные оценки обычно визуализируют для удобства восприятия. Одним из эффективных методов графического представления информации о нескольких моделях является диаграмма Тейлора. Она широко используется в оценке и тестировании моделей в последние годы [21, 23]. Диаграмма Тейлора объединяет в себе три показателя: стандартное отклонение, среднеквадратическую ошибку и коэффициент корреляции [26]. Стандартное отклонение позволяет оценить вариативность предсказанных данных, которая, как предполагается, не должна превышать вариативность наблюдаемых данных. По коэффициенту корреляции возможно оценить согласованность изменчивости между наблюдаемыми и предсказанными наборами данных. Среднеквадратическая ошибка показывает среднеквадратическое расхождение между наблюдаемыми и предсказанными значениями. Различные виды показателей точности (или ошибок) имеют смысл расстояния (или обратного расстояния) между наблюдаемыми и предсказанными наборами данных. Использование комплекса этих показателей, безусловно, дает дополнительную информацию о модели, но, поскольку эти показатели коррелированы друг с другом, их применение не сильно увеличивает нашу информированность о способности модели предсказывать.

Кроме того, хотелось бы знать, насколько значимы в некотором вероятностном смысле наблюдаемые показатели производительности. Чтобы ответить на этот вопрос, в настоящей работе предложено использовать перестановочную идеологию по отношению к подходящим статистикам [4, 7, 11, 15, 19]. Для апробации этого метода мы сравнили производительность трех моделей, основанных на искусственных нейронных сетях: многослойный

персептрон (multilayer perceptron, MLP), сеть радиально-базисных функций (radial basis function network, RBF), нейронная сеть с обобщенной регрессией (generalized regression neural network, GRNN).

Материалы и методы

Исследование было проведено на территории субарктического города Ноябрьск, Ямало-Ненецкий автономный округ, Россия (рис. 1). Ноябрьск (дата основания — 1976 г.) расположен в центральной части Сибирских Увалов, на водоразделе рек Обь и Пур.

Исследовались концентрации двух металлов (медь и железо). Отбор проб почвы производился на ненарушенных, естественных участках в узлах квадратной сетки с шагом 250 м. Образцы были извлечены из верхних слоев почвы (глубина 0,05 м). Затем отобранные пробы были доставлены в сертифицированную лабораторию в соответствии с ISO/IEC 17025:2005. Подготовка и химический анализ проб проведены в соответствии с действующими стандартными требованиями Российской федеральной системы сертификации.

Всего было отобрано 237 проб почвы. Для моделирования данные по каждому элементу были разбиты на два подмножества: тренировочное и тестовое [6, 14, 17, 33]. Первое подмножество использовалось для обучения сети. Традиционно в это подмножество входит большая часть тестовых точек (в нашем случае это 205 точек). Оставшиеся точки составили тестовое подмножество (32 точки). Оно использовалось только для тестирования нейронной сети (рис. 1).

Для каждого элемента строились прогнозы с использованием нейронных сетей: многослойного персептрона, сети радиально-базисных функций и сети с обобщенной регрессией.

Многослойный персептрон

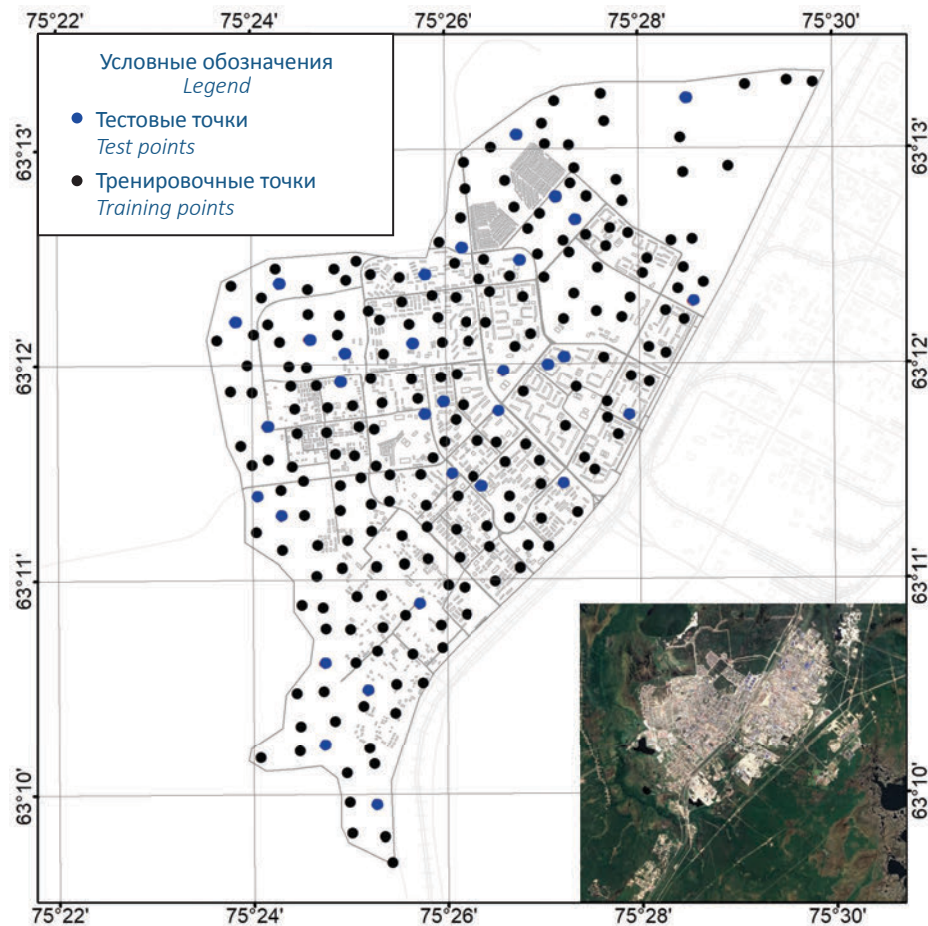
Многослойный персептрон — это сеть прямого распространения сигнала [11, 18]. Для ее обучения чаще всего используется алгоритм обратного распространения ошибки (в нашем случае использовался алгоритм Левенберга-Марквардта [8, 16]). Многослойный персептрон может использоваться для любых входных и выходных данных. Сеть состоит из входного, выходного и одного или нескольких скрытых слоев. Для активации нейронов скрытого слоя обычно используется нелинейная функция активации (в нашем случае гиперболический тангенс):

$$tansig = \frac{2}{1 + e^{-2n}} - 1, \tag{1}$$

где n — входные векторы сети.

При обучении сети MLP основная задача состоит в подборе количества нейронов в скрытом слое. Однозначного способа определения этого параметра не существует. Для каждого конкретного случая при-

Рис. 1. Место отбора проб
Fig. 1. Sampling areak



ходится заново подбирать конфигурацию нейронной сети для получения оптимального результата.

Сеть радиально-базисных функций

Сеть радиально-базисных функций (впервые сформулированы Брумхедом и Лоу в 1988 г., [5]) — это искусственная нейронная сеть, которая в качестве функции активации использует радиальные базисные функции:

$$radbas(n) = e^{-n^2}, \tag{2}$$

где n — входные векторы сети.

Сеть радиально-базисных функций строится на прямой передаче сигнала и обратном методе распространения ошибки, но при этом, в отличие от перцептрона, состоит из значительно большего количества нейронов. Кроме того, на создание RBF требуется меньше времени.

RBF состоит из двух слоев: скрытого радиального базисного слоя и выходного линейного слоя. Существует два вида сетей радиально-базисных функций: сеть с нулевой ошибкой и итерационная сеть. В настоящем исследовании использовалась итерационная сеть, в которой нейроны добавлялись к скрытому слою до тех пор, пока ошибка сети не стала меньше заданного значения или не было использовано максимальное количество нейронов.

Входными параметрами сети радиально-базисных функций являются входные и выходные значения, заданная допустимая среднеквадратическая ошибка и параметр влияния SPREAD. Параметр SPREAD подбирался в пределах диапазона значений, при котором значение среднеквадратической ошибки было минимальным.

Нейронная сеть с обобщенной регрессией

Нейронная сеть с обобщенной регрессией является разновидностью сетей радиально-базисных функций, основанной на непараметрической регрессии [24]. GRNN характеризуется простотой и высокой скоростью обучения. К преимуществам сети также относятся однопроходное обучение (обучение без использования метода обратного распространения ошибки) и высокая точность прогноза.

Архитектура сети с обобщенной регрессией аналогична сети радиальных базисных функций. Первым слоем является радиальный базисный слой, в котором число нейронов равно количеству элементов обучающего множества. Основное отличие заключается в структуре радиального базисного слоя: в нем дополнительно используется блок для вычисления нормированного скалярного произведения весов и вектора входа. Второй слой является

линейным, в котором число нейронов также равно количеству элементов обучающего множества.

Как и в случае радиальных базисных сетей, параметрами GRNN являются векторы входа и выхода, а также параметр влияния (или параметр сглаживания) SPREAD. Параметр SPREAD является единственным настраиваемым параметром сети. При увеличении этого параметра функция, которая генерируется нейронной сетью, становится более гладкой. Значение параметра SPREAD подбирается экспериментально. В настоящей работе SPREAD подбирался по минимальной среднеквадратической ошибке (табл. 1).

Табл. 1. Подобранные значения параметра SPREAD for GRNN and RBF

Tab. 1. SPREAD parameter for GRNN and RBF

я	Элемент	SPREAD
GRNN	Cu	0.005
	Fe	0.002
RBF	Cu	0.048
	Fe	0.63

Оценка производительности

Прогнозируемые пространственные наборы данных элементов Fe и Cu сравнивались с наблюдаемыми значениями тестового фрагмента. Для оценки продуктивности прогноза использовались три подхода. В первом подходе рассчитывались коэффициенты корреляции, показатели ошибок или индексы согласия (MAE, RMSE, коэффициент корреляции Пирсона и т.д.). Второй подход – графический – это построение диаграммы Тейлора, которая на одном графике объединяет три показателя: стандартное отклонение, корреляцию и среднеквадратическую ошибку, и дает векторнозначную оценку производительности модели. Эти два подхода имеют смысл расстояния между наблюдаемыми и предсказанными наборами данных. Третий подход основан на рандомизационной оценке вероятности получить наблюдаемое значение выбранной статистики как меру согласия/несогласия наблюдаемых и предсказанных наборов данных в предположении, что оба этих набора данных извлечены из одной и той же популяции.

Традиционные показатели оценки

Для оценки точности прогноза моделей использовались некоторые критерии и ошибки прогноза (коэффициент корреляции Пирсона, средняя абсолютная ошибка (MAE), среднеквадратическая ошибка (RMSE), среднеквадратическая относительная ошибка (RMSRE) и индекс согласия IA2). Коэффициент корреляции Пирсона (3) демонстрирует взаимосвязь двух переменных и то, насколько про-

порциональна их изменчивость. Он характеризует существование линейной связи между двумя переменными:

$$Corr = \frac{\sum(P - \bar{P})(O - \bar{O})}{\sqrt{\sum(P - \bar{P})^2 \sum(O - \bar{O})^2}}, \tag{3}$$

где P – предсказанный набор данных, O – наблюдаемый набор данных, \bar{P} , \bar{O} – средние значения наборов P и O соответственно.

MAE (4), RMSE (5) и RMSRE (6) показывают, насколько велики ошибки предсказанного ряда в сравнении со значениями наблюдаемого ряда:

$$MAE = \frac{\sum|P - O|}{n}, \tag{4}$$

$$RMSE = \sqrt{\frac{\sum(P - O)^2}{n}}, \tag{5}$$

$$RMSRE = \sqrt{\frac{1}{n} \sum \left(\frac{P - O}{O} \right)^2}, \tag{6}$$

где P – предсказанный набор данных, O – наблюдаемый набор данных, n – размер наборов данных.

Индекс согласия IA2 (7) является стандартизированной мерой степени ошибки предсказания модели и варьируется от -1 до 1, где значение IA2=0 соответствует полному согласию модели с исходной выборкой, значение IA2=0,5 показывает, что сумма ошибок модели соответствует половине наблюдаемых отклонений, IA2=-0,5 дает превышение ошибок модели над наблюдаемыми в два раза и при IA2 близком к -1, может означать, что модель является плохой оценкой наблюдаемых отклонений [30, 31]:

$$IA2 = \begin{cases} 1 - \frac{\sum|P - O|}{c \sum(|O - \bar{O}|)}, & \text{если } \sum|P - O| \leq c \sum|O - \bar{O}| \\ \frac{c \sum(|O - \bar{O}|)}{\sum|P - O|} - 1, & \text{если } \sum|P - O| > c \sum|O - \bar{O}| \end{cases}, \tag{7}$$

где P – предсказанный набор данных, O – наблюдаемый набор данных, \bar{P} , \bar{O} – средние значения наборов P и O соответственно, c – константа.

Диаграмма Тейлора

Для оценки прогностической способности моделей (в нашем случае RBF, GRNN и MLP) использовалась диаграмма Тейлора [26, 27]. Диаграмма Тейлора объединяет в себе три показателя: коэффициент корреляции, стандартное отклонение и центрированное среднеквадратическое отклонение. Маркеры для каждой модели нанесены на диаграмму в полярных координатах. Радиальные расстояния от начала координат представляют собой стандартные отклонения, а азимутальные углы – коэффициенты корреляции между исходным (наблюдённым) и прогнозируемым распределениями.

Диаграмма Тейлора дает более наглядное различие между построенными моделями по сравнению с классическими графическими представлениями результатов прогнозирования.

Рандомизационный алгоритм

В настоящей работе мы использовали рандомизационный алгоритм, основанный на перестановочном методе. Рандомизационный алгоритм предполагал полную (если это возможно) кросс-перестановку всех значений между наблюдаемым и прогнозируемым наборами данных. Рандомизация представляла собой n -разрядное двоичное число, где n — количество значений в наборе данных. Таким образом, всего возможно 2^n перестановок. На основе вычисленных 2^n рандомизаций строится точное распределение выбранных статистик при справедливости нулевой гипотезы. Единицей обозначалась позиция, по которой между двумя наборами данных была совершена перестановка, нулем — отсутствие перестановки. Для каждой рандомизации вычислялись две статистики: разность средних и коэффициент корреляции.

Под p -уровнем значимости в настоящей работе понималась вероятность получить наблюдаемое или более экстремальное значение статистики (в нашем случае, разности средних и коэффициента корреляции), при условии, что нулевая гипотеза верна. Мы вычисляли рандомизационные p -значения, ге-

нерируя рандомизационные распределения этих двух статистик. При расчете разности средних оценивалась значимость расхождения, при расчете коэффициентов корреляции — значимость коррелированности наблюдаемого и предсказанного наборов данных. Формулы расчета, нулевой гипотезы, p -значения и комментарии к ним для каждой из статистик приведены в таблице 2.

В таблице 3 и на рисунке 2 приведена демонстрация применения рандомизационного алгоритма для расчета разности средних и коэффициента корреляции на примере перестановок наборов данных длиной 4. Для 4 значений количество перестановок равно $2^4=16$. Схема рандомизаций показывает, на какой позиции между наборами данных производится перестановка: 0 — перестановки нет, 1 — перестановка есть. Затем, согласно рандомизационной схеме, строились два набора данных, для которых рассчитывались разность средних и коэффициент корреляции. Затем строились гистограммы распределений для каждого показателя. Чтобы определить p -значение, измерялась площадь под гистограммой значений перестановки выше наблюдаемого или более экстремального значения.

В настоящей работе предсказанный набор данных составлял 32 значения = 232 (4294967296) перестановок. На практике такое количество перестановок выполнить затруднительно, так как это требует больших вычислительных мощностей и временных

Табл. 2. Используемые статистики

Tab. 2. Used statistics

Статистика/Формула/Нулевая гипотеза	P -значение	Комментарий к p -значению
Разность средних $Diff(X, Y) = \bar{Y} - \bar{X} = \frac{\sum Y(j)}{n} - \frac{\sum X(j)}{n} =$ $= \frac{\sum Y(j) - \sum X(j)}{n} = \frac{\sum (Y(j) - X(j))}{n}$ $H_0 : E(Diff(X, Y)) = 0$	Если p -значение высокое, модель обладает высокой прогностической способностью	Изменялось одностороннее p -значение в зависимости от положения, в которое попадало наблюдаемое значение статистики. P -value определялся как минимальный из двух односторонних (слева или справа)
Корреляция $Corr(X, Y) = \frac{\sum((X(j) - \bar{X})(Y(j) - \bar{Y}))}{\sqrt{\sum(X(j) - \bar{X})^2 \sum(Y(j) - \bar{Y})^2}}$ $H_0 : E(Corr(X, Y)) = 0$	Если p -значение низкое, модель обладает высокой прогностической способностью	Изменялось правостороннее p -значение

Примечание к таблице: X — вектор «наблюдаемых» данных, Y — вектор «предсказанных» данных, X и Y — фактически наблюдаемые и предсказанные данные соответственно, для случая, когда ни одна перестановка еще не производилась, j — номер позиции в наборе данных, \bar{X} , \bar{Y} — средние значения наборов данных X и Y , $Diff(X, Y)$ — разность средних наборов данных X и Y , $Corr(X, Y)$ — корреляция между наборами данных X и Y , $E(Diff(X, Y))$ и $E(Corr(X, Y))$ — математические ожидания разности средних и корреляции соответственно.

Табл. 3. Пример реализации перестановочного метода

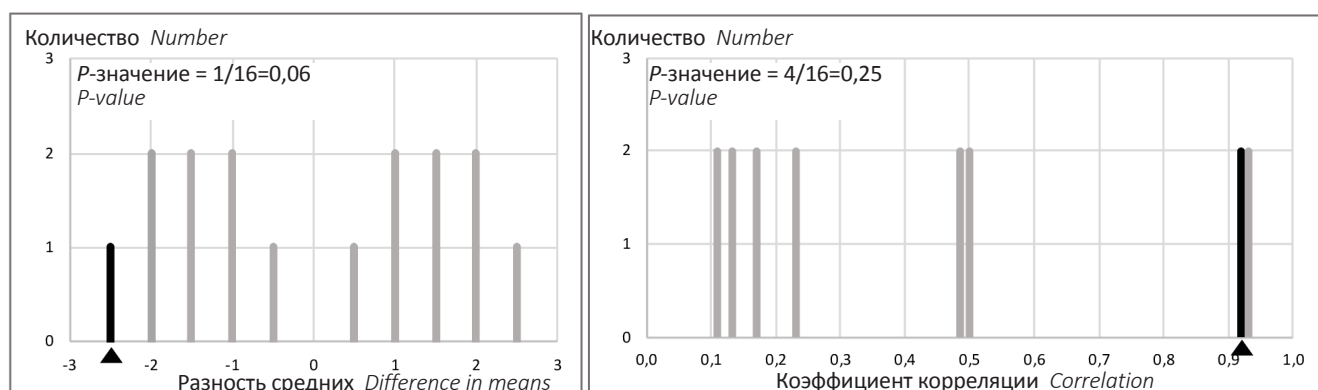
Tab. 3. An example of the implementation of the permutation method

Номер рандомизации (i)	Рандомизация				Набор данных 1				Набор данных 2				Разность средних	Корреляция
	0	0	0	0	5	10	3	4	3	4	2	3		
0*	0	0	0	0	5	10	3	4	3	4	2	3	2,5	0,919
1	0	0	0	1	5	10	3	3	3	4	2	4	2,0	0,501
2	0	0	1	0	5	10	2	4	3	4	3	3	2,0	0,930
3	0	0	1	1	5	10	2	3	3	4	3	4	1,5	0,487
4	0	1	0	0	5	4	3	4	3	10	2	3	-0,5	0,110
5	0	1	0	1	5	4	3	3	3	10	2	4	-1,0	0,170
6	0	1	1	0	5	4	2	4	3	10	3	3	-1,0	0,132
7	0	1	1	1	5	4	2	3	3	10	3	4	-1,5	0,230
8	1	0	0	0	3	10	3	4	5	4	2	3	1,5	0,230
9	1	0	0	1	3	10	3	3	5	4	2	4	1,0	0,132
10	1	0	1	0	3	10	2	4	5	4	3	3	1,0	0,170
11	1	0	1	1	3	10	2	3	5	4	3	4	0,5	0,110
12	1	1	0	0	3	4	3	4	5	10	2	3	-1,5	0,487
13	1	1	0	1	3	4	3	3	5	10	2	4	-2,0	0,930
14	1	1	1	0	3	4	2	4	5	10	3	3	-2,0	0,501
15	1	1	1	1	3	4	2	3	5	10	3	4	-2,5	0,919

* — наблюдаемые vs предсказанные наборы данных

Рис. 2. Пример

Fig. 2. The example



затрат. Поэтому нами был использован алгоритм генерирования 100000 случайных рандомизаций.

Результаты и обсуждение

Критерии и ошибки прогноза

В таблице 4 представлены результаты сравнения производительности моделей. Наилучшие значения показателей точности выделены жирным шрифтом. Коэффициент корреляции, MAE, RMSE, RMSRE и IA2 показали, что прогностическая точность модели MLP выше, чем у других представленных моделей.

Диаграмма Тейлора

На рисунке 3 видно, что модели RBF и MLP в целом лучше всего согласуются с наблюдениями. Коэффициент корреляции говорит о синхронности предсказанных и наблюдаемых значений. Большое стандартное отклонение в предсказанных данных говорит об обнаружении эффектов, которых нет в исходных данных.

Среди трех предложенных моделей MLP имела наибольшую корреляцию с наблюдаемыми значениями меди и железа. В случае с медью модель RBF имела несколько более близкое к наблюдаемому

Рис. 3. Диаграммы Тейлора

Fig. 3. Taylor diagrams

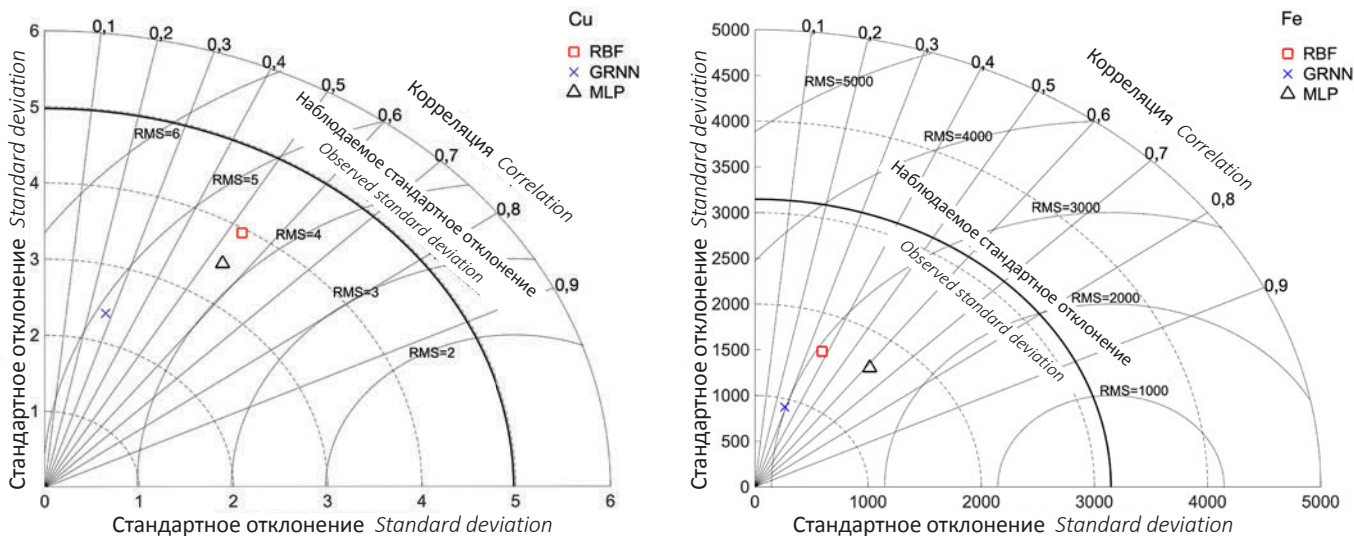


Рис. 4. Корреляция

Fig. 4. Correlation

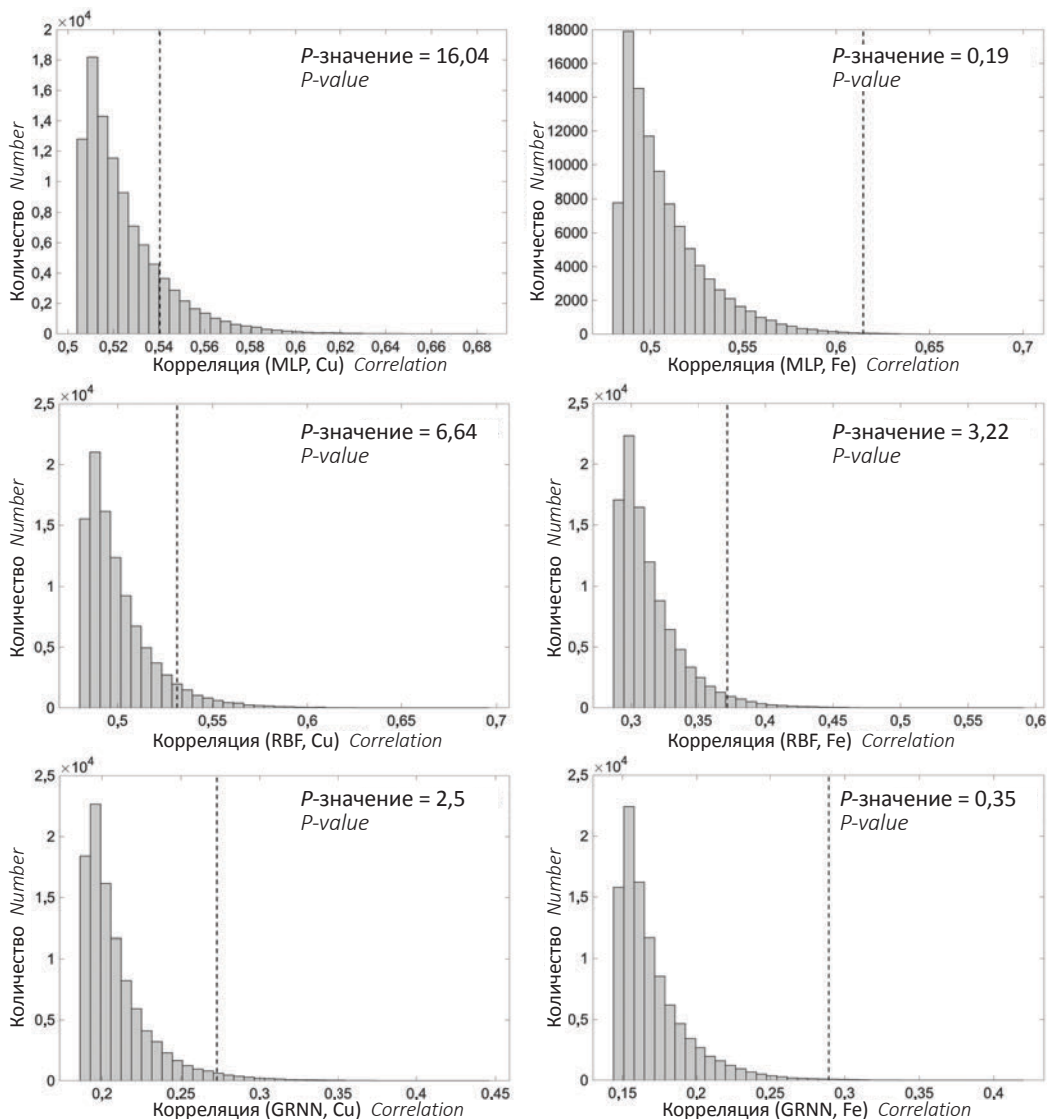


Рис. 5. Разность средних
Fig. 5. Difference in means

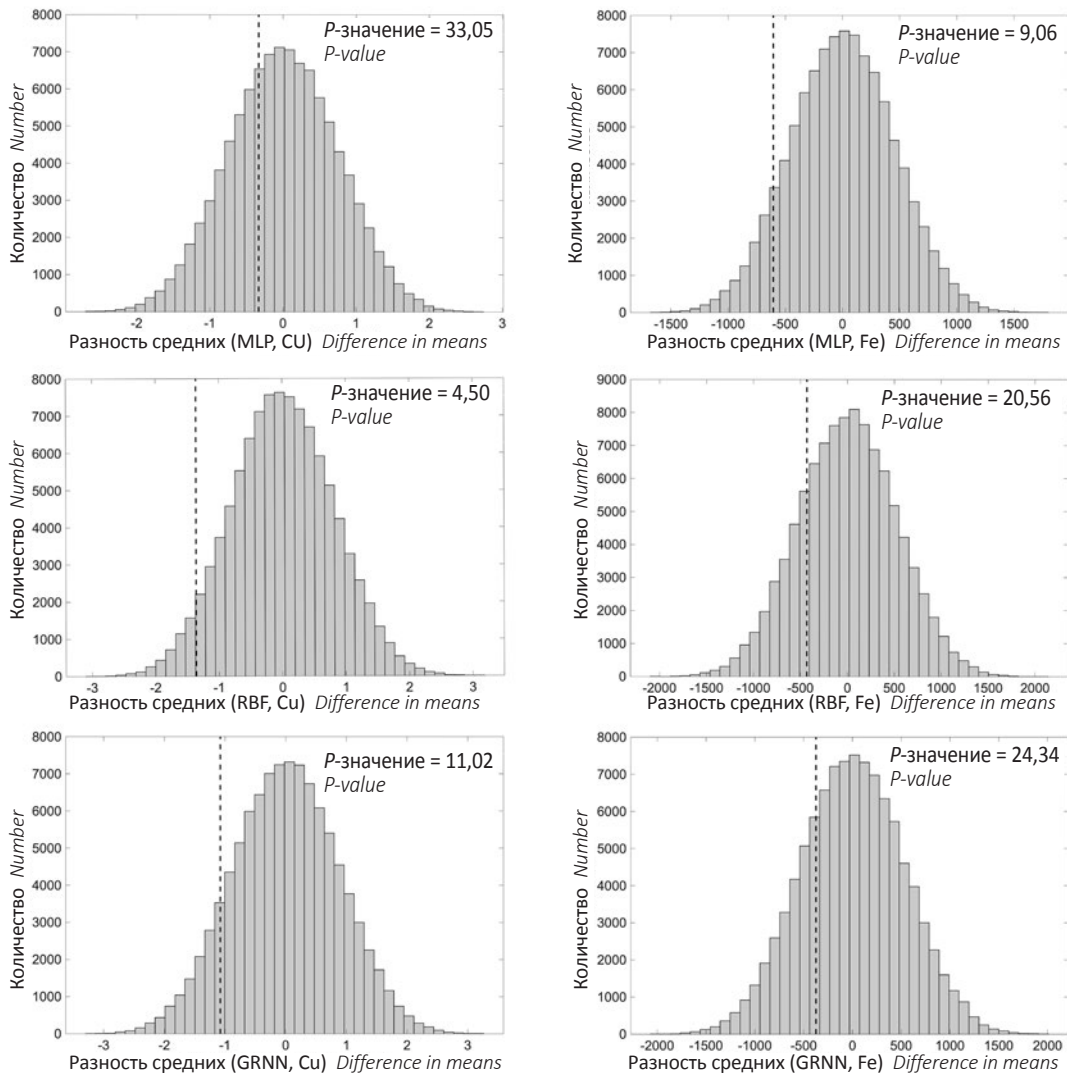


Табл. 4. Показатели оценки точности прогноза
Tab. 4. Predictive accuracy assessment indexes

Показатель точности	Cu			Fe		
	RBF	GRNN	MLP	RBF	GRNN	MLP
Корреляция	0,53	0,27	0,54	0,37	0,29	0,61
MAE	3,69	3,87	3,23	2266	2291	2026
RMSRE	0,39	0,47	0,38	0,26	0,27	0,21
RMSE	4,55	4,94	4,21	2937	2988	2532
IA2	0,08	0,03	0,19	0,07	0,06	0,17
P-значение (разность средних)	4,50	11,02	33,05	20,56	24,34	9,06
P-значение (корреляция)	6,64	2,50	16,04	3,22	0,35	0,19

стандартное отклонение, чем MLP, однако для железа наиболее близкое к наблюдаемому стандартное отклонение показала модель MLP.

Рандомизационная оценка

Для всех новых распределений (перестановок) прогноза и наблюдаемых значений были найдены два показателя: корреляция между ними и разность средних. Построены гистограммы распределений для каждого показателя. На каждой гистограмме добавлена вертикальная линия, показывающая значение показателя для наблюдаемых данных и прогноза (рис. 4, 5).

Для разности средних самое высокое p -значение было получено у MLP для Cu (33,05) и GRNN для Fe (24,34), поэтому прогностическая способность моделей, основанных на MLP и GRNN, в этом случае была высокой. Для коэффициента корреляции наименьшее p -значение было получено у GRNN для Cu (2,50) и MLP для Fe (0,19), поэтому прогностическая способность моделей, основанных на GRNN и MLP, в этом случае была высокой.

Заключение

Проблема оценки производительности моделей является не менее важной, чем их построение. В настоящей работе предлагается использовать перестановочный метод для оценки предсказательной способности моделей на основе искусственных ней-

ронных сетей. Для апробации этого метода были реализованы три модели на основе искусственных нейронных сетей: многослойный персептрон, сеть с радиальной базисной функцией и обобщенная ре-рессионная нейронная сеть.

P -значения для двух разных элементов дали разные результаты. Это подчеркивает необходимость использования предложенного рандомизационного алгоритма для оценки производительности моделей с использованием нескольких методов.

В целом в каждом методе модель MLP показала наибольшую прогностическую точность.

Перестановочный метод показал свою продуктивность, так как позволил оценить значимость расхождения между наблюдаемыми и предсказанными наборами данных. Перестановочный метод может быть использован для оценки любых моделей в задачах прогнозирования, в том числе экологического, поскольку он дополняет традиционные способы оценивания точности моделей вероятностными величинами, имеющими смысл достигнутых уровней значимости. Еще одним преимуществом, которое дает перестановочный метод, является то обстоятельство, что исследователь сам может выбирать выборочные статистики, имеющие предметное обоснование. Кроме того, перестановочный метод является точным в том смысле, что не требует априорных знаний о вероятностных распределениях ошибок прогноза.

Список источников

1. Субботина И.Е., Бувевич А.Г., Сергеев А.П., Шичкин А.В., Баглаева Е.М., Ремезова М.С. Двухшаговый комбинированный алгоритм повышения точности прогнозирования концентрации метана в атмосферном воздухе на основе нейронной сети NARX и последующего прогнозирования невязок // Арктика: Экология и экономика. – 2020. – № 2 (38). – С. 59–67. DOI:10.25283/2223-4594-2020-2-59-67.
2. Ali I., Alharbi O.M.L., AlOthman Z.A., Al-Mohaimed A.M., Alwarthan A. Modeling of fenuron pesticide adsorption on CNTs for mechanistic insight and removal in water // Environmental Research. – 2019. – Vol. 170. – P. 389–397. DOI:10.1016/j.envres.2018.12.066.
3. Baglaeva E.M., Sergeev A.P., Shichkin A.V., Buevich A. G. The extraction of the training subset for the spatial distribution modelling of the heavy metals in topsoil // Catena. – 2021. – Vol. 207. – 105699. DOI:10.1016/j.catena.2021.105699.
4. Berry K.J., Johnston J.E., Mielke P.W., Johnston L.A. Permutation methods. Part II // Wiley Interdisciplinary Reviews: Computational Statistics. – 2018. – Vol. 10. – Iss. 3. – e1429. DOI:10.1002/wics.1429.
5. Broomhead D.S., Lowe D. Radial basis functions, multi-variable functional interpolation and adaptive networks. – Malvern : Royals Signals & Radar Establishment, 1988. – 34 p.
6. Fernández Jaramillo J.M., Mayerle R. Sample selection via angular distance in the space of the arguments of an artificial neural network // Computers and Geosciences. – 2018. – Vol. 114. – P. 98–106. DOI:10.1016/j.cageo.2018.02.003.
7. Fişek M.H., Barlas Z. Permutation tests for goodness-of-fit testing of mathematical models to experimental data // Social Science Research. – 2013. – Vol. 42. – Iss. 2. – P. 482–495. DOI:10.1016/j.ssresearch.2012.09.010.
8. Goovaerts P. Geostatistics in soil science: state-of-the-art and perspectives // Geoderma. – 1999. – Vol. 89. – P. 1–45. DOI:10.1016/S0016-7061(98)00078-0.
9. Graf R., Zhu S., Sivakumar B. Forecasting river water temperature time series using a wavelet–neural network hybrid modelling approach // Journal of Hydrology. – 2019. – Vol. 578. – 124115. DOI:10.1016/j.jhydrol.2019.124115.
10. Hamilton S.H., Fu B., Guillaume J.H.A., Badham J., Elsawah S., Gober P., Hunt R.J., Iwanaga T., Jakeman A.J., Ames D.P., Curtis A., Hill M.C., Pierce S., Zare F. A framework for characterising and evaluating the effectiveness of environmental modelling // Environmental Modelling & Software. – 2019. – Vol. 118. – P. 83–98. DOI:10.1016/j.envsoft.2019.04.008.
11. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. – New York : Springer, 2009. – 745 p.
12. Lambiotte R., Rosvall M., Scholtes I. From networks to optimal higher-order models of complex systems // Nature Physics. – 2019. – Vol. 15. – P. 313–320. DOI:10.1038/s41567-019-0459-y.

13. *Liodakis S., Kyriakidis P., Gaganis P.* Conditional Latin Hypercube Simulation of (Log) Gaussian Random Fields // *Mathematical Geosciences*. – 2018. – Vol. 50. – P. 127–146. DOI:10.1007/s11004-017-9715-9.
14. *McKay M.D., Beckman R.J., Conover W.J.* Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code // *Technometrics*. – 1979. – Vol. 21. – Iss. 2. – P. 239–245. DOI: 10.1080/00401706.1979.10489755.
15. *Mielke P.W., Berry K.J.* Permutation Methods. A Distance Function Approach. – New York : Springer, 2001. DOI:10.1007/978-1-4757-3449-2.
16. *Moré J.J.* The Levenberg-Marquardt algorithm: Implementation and theory // *Numerical Analysis. (Lecture Notes in Mathematics. Vol. 630.)* / Ed. G.A. Watson. – Berlin, Heidelberg : Springer, 1978. – P. 105–116. DOI:10.1007/BFb0067700.
17. *Nath A., Subbiah K.* The role of pertinently diversified and balanced training as well as testing data sets in achieving the true performance of classifiers in predicting the antifreeze proteins // *Neurocomputing*. – 2018. – Vol. 272. – P. 294–305. DOI:10.1016/j.neucom.2017.07.004.
18. *Rosenblatt F.* Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. – Washington : Spartan Books, 1962. – 616 p.
19. *Ross S.M.* Simulation, bootstrap statistical methods, and permutation tests // *Introduction to Probability and Statistics for Engineers and Scientists*. – Amsterdam : Academic Press; Elsevier, 2021. – P. 619–647. DOI:10.1016/B978-0-12-824346-6.00024-7.
20. *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.
21. *Samadianfard S., Asadi E., Jarhan S., Kazemi H., Kheshtgar S., Kisi O., Sajjadi S., Manaf A.A.* Wavelet neural networks and gene expression programming models to predict short-term soil temperature at different depths // *Soil and Tillage Research*. – 2018. – Vol. 175. – P. 37–50. DOI:10.1016/j.still.2017.08.012.
22. *Sergeev A.P., Buevich A.G., Baglaeva E.M., Shichkin A.V.* Combining spatial autocorrelation with machine learning increases prediction accuracy of soil heavy metals // *Catena*. – 2019. – Vol. 174. – P. 425–435. DOI: 10.1016/j.catena.2018.11.037.
23. *Simão M.L., Videiro P.M., Silva P.B.A., Assad L.P.F., Sagrilo L.V.S.* Application of Taylor diagram in the evaluation of joint environmental distributions' performances // *Marine Systems & Ocean Technology*. – 2020. – Vol. 15. – P. 151–159. DOI:10.1007/s40868-020-00081-5.
24. *Specht D.F.* A general regression neural network // *IEEE Transactions on Neural Networks*. – 1991. – Vol. 2. – No. 6. – P. 568–576.
25. *Tan A.P., Cheong C.H., Lee T., Seng K.Y., Teo C.J.* Computer modelling of heat strain responses of exercising personnel in tropical climate // *Computers in Biology and Medicine*. – 2021. – Vol. 134. – 104530. DOI:10.1016/j.combiomed.2021.104530.
26. *Taylor K.E.* Summarizing multiple aspects of model performance in a single diagram // *Journal of Geophysical Research Atmospheres*. – 2001. – Vol. 106. – Iss. D7. – P. 7183–7192. DOI:10.1029/2000JD900719.
27. *Taylor A.B., MacKinnon D.P.* Four applications of permutation methods to testing a single-mediator model // *Behavior Research Methods*. – 2012. – Vol. 44. – Iss. 3. – P. 806–844. DOI:10.3758/s13428-011-0181-x.
28. *Ucun Ozel H., Gemici B. T., Gemici E., Ozel H.B., Cetin M., Sevik H.* Application of artificial neural networks to predict the heavy metal contamination in the Bartın River // *Environmental Science and Pollution Research*. – 2020. – Vol. 27. – P. 42495–42512. DOI:10.1007/s11356-020-10156-w.
29. *Wang F., Gao J., Zha Y.* Hyperspectral sensing of heavy metals in soil and vegetation: Feasibility and challenges // *ISPRS Journal of Photogrammetry and Remote Sensing*. – 2018. – Vol. 136. – P. 73–84. DOI:10.1016/j.isprsjprs.2017.12.003.
30. *Willmott C.J.* On the validation of models // *Physical Geography*. – 1981. – Vol. 2. – P. 184–194. DOI:10.1080/02723646.1981.10642213.
31. *Willmott C.J., Robeson S.M., Matsuura K.* A refined index of model performance // *International Journal of Climatology*. – 2012. – Vol. 32. – Iss. 13. – P. 2088–2094. DOI:10.1002/joc.2419.
32. *Xu W., Peng H., Zeng X., Zhou F., Tian X., Peng X.* Deep belief network-based AR model for nonlinear time series forecasting // *Applied Soft Computing*. – 2019. – Vol. 77. – P. 605–621. DOI:10.1016/J.ASOC.2019.02.006.
33. *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: A Case Study , Ghana // *Earth Sciences Research Journal*. – 2019. – Vol. 23. – Iss. 1. – P. 67–77. doi:10.15446/esrj.v23n1.63860.

References

1. *Subbotina I.E., Buevich A.G., Sergeev A.P., Shichkin A.V., Baglaeva E.M., Remezova M.S.* Two-step combined algorithm for improving the accuracy of predicting methane concentration in atmospheric air based on the NARX neural network and subsequent prediction of residuals. *Arctic: ecology and economy*. 2020;38:59–67. DOI:10.25283/2223-4594-2020-2-59-67.
2. *Ali I., Alharbi O.M.L., AlOthman Z.A., Al-Mohaimed A.M., Alwarthan A.* Modeling of fenuron pesticide adsorption on CNTs for mechanistic insight and removal in water. *Environmental Research*. 2019;170:389–397. DOI:10.1016/j.envres.2018.12.066.
3. *Baglaeva E.M., Sergeev A.P., Shichkin A.V., Buevich A. G.* The extraction of the training subset for the spatial distribution modelling of the heavy metals in topsoil. *Catena*. 2021;207:105699. DOI:10.1016/j.catena.2021.105699.
4. *Berry K.J., Johnston J.E., Mielke P.W., Johnston L.A.* Permutation methods. Part II. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2018;10(3):e1429. DOI:10.1002/wics.1429.
5. *Broomhead D.S., Lowe D.* Radial basis functions, multi-variable functional interpolation and adaptive networks. Malvern: Royals Signals & Radar Establishment; 1988. 34 p.
6. *Fernández Jaramillo J.M., Mayerle R.* Sample selection via angular distance in the space of the arguments of an artificial neural network. *Computers and Geosciences*. 2018;114:98–106. DOI:10.1016/j.cageo.2018.02.003.
7. *Fişek M.H., Barlas Z.* Permutation tests for goodness-of-fit testing of mathematical models to experimental data. *Social Science Research*. 2013;42(2):482–495. DOI:10.1016/j.ssresearch.2012.09.010.
8. *Goovaerts P.* Geostatistics in soil science: state-of-the-art and perspectives. *Geoderma*. 1999;89:1–45. DOI:10.1016/S0016-7061(98)00078-0.

9. Graf R., Zhu S., Sivakumar B. Forecasting river water temperature time series using a wavelet–neural network hybrid modelling approach. *Journal of Hydrology*. 2019;578:124115. DOI:10.1016/j.jhydrol.2019.124115.
10. Hamilton S.H., Fu B., Guillaume J.H.A., Badham J., Elsworth S., Gober P., Hunt R.J., Iwanaga T., Jakeman A.J., Ames D.P., Curtis A., Hill M.C., Pierce S., Zare F. A framework for characterising and evaluating the effectiveness of environmental modelling. *Environmental Modelling & Software*. 2019;118:83–98. DOI:10.1016/j.envsoft.2019.04.008.
11. Hastie T., Tibshirani R., Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer; 2009. 745 p.
12. Lambiotte R., Rosvall M., Scholtes I. From networks to optimal higher-order models of complex systems. *Nature Physics*. 2019;15:313–320. DOI:10.1038/s41567-019-0459-y.
13. Liodakis S., Kyriakidis P., Gaganis P. Conditional Latin Hypercube Simulation of (Log) Gaussian Random Fields. *Mathematical Geosciences*. 2018;50:127–146. DOI:10.1007/s11004-017-9715-9.
14. McKay M.D., Beckman R.J., Conover W.J. Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*. 1979;21(2):239–245. DOI: 10.1080/00401706.1979.10489755.
15. Mielke P.W., Berry K.J. *Permutation Methods. A Distance Function Approach*. New York: Springer; 2001. DOI:10.1007/978-1-4757-3449-2.
16. Moré J.J. The Levenberg-Marquardt algorithm: Implementation and theory. In: *Numerical Analysis*. (Lecture Notes in Mathematics. Vol. 630.) Watson G.A., ed. Berlin, Heidelberg: Springer; 1978. p. 105–116. DOI:10.1007/BFb0067700.
17. Nath A., Subbiah K. The role of pertinently diversified and balanced training as well as testing data sets in achieving the true performance of classifiers in predicting the antifreeze proteins. *Neurocomputing*. 2018;272:294–305. DOI:10.1016/j.neucom.2017.07.004.
18. Rosenblatt F. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington: Spartan Books; 1962. 616 p.
19. Ross S.M. Simulation, bootstrap statistical methods, and permutation tests. In: *Introduction to Probability and Statistics for Engineers and Scientists*. Amsterdam: Academic Press, Elsevier; 2021. p. 619–647. DOI:10.1016/B978-0-12-824346-6.00024-7.
20. 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.
21. Samadianfard S., Asadi E., Jarhan S., Kazemi H., Kheshtgar S., Kisi O., Sajjadi S., Manaf A.A. Wavelet neural networks and gene expression programming models to predict short-term soil temperature at different depths. *Soil and Tillage Research*. 2018;175:37–50. DOI:10.1016/j.still.2017.08.012.
22. Sergeev A.P., Buevich A.G., Baglaeva E.M., Shichkin A.V. Combining spatial autocorrelation with machine learning increases prediction accuracy of soil heavy metals. *Catena*. 2019;174:425–435. DOI: 10.1016/j.catena.2018.11.037.
23. Simão M.L., Videiro P.M., Silva P.B.A., Assad L.P.F., Sagrilo L.V.S. Application of Taylor diagram in the evaluation of joint environmental distributions' performances. *Marine Systems & Ocean Technology*. 2020;15:151–159. DOI:10.1007/s40868-020-00081-5.
24. Specht D.F. A general regression neural network. *IEEE Transactions on Neural Networks*. 1991;2(6):568–576.
25. Tan A.P., Cheong C.H., Lee T., Seng K.Y., Teo C.J. Computer modelling of heat strain responses of exercising personnel in tropical climate. *Computers in Biology and Medicine*. 2021;134:104530. DOI:10.1016/j.combiomed.2021.104530.
26. Taylor A.B., MacKinnon D.P. Four applications of permutation methods to testing a single-mediator model. *Behavior Research Methods*. 2012;44(3):806–844. DOI:10.3758/s13428-011-0181-x.
27. Taylor K.E. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research Atmospheres*. 2001;106(D7):7183–7192. DOI:10.1029/2000JD900719.
28. Uzun Ozel H., Gemici B. T., Gemici E., Ozel H.B., Cetin M., Sevik H. Application of artificial neural networks to predict the heavy metal contamination in the Bartın River. *Environmental Science and Pollution Research*. 2020;27:42495–42512. DOI:10.1007/s11356-020-10156-w.
29. Wang F., Gao J., Zha Y. Hyperspectral sensing of heavy metals in soil and vegetation: Feasibility and challenges. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2018;136:73–84. DOI:10.1016/j.isprsjprs.2017.12.003.
30. Willmott C.J. On the validation of models. *Physical Geography*. 1981;2:184–194. DOI:10.1080/02723646.1981.10642213.
31. Willmott C.J., Robeson S.M., Matsuura K. A refined index of model performance. *International Journal of Climatology*. 2012;32(13):2088–2094. DOI:10.1002/joc.2419.
32. Xu W., Peng H., Zeng X., Zhou F., Tian X., Peng X. Deep belief network-based AR model for nonlinear time series forecasting. *Applied Soft Computing*. 2019;77:605–621. DOI:10.1016/j.asoc.2019.02.006.
33. 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: A Case Study, Ghana. *Earth Sciences Research Journal*. 2019;23(1):67–77. doi:10.15446/esrj.v23n1.63860.

Статья поступила в редакцию 05.05.2022, одобрена после рецензирования 12.05.2022, принята к публикации 25.05.2022.
The article was submitted 05.05.2022; approved after reviewing 12.05.2022; accepted for publication 25.05.2022.

Информация об авторах

Сергеев Александр Петрович

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

Information about authors

Alexander P. Sergeev

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

Буторова Анастасия Сергеевна

Инженер-исследователь Института промышленной экологии
УрО РАН

620990 г. Екатеринбург, ул. С. Ковалевской, д. 20;

Аспирант 1 курса Института радиоэлектроники и
информационных технологий – РтФ Уральского федерального
университета им. Б.Н. Ельцина

620002 г. Екатеринбург, ул. Мира, д. 19

e-mail: amoskalyova11@yandex.ru

Anastasia S. Butorova

Research engineer 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: amoskalyova11@yandex.ru

Шичкин Андрей Васильевич

Научный сотрудник Института промышленной экологии
УрО РАН

620990 г. Екатеринбург, ул. С. Ковалевской, д. 20

e-mail: and@ecko.uran.ru

Andrey V. Shichkin

Researcher of the Institute of Industrial Ecology, UB RAS
20, S. Kovalevskoy str., Ekaterinburg, 620990, Russia

e-mail: and@ecko.uran.ru

Бувич Александр Геннадьевич

Научный сотрудник Института промышленной экологии
УрО РАН

620990 г. Екатеринбург, ул. С. Ковалевской, д. 20

e-mail: bag@ecko.uran.ru

Alexander G. Buevich

Researcher of the Institute of Industrial Ecology, UB RAS
20, S. Kovalevskoy str., Ekaterinburg, 620990, Russia

e-mail: bag@ecko.uran.ru

Баглаева Елена Михайловна

Кандидат физико-математических наук

Старший научный сотрудник Института промышленной
экологии УрО РАН

620990 г. Екатеринбург, ул. С. Ковалевской, д. 20

e-mail: e.m.baglaeva@urfu.ru

Elena M. Baglaeva

Candidate of Physical and Mathematical Sciences

Senior Researcher of the Institute of Industrial Ecology, UB RAS
20, S. Kovalevskoy str., Ekaterinburg, 620990, Russia

e-mail: e.m.baglaeva@urfu.ru

Субботина Ирина Евгеньевна

Кандидат физико-математических наук

Научный сотрудник Института промышленной
экологии УрО РАН

620990 г. Екатеринбург, ул. С. Ковалевской, д. 20

e-mail: iesub@mail.ru

Irina Ev. Subbotina

Candidate of Physical and Mathematical Sciences

Researcher of the Institute of Industrial Ecology, UB RAS
20, S. Kovalevskoy str., Ekaterinburg, 620990, Russia

e-mail: iesub@mail.ru