

УДК 504.064.2.001.18

ОЦЕНКА ПРОСТРАНСТВЕННОГО РАСПРЕДЕЛЕНИЯ ХРОМА В СУБАРКТИЧЕСКОМ НОЯБРЬСКЕ С ИСПОЛЬЗОВАНИЕМ КОКРИГИНГА, ГЕНЕРАЛИЗОВАННОЙ РЕГРЕССИОННОЙ НЕЙРОННОЙ СЕТИ, МНОГОСЛОЙНОГО ПЕРСЕПТРОНА И ГИБРИДНОЙ ТЕХНИКИ

© 2019 г. А. Г. Бувевич^{1,*}, И. Е. Субботина^{1,**}, А. В. Шичкин^{1,***}, А. П. Сергеев^{1,****}, Е. М. Баглаева^{1,*****}

¹Институт промышленной экологии Уральского отделения Российской академии наук (ИПЭ УрО РАН), ул. С. Ковалевской, 20, г. Екатеринбург, 620219 Россия

*E-mail: bagalex3@gmail.com

**E-mail: iesub@mail.ru

***E-mail: and@ecko.uran.ru

****E-mail: alexanderpsergeev@gmail.com

*****E-mail: sem@ecko.uran.ru

Поступила в редакцию 22.11.2018 г.

Сочетание методов геостатистической интерполяции (кригинга) и машинного обучения (искусственных нейронных сетей – ИНС) приводит к повышению точности прогнозирования. В работе рассматривается применение кригинга остатков искусственной нейронной сети для прогнозирования пространственного загрязнения поверхностного слоя почвы хромом (Cr). Рассмотрены две нейронные сети: обобщенная регрессионная нейронная сеть (GRNN) и многослойный перцептрон (MLP), а также комбинированный метод – кригинг остатков многослойного перцептрона (MLPRK). Исследование основано на результатах скрининга поверхностного слоя почвы в субарктическом Ноябрьске (Россия). Модели разработаны на основе компьютерного моделирования с минимизацией среднеквадратической ошибки (RMSE). Лучшую прогностическую точность показала модель MLPRK.

Ключевые слова: искусственные нейронные сети, хром, кригинг остатков, кокригинг.

DOI: <https://doi.org/10.31857/S0869-78092019277-86>

ВВЕДЕНИЕ

Значительная часть исследований экологических проблем заключается в изучении загрязнения почв. Различные факторы, такие как климат, градостроительные, метеорологические, гидрологические условия, влияют на результаты мониторинга. Часто эти факторы приводят к неоднородностям и в некоторых случаях к аномалиям в распределении загрязняющих веществ [13, 24, 32]. Это усложняет процесс прогнозирования, поэтому требуется комплексный и долгосрочный мониторинг с использованием менее трудоемких, но не менее точных методов. Пространственная интерполяция является одним из методов, который способствует определению расположения источников загрязнения с достаточной точностью.

В геостатистике (например, в кригинге) предсказания производятся в произвольных точках пространства с учетом пространственной корре-

ляции значений выборки на основе предположения о средней и ковариационной стационарности [12, 31]. Кригинг известен как метод оптимальной интерполяции, который генерирует наилучшие линейные несмещенные оценки в каждой точке, используя модель вариограммы [19].

Методы кригинга продемонстрировали преимущества в прогнозировании свойств почв и количества осадков по сравнению с детерминированными методами [10, 23, 30]. Кокригинг – многомерный вариант кригинга, который вычисляет оценки элемента с помощью других элементов при наличии сильной корреляции между ними [20].

Точность кригинга (кокригинга) зависит от плотности и размера сетки выборки, поскольку метод основан на интерполяции. Однако не всегда возможно отобрать необходимое количество образцов из-за ограничений, связанных с временем

или ресурсом. Для преодоления этих недостатков и повышения точности существующих методов требуется более эффективный метод.

В настоящее время подходящим методом прогнозирования являются искусственные нейронные сети (ИНС), которые предоставляют множество мощных методов прогнозирования, распознавания образов, анализа данных и многих других. В традиционной модели ИНС известные переменные (пространственные координаты) используются в качестве входных сигналов, а предсказательные переменные (содержание элементов в анализируемой среде) используются в качестве выходов. Функциональная связь между ними устанавливается через сеть синаптических весов, которые определяются посредством итеративного процесса обучения ИНС. Обзор [5] показал высокую универсальность ИНС. В последнее время этот метод широко используется при решении экологических проблем. Методы ИНС могут помочь прогнозировать загрязнение в сложных нелинейных системах; также успешно прогнозируют содержание загрязняющих веществ в тех местах, где не проводятся наблюдения [18, 26, 27]. ИНС показали более высокую точность прогноза, чем другие методы или экспертные оценки [1, 2].

Существует большое количество различных типов нейронных сетей, каждый из которых имеет свои преимущества и недостатки. Часто используемый тип ИНС в исследованиях окружающей среды – многослойный персептрон (MLP). Персептроны используются в исследованиях по оценке содержания химических веществ в почве [4, 7, 9, 17]. Ряд исследователей применяли персептроны для оценки ресурсов [15, 22], и в большинстве случаев было показано превосходство MLP над геостатистическими и детерминированными методами.

Обобщенные регрессионные нейронные сети (GRNN) представляют собой вариацию нейронных сетей радиальных базисных функций (RBF). GRNN используются как интерполяторы и известны как универсальные аппроксиматоры функций, которые могут аппроксимировать любую непрерывную нелинейную функцию [21]. Ключевое различие между GRNN и MLP заключается в том, что GRNN не требуют процесса обучения с использованием продолжительных итерационных процедур, применяемых для сетей обратного распространения ошибки (MLP и т.д.). GRNN аппроксимирует функцию за один проход, получая оценки непосредственно из набора данных обучения, минимизируя ошибку оценки так называемым параметром SPREAD. Подобно RBF, GRNN имеет радиально базисный слой, где число нейронов равно или меньше числа элементов набора тренировок.

Чтобы обойти слабые места и усилить достоинства упомянутых моделей, было предложено комбинировать различные методы. Исследователи успешно используют гибридизацию геостатистических и нейронных подходов, которая уменьшает ошибки моделей [8, 14, 16, 29].

В этой работе рассмотрены две прогностические модели на основе только ИНС (MLP, GRNN), а также гибридная модель, объединяющая прогнозирование на основе ИНС и кригинга (MLPRK) для прогнозирования содержания Cr в верхнем слое почвы в г. Ноябрьск (Ямало-Ненецкий автономный округ, Россия). Проведено сравнение результатов, полученных с применением этих моделей.

МАТЕРИАЛЫ И МЕТОДЫ

Область исследования

Данные для исследования были получены по результатам почвенной съемки в г. Ноябрьске (N63.193°, E75.507°). Площадь отбора проб составила примерно 16.5 км². Детальное пространственное расположение точек выборки показано на рис. 1. Ландшафт исследуемой территории плоский, грунт состоит из торфяно-подзолисто-глейвиально-гумусового песка. Всего было отобрано 237 образцов верхнего слоя почвы на глубину 0.05 м. Значения содержания химических элементов были получены химическим анализом.

Химический анализ

Подготовка образцов почвы и химический анализ проводились в соответствии со стандартными требованиями в сертифицированной химико-аналитической лаборатории, которая отвечает всем требованиям, предъявляемым к компетенции испытательных и калибровочных лабораторий ИСО/МЭК 17025:2005.

Для анализа применялся метод атомно-абсорбционной спектроскопии (ААС) с распылителями пламени. Высушенные и измельченные образцы почвы разлагали 1 см³ концентрированной хлорной кислоты, 1 см³ концентрированной азотной кислоты и 3 см³ концентрированной фтористоводородной кислоты. После смешивания и нагрева до 50°C в течение 1 часа раствор расщепляли дополнительными 2 см³ концентрированной азотной кислоты и 3 см³ концентрированной фтористоводородной кислоты и выдерживали в течение 1 часа при температуре 90°C. Затем смесь выпаривали досуха, добавляли 1 см³ концентрированной азотной кислоты и 20 см³ бидистиллированной воды и кипятили раствор для растворения осадка. Далее раствор нагревали, чтобы уменьшить

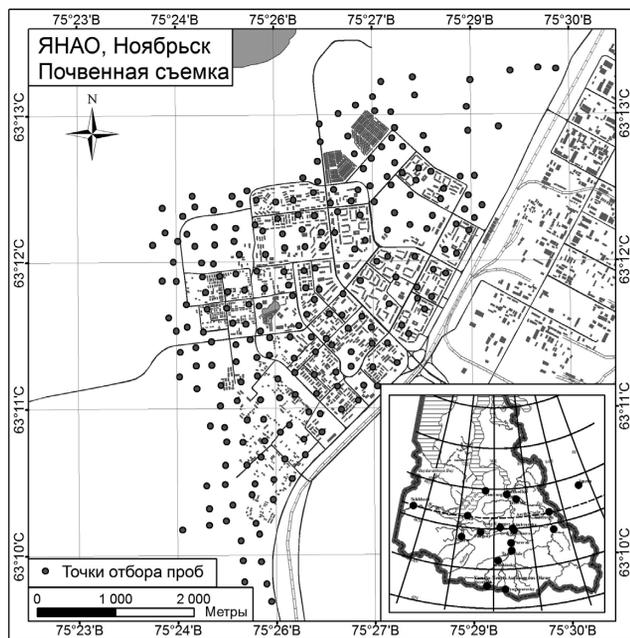


Рис. 1. Карта-схема места отбора проб: г. Ноябрьск, Ямало-Ненецкий автономный округ, Россия; точки обозначены места отбора проб.

объем в 2 раза, и фильтровали через предварительно промытый и предварительно подготовленный фильтр. Раствор анализировали на приборе Perkin Elmer AAS-3 методом атомно-абсорбционной спектроскопии. Предел обнаружения для Cr составлял 10 мг/кг.

В пробах определено содержание восьми элементов (Al, Cr, Mn, Fe, Co, Ni, Zn, Pb).

Пространственное прогнозирование содержания Cr

Алгоритм исследования

Для оценки содержания загрязняющих веществ и прогноза их распределения в неизвестных местах авторы применили три конкурирующих метода: два метода на основе ИНС (MLP и GRNN), а также гибридную модель MLPRK. Программирование ИНС проводилось в среде MATLAB.

MLPRK – это трехступенчатый алгоритм, объединяющий два различных метода интерполяции. Первым шагом является оценка крупномасштабных нелинейных трендов с использованием нейронных сетей (MLP). Второй шаг – анализ остатков ординарным кригингом; последний шаг – оценка, полученная как сумма предсказаний ИНС и результата анализа остатков ординарным кригингом.

В качестве входных параметров моделей используются пространственные координаты, в качестве выходных – концентрации Cr. Кокригинг

использует концентрации трех коэлементов (Fe, Co, Ni) в качестве входных данных.

Все образцы были случайным образом разделены на независимые обучающие и тестовые наборы данных. Набор обучающих данных (165 образцов) использовался для обучения ИНС. Набор тестовых данных (72 образца) использовался только для проверки и оценки точности моделей.

Изотропия данных

Для определения зависимости пространственной структуры от направления экспериментальные вариограммы строятся в разных направлениях. Пары выбираются из соответствующего сектора с заданным направлением, углом раствора и шириной полосы, что сужает поле поиска на больших расстояниях.

Инструментом, дающим представление о поведении пространственной структуры, является поверхность вариограммы. Для ее построения вектор h представляется в виде проекции на координатные оси Δx и Δy . Поверхность вариограммы – это поверхность значений, вычисленных на регулярной сетке в пространстве лага. Поверхность вариограммы имеет точечную симметрию относительно точки (0, 0). В случае зависимости от пространственной ориентации пар точек на поверхности вариограммы это может указывать на наличие анизотропии в данных.

Создание ИНС

В качестве первого типа ИНС использовался многослойный перцептрон прямого распространения с методом обучения Левенберга-Маркварта [28]. Структура сети определялась методом компьютерного моделирования. Скрытый слой состоял из нескольких нейронов, выходной слой представлял собой содержание элемента в соответствующем образце. Выбор количества нейронов в скрытом слое осуществлялся путем вычисления RMSE предсказания концентрации элемента (Cr) для обучающего (165 точек), тестового (72 точек) и полного набора данных (237 точек). Количество нейронов скрытого слоя изменялось от 2 до 20. Каждая сеть обучалась 500 раз, и выбиралась лучшая из них. Качество сетевого обучения проверялось коэффициентом корреляции Спирмена, средней абсолютной ошибкой (MAE) (4) и RMSE (5) между результатами прогнозов сетей и набором данных для обучения.

В качестве второй ИНС была выбрана GRNN. Основным параметром, характеризующий аппроксимационные свойства и точность интерполяции для сети GRNN, – параметр SPREAD. Во время

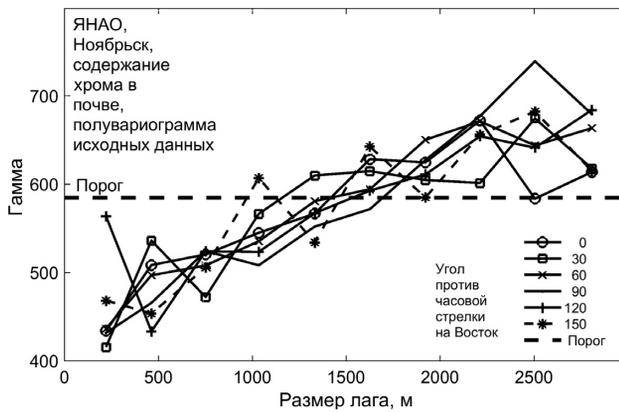


Рис. 2. Полувариограммы в шести направлениях.

моделирования параметр SPREAD изменялся от 0.001 до 0.30 с шагом 0.001. Сеть GRNN имела 165 входных нейронов в соответствии со 165 точками обучающей подвыборки.

Оценка остатков ординарным кригингом

Процедура оценки остатков ординарным кригингом (ОК) может быть определена следующим образом:

$$r(x_i) = Z(x_i) - Z_{\text{инс}}(x_i), \quad (1)$$

где $r(x_i)$ – остатки набора данных x_i , $Z(x_i)$ – измеренные значения, $Z_{\text{инс}}(x_i)$ – значения, оцененные нейронной сетью. Полученные остатки оценивали с использованием ОК. Оценка в ОК построена как линейная комбинация входных данных:

$$r_{\text{ОК}}(x) = \sum \lambda_i r(x_i), \quad (2)$$

где $r_{\text{ОК}}$ – оценочное значение в точке x с использованием ОК, λ_i – оптимальный вес с условием $\sum \lambda_i = 1$. Для прогнозирования остатков использовался ОК в приложении ArcGIS. Окончательная оценка содержания загрязняющих веществ $Y(x_i)$ была получена как сумма оценки нейронной сети и оценки остатков ОК.

$$Y_{(x_i)} = Z_{\text{инс}(x_i)} + r_{\text{ОК}(x_i)}. \quad (3)$$

Оценка точности методов интерполяции

Прогностическая точность каждого выбранного подхода была определена ошибками MAE (4) и RMSE (5) между предсказанием модели и данными из тестового набора данных.

$$MAE = \frac{1}{n} \sum_{i=1}^n |z_{\text{mod}}(x_i) - z(x_i)|, \quad (4)$$

Таблица 1. Корреляционная матрица содержаний элементов

	Al	Cr	Mn	Fe	Co	Ni	Zn	Pb
Al	1							
Cr	-0.01	1						
Mn	0.39	0.54	1					
Fe	0.16	0.83	0.69	1				
Co	0.04	0.65	0.59	0.64	1			
Ni	-0.01	0.72	0.60	0.75	0.70	1		
Zn	0.34	0.10	0.35	0.18	0.23	0.36	1	
Pb	0.34	0.07	0.14	0.07	-0.03	0.17	0.50	1

Таблица 2. Описательные статистики моделируемого элемента (Cr) и его коэлементов (Fe, Co, Ni)

Элемент	Минимальное значение, мг/кг	Максимальное значение, мг/кг	Среднее значение, мг/кг	Стандартное отклонение, мг/кг	CV*	Асимметрия	Экссесс	Медиана, мг/кг
Cr	16.6	140	62.4	24.2	0.39	0.81	0.44	58.8
Fe	4751	28270	132678	43967	0.33	0.79	0.75	12673
Co	2.00	11.4	4.50	1.49	0.33	0.99	2.13	4.42
Ni	3.58	41.9	11.4	4.52	0.40	2.24	12.0	11.0

*CV – коэффициент вариации

$$RMSE = \frac{1}{\sqrt{n}} \sqrt{\sum_{i=1}^n (z_{mod}(x_i) - z(x_i))^2}, \quad (5)$$

где $z_{mod}(x_i)$ – предсказанная концентрация (ИНС, кригинг), $z(x_i)$ – измеренная концентрация, n – количество точек.

РЕЗУЛЬТАТЫ

Описательная статистика

Корреляционный анализ (табл. 1) выявил возможные коэлементы для Сг. Три из них (Fe, Co, Ni) (выделены жирным) были использованы для кокригинга. Описательная статистика Сг и его коэлементов показана в табл. 2.

Общая концентрация элементов в городских условиях не превышает значения фоновых концентраций в Уральском регионе (Уральский Кларк) [2, 3] и в мировых почвах (World Clarke). Известно, что общее содержание Сг в подзолах находится в диапазоне от 2.6 до 34 мг/кг в Канаде [11] и от 3 до 200 мг/кг в США [25]. Концентрации Сг во всех точках отбора проб составляли от 16.6 до 140 мг/кг со средним значением 62.4 мг/кг и стандартным отклонением 24.2 мг/кг. Из-за асимметрии распределения медианное значение (58.8 мг/кг) более характерно для среднего содержания Сг в исследуемой области, чем среднее арифметическое. Коэлементы демонстрируют сходные характеристики, когда медианные значения (для Fe – 12 673 мг/кг, Co – 4.4 мг/кг, Ni – 11.0 мг/кг) более репрезентативны, чем средние.

Пространственное прогнозирование концентрации Сг

Результат теста хи-квадрат показывает, что распределение этой переменной близко к нормальному ($p=0.18$). Для демонстрации различий пространственных структур в зависимости от направления построены вариограммы в шести направлениях ($0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ$ и 150°) (рис. 2). Анизотропия исходных данных во всех этих направлениях не обнаруживается на вариограммах направления (см. рис. 2) и поверхности вариограммы (рис. 3).

Окончательная конфигурация выбранной сети MLP составила 2-6-1, т.е. скрытый слой содержит 6 нейронов (рис. 4). В нашем случае 165 точек выборки сформировали набор обучающих данных, который применяется для обучения сети.

Во время моделирования для GRNN параметр SPREAD изменяется с 0.001 до 0.30 с шагом 0.001. Минимальное значение RMSE достигается с параметром SPREAD 0.035 (рис. 5).

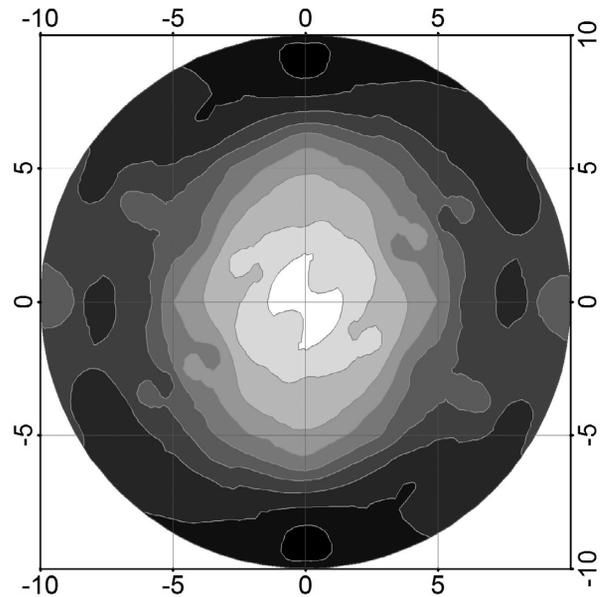


Рис. 3. Вариограммная поверхность.

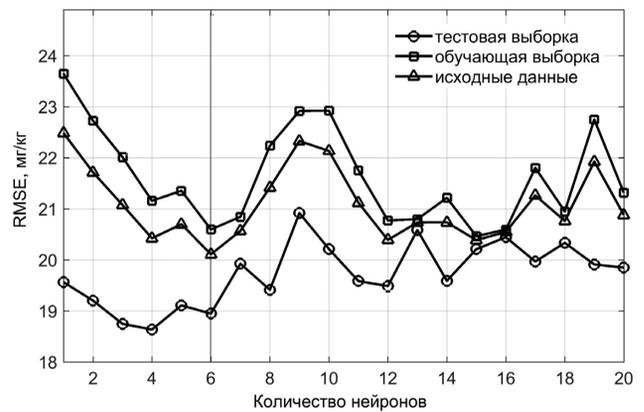


Рис. 4. Выбор структуры MLP на основе минимизации RMSE: среднеквадратичная ошибка (RMSE) нейронной сети для тестирования, обучения и общих данных для разного количества нейронов в скрытом слое для Сг.

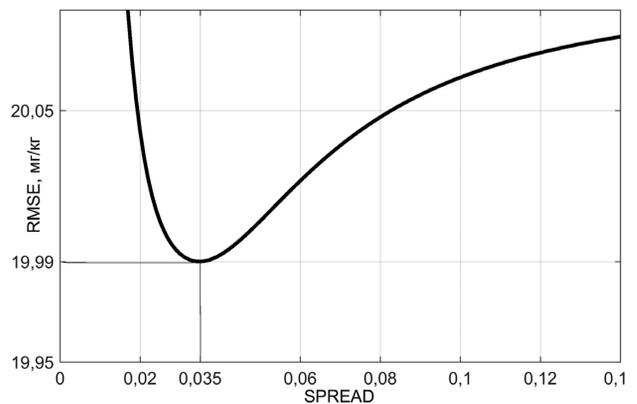


Рис. 5. Выбор параметра SPREAD на основе минимизации RMSE.

Таблица 3. Показатели точности моделей

Метод	Коэффициент ранговой корреляции Спирмена	RMSE, мг/кг	MAE, мг/кг
Кокригинг	0.15	20.7	15.3
MLP	0.42	18.9	14.7
GRNN	0.09	20.0	15.0
MLPRK	0.41	18.8	14.8

Оценка точности методов интерполяции

MLP и MLPRK показали увеличение точности моделирования по сравнению с геостатистическим методом (кокригинг) и GRNN. Как видно из табл. 3, модели на основе MLP имели меньшую RMSE, чем кокригинг (улучшение на 9.5%). MAE у моделей на основе MLP примерно на 4% лучше, чем у кокригинга. Базовая модель GRNN продемонстрировала неожиданно низкий коэффициент корреляции. Это означает, что метод не может применяться для моделирования поставленной задачи.

ОБСУЖДЕНИЕ

Авторы сравнили разные подходы к моделированию пространственного распределения концентраций химических элементов в поверхностном слое почвы (геостатистика, ИНС и гибридная модель MLPRK). Качество предсказания моделей проанализировано с помощью набора тестовых данных, который не применялся для обучения сетей или оценок кригинга. Направленные вариограммы исходных данных и поверхность вариограммы обучающих данных показаны на рис. 2, 3. Они подтверждают отсутствие какой-либо анизотропии. В табл. 3 приведены статистические параметры, используемые для оценки эффективности различных методов (лучшие значения, продемонстрированные на основе MLP, выделены жирным шрифтом).

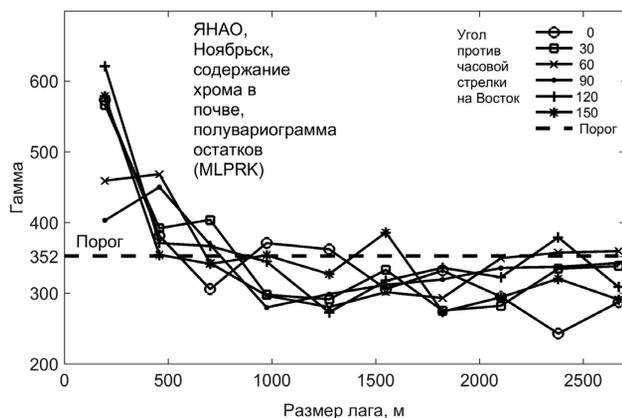


Рис. 6. Полувариограммы для остатков по различным направлениям для MLPRK.

Модель MLPRK хорошо воспроизводит пространственную структуру распределения Cr. Она выделяет структурированную информацию, сглаживая шум и локальную изменчивость. Изучение остатков подтверждает важность вариографии для анализа и моделирования пространственных данных с использованием алгоритмов нейронной сети. Направленные вариограммы для остатков демонстрируют чистый эффект самородка, который свидетельствует об отсутствии пространственной корреляции (рис. 6).

На рис. 7 показано сравнение карт распределения пространственного содержания Cr для различных моделей. Распределения были построены с использованием программного обеспечения ArcGIS. Оценка остатков MLP ординарным кригингом позволяет сгладить высокие и низкие значения концентраций химического элемента, что улучшает точность прогнозирования.

Одним из возможных объяснений относительно низкой эффективности методов геостатистики может быть сильная зависимость концентрации Cr от локальных изменений условий окружающей среды в исследуемой области. Причиной вариативности содержания Cr могут быть как внутренние факторы (такие как материал и текстуры почвы), так и внешние (такие как антропогенное воздействие, климат и т.д.). Как правило, сильно пространственно зависимые свойства могут определяться внутренними изменениями характеристик почвы, тогда как внешние изменения могут контролировать изменчивость слабо пространственно зависимых параметров [6].

ЗАКЛЮЧЕНИЕ

Цель выполненного исследования – разработка модели, которая может оценить пространственное распределение интересующего элемента верхнего слоя почвы (Cr) в г. Ноябрьске. В настоящей работе показано, что оптимальными показателями обладают модели с использованием MLP и MLPRK. Эффективность этих моделей определяется посредством их сравнительной оценки с другими моделями (кокригинг, GRNN).

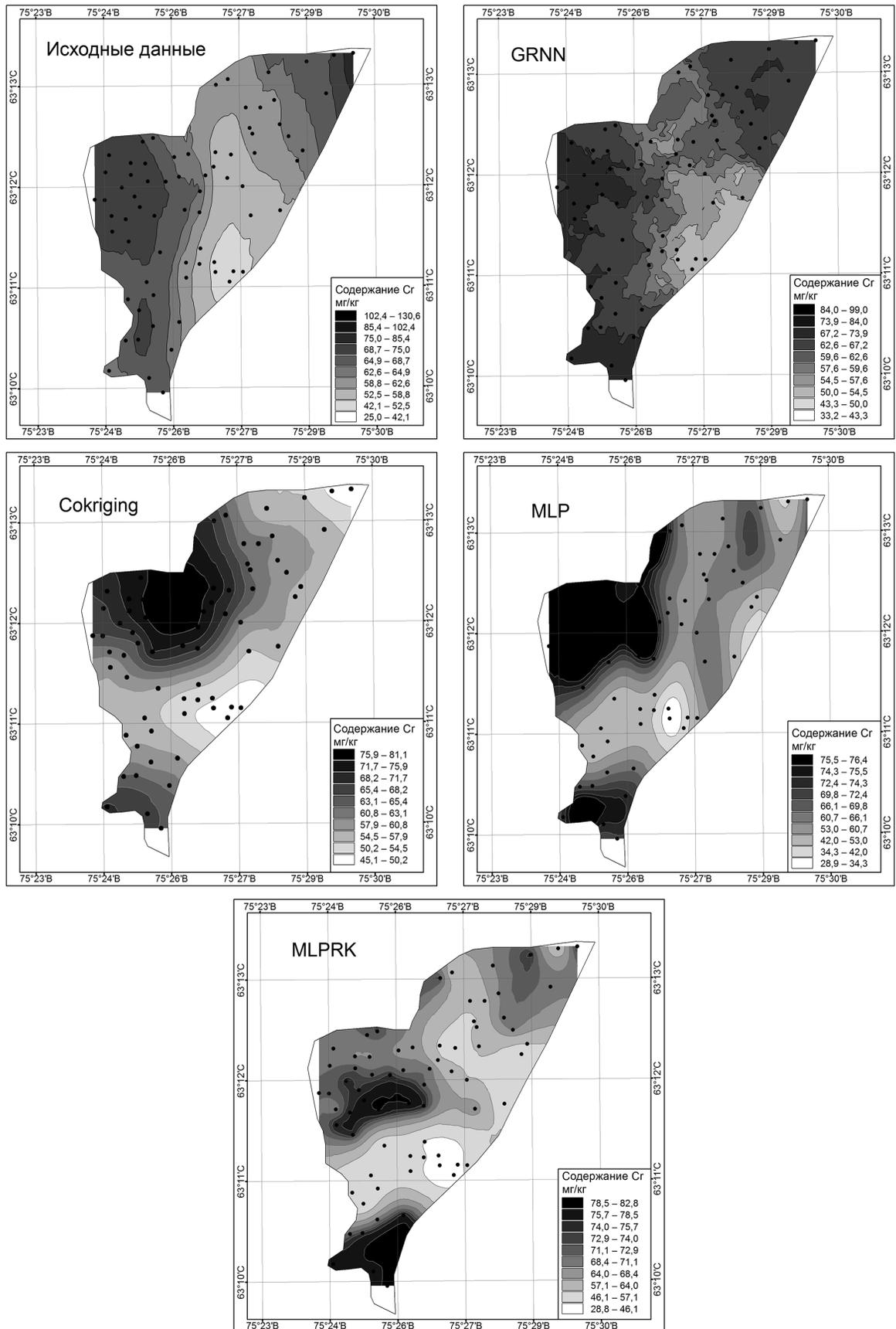


Рис. 7. Карты пространственного распределения содержания Cr на исследуемой области для различных моделей.

Таким образом, проводится сравнение различных подходов к прогнозированию распределения химических элементов в поверхностном слое почвы. Оценка ИНС с остатками предсказания ординарным кригингом уменьшает ошибки предсказания ИНС, что повышает точность моделей. Результаты показывают, что построенные модели на основе MLP более точны, чем основанные на геостатистике. Наибольшее уменьшение RMSE (9.5%) по сравнению с кокригингом наблюдается в модели MLPK.

Полученные результаты подтверждают обширные возможности гибридных методов (ИНС + кригинг), которые могут быть использованы для повышения точности моделирования пространственного распределения концентраций химических элементов в верхнем слое почвы городских районов, которое характеризуется высокой гетерогенностью.

СПИСОК ЛИТЕРАТУРЫ

1. Бувевич А.Г., Сергеев А.П., Тарасов Д.А., Рахматова А.Ю., Крамаренко А.А. Искусственные нейронные сети и геостатистика в прогнозировании распределения химических элементов на фоновой площадке // *Геоэкология*. 2017. №2. С. 74–82.
2. Войткевич Г.В., Мирошников А.Е., Поваренных А.С., Прохоров В.Г. Краткий справочник по геохимии. М.: Недра, 1977.
3. Саит Ю.Е., Ревич Б.А., Янин Е.П. и др. Геохимия окружающей среды. М.: Недра, 1990. 335 с.
4. Anagu I., Ingwersen J., Utermann J., Streck T. Estimation of heavy metal sorption in German soils using artificial neural networks // *Geoderma*, 2009, 152, pp.104–112.
5. Bishop C. Neural networks for pattern recognition // Clarendon, Oxford, 1995, 504p.
6. Cambardella C., Moorman T., Novak J., Parkin T., Karlen D., Turco R., Konopka A. Field-scale variability of soil properties in central Iowa soils // *Soil Sci. Soc. Am. J.*, 1994, 58, pp. 1501–1510.
7. Dai F., Zhou O., Lva 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, No 45, pp. 184–194.
8. Demyanov V., Kanevsky M., Chernov S., Savelieva E., Timonin V. Neural Network Residual Kriging Application for Climatic Data // *Journal of Geographic Information and Decision Analysis*, 1998, 2, pp. 215–232.
9. Falamaki A. Artificial neural network application for predicting soil distribution coefficient of nickel // *Journal of Environmental Radioactivity*, 2013, 115, pp. 6–12.
10. Forsythe K. W., Marvin C. H., Valancius C. J., Watt J. P., Aversa J. M., Swales S. J., Jakubek D.J., Shaker R. R. Geovisualization of Mercury Contamination in Lake St. Clair Sediments // *Journal of Marine Science and Engineering*, 2016, 4(1), 19.
11. Frank R., Ishida K., Suda P. Metals in agricultural soils of Ontario // *Canadian Journal of Soil Science*, 1976, 56, 181-196.
12. Goovaerts P. Geostatistics in soil science: State of the art and perspectives // *Geoderma*, 1999, 89, pp.1–45.
13. Guo G.H., 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, issue 3, pp. 410–418.
14. Kanevski M., Pozdnoukhov A., Timonin V. Machine Learning for Spatial Environmental Data. Theory // Applications and Software. 2009, EPFL Press.
15. Koike K., Matsuda S., Suzuki T., Ohmi M. Neural network-based estimation of principal metal contents in the Hokuroku district, northern Japan, for exploring Kuroko-type deposits // *Natural Resources Research*, 2002, 11, (2), pp. 135–156.
16. Lakes T., Müller D., Krüger C. Cropland change in southern Romania: A comparison of logistic regressions and artificial neural networks // *Landscape Ecology*, 2009, 24(9), pp. 1195–1206.
17. Li Y., Li C., Tao J.J., Wang L.D. Study on Spatial Distribution of Soil Heavy Metals in Huizhou City Based on BP-ANN Modeling and GIS // *Proc. Environmental Sciences*, 2011, 10, p. 1953–1960.
18. Liu F., He X., Zhou L. Application of generalized regression neural network residual kriging for terrain surface interpolation // *Proc. SPIE 7492, Int. Symposium on Spatial Analysis, Spatial-Temporal Data Modeling, and Data Mining*, 2009, 74925F.
19. Matheron G. Principles of geostatistics // *Economic Geology*, 1963, 58, pp. 1246–66.
20. Meer F. D. Introduction to geostatistics // *ITC Lecture Notes*, 1993, 72 pp.
21. Mohanty K., Majumdar T.J. Using artificial neural networks for synthetic surface fitting and the classification of remotely sensed data // *Int. Journal of Applied Earth Observation and Geoinformation*, 1999, 1(1), pp. 78–84.
22. Samanta B., Ganguli R., Bandopadhyay S. Comparing the Predictive Performance of Neural Networks with Ordinary Kriging in a Bauxite Deposit // *Transactions of the Institution of Mining and Metallurgy, Section A, Mining Technology*, 2005, 114, No. 3, pp. 129–139.
23. Schloeder C.A., Zimmerman N.E., Jacobs M.J. Comparison of methods for interpolating soil properties using limited data // *Soil Sci. Soc. Am. J.*, 2001, 65, pp. 470-479.
24. Sergeev A.P., Baglaeva E.M., Shichkin A.V. Case of soil surface chromium anomaly of a northern

- urban territory – preliminary results // *Atmospheric Pollution Research*, 2010, V. 1, pp. 44-49.
25. Shackle H. T., Boerngen J. G. Element concentrations in soils and other surficial materials of the conterminous United States, U.S. Geological Survey professional paper, 1984, 1270, 105 p.
 26. Shaker R., Tofan L., Bucur M., Costache S., Sava D., Ehlinger T. Land cover and landscape as predictors of groundwater contamination: a neural-network modelling approach applied to Dobrogea, Romania // *Journal of Environmental Protection and Ecology*, 2010, 11(1), pp. 337-348.
 27. Shaker R.R., Ehlinger T.J. Exploring non-linear relationships between landscape and aquatic ecological condition in southern Wisconsin: A GWR and ANN approach // *Int. Journal of Applied Geospatial Research*, 2014, 5(4), pp. 1-20.
 28. Shepherd A.J. Second-Order Methods for Neural Networks: Fast and Reliable Training Methods for Multi-Layer Perceptrons. Springer-Verlag, 1997, 145 p.
 29. Tarasov D.A., Buevich A.G., Sergeev A.P., Shichkin A.V. High Variation Topsoil Pollution Forecasting in the Russian Subarctic: Using Artificial Neural Networks Combined with Residual Kriging // *Applied Geochemistry*, 2018, vol. 88, part B, pp. 188-197.
 30. Worsham L., Markewitz D., Nibbelink N. Incorporating spatial dependence into estimates of soil carbon contents under different land covers // *Soil Sci. Am. J.*, 2010, 74, pp. 635–646.
 31. Yfantis E. A., Flatman, G. T., Behar, J. V. Efficiency of kriging estimation for square, triangular, and hexagonal grids // *Math. Geol.*, 1987, 19, pp. 183-205.
 32. Zhang C., Fay D., McGrath D., Grennan E., Carton O.T. Use of trans-Gaussian kriging for national soil geochemical mapping in Ireland // *Geochemistry: Exploration Environment Analysis*, 2008, vol. 8, pp. 255–265.
 4. Anagu, I., Ingwersen, J., Utermann, J., Streck T. Estimation of heavy metal sorption in German soils using artificial neural networks. *Geoderma*, 2009, 152, pp.104-112.
 5. Bishop, C. Neural networks for pattern recognition. Clarendon, Oxford, 1995, 504 p.
 6. Cambardella, C., Moorman, T., Novak, J., Parkin, T., Karlen, D., Turco, R., Konopka, A. Field-scale variability of soil properties in central Iowa soils. *Soil Sci. Soc. Am. J.*, 1994, no. 58, pp. 1501-1510.
 7. Dai, F., Zhou, O., Lva, 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, no. 45, pp. 184-194.
 8. Demyanov, V., Kanevsky, M., Chernov, S., Savelieva, E., Timonin, V. Neural Network Residual Kriging Application for Climatic Data. *Journal of Geographic Information and Decision Analysis*, 1998, no. 2, pp. 215-232.
 9. Falamaki, A. Artificial neural network application for predicting soil distribution coefficient of nickel. *Journal of Environmental Radioactivity*, 2013, no. 115, pp. 6-12.
 10. Forsythe, K.W., Marvin, C.H., Valancius, C. J., Watt, J.P., Aversa J.M., Swales, S.J., Jakubek, D.J., Shaker, R.R. Geovisualization of Mercury Contamination in Lake St. Clair Sediments. *Journal of Marine Science and Engineering*, 2016, no. 4(1), p. 19.
 11. Frank, R., Ishida, K., Suda, P. Metals in agricultural soils of Ontario. *Canadian Journal of Soil Science*, 1976, no. 56, pp. 181-196.
 12. Goovaerts, P. Geostatistics in soil science: State of the art and perspectives. *Geoderma*, 1999, no. 89, pp.1-45.
 13. Guo, G.H., 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, issue 3, pp. 410-418.
 14. Kanevski, M., Pozdnoukhov, A., Timonin, V. Machine Learning for Spatial Environmental Data Theory. Applications and Software, 2009, EPFL Press.
 15. Koike K., Matsuda S., Suzuki T., Ohmi M. Neural network-based estimation of principal metal contents in the Hokuroku district, northern Japan, for exploring Kuroko-type deposits . *Natural Resources Research*, 2002, no. 11(2), pp. 135-156.
 16. Lakes, T., Müller, D., Krüger, C. Cropland change in southern Romania: A comparison of logistic regressions and artificial neural networks. *Landscape Ecology*, 2009, no. 24(9), pp. 1195-1206.
 17. Li, Y., Li, C., Tao, J.J., Wang, L.D. Study on Spatial Distribution of Soil Heavy Metals in Huizhou City Based on BP-ANN Modeling and GIS. *Procedia Environmental Sciences*, 2011, no. 10, pp. 1953-1960.
 18. Liu, F., He, X., Zhou, L. Application of generalized regression neural network residual kriging for terrain surface interpolation. Proc. SPIE 7492, International

REFERENCES

1. Buevich, A.G., Sergeev, A.P., Tarasov, D.A., Rakhmatova, A.Y., Kramarenko, A.A. *Iskusstvennyye neironnye seti i geostatistika v prognozirovanii raspredeleniya khimicheskikh elementov na fonovoi ploshchadke* [Artificial neural networks and geostatistics in forecasting the distribution of chemical elements in the background area], *Geoekologiya*, 2017, 2, pp. 74-82 (in Russian).
2. Voitkevich, G.V., Miroshnikov, A.E., Povarennykh, A.S., Prokhorov, V.G. *Kratkii spravochnik po geokhimi* [Short guidebook in geochemistry], Nedra Publ., 1977 (in Russian).
3. Saet, Yu.E., Revich, B.A., Yanin, E.P. et al. *Geokhimiya okruzhayushchei sredy* [Environment geochemistry] Moscow, Nedra Publ., 1990, pp. 84-108 (in Russian).

- al Symposium on Spatial Analysis, Spatial-Temporal Data Modeling, and Data Mining, 2009, 74925F.
19. Matheron, G. Principles of geostatistics. *Economic Geology*, 1963, no. 58, pp. 1246-1266.
 20. Meer, F. D. Introduction to geostatistics. *ITC Lecture Notes*, 1993, 72 p.
 21. Mohanty, K., Majumdar, T.J. Using artificial neural networks for synthetic surface fitting and the classification of remotely sensed data. *International Journal of Applied Earth Observation and Geoinformation*, 1999, no. 1(1), pp. 78-84.
 22. Samanta, B., Ganguli, R., Bandopadhyay, S. Comparing the Predictive Performance of Neural Networks with Ordinary Kriging in a Bauxite Deposit. *Transactions of the Institution of Mining and Metallurgy*, 2005, no. 114, pp. 129-139.
 23. Schloeder, C.A., Zimmerman, N.E., Jacobs, M.J. Comparison of methods for interpolating soil properties using limited data. *Soil Sci. Soc. Am. J.*, 2001, no. 65, pp. 470-479.
 24. Sergeev, A.P., Baglaeva, E.M., Shichkin, A.V. Case of soil surface chromium anomaly of a northern urban territory – preliminary results. *Atmospheric Pollution Research*, 2010, vol. 1, pp. 44-49.
 25. Shacklette, H.T., Boerngen, J. G. Element concentrations in soils and other surficial materials of the conterminous United States, *U.S. Geological Survey professional paper*, 1984, no. 1270, 105 p.
 26. Shaker, R., Tofan, L., Bucur, M., Costache, S., Sava, D., Ehlinger, T. Land cover and landscape as predictors of groundwater contamination: a neural-network modelling approach applied to Dobrogea, Romania. *Journal of Environmental Protection and Ecology*, 2010, no. 11(1), pp. 337-348.
 27. Shaker, R.R., Ehlinger, T.J. Exploring non-linear relationships between landscape and aquatic ecological condition in southern Wisconsin: A GWR and ANN approach. *International Journal of Applied Geospatial Research*, 2014, no. 5(4), pp. 1-20.
 28. Shepherd, A.J. Second-Order Methods for Neural Networks: Fast and Reliable Training Methods for Multi-Layer Perceptrons. Springer-Verlag Publ., 1997, 145 p.
 29. Tarasov, D.A., Buevich, A.G., Sergeev, A.P., Shichkin, A.V. High Variation Topsoil Pollution Forecasting in the Russian Subarctic: Using Artificial Neural Networks Combined with Residual Kriging. *Applied Geochemistry*, 2018, vol. 88, part B, pp.188-197.
 30. Worsham, L., Markewitz, D., Nibbelink, N. Incorporating spatial dependence into estimates of soil carbon contents under different land covers. *Soil Sci. Am.*, 2010, no. 74, pp. 635–646.
 31. Yfantis, E.A., Flatman, G.T., Behar, J.V. Efficiency of kriging estimation for square, triangular, and hexagonal grids. *Math. Geol.*, 1987, no. 19, pp. 183-205.
 32. Zhang, C., Fay, D., McGrath, D., Grennan, E., Carton, O.T. Use of trans-Gaussian kriging for national soil geochemical mapping in Ireland. *Geochemistry: Exploration Environment Analysis*, 2008, vol. 8, pp. 255–265.

PREDICTION OF THE CHROME DISTRIBUTION IN SUBARCTIC NOYABRSK USING CO-KRIGING, GENERALIZED REGRESSION NEURAL NETWORK, MULTILAYER PERCEPTRON, AND HYBRID TECHNICIS

A. G. Buevich^{1,*}, I. E. Subbotina^{1,**}, A. V. Shichkin^{1,***}, A. P. Sergeev^{1,****}, E. M. Baglaeva^{1,*****}

¹*Institute of Industrial Ecology, Ural Branch, Russian Academy of Sciences,
ul. S. Kovalevskoi, 20, Yekaterinburg, 620219 Russia*

*E-mail: bagalex3@gmail.com

**E-mail: iesub@mail.ru

***E-mail: and@ecko.uran.ru

****E-mail: alexanderpsergeev@gmail.com

*****E-mail: sem@ecko.uran.ru

Combination of geostatistical interpolation (kriging) and machine learning (artificial neural networks, ANN) methods leads to an increase in the accuracy of forecasting. The paper considers the application of residual kriging of an artificial neural network to predicting the spatial contamination of the surface soil layer with chromium (Cr). We reviewed and compared two neural networks: the generalized regression neural network (GRNN) and multilayer perceptron (MLP), as well as the combined method: multilayer perceptron residual kriging (MLPRK). The study is based on the results of the screening of the surface soil layer in the subarctic Noyabrsk, Russia. The models are developed based on computer modeling with minimization of the RMSE. The MLPRK model showed the best prognostic accuracy.

Keywords: artificial neural networks, chromium, residual kriging, co-kriging.

DOI: <https://doi.org/10.31857/S0869-78092019277-86>