

МЕТОДОЛОГИЯ  
И МЕТОДИКА ИССЛЕДОВАНИЙ

УДК 504.064.2.001.18

ГИБРИДНОЕ МОДЕЛИРОВАНИЕ НА ОСНОВЕ LUR ПОДХОДА ВАРИАЦИЙ  
РАСПРЕДЕЛЕНИЯ ТЯЖЕЛЫХ МЕТАЛЛОВ В ВЕРХНЕМ СЛОЕ ПОЧВЫ  
НА ПРИМЕРЕ Г. ТАРКО-САЛЕ

© 2025 г. Е. М. Баглаева<sup>1,\*</sup>, А. Г. Буевич<sup>1,\*\*\*</sup>, А. В. Шичкин<sup>1,\*\*\*</sup>, А. П. Сергеев<sup>1,\*\*\*\*</sup>,  
А. С. Буторова<sup>1,\*\*\*\*\*</sup>

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

\*E-mail: e.m.baglaeva@urfu.ru

\*\*E-mail: bag@ecko.uran.ru

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

\*\*\*\*E-mail: sergeev@ecko.uran.ru

\*\*\*\*\*E-mail: a.s.butorova@urfu.ru

Поступила в редакцию 25.03.2024 г.

После доработки 23.08.2024 г.

Принята к публикации 10.10.2024 г.

Для оценки вариаций пространственного загрязнения тяжелыми металлами верхнего слоя городской почвы использована гибридная модель Land Use Regression (LUR) и регрессионный кrigинг (RK). Модель построена на основе данных мониторинга содержания никеля и марганца в верхнем слое почвы арктического г. Тарко-Сале. Предложенный LUR и RK метод моделирования загрязнения автотранспортом верхнего слоя почвы может быть использован для любой территории. Применение RK улучшает точность LUR модели: корреляция тестового и предсказанного наборов увеличилась на 7% для никеля и на 17% для марганца, а относительная среднеквадратическая ошибка RRMSE уменьшилась на 10% для обоих элементов. Результаты гибридного моделирования LUR с регрессионным кrigингом позволили определить, что распределение марганца и никеля в верхнем слое почвы города не зависит от автотранспорта. Это свидетельствует о природном происхождении марганца и никеля в городской почве в отсутствие других источников загрязнения. Последовательное включение различных источников загрязнения в модель LUR является способом оценки вклада каждого из выбранных источников в загрязнение выбранным элементом.

**Ключевые слова:** тяжелые металлы, верхний слой почвы, LUR (land use regression), пространственное распределение, гибридное моделирование

DOI: 10.31857/S0869780925010097 EDN: DOFNYZ

ВВЕДЕНИЕ

При описании пространственного распределения примесей в верхнем слое почвы в городах и промышленных центрах возникают сложности, связанные с существенным вкладом локальных источников (автотранспорт, карьеры и т.п.) в загрязнение компонентов окружающей среды [11]. Это приводит к неравномерности распределения примесей в почвенном покрове в пределах города, поэтому требуются особые (по сравнению с существующими) методы описания. Для оценки концентраций примесей можно использовать метод Land Use Regression (далее LUR), который имеет значительное преимущество перед классом

геостатистических кrigинг-методов, состоящее в учете взаимного расположения источников загрязнения (дороги, парковки, промзоны и т.д.) по отношению к точкам наблюдения [1–3, 7]. Метод заключается в построении математических моделей распределения примесей в природных средах на основе экспериментальных данных об их содержании и данных геоинформационных систем (ГИС). Полученные модели применяются для создания карт распределения примесей [9, 15]. LUR подход позволяет получать оценки распределения примесей для значительной по площади территории (целый город) при относительно небольших материальных и временных затратах [6, 9]. Еще одной перспективной для решения с помощью

LUR подхода проблемой может быть оценка вклада различных по природе источников загрязнения в поверхностное распределение примесей.

В большинстве исследований, описанных в литературе, методом LUR моделируют атмосферное загрязнение оксидов азота ( $\text{NO}_2$ ,  $\text{NO}$ ) [5–8], пыли (PM2.5, PM10) [6, 12–13], диоксида серы [7], нефтепродуктов (BTEX — Benzene, toluene, ethylbenzene, xylenes) [5], летучих органических соединений (VOC) [14], сажи [8, 10] и др. Гораздо реже встречаются оценки распределения примесей в почвенном [11] или снежном [3, 4] покровах.

Для построения модели LUR для оценки загрязнения некоторой природной среды используют данные о концентрациях поллютантов, полученные в ходе регулярного мониторинга или запланированной съемки. Количество проб и выбор мест их отбора влияет на качество модели. При построении модели предпочтительны данные запланированной съемки, но в отсутствии возможности выполнить отбор проб и их анализ, используют имеющиеся данные. Модели LUR, построенные по результатам наблюдений на стационарной станции, использовали систему нормативного мониторинга в Сидзуоке (Япония) [10]. Модели в этой работе были построены с четырьмя различными типами географических переменных, включая тип дороги, интенсивность движения, землепользование и физическую географию, относительно хорошо смогли оценить содержание  $\text{NO}_2$ .

На качество модели LUR также влияют географические данные исследуемой территории и выбор географических переменных, участвующих в модели. Существуют достаточно подробные модели разных городов, также для создания географических информационных моделей можно использовать открытый источник географических данных программу Google Earth. Однако единого подхода к реализации метода LUR не существует, и способы его применения различаются в каждом исследовании. Это связано, в первую очередь, со значительными неоднородностями исследуемых территорий городов, особенности которых не позволяли создать единый способ моделирования. Во-вторых, метод со дня его первого применения и до сих пор продолжает развиваться: улучшается качество получаемых результатов и разрешаются трудности и проблемы моделирования. В России результаты моделирования методом LUR представлены работами [1–3].

В данной работе предложено использовать метод LUR и регрессионный кrigинг (далее RK) для оценки загрязнения верхнего слоя почвы арктического г. Тарко-Сале. В работе на основе данных мониторинга концентраций примесей в верхнем слое почвы и карты города были выбраны и рассчитаны географические переменные, проведен статистический анализ экспериментальных и расчетных данных и построена модель карты загрязнения марганцем и никелем территории города.

Городская почва — долговременная депонирующая среда, картина ее загрязнения формируется десятилетиями, поэтому связи с техногенными источниками загрязнения не всегда могут быть установлены, и/или мощность этих источников оказывается недостаточной для формирования устойчивого следа загрязнения в почве. Метод LUR позволяет прозрачно интерпретировать связь между пространственным расположением источников загрязнения (географическими переменными), точек наблюдения и распределением примесей в верхнем слое почвы. Кригинг остатков LUR модели позволяет предметно извлекать дополнительную информацию из географических переменных. Окончательная картина элементного загрязнения верхнего слоя почвы формируется суммой результатов LUR модели и прогноза остатков.

## МАТЕРИАЛЫ И МЕТОДЫ

### Исследуемая территория

Для прогноза загрязнения верхнего слоя почвы с помощью LUR и регрессионного кригинга выбрана зона жилой застройки г. Тарко-Сале Ямало-Ненецкого автономного округа. Территория муниципального образования г. Тарко-Сале располагается на низкой надпойменной террасе правого берега р. Пяку-Пур. В геоморфологическом отношении долина р. Пяку-Пур на участке г. Тарко-Сале представляет собой пойму реки, возывающуюся над урезом воды при средних уровнях на 2–5 м и сложенную суглинисто-песчаными породами. Поверхность равнинная, уклон территории в сторону р. Пяку-Пур не более 0.5%. Среди аллювиальных отложений рассматриваемой территории преобладают пески различной крупности, которая увеличивается с глубиной от пылеватых до средней крупности. В песчаных толщах наблюдаются прослои и линзы супесей и суглинков, имеющих консистенцию от пластичной до текучей.

Зональные подзолистые почвы развиваются только на песчаных почвообразующих породах. Наиболее распространены следующие подтипы почв: таежно-поверхностно-глеевые и охристо-элювиально-глеевые. Почва городских кварталов Тарко-Сале по гранулометрическому составу представляет собой мелкий гомогенный песок диаметром зерен меньше 1 мм. При оседании примеси на поверхность почвы одновременно идут разнонаправленные процессы. С одной стороны, городская почва длительное время накапливает загрязнение, с другой — примеси могут вымываться как вертикально в более глубокие слои, так и переноситься по латерали в горизонтах.

На рис. 1 приведена карта города, включающая жилые и промышленные здания, дворовые территории, дороги, автостоянки и точки отбора проб верхнего слоя почвы. Всего была отобрана



Рис. 1. Карта исследуемой территории.

и проанализирована на содержание хрома, марганца, меди, никеля и других элементов 101 проба. Подробно отбор, подготовка и химический анализ проб описан в работе [4].

Для исследования были выбраны данные о валовом содержании марганца и никеля в верхнем слое почвы. Марганец и никель не содержатся в материнских породах, характерных для этой территории, и вероятно имеют техногенное происхождение. Их следы в почве могут быть связаны с деятельностью человека, как-то промышленное загрязнение, строительство, перемещение грунта и т. д., этим и обусловлен выбор техногенных объектов в качестве источников загрязнения. Все пробы (101 шт.) были разбиты на тренировочные (80 шт.) и тестовые (21 шт.).

## Метод Land Use Regression (LUR)

**Подготовка географических данных.** Для моделирования методом LUR использовали растровую карту г. Тарко-Сале, включающую данные о распределении автомобильных дорог и парковок. Эти характеристики рассчитываются внутри так называемых буферных зон, построенных вокруг каждой точки наблюдения. Буферная зона (circular buffers) — круг заданного радиуса с центром в точке отбора проб, внутрь которого попадают выбранные объекты. В литературе встречаются расчет средства-ми ГИС линейных источников загрязнения (длины дорог, попавших внутрь данного круга) и площад-ных источников (площадей парковок, карьеров и т.д.). Размер буфера определяет пространственное разрешение модели в предположении о влиянии изучаемых предикторов. Снизу размер буфера ограничен пространственным разрешением растра,

сверху — размерами и пространственными особенностями исследуемой территории (ширина и длина исследуемой территории не превышает 3 км). Кроме того, во многих исследованиях указано, что концентрация связанных с автотранспортом загрязнителей убывает экспоненциально с расстоянием от дорог. Считается, что влияние дорог на концентрацию этих веществ ограничено расстояниям около 100 м от основных городских дорог и около 500 м от за-городных шоссе [7, 9]. Географические переменные-предикторы для LUR модели вычисляются как мера пересечения буферной зоны и выбранных источников загрязнения. Единицы измерения этих предикторов совпадают с размерностями длины для линейных источников и площади для площадных источников. Современные ГИС позволяют получить множество переменных с любыми размерами буферных зон и любыми типами источников загрязнения. Для дорог с разной плотностью транспортного потока вводят категории, что увеличивает количество географических переменных. Однозначного алгоритма определения размера буферных зон и расчета переменных-предикторов в литературе не представлено. В данном исследовании мы следуем работе А.Г. Буевича и др. [2], в которой было предложено рассматривать дороги с разной интенсивностью движения как площадные источники, что позволило уйти от разделения дорог на категории, упростить проведение расчетов и повысить качество модели.

В предположении, что основное влияние на распределение марганца и никеля в почве Тарко-Сале оказывает интенсивность транспортных потоков, ограничились источниками загрязнения, связанными с автотранспортом: автодорогами и парковками. В данной работе выбраны буферные зоны радиусом 100, 200 и 500 м. На рис. 2 представлен фрагмент карты города с произвольно выбранной точкой пробоотбора и географические предикторы дорог и парковок, рассчитанные для трех буферных зон. Предикторы для каждой точки — это площади дорог и парковок для буферных зон разных размеров.

Всего использовано шесть географических переменных (см. рис. 2): R100, R200 и R500 — площади дорог; SR100, SR200 и SR500 — площади парковок, попавших в буферные зоны 100, 200 и 500 м, соответственно. Такой выбор площадных географических переменных позволяет унифицировать источники загрязнения. Переменная с меньшим масштабом R100 описывает влияние ближайших к точке наблюдения дорог от 0 до 100 м. Переменные с разным радиусом буферной зоны позволяют впоследствии в ходе статистического анализа найти наилучшее описание изменения величины загрязнения с расстоянием от дорог.

Для выбора регрессионных переменных был проведен анализ парных коэффициентов корреляции всех географических переменных

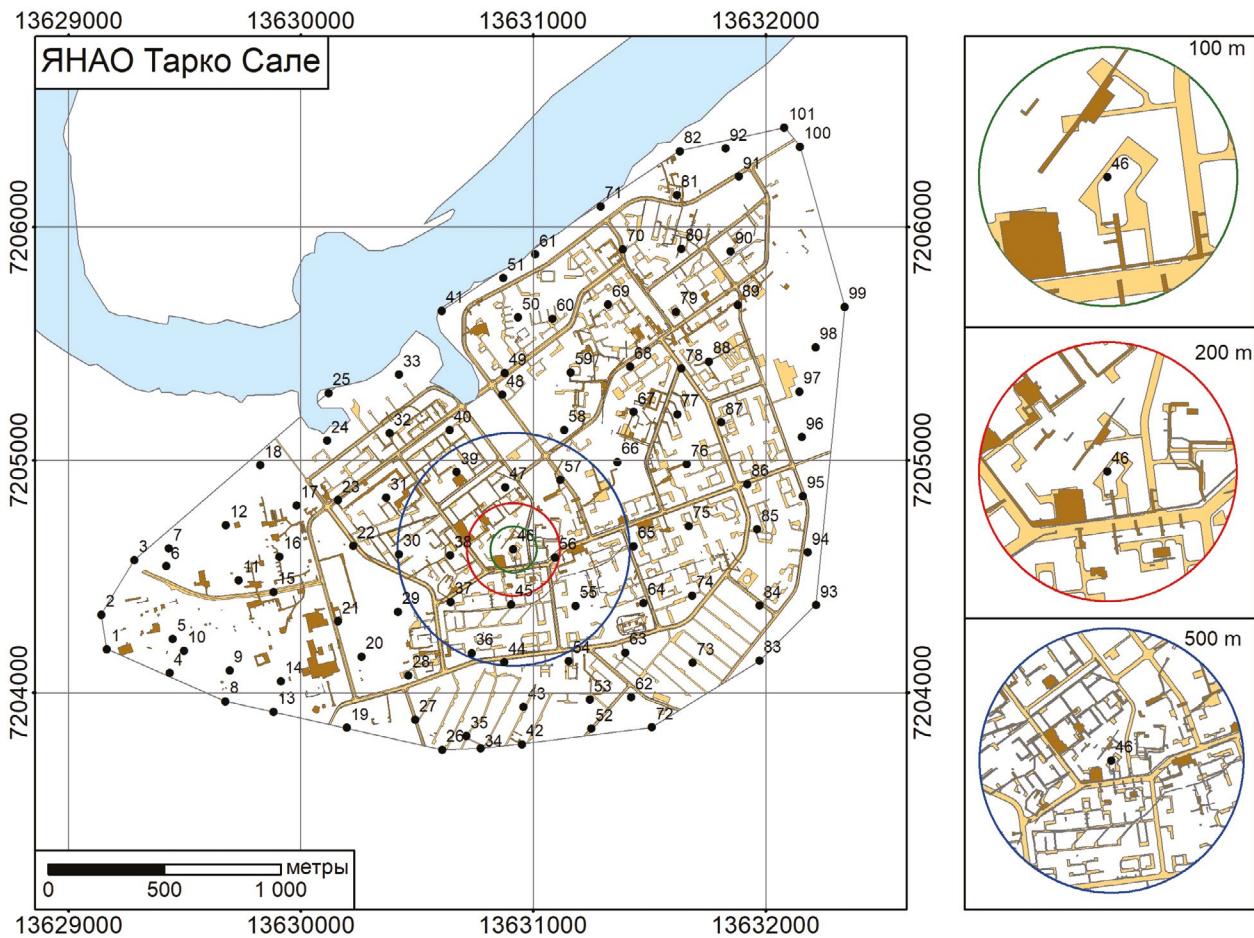


Рис. 2. Построение переменных LUR модели.

(из 6 географических переменных можно построить 15 пар). Для построения модели из каждой пары географических переменных с сильной корреляционной связью (коэффициент корреляции больше 0.8) выбирали одну переменную, влияющую на загрязнение. Перебор всех вариантов некоррелированных переменных позволил построить оптимальный набор географических переменных для построения LUR модели с наименьшей ошибкой.

**Особенности пробоотбора для модели LUR.** От количества и расположения проб зависит качество модели LUR. В обзоре G. Hoek et al. [9] рекомендуют использовать от 40 до 80 проб для построения модели. Здесь нижняя граница определена точностью модели (ошибка не превышает 10%), а верхняя обусловлена высокими затратами на пробоотбор и последующий химический анализ. Для построения LUR модели г. Тарко-Сале было использовано 80 проб верхнего слоя почвы, отобранных по нерегулярной сетке с ненарушенных участков почвенного покрова, для оценки состояния окружающей природной среды города. 21 проба верхнего слоя почвы была использована для тестирования модели.

**LUR модель.** Метод LUR использует регрессионный анализ для создания моделей распределения примесей в природных средах, чаще всего множественную регрессию. В данной работе использована множественная линейная регрессия, уравнение которой связывает измеренные концентрации примесей с рассчитанными географическими переменными и их комбинациями:

$$y = a_0 + \sum_{i=1}^N a_i x_i, \quad (1)$$

где  $y$  — предсказанная концентрация исследуемого элемента, мг/кг;  $a_0$  — свободный член регрессии (фоновая концентрация элемента), мг/кг;  $a_i$  — коэффициенты линейной регрессии (физический смысл поверхностной плотности исследуемого элемента), мг/(кг·м<sup>2</sup>) или мг/(кг·м<sup>4</sup>);  $x_i$  — переменные регрессии, м<sup>2</sup> или м<sup>4</sup>.

**Оценка точности модели.** Для оценки точности модели были использованы коэффициент корреляции  $Corr$  (2) предсказанных и измеренных значений концентрации примесей; среднеквадратическая ошибка  $RMSE$  (3); относительная среднеквадратическая ошибка  $RRMSE$  (4); нормированная среднеквадратическая ошибка  $NRMSE$  (5).

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (p(i) - o(i))^2}{n}}, \quad (3)$$

$$RRMSE = \sqrt{\frac{\sum_{i=1}^n (p(i) - o(i))^2}{n \sum_{i=1}^n (o(i))^2}}, \quad (4)$$

$$NRMSE = \sqrt{\frac{\sum_{i=1}^n (p(i) - o(i))^2}{n \sum_{i=1}^n (o(i) - \bar{o})^2}}, \quad (5)$$

где  $p(i)$  — прогнозируемые данные в  $i$ -точке тестового набора;  $o(i)$  — измеренные данные в  $i$ -точке;  $n$  — количество точек в тестовом наборе;  $\bar{o}$  — среднее измеренных данных тестового набора;  $\bar{p}$  — среднее предсказанных данных тестового набора.

### Алгоритм, объединяющий LUR и регрессионный кригинг

Предложенный в данной работе подход представлен на рис. 3 и включает несколько шагов. На первом шаге выбирали набор наилучших переменных — предикторов. В модель LUR включили три географических предиктора R100, R500 и SR500, исключив одну из каждой пары географических переменных с сильной корреляционной связью.

Первый шаг для LUR — построение буферных зон радиусами 100, 200 и 500 м вокруг каждой точки пробоотбора. Далее инструментами GIS находили пересечения буферных зон с дорогами и парковками. Для каждого типа переменных найдено пересечение со всеми тремя буферными зонами и найдены площади, занятые дорогами и парковками для каждой буферной зоны. В результате получили шесть предикторов, содержащих площади дорог или парковок для каждой буферной зоны.

Следующий шаг — построение модели LUR для тренировочного набора данных и предсказание значений концентраций примесей в тестовом наборе данных. Дальнейшие действия — получение

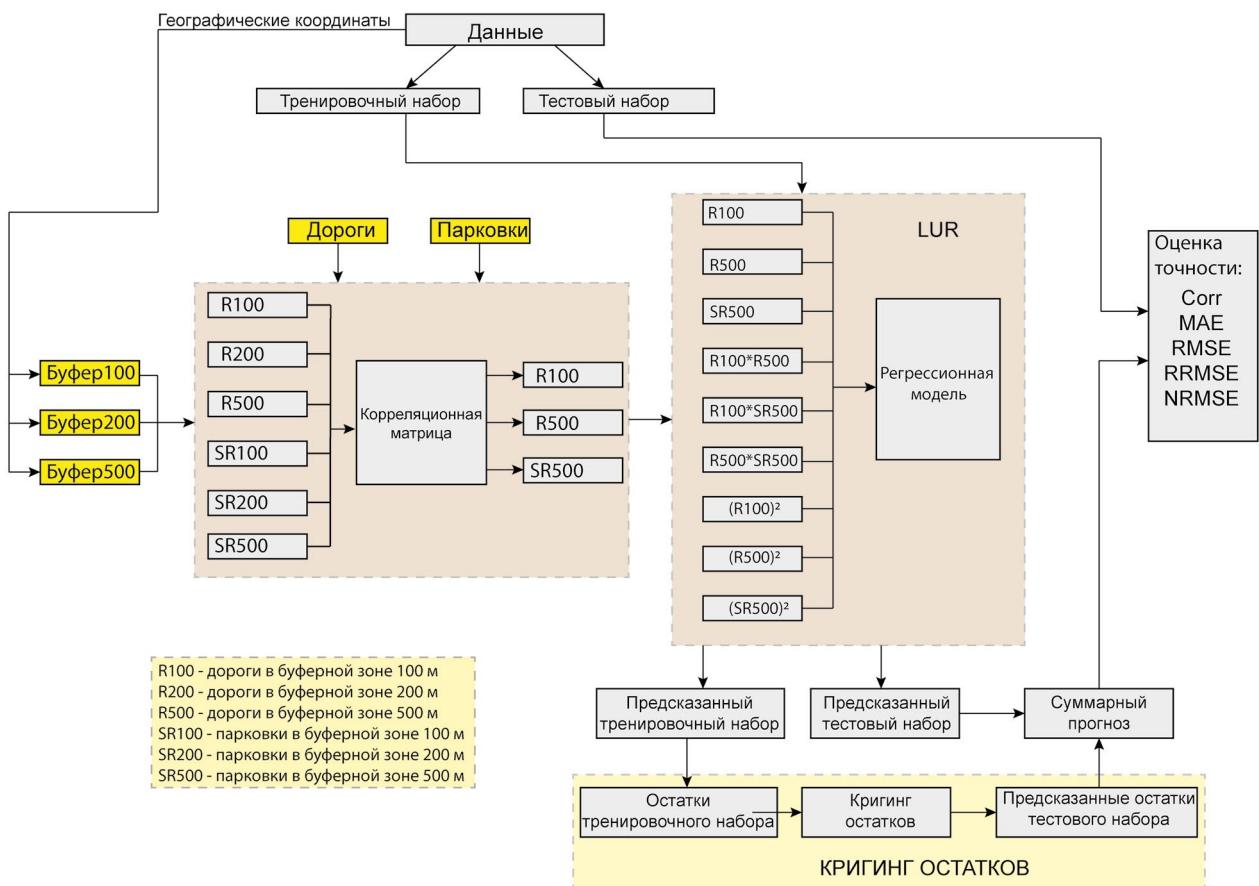


Рис. 3. Алгоритм, объединяющий LUR и регрессионный кригинг.

остатков (разница между измеренными данными и прогнозом) для тренировочного набора и построение кригинга остатков. На основе кригинга остатков тренировочного набора был получен прогноз остатков для тестового набора. В итоге суммируем значения концентрации, предсказанные LUR моделью и предсказанные методом кригинга остатков для тестового набора данных.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Статистика измеренных значений содержаний никеля и марганца на исследуемой территории приведена в табл. 1. Парные коэффициенты корреляции переменных R100, R200, R500, SR100, SR200 и SR500 приведены в табл. 2 (жирным шрифтом выделены коэффициенты корреляции больше 0.8).

Исключение переменных с сильной корреляционной связью и перебор всех вариантов некоррелированных переменных выявил оптимальный набор географических переменных для построения LUR модели. Для марганца и никеля оптимальными параметрами регрессии оказалось использование трех предикторов R100, R500 и SR500. Регрессионная модель содержит линейные, квадратичные и произведения пар различных предикторов (соответствие географических переменных переменным регрессии приведено в табл. 3). В табл. 4 приведены переменные LUR модели и их коэффициенты, рассчитанные в пакете Matlab 2021b.

Для улучшения качества прогноза применялся кригинг остатков тренировочного набора данных, на его основе были предсказаны остатки тестового набора. Модель кригинга для обоих элементов была построена с использованием экспоненциальной вариограммы и с учетом анизотропии. На рис. 4 показаны результаты моделирования содержания марганца и никеля в верхнем слое почвы г. Тарко-Сале.

Как видно из рис. 4, картины загрязнения верхнего слоя почвы города марганцем и никелем, восстановленные при использовании моделей LUR или LUR с RK, улавливают только некоторые особенности поверхностных распределений измеренных содержаний марганца и никеля, полученные геостатистическим методом.

Результаты оценки точности модели LUR и модели LUR с кригингом на тестовом наборе приведены в табл. 5 для обоих элементов. Прогноз остатков суммировали с результатами регрессионной модели на тестовом наборе.

Из табл. 5 видно, что применение кригинга остатков улучшает точность LUR модели. Корреляция тестового и предсказанного набора увеличилась на 7% для никеля и на 17% для марганца. RRMSE уменьшалась на 10% и для никеля, и для марганца.

Для визуализации прогнозирующей способности между смоделированными и наблюдаемыми значениями содержаний элементов в верхнем слое почве на тестовом наборе данных использовали

**Таблица 1.** Описательная статистика содержаний никеля и марганца

Элемент	Содержание, мг/кг					Коэффициент вариации, %	Асимметрия	Эксцесс
	Минимум	Максимум	Среднее	СО*	Медиана			
Mn	38.8	581.5	125.3	66.4	115.8	53	3.5	21
Ni	4.6	41.3	14.3	7.5	11.3	53	1.2	1.2

\* СО – стандартное отклонение

**Таблица 2.** Парные коэффициенты корреляции географических переменных

Переменная 1	Переменная 2	Corr	Переменная 1	Переменная 2	Corr	Переменная 1	Переменная 2	Corr
R100	R200	<b>0.88</b>	R500	SR100	0.37	R200	SR500	0.42
R200	R500	<b>0.92</b>	R100	SR200	0.29	R500	SR500	0.46
R500	R100	0.75	R200	SR200	0.39	SR100	SR200	<b>0.86</b>
R100	SR100	0.33	R500	SR200	0.41	SR200	SR500	<b>0.81</b>
R200	SR100	0.29	R100	SR500	0.32	SR100	SR500	0.71

**Таблица 3.** Связь географических переменных и переменных LUR

Переменные регрессии	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$
Предикторы	R100	R500	SR500	R100·R500	R100·SR500	R500·SR500	(R100) <sup>2</sup>	(R500) <sup>2</sup>	(SR500) <sup>2</sup>

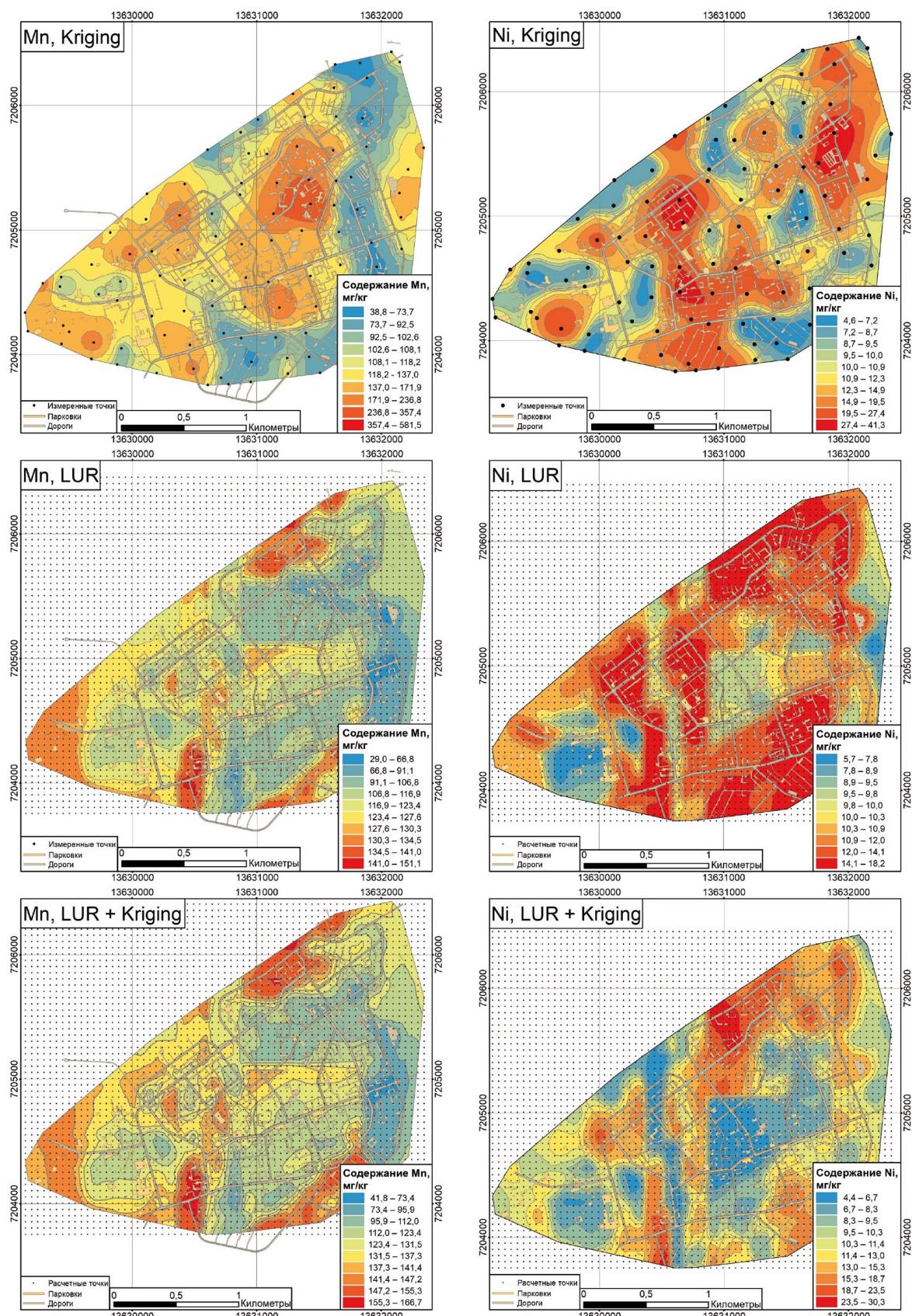


Рис. 4. Результаты моделирования содержания марганца и никеля в верхнем слое почвы г. Тарко-Сале.

Таблица 4. Коэффициенты LUR модели

Элемент	Коэффициенты регрессии				
	$a_0$ , мг/кг	$a_1$ , мг/(кг·м <sup>2</sup> )	$a_2$ , мг/(кг·м <sup>2</sup> )	$a_3$ , мг/(кг·м <sup>2</sup> )	$a_4$ , мг/(кг·м <sup>4</sup> )
Mn	127.93	0.001	-0.001	0.001	$10^{-8}$
Ni	9.83	0.001	$-3 \cdot 10^{-6}$	0.0001	$4 \cdot 10^{-9}$
Коэффициенты регрессии					
Элемент	$a_5$ , мг/(кг·м <sup>4</sup> )	$a_6$ , мг/(кг·м <sup>4</sup> )	$a_7$ , мг/(кг·м <sup>4</sup> )	$a_8$ , мг/(кг·м <sup>4</sup> )	$a_9$ , мг/(кг·м <sup>4</sup> )
	$10^{-7}$	$10^{-8}$	$10^{-6}$	$10^{-9}$	$3 \cdot 10^{-8}$
Mn	$4 \cdot 10^{-11}$	$3 \cdot 10^{-9}$	$-10^{-7}$	$-5 \cdot 10^{-10}$	$-10^{-8}$
Ni					

Таблица 5. Оценка точности модели

	Corr	MAE	RRMSE	RMSE	NRMSE
Mn_LUR	0.4437	74.2276	0.6603	118.0850	1.0522
Mn_LUR_Kriging	0.6162	64.5473	0.6031	112.4673	1.0022
Ni_LUR	0.5237	9.3108	0.4697	11.7698	1.1523
Ni_LUR_Kriging	0.5923	8.5602	0.4167	11.4739	1.1234

диаграмму Тейлора [16]. Диаграмма объединяет среднеквадратическое значение, коэффициент корреляции и стандартное отклонение прогнозных моделей LUR и LUR с RK для обоих элементов. Так как параметры диаграммы для исследуемых элементов очень сильно отличаются друг от друга, для лучшей визуализации была построена нормализованная диаграмма Тейлора, для которой стандартные отклонения обоих элементов приведены к единице. Такая нормализованная диаграмма позволяет сравнить качество прогноза всех моделей. Статистические данные позволяют оценить, насколько точно предложенные методы прогноза моделируют наблюдаемые данные. Маркеры для каждой модели нанесены на диаграмме в полярных координатах. Радиальные расстояния от начала координат представляют собой нормализованные стандартные отклонения, а азимутальные положения представляют собой коэффициенты корреляции между измеренным и прогнозируемым распределением. Концентрические круги представляют RMSE. Диаграмма Тейлора для марганца и никеля для LUR и LUR с RK представлена на рис. 5.

Лучшей предсказательной способностью обладает модель LUR с RK для никеля. Модели LUR с RK, как видно из рис. 5, обладают большим коэффициентом корреляции и меньшим среднеквадратичным значением по сравнению с моделями LUR.

Авторы предлагают следующую интерпретацию коэффициентов и переменных построенной LUR модели. Это линейная регрессионная модель для 9 переменных, уравнение которой связывает измеренные концентрации примесей с рассчитанными географическими переменными и их комбинациями. Свободный член регрессии может быть характеристикой фоновой концентрации исследуемого элемента; переменные регрессии  $x_i$ ,

при  $i = 1, 2, 3$  характеризуют площади источников загрязнения для разных зон влияния (буферных зон); коэффициенты  $a_i$  при  $i = 1, \dots, 9$  — степень влияния соответствующих географических переменных на поверхностную плотность загрязнения почвы исследуемым элементом.

LUR имеет следующие недостатки: зависимость от способа построения буферных зон, от выбора географических предикторов и отсутствие алгоритма применения. Для снижения подобных недостатков полезно дополнять LUR кригингом остатков, что увеличивает точность модели. Лучшая модель построена для распределения никеля в верхнем слое почвы, что может быть связано с формой распределения элемента. Распределение марганца имеет

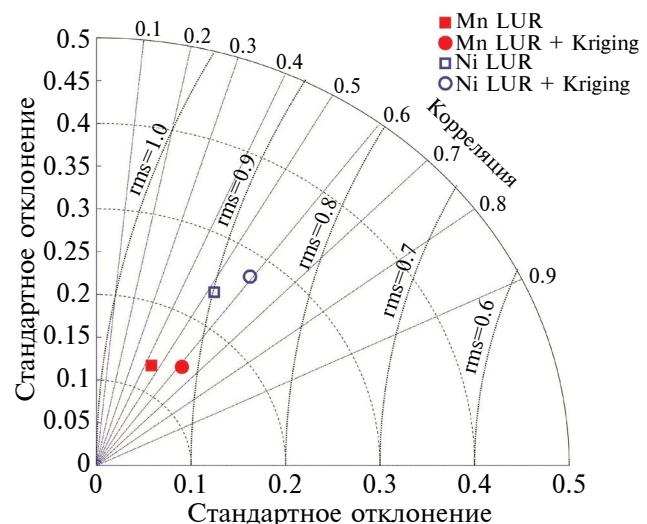


Рис. 5. Диаграмма Тейлора.

высокую дисперсию и тяжелый правый хвост, о чем свидетельствует значение стандартного отклонения и высокий коэффициент асимметрии (см. табл. 1). Полученные результаты свидетельствуют об отсутствии связи содержания никеля и марганца в верхнем слое почвы с автотранспортом.

Ранее авторы изучали загрязнение верхнего слоя г. Тарко-Сале, в частности, хрома и марганца [4]. Статистической связи распределения содержания тяжелых металлов в поверхностном слое почвы города с антропогенными источниками загрязнения обнаружено не было. Аномалии высокого содержания хрома в почве скорее всего связаны с концентрированием либо с отсыпкой привозным грунтом; никель и марганец имеют природное происхождение. В песчаных почвах Тарко-Сале одновременно идут длительные процессы накопления примеси в верхнем слое почвы и вымывание ее в более глубокие слои. Метод LUR дает хороший результат построения картины загрязнения почвы мощными техногенными источниками, если почвы не изменяются при функционировании города.

## ВЫВОДЫ И ЗАКЛЮЧЕНИЕ

Построенная в данной работе гибридная модель, использующая метод LUR и RK, характеризует загрязнение марганцем и никелем верхнего слоя почвы г. Тарко-Сале. Объединяющий LUR и RK метод моделирования загрязнения автотранспортом верхнего слоя почвы может быть использован для любой территории.

Данные техногенных источников, которые мы использовали для построения регрессионной модели, не обнаружили связи с картиной загрязнения марганцем и никелем верхнего слоя почвы. Картина распределения марганца и никеля в верхнем слое почвы обусловлена скорее природными факторами.

Результаты моделирования LUR позволяют делать выводы о происхождении тяжелых металлов в почве. Прошлые основанные на статистическом анализе результаты не выявили связи загрязнения хрома с антропогенными источниками (дороги, промзоны), никель и марганец также не связаны с антропогенными источниками. Последовательное включение различных источников загрязнения в модель LUR может позволить оценить вклад источника в загрязнение выбранным металлом.

*Исследование выполнено за счет субсидий Минобрнауки РФ на выполнение научной темы FUMN-2024-0003.*

## СПИСОК ЛИТЕРАТУРЫ

1. Антропов К.М., Вараксин А.Н. Оценка загрязнения атмосферного воздуха г. Екатеринбурга диоксидом азота методом Land Use Regression // Экологические системы и приборы. 2011. №8. С. 47–54.
2. Буевич А.Г., Сафина А.М., Сергеев А.П. и др. Анализ статистических зависимостей распределения загрязняющих веществ в поверхностном слое почвы урбанизированных территорий с применением математических моделей (LUR метод) // Геоэкология. 2015. №3. С. 268–279.
3. Медведев А.Н., Медведев М.А. О применении подхода Land Use Regression для моделирования площадного загрязнения снега при малом количестве точек наблюдения // XI Междунар. конф. «Российские регионы в фокусе перемен». Екатеринбург, 17–19 ноября 2016 г.: сб. докладов. Екатеринбург: Издательство УМЦ УПИ, 2016. Ч. 1. С. 487–494.
4. Сергеев А.П., Баглаева Е.М., Субботина И.Е. Загрязнение почв города Тарко-Сале тяжелыми металлами // Геоэкология. 2014. №1. С. 28–36.
5. Aguilera I., Sunyer J., Fernandez-Patier R., Hoek G. et al. Estimation of outdoor  $\text{NO}_x$ ,  $\text{NO}_2$  and BTEX exposure in a cohort of pregnant women using land use regression modeling // Environ. Sci. Technol. 2008. V. 42. P. 815–821.
6. Brauer M., Hoek G., van Vliet P., Melfi K. et al. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems // Epidemiology. 2003. V. 14. P. 228–239.
7. Briggs D.J., de Hoogh C., Gulliver J., Wills J. et al. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments // Sci. Total. Environ. 2000. V. 253(1–3). P. 151–167.
8. Carr D., von Ehrenstein O., Weiland S., Wagner C. et al. Modeling annual benzene, toluene,  $\text{NO}_2$ , and soot concentrations on the basis of road traffic characteristics // Environ. Res. 2002. V. 90. P. 111–118.
9. Hoek G., Beelen R., de Hoogh K., Vienneau D. et al. A review of land-use regression models to assess spatial variation of outdoor air pollution // Atmos. Environ. 2008. V. 36. P. 4077–4088.
10. Kashima S., Yorifuji T., Tsuda T., Doi H. Application of land use regression to regulatory air quality data in Japan // Sci Total Environ. 2009. V. 407(8). P. 3055–3062.
11. Liu Y., Song S., Bi C., Zhao J., Xi D., Su Z. Occurrence, Distribution and Risk Assessment of Mercury in Multimedia of Soil-Dust-Plants in Shanghai, China // Int. J. Environ. Res. Public Health. 2019. V. 16. 3028. <https://doi.org/10.3390/ijerph16173028>
12. Moore D.K., Jerrett M., Mack W.J., Kunzli N. A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA // J. Environ. Monitor. 2007. V. 9. P. 246–252.
13. Ross Z., English P.B., Scalf R., Gunier R. et al. Nitrogen dioxide prediction in Southern California using land use regression modeling: potential for environmental health analyses // J. Expo. Sci. Environ. Epidemiol. 2006. V. 16. P. 106–114.
14. Smith L., Mukerjee S., Gonzales M., Stallings C. et al. Use of GIS and ancillary variables to predict volatile organic compound and nitrogen dioxide levels at unmonitored locations // Atmos. Environ. 2006. V. 40. P. 3773–3787.
15. Stedman J., Vincent K., Campbell G., Goodwin J., Downing C. New high resolution maps of estimated background ambient  $\text{NO}$  and  $\text{NO}_2$  concentrations in the U.K. // Atmos. Environ. 1997. V. 31. P. 3591–3602.
16. Taylor K. Summarizing multiple aspects of model performance in a single diagram // J. Geophys. Res. 2001. V. 106. P. 7183–7192. <https://doi.org/10.1029/2000JD900719>

## LAND-USE REGRESSION MODEL TO ASSESS SPATIAL VARIATION OF TOPSOIL POLLUTION IN TARKO-SALE

E. M. Baglaeva<sup>a, #</sup>, A. G. Buevich<sup>a, ##</sup>, A. V. Shichkin<sup>a, ###</sup>, A. P. Sergeev<sup>a, #####</sup>, A. S. Butorova<sup>a, #####</sup>

<sup>a</sup>*Institute of Industrial Ecology, Ural Branch, Russian Academy of Sciences,  
ul. S. Kovalevskoi 20, Yekaterinburg, 620219 Russia*

<sup>#</sup>*E-mail: e.m.baglaeva@urfu.ru*

<sup>##</sup>*E-mail: bag@ecko.uran.ru*

<sup>##</sup>*E-mail: and@ecko.uran.ru*

<sup>###</sup>*E-mail: sergeev@ecko.uran.ru*

<sup>#####</sup>*E-mail: a.s.butorova@urfu.ru*

A hybrid model combining land use regression (LUR) and regression kriging (RK) methods is constructed to assess the variation in spatial pollution of urban topsoil by heavy metals. The environmental monitoring data of nickel and manganese content in the topsoil of the Arctic town Tarko-Sale were used. This hybrid method of modelling topsoil pollution is suitable for all pollutants, for different territories and types of human-induced pollution sources. The use of RK improves the LUR model accuracy: the correlation between test and predicted sets increased by 7 and 17% for nickel and manganese, respectively; and the relative root mean squared error (RRMSE) decreased by 10% for both elements. The results of hybrid modeling of LUR with RK showed that the spatial distribution of manganese and nickel content in topsoil of the city does not depend on city vehicles. This points to the natural origin of manganese and nickel in urban soil in the absence of other pollution sources. The sequential inclusion of different pollution sources in the LUR model is a way to assess the contribution of each of the selected sources to pollution by the selected element. The data from technogenic sources used in the regression model did not show relationship with the pattern of manganese and nickel contamination in the topsoil. The spatial distribution of manganese and nickel in the top layer of soil is controlled rather by natural factors and is not associated with anthropogenic activities. The results of modelling LUR with RK allow us to draw conclusions about the origin of heavy metals in the soil. Previous results based on statistical analysis have shown no association between chromium pollution and anthropogenic sources (roads, industrial areas), and nickel and manganese are also not associated with anthropogenic sources. The sequential inclusion of various sources of pollution makes it possible to evaluate the source contribution to the pollution by certain metal.

**Keywords:** *heavy metals, topsoil, land use regression (LUR), regression kriging, spatial distribution, hybrid modelling*

### REFERENCES

1. Antropov, K.M., Varaksin, A.N. [Assessment of atmospheric air pollution in Yekaterinburg with nitrogen dioxide using the Land Use Regression method]. *Ekologicheskie sistemy i pribory*, 2011, no. 8, pp. 47–54. (in Russian)
2. Buevich, A.G., Safina, A.M., Sergeev, A.P., et al. [Analysis of statistical dependencies of the distribution of pollutants in the surface layer of soil in urbanized areas using mathematical models (LUR method)]. *Geoekologiya*, 2015, no. 3, pp. 268–279. (in Russian)
3. Medvedev, A.N., Medvedev, M.A. [On the application of the Land Use Regression approach for modeling areal snow pollution with a small number of observation points]. Proc. XI Int. Conf. “Russian Regions in the Focus of Changes”. Yekaterinburg, November 17–19, 2016. Yekaterinburg, UMTs UPI Publ. 2016, part 1, pp. 487–494. (in Russian)
4. Sergeev, A.P., Baglaeva, E.M., Subbotina, I.E. [Soil contamination of the city of Tarko-Sale with heavy metals]. *Geoekologiya*, 2014, no. 1, pp. 28–36. (in Russian)
5. Aguilera, I., Sunyer, J., Fernandez-Patier, R., Hoek, G., et al. Estimation of outdoor NO<sub>x</sub>, NO<sub>2</sub> and BTEX exposure in a cohort of pregnant women using land use regression modeling. *Environ. Sci. Technol.*, 2008, vol. 42, pp. 815–821.
6. Brauer, M., Hoek, G., van Vliet, P., Meliefste, K., et al. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology*, 2003, vol. 14, pp. 228–239.
7. Briggs, D.J., de Hoogh, C., Gulliver, J., Wills, J., et al. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Sci. Total. Environ.*, 2000, vol. 253(1–3), pp. 151–167.
8. Carr, D., von Ehrenstein, O., Weiland, S., Wagner, C., et al. Modeling annual benzene, toluene, NO<sub>2</sub>, and soot concentrations on the basis of road traffic characteristics. *Environ. Res.*, 2002, vol. 90, pp. 111–118.
9. Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D. et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.*, 2008, vol. 36, pp. 4077–4088.
10. Kashima, S., Yorifuji, T., Tsuda, T., Doi, H. Application of land use regression to regulatory air quality data in Japan. *Sci. Total Environ.*, 2009, vol. 407(8), pp. 3055–3062.
11. Liu, Y., Song, S., Bi, C., Zhao, J., Xi, D., Su Z. Occurrence, distribution and risk assessment of mercury in multimedia of soil-dust-plants in Shanghai, China. *Int. J. Environ. Res. Public Health.*, 2019, vol. 16: 3028. <https://doi.org/10.3390/ijerph16173028>
12. Moore, D.K., Jerrett, M., Mack, W.J., Kunzli, N. A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA. *J. Environ. Monitor.*, 2007, vol. 9, pp. 246–252.
13. Ross, Z., English, P.B., Scalf, R., Gunier, R., et al. Nitrogen dioxide prediction in Southern California using land use regression modeling: potential for environmental health analyzes. *J. Expo. Sci. Environ. Epidemiol.*, 2006, vol. 16, pp. 106–114.
14. Smith, L., Mukerjee, S., Gonzales, M., Stallings, C., et al. Use of GIS and ancillary variables to predict volatile organic compound and nitrogen dioxide levels at unmonitored locations. *Atmos. Environ.*, 2006, vol. 40, pp. 3773–3787.
15. Stedman, J., Vincent, K., Campbell, G., Goodwin, J., Downing, C. New high resolution maps of estimated background ambient NO<sub>x</sub> and NO<sub>2</sub> concentrations in the U.K. *Atmos. Environ.*, 1997, vol. 31, pp. 3591–3602.
16. Taylor, K. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, 2001, vol. 106, pp. 7183–7192. <https://doi.org/10.1029/2000JD900719>