

---

---

**МЕТОДОЛОГИЯ  
И МЕТОДИКА ИССЛЕДОВАНИЙ**

---

---

УДК 504.064.2.001.18:551.5 78.4

**АНАЛИЗ СТАТИСТИЧЕСКИХ ЗАВИСИМОСТЕЙ  
РАСПРЕДЕЛЕНИЯ ЗАГРЯЗНЯЮЩИХ ВЕЩЕСТВ  
В ПОВЕРХНОСТНОМ СЛОЕ ПОЧВЫ УРБАНИЗИРОВАННЫХ  
ТЕРРИТОРИЙ С ПРИМЕНЕНИЕМ МАТЕМАТИЧЕСКИХ  
МОДЕЛЕЙ (LUR МЕТОД)**

© 2015 г. А. Г. Бувич, А. М. Сафина, А. П. Сергеев, А. Н. Вараксин,  
А. Н. Медведев

*Институт промышленной экологии Уральского отделения Российской академии наук  
(ИПЭ УрО РАН),  
ул. С. Ковалевской, д. 20, г. Екатеринбург, 620219 Россия. E-mail: iie@esko.uran.ru*

Поступила в редакцию 8.05.2014 г.

После исправления 8.09.2014 г.

Работа посвящена применению метода Land Use Regression (LUR) для построения полей распределения загрязняющих веществ в верхнем слое почвенного покрова урбанизированных территорий на примере микрорайона г. Екатеринбурга. LUR – новый подход для описания распределения загрязняющих веществ в поверхностном слое почвы. Метод сочетает инструментальные измерения распределения загрязняющих веществ, построение статистических моделей на основе экспериментальных данных о распределении загрязняющих веществ и данных географических информационных систем (ГИС). Для исследования отобраны пробы почвы и произведен их химический анализ. Проведен анализ регрессионных моделей для описания пространственного распределения загрязняющих веществ на основе экспериментальных данных о загрязнении почвы и географической информации микрорайона Втузгородок г. Екатеринбурга.

**Ключевые слова:** *Land Use Regression, регрессионная модель, методы интерполяции, методы моделирования рассеяния, однофакторный анализ.*

## ВВЕДЕНИЕ

Применение математического моделирования позволяет решать задачи описания распределения загрязняющих веществ (ЗВ) в почвенном покрове населенных пунктов. При исследовании распределения ЗВ в крупных городах и промышленных центрах возникают значительные сложности, связанные с существенным вкладом в общее загрязнение локальных источников (автотранспорт, промышленные зоны и т.п.), что обуславливает неравномерность их распределения в почвенном покрове и требует особых методов описания.

Цель настоящего исследования – применение LUR метода для описания распределения загрязняющих веществ в почвенном покрове урбанизированной территории на примере микрорайона Втузгородок г. Екатеринбурга.

## МЕТОД LAND USE REGRESSION

### Описание метода LUR

Для составления карт распределения ЗВ в почвенном покрове традиционно используются два типа методов: интерполяции измерений и моделирования рассеивания.

Эффективность различных методов пространственной интерполяции зависит от реальной пространственной неоднородности распределения моделируемого загрязняющего вещества, а также качества первичных экспериментальных данных (например, плотность отбора проб, выбор мест отбора проб). В большинстве случаев методы пространственной интерполяции могут быть использованы только для регионального масштаба.

Land Use Regression (LUR) – метод построения математических моделей распределения ЗВ. Ин-

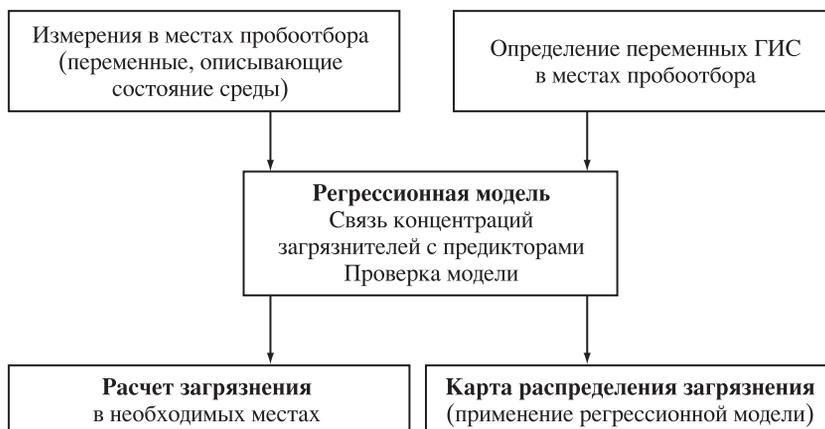


Рис. 1. Алгоритм моделирования LUR.

формационная основа метода – географические данные исследуемой территории, данные специальной измерительной кампании о концентрациях ЗВ в почвенном покрове в определенном числе мест города и регрессионный анализ для определения взаимосвязей измеренных концентраций и полученных географических переменных. В различных исследованиях отличаются используемые исходные географические данные, методы измерения концентраций ЗВ, а также особенности применения регрессионного анализа для построения моделей распределения ЗВ в почвенном покрове. Из-за значительных местных особенностей способы применения метода различны в каждом исследовании. В настоящее время метод LUR не стандартизирован и не оформлен в виде строгого алгоритма.

Основные этапы работы по построению модели LUR:

- измерения концентраций загрязняющих веществ в почве,
- поиск источников географической информации,
- статистический анализ полученных экспериментальных и расчетных данных,
- создание на основе полученной статистической модели карты распределения загрязняющих веществ в почвенном покрове исследуемого города.

Метод LUR впервые был применен в 1996 г., первое название метода Regressionmapping (регрессионная картография) [5]. В настоящее время этот метод чаще называют Land Use Regression. Применяют данный метод, который продолжает развиваться и изменяться, по-разному, но в каж-

дом исследовании имеются общие основные этапы работы (рис. 1).

В исследовании места отбора проб выбираются таким образом, чтобы в результате мониторинга получить весь спектр возможных концентраций ЗВ в городе. Далее для каждого места измерения вычисляется ряд географических переменных, которые, предположительно, могут быть связаны с распределением ЗВ. Используются географические переменные, описывающие местоположение точек измерения, тип землепользования, плотность застройки, показатели интенсивности автомобильного движения и другие. Затем проводится регрессионный анализ для определения взаимосвязей измеренных концентраций и полученных географических переменных. В результате получается регрессионное уравнение, которое может быть использовано для оценки концентраций ЗВ в любой точке города.

В отличие от методов интерполяции метод LUR, кроме данных об измеренных концентрациях, для построения поверхности распределения загрязняющих веществ использует различные доступные данные ГИС, включение которых в анализ значительно улучшает модель, пространственное разрешение результатов моделирования, позволяя ограничиться лишь небольшим количеством измерений. Выбор числа мест отбора проб ограничен физическими и материальными возможностями исследователя, наличием необходимого измерительного оборудования. Строгой методики для определения количества мест отбора проб не разработано. Среди ранее проведенных исследований количество проб варьировалось от 18 до 120. Число проб должно зависеть от местных особенностей, ожидаемых вариаций измеряемых концентраций, а также размера области, в которой проводится исследование. В предыдущих

работах, использующих LUR, анализ для моделирования уровней распределения загрязняющих веществ крупного города считалось целесообразным использовать от 40 до 80 мест отбора проб [3]. Такое число мест измерений обусловлено необходимым объемом данных для статистического анализа и, по мнению исследователей [3], не зависит от размеров города.

В Канаде [4, 8] был применен алгоритм для определения мест отбора проб, в котором учитывалась информация о транспортной сети и местах жительства изучаемой группы населения. Идея заключалась в том, чтобы расположить больше точек пробоотбора в местах, где предположительно больше пространственная неоднородность концентрации ЗВ (которая оценивалась по плотности загруженных дорог), и выделить районы города с наибольшим числом проживающих там участников исследования.

Для исследуемого города создается геоинформационная модель, база географических данных, в которую включают данные о распределении автомобильных дорог, различных типов землепользования, топографических параметров местности, плотности населения и другие. С помощью технологий ГИС на основе полученной базы географических данных вычисляются переменные – предикторы, описывающие места измерений.

Обычно удается преобразовать собранную информацию в 50–150 различных географических переменных. В окончательном варианте в модели LUR, как правило, включается от 2 до 8 переменных. Поскольку априори неизвестно, какие переменные имеют наибольшую взаимосвязь с уровнем загрязнения данного конкретного населенного пункта, неизвестно какого размера должна быть буферная зона предиктора для лучшего описания взаимосвязи с загрязнением, то изначально рассчитывается большой набор потенциальных переменных – предикторов. Например, исследование [7] включало 55 потенциальных предикторов, исследование [9] – 140 предикторов.

Географические данные редко находятся в необходимом для моделирования формате. Для моделирования LUR используются растровые форматы географических данных, все векторные форматы преобразуются в растровые. Кроме прямого преобразования вектор – растр используются различные инструменты ГИС для пространственной обработки данных: оценка расстояния до объектов, оценка плотности объектов и другие.

Для характеристики плотности дорог различных типов, площади парковых, промышленных, жилых зон и других параметров в координатах мест измерений формируются буферные зоны – круги разных радиусов. Затем средствами ГИС производится расчет длины дорог, попавших внутрь круговой буферной зоны, площади парковых зон и т.п. Таким образом, современные технологии ГИС позволяют получить множество переменных с любыми размерами буферных зон. Выбор размеров должен быть основан на информации о вероятном распределении ЗВ. Однако, когда нет четких данных о требуемых размерах буферных зон для характеристики определенных параметров, то можно вычислить большое количество переменных с различными размерами буферных зон. Размеры могут быть от самых малых, например, несколько метров, насколько позволяет пространственное разрешение входных данных, до очень больших до 1000 м. Строить буферные зоны более 1 км нет необходимости, так как теряется предметный смысл переменных.

#### **Направления развития метода Land Use Regression**

Метод LUR был разработан и впервые применен в Европе [5, 7] для оценки персональных экспозиций к автотранспортным выбросам для эпидемиологических исследований с большим числом участников. Несмотря на частое применение на западе, метод Land Use Regression в Российской Федерации не использовался до 2012 г.

Для создания карты распределения ЗВ конкретной территории необходимо провести поиск доступных источников информации. Каждому источнику информации необходимо дать оценку возможности взаимосвязи с загрязнением, а затем на основе собственных предположений создать максимально возможное количество вариантов переменных, отражающих рассматриваемые взаимосвязи. Каждая созданная переменная, по сути, является гипотезой о форме и направленности взаимосвязи географических данных с распределением ЗВ. В процессе статистического анализа эти гипотезы либо подтверждаются (тогда переменная попадает в модель), либо отвергаются (тогда переменная отбрасывается). Наличие подробного описания исследуемого города и верное преобразование исходных данных в переменные – предикторы – это наиболее важное условие для успешного моделирования распределения ЗВ методом LUR.

В предыдущих исследованиях европейских и американских городов был определен ряд нерешенных задач для развития LUR.

1. Метеорологические процессы играют очень важную роль в формировании и переносе ЗВ. Однако на данный момент не решена задача использования метеорологических параметров в моделях LUR. В некоторых исследованиях [6, 10] были сделаны попытки учитывать направление ветра, но улучшить характеристики получаемой модели распределения ЗВ не удалось. Был сделан вывод о том, что местные метеорологические условия сложились таким образом, что их учет не приводит к значительному улучшению модели (речь идет не о краткосрочных данных, на которые мог бы повлиять ветер, а о долгосрочных).

2. Учет структуры застройки территории, при которой здания влияют на распространение ЗВ (например, образование улицами искусственных каньонов). Как правило, такие условия достаточно сложно включить в модель с помощью доступных данных ГИС, а ручная классификация неприемлема для больших территорий целого города.

3. Метод LUR предназначен для моделирования суммарного распределения ЗВ от всех источников, моделирование автотранспорта не представляет особых сложностей. Однако LUR уступает современным расчетным моделям в точности моделирования выбросов стационарных источников ЗВ. В некоторых случаях могут возникнуть искажения модели и карты распределения ЗВ. Сочетание LUR и одного из методов моделирования рассеивания позволит создавать более качественные карты распределения ЗВ в городах со значительным количеством стационарных источников загрязнения.

### **Построение моделей (регрессионный анализ)**

Традиционный метод LUR использует регрессионный анализ для создания моделей распределения ЗВ. В большинстве исследований для создания прогностических моделей распределения ЗВ в почвенном покрове использовался метод множественной линейной регрессии, уравнение которой связывает измеренные концентрации и рассчитанные географические переменные.

Регрессионный анализ – один из основных методов современной прикладной математической статистики [2]. Метод заключается в механизмах усреднения, процедурах сглаживания, принципах согласования противоречивых позиций, концепциях оптимальности.

Выбор переменных, тесно связанных с построением модели объекта, – одна из известных и сложных проблем. Стандартных общих решений для всех случаев не существует.

Существенного сокращения числа генерируемых для сравнения наборов предикторов можно добиться с помощью пошаговых процедур отбора [1]. Пошаговая схема отбора – один из лучших инструментов для нахождения оптимальной модели, которая позволяет при большом числе предикторов построить модель регрессии и избежать большого числа вычислений. Но, как и во всех других методах регрессионного анализа, требуются разумные суждения при первоначальном выборе переменных и критическом анализе модели путем исследования остатков.

Поиск оптимальной модели с помощью методов регрессионного анализа может обнаружить ситуацию, когда имеются два или более предикторов – “кандидатов для включения в модель” вместо одного. Если это имеет место, то можно сказать, что данные содержат недостаточно информации для однозначного выбора. Чтобы окончательно выбрать модель, требуются дополнительные априорные соображения.

### **МЕТОДИКА ИССЛЕДОВАНИЙ**

**Отбор проб почвы.** Фактическое расположение точек определялось при проведении опробования непосредственно на местности, исходя из необходимости отбора проб почвы на ненарушенных и естественных участках исследуемой зоны. Всего в рамках настоящей почвенной съемки была отобрана 81 проба вблизи перекрестков значимых автомобильных дорог.

Поверхность места предполагаемого отбора пробы почвы размечалась в виде квадрата со стороной около 1 м. В вершинах, центре и внутри размеченного квадрата пробоотборником из нержавеющей стали диаметром порядка 0.05 м отбирались 5 кернов почвы на глубину 0.05 м. Координаты центра квадрата записывались в GPS приемник Garmin GPSmap 60Сх. Отобранные керны объединялись в одну пробу и запаковывались в двойные полиэтиленовые пакеты. На внутреннем пакете маркером наносился уникальный идентификатор пробы. Суммарная площадь пяти кернов составляла 0.01 м<sup>2</sup>, а суммарный объем – 0.0005 м<sup>3</sup>, масса высушенной пробы 0.9–1.2 кг.

**Рентгенофлуоресцентный анализ.** Для анализа проб почвы был выбран метод рентгенофлуоресцентной спектроскопии, не требующий

Коэффициенты корреляции между концентрациями элементов, выявленных при анализе проб почвы и предикторами (обозначение предиктора см. в тексте)

	Ca	K	Ti	Cr	V	Fe	Mn	Cu	Zn	Sr	Zr	Ba	Bi	Pb
S0008_R1	0,03	0,17	0,09	-0,09	<b>0,28</b>	-0,02	-0,06	-0,12	-0,02	0,08	0,02	0,19	0,12	-0,07
S0008_R2	-0,01	0,09	0,18	-0,03	0,16	-0,11	-0,07	-0,25	-0,06	-0,07	0,23	0,07	0,00	-0,17
S0008_R3	-0,10	-0,06	<b>0,28</b>	-0,00	0,11	-0,00	0,17	-0,20	-0,09	-0,19	<b>0,35</b>	-0,10	-0,00	-0,18
S0016_R1	-0,01	0,06	0,24	-0,03	0,22	-0,02	-0,03	-0,25	-0,08	-0,08	<b>0,26</b>	0,00	0,06	-0,21
S0016_R2	0,09	0,01	0,22	-0,05	0,23	-0,04	0,06	-0,26	-0,07	-0,01	0,15	-0,03	0,03	-0,13
S0016_R3	0,07	0,09	0,24	-0,02	0,23	0,00	0,05	-0,22	-0,05	-0,13	0,24	0,10	0,08	-0,18
S0032_R1	0,20	0,12	0,10	-0,00	0,20	-0,08	-0,02	-0,17	0,01	-0,04	0,10	0,14	0,02	0,01
S0032_R2	0,12	0,01	<b>0,31</b>	0,14	0,10	0,08	0,08	-0,15	-0,05	-0,11	<b>0,29</b>	-0,01	0,02	-0,14
S0032_R3	0,07	0,10	0,23	-0,02	0,18	-0,11	0,02	-0,29	-0,05	-0,05	0,22	0,05	-0,01	-0,06
S0064_IZ	-0,06	-0,08	0,12	-0,00	0,05	-0,02	-0,05	-0,14	-0,07	-0,12	0,19	-0,12	-0,05	-0,13
S0064_R1	0,12	-0,05	0,09	0,14	-0,05	0,02	-0,17	-0,13	0,13	0,00	0,05	-0,04	0,02	0,01
S0064_R2	0,13	-0,14	0,28	<b>0,31</b>	0,02	0,17	0,03	-0,18	0,26	-0,25	0,16	-0,10	0,11	-0,08
S0064_R3	0,17	-0,01	0,17	0,20	0,03	0,10	0,12	-0,16	-0,09	-0,17	0,16	-0,07	-0,01	-0,16
S0128_IZ	-0,16	0,10	0,17	-0,05	0,07	0,14	0,12	0,06	-0,16	-0,10	0,18	0,02	0,07	-0,01
S0128_R1	0,18	-0,06	0,13	0,11	0,04	0,04	-0,05	-0,09	0,06	0,03	0,06	0,00	0,08	-0,02
S0128_R2	0,15	-0,19	0,23	<b>0,28</b>	0,08	0,24	0,09	-0,00	0,02	-0,12	0,02	-0,14	0,15	-0,15
S0128_R3	0,14	-0,13	0,06	0,23	-0,07	0,03	0,01	-0,22	0,30	-0,14	0,13	0,00	0,02	-0,14
S0256_IZ	0,16	0,10	0,12	0,01	0,17	-0,08	-0,05	-0,14	0,02	-0,02	0,12	0,07	0,06	0,07
S0256_R1	0,17	0,13	-0,01	0,16	-0,01	-0,05	-0,02	-0,19	-0,03	0,04	-0,01	0,13	0,05	-0,01
S0256_R2	-0,14	-0,04	0,01	0,10	-0,15	-0,18	-0,09	-0,18	0,10	-0,17	0,18	0,08	-0,02	-0,02
S0256_R3	0,20	-0,02	-0,08	0,07	-0,04	-0,07	-0,04	-0,13	0,08	0,11	-0,02	0,11	-0,02	0,07
S0512_IZ	0,09	0,09	0,11	0,03	0,08	-0,09	-0,05	-0,16	-0,01	-0,13	0,12	0,05	-0,05	-0,04
S0512_R1	0,11	0,14	-0,05	-0,04	-0,05	-0,23	-0,26	-0,21	-0,03	0,23	0,26	0,22	-0,08	0,02
S0512_R2	-0,00	0,06	-0,05	-0,20	0,03	-0,13	-0,11	-0,01	-0,13	0,07	-0,08	0,05	-0,11	0,07
S0512_R3	-0,00	0,27	0,10	-0,11	0,14	-0,22	-0,06	-0,22	0,02	-0,01	0,23	0,27	-0,06	0,04
S1024_IZ	0,11	0,01	0,26	-0,06	0,20	0,06	0,09	-0,10	-0,05	0,03	0,24	0,05	0,09	0,01
S1024_R1	0,21	<b>0,42</b>	-0,05	-0,05	0,13	-0,13	0,06	-0,05	-0,06	0,20	0,03	0,27	0,05	-0,02
S1024_R2	-0,02	<b>0,39</b>	-0,14	-0,20	0,13	-0,27	-0,04	-0,05	-0,15	0,25	0,12	0,27	-0,03	-0,02
S1024_R3	0,02	<b>0,43</b>	-0,09	-0,26	0,12	-0,27	-0,01	-0,05	-0,12	0,19	0,14	0,30	-0,06	0,01

сложной пробоподготовки и дорогостоящих реагентов, позволяющий проводить анализ образца многократно.

Пробы проанализированы на рентгенофлуоресцентном спектрометре Innov XX 5000. Образец почвы, помещенный в тонкий полиэтиленовый пакет, анализировался в подходящем для экологических исследований режиме "Почва", позволяющем использовать для анализа лучи различной мощности (от 1 до 50 кЭВ), для распознавания отдельных элементов в минимальных концентрациях (ppm). Время экспозиции каждой пробы составляло 90 с. Каждый образец экспонировался 5 раз, для дальнейшей обработки использовалось среднее значение. Полученные результаты записывались прибором в файл приложения MS Excel для дальнейшего анализа. Перечень элементов, определенных спектрометром в пробах почвы представлен в таблице.

## Построение геоинформационной базы данных

**Создание таблицы полевых и аналитических результатов.** Полученные при проведении полевых работ данные и результаты рентгенофлуоресцентного анализа проб были объединены в общую таблицу базы данных SoilXY в приложении MS Access.

**Создание таблиц топоосновы и географических переменных.** Средствами приложения ArcCatalog на основе таблицы базы данных SoilXY созданы таблицы топоосновы и геоданных, в которые были записаны результаты вычислений и из которых с помощью программы ArcGIS извлекалась информация для построения полей распределения ЗВ.

Для создания географических информационных моделей использовался открытый источник географических данных: программа Google Earth.

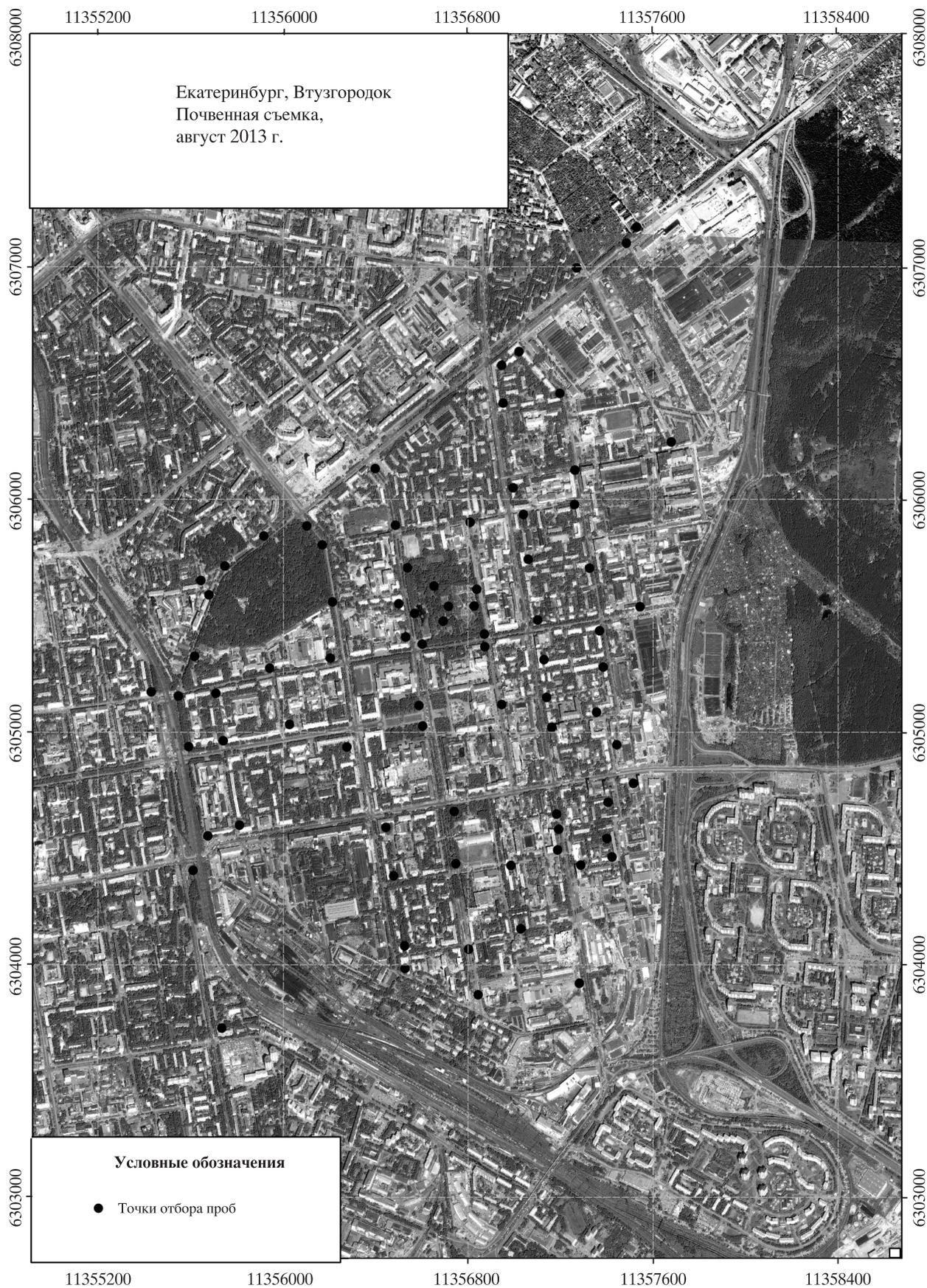


Рис. 2. Места отбора проб почв.

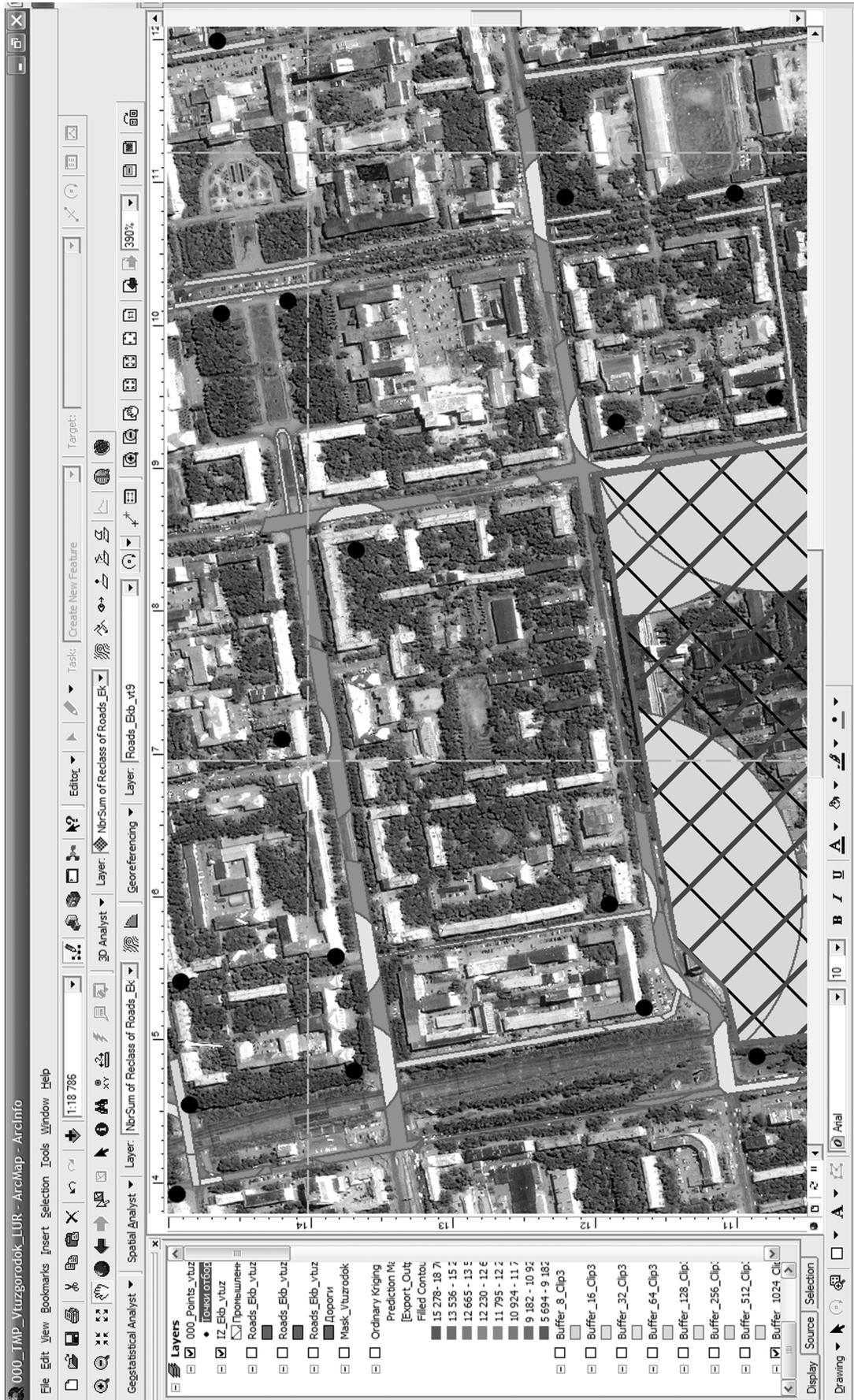


Рис. 3. Пересечение объекта с буферной зоной

При помощи этой программы в приложение ArcGIS загружен растр исследуемого района.

Затем в приложении ArcGIS координаты точек отбора из таблицы SoilXY были нанесены на растр (рис. 2).

Предполагая, что основное влияние на распределение ЗВ в почве исследуемого района оказывают интенсивность транспортных потоков и воздействие близко расположенных промышленных зон, в настоящем исследовании было решено ограничиться анализом переменных, описывающих эти факторы.

В отличие от предыдущих исследований, в которых дороги определялись как линейные источники выбросов, было предложено рассматривать дороги как площадные источники, что при определенных допущениях должно позволить уйти от предварительного разделения дорог на категории, что может повысить качество моделей и упростить проведение анализа. «Категорийность» дороги в этом случае будет определять ее площадь, попадающая в соответствующую буферную зону. В настоящем исследовании было решено сохранить разбиение дорог на категории для проверки этого предположения последующим анализом.

Была нанесена сетка дорог и разделена на категории в зависимости от плотности транспортного потока, исходя из визуального наблюдения во время проведения отбора проб. Также нанесены промышленные зоны, которые оказывают наибольшее влияние на выбранный полигон.

#### **Группы переменных:**

первая – промышленные площадки микрорайона Втузгородок);

вторая – плотность дорог R1 (главные транспортные артерии микрорайона Втузгородок);

третья – плотность дорог R2 (второстепенные дороги микрорайона Втузгородок с большим трафиком);

четвертая – плотность дорог R3 (все прочие незначительные улицы и дороги с небольшим движением автотранспорта и т.п.).

Всего было получено 32 переменных-предиктора. Название предиктора составлено из двух частей: первая часть, начинающаяся с буквы S, – радиус буферной зоны в метрах; вторая – указывает на принадлежность к дорогам соответствующей категории R1, R2, R3, либо к промышленным зонам IZ. Например, предиктор S0128\_2 обозначает

плотность дорог второй категории в буферной зоне радиусом 128 м, предиктор S0512\_IZ обозначает плотность промышленных зон в буферной зоне 512 м.

Разнообразие переменных, а также масштабов, в которых они были рассчитаны (от 8 до 1024 м), дало возможность определить те из них, которые имеют наибольшую взаимосвязь с уровнем загрязнения, а затем построить основанную на них модель распределения ЗВ населенного пункта.

**Создание таблицы пересечений.** В приложении ArcGIS были построены пересечения дорог и промышленных зон с буферными зонами (рис. 3). При построении автоматически создавался файл баз данных, состоящий из большого числа таблиц.

Средствами приложения MS Access из таблицы SoilXY и таблиц геоданных с результатами вычисления площадей пересечений буферных зон и соответствующих групп переменных была создана единая таблица баз географических и аналитических данных SoilXY\_Int. Далее таблица была импортирована в программу Statistica для осуществления LUR анализа.

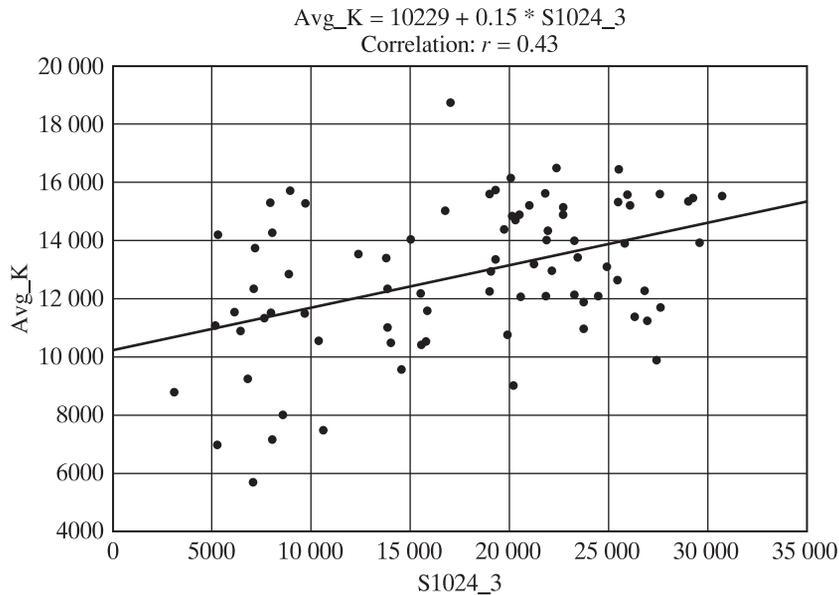
#### **LUR анализ экспериментальных данных**

С помощью метода пошаговой регрессии уравнение регрессии формируется последовательно – на каждом шаге в модель распределения ЗВ добавляется одна переменная в порядке статистической значимости.

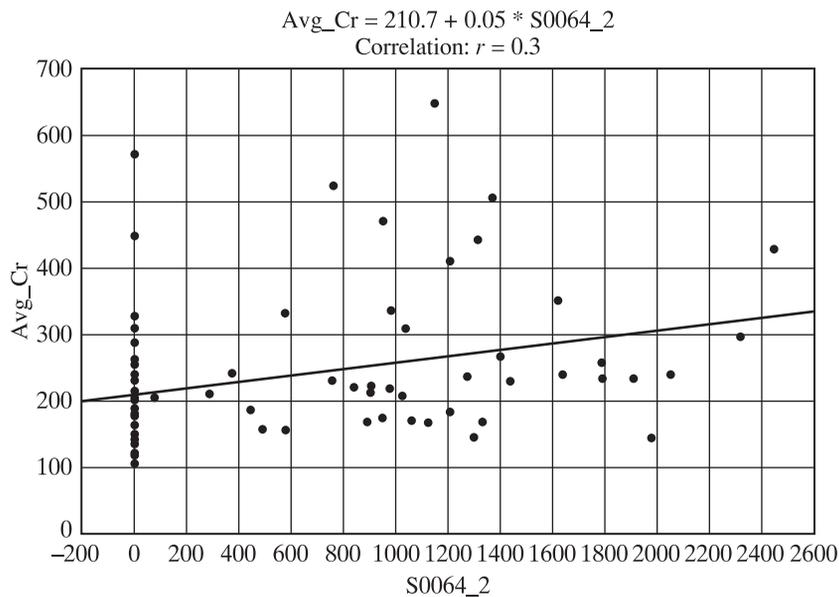
Формирование регрессионного уравнения всегда начинается с наиболее важной переменной, имеющей наибольшую статистическую значимость. Более сложный вопрос заключается в выборе второй и последующих переменных. Остальные переменные не всегда имеют высокую статистическую значимость, возможны ситуации, когда несколько разных переменных могут быть выбраны на роль второй переменной в модели. Это может привести к появлению нескольких вариантов моделей распределения ЗВ, каждую из которых необходимо изучить отдельно и сделать вывод о применимости и качестве с предметной точки зрения.

#### **РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ**

Анализ парных коэффициентов корреляции между концентрациями ЗВ в почве и предикторами показал, что наибольшие коэффициенты корреляции у предикторов, построенных на основе



**Рис. 4.** Зависимость между концентрацией калия и плотностью дорог третьей категории в буферной зоне радиусом 1024 м.



**Рис. 5.** Зависимость между концентрацией хрома и плотностью дорог второй категории в буферной зоне радиусом 64 м.

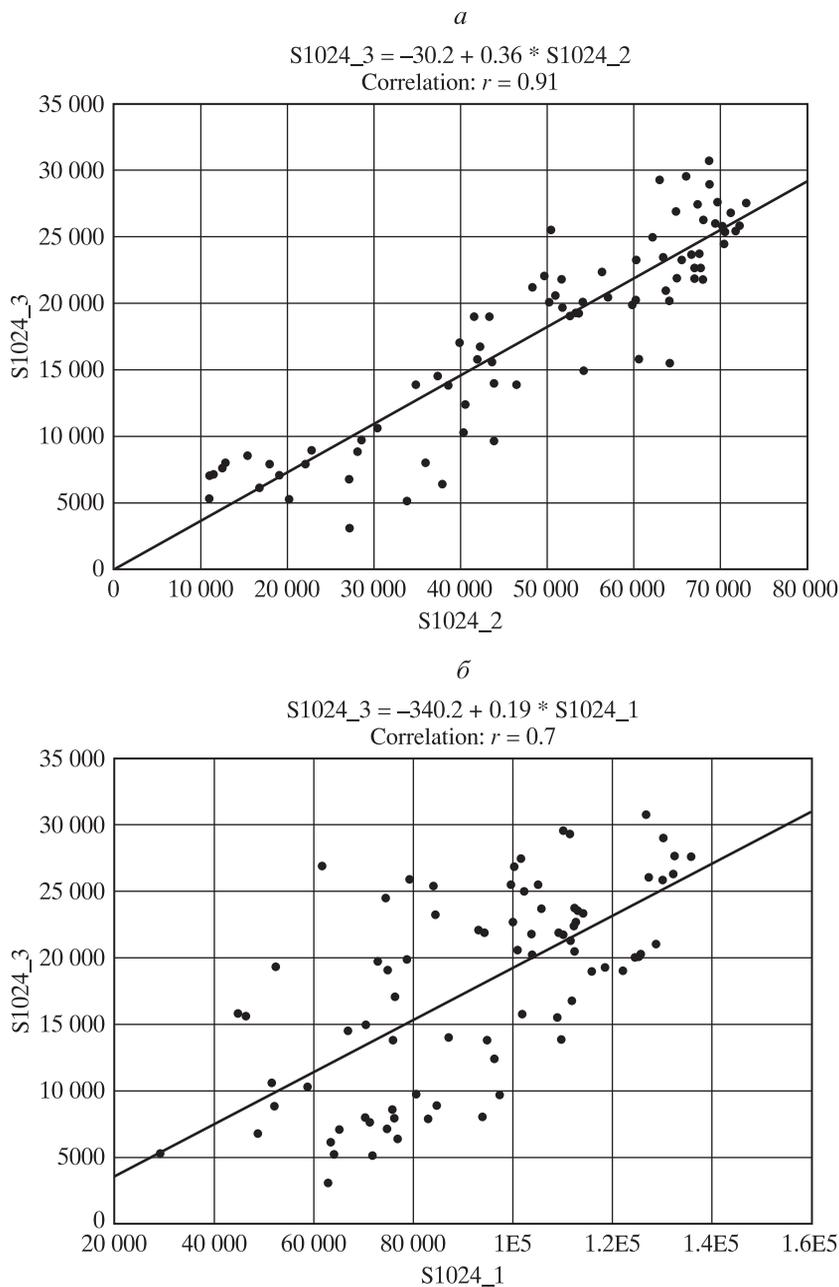
дорог и пяти элементов: калия, титана, ванадия, хрома и циркония (см. табл.).

Влияние переменных, связанных с промышленными зонами, на концентрации ЗВ минимально. Максимальные коэффициенты корреляции Пирсона (0.42, 0.39 и 0.43) отмечены между концентрацией калия и предикторами S1024\_3, S1024\_1 и S1024\_2, описывающими плотность дорог 3-й, 1-й и 2-й категории в буферной зоне 1024 м (рис. 4).

Объяснить эту связь с позиций предметной области в рамках настоящего исследования не

представляется возможным. Концентрации титана, ванадия, хрома и циркония коррелируют с предикторами, описывающими плотность дорог в буферных зонах от 8 до 512 м, причем максимальные коэффициенты корреляции (до 0.35 у циркония) наблюдаются для плотности дорог в радиусе от 8 до 64 м (рис. 5).

Однако построить минимально качественную многофакторную модель распределения концентраций не удалось ни для одного из элементов. Гетерогенная среда, подобная исследуемой в настоящей работе, для построения адекватной модели требует учета большого числа факторов, влияющих на ее



**Рис. 6.** Связь между предикторами, описывающими плотность дорог: *a* – 2-й и 3-й категории, *б* – 1-й и 3-й категории, в буферной зоне 1024 м.

состояние, для чего следует вводить в анализ новые переменные (метеоданные, плотность и характер застройки, зеленые насаждения и парковые зоны, водоемы и т.п.). Также необходим более тщательный подход к местам отбора проб из-за значительной площади почвенного покрова, нарушенного антропогенным воздействием.

Выявленные зависимости концентраций титана, ванадия, хрома и циркония от плотности дорог в радиусе от 8 до 64 м свидетельствуют о возможном влиянии последних на распределение

данных элементов в поверхностном слое почвы. Вероятно, что здесь присутствует эффект вторичного приземного массопереноса, который формирует интенсивные локальные пятна высокой концентрации элементов на исследуемой территории.

Из зависимостей, представленных на рис. 6, видно, что существует значимая связь между предикторами, описывающими плотность дорог различных категорий в буферной зоне 1024 м.

Учитывая масштаб исследуемого района (максимальная длина около 4000 м, ширина около 2000 м) в буферные зоны 1024 м попали все дороги, участвующие в анализе. Поскольку дороги в настоящей работе рассматриваются как площадные источники, информация об их “категорийности” вносится во время создания таблицы пересечений (большая площадь дороги – большая площадь, попавшая в буферную зону). Принимая во внимание, что интенсивность транспортных потоков на исследуемой территории примерно одинакова и зависит именно от площади проезжей части, и учитывая сильную связь между предикторами, описывающими плотность дорог всех трех категорий, от предварительного разбиения дорог на категории можно отказаться, при этом анализ упростится из-за существенного сокращения количества переменных.

### ВЫВОДЫ

В настоящей работе применен метод Land Use Regression (LUR) для построения полей распределения загрязняющих веществ в верхнем слое почвенного покрова урбанизированных территорий на примере микрорайона г. Екатеринбурга.

Предложено рассматривать дороги как площадной источник выбросов загрязняющих веществ, что позволяет предварительно не разбивать их на категории и упростить процедуру LUR анализа (в предыдущих работах дороги рассматривались как линейные источники с обязательным разбиением на категории по интенсивности транспортных потоков). Данное допущение для исследуемой территории подтвердилось статистическим анализом.

Выявлены зависимости концентраций титана, ванадия, хрома и циркония от плотности дорог в радиусе от 8 до 64 м (коэффициенты корреляции от 0.28 до 0.35), но построить многофакторную модель с высоким коэффициентом детерминации ( $R^2 > 0.5$ ) не удалось ни для одного из элементов.

Метод LUR в гетерогенной среде почвенного покрова урбанизированных территорий оказался чувствителен к выбору местоположения отбора проб, необходимо заранее представлять уровень загрязнения и пространственное распределение загрязнителей, для чего требуется проведение предварительного исследования с использованием других методов интерполяции, на основании которых откорректировать места отбора. Для получения точных моделей пространственного распределения загрязняющих веществ необходимо рассмотреть значительное число переменных (плотность и характер застройки, зеленые насаж-

дения и т.д.), некоторые из которых трудно учесть (например, метеорологические данные), что заметно усложняет проведение исследования.

### ЗАКЛЮЧЕНИЕ

Для моделирования распространения загрязняющих веществ в почвенном покрове методом регрессионной картографии целесообразно создать комплекс программ автоматизированного расчета переменных модели, выбора и формирования адекватной модели, расширяющих возможности стандартного программного обеспечения.

Подробная картина пространственного распределения загрязняющих веществ в почвенном покрове необходима для решения многих задач: идентификации конкретных источников загрязнения данной территории, разработки мероприятий по эффективному сокращению выбросов; количественной оценки влияния загрязняющих веществ на здоровье населения. Информация о пространственном распределении концентрации загрязняющих веществ необходима при планировании строительства транспортных узлов и социально-значимых учреждений.

*Исследование выполнено по проекту 15-ИПЭ-02.*

### СПИСОК ЛИТЕРАТУРЫ

1. Айвазян С.А., Енюков И.С. и др. Прикладная статистика. Исследование зависимостей. М.: Финансы и статистика, 1985. 487 с.
2. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. М.: Финансы и статистика, 1986. Кн. 1. 366 с.; Кн. 2. 351 с.
3. Brauer M., Henderson S. Land Use Regression Road Map for the Burrard Inlet Area Local Air Quality Study. Final report. 2006. P. 47.
4. Brauer M., Hoek G. Prediction of long term average particulate air pollution concentrations by traffic indicators for epidemiological studies // *Epidemiol.* 2003. V. 14. № 2. P. 228–239.
5. Briggs D., Collins S. Avd. Veen. Mapping urban air pollution using GIS: a regression-based approach // *J. Geogr. Inf. Sci.* 1997. V. 11. № 7. P. 699–718.
6. Jerrett M., Arain M. Modelling the intra-urban variability of ambient traffic pollution in Toronto, Canada // *J. Toxicol. Environ. Health.* 2007. V. 70. № 3–4. P. 200–212.
7. Henderson S., Beckerman B. Application of LUR to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter // *Environ. Sci. Technol.* 2007. V. 41. (7). P. 2422–2428.

8. *Kanaroglou P.S., Jerrett M.* Establishing an air pollution monitoring network for intra-urban population exposure assessment: a location-allocation approach // *Atmospheric Environment*. 2005. V. 39. № 13. P. 2399–2409.
9. *Moore D.K., Jerrett M.* A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA // *J. Environ. Monitor.* 2007. V. 9. № 3. P. 246–252.
10. *Sahsuvaroglu T., Arain A.* A LUR model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Canada // *J. Air Waste Manage. Assoc.* 2006. V. 56. № 8. P. 1059–1069.

## STATISTICAL ANALYSIS OF POLLUTANT DISTRIBUTION BY THE SURFACE SOIL LAYER IN URBAN AREAS USING MATHEMATICAL MODELS (LUR METHOD)

**A. G. Buevich, A. M. Safina, A. P. Sergeev, A. N. Varaksin, A. N. Medvedev**

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

The applicability of Land Use Regression (LUR) method for pollutant distribution analysis in the upper layer of soil cover in urban areas is analyzed by the example of Yekaterinburg city. LUR is a new approach to describing the distribution of contaminants in the surface soil layer. The method combines instrumental measurement of the pollutant distribution with the compilation of statistical models on the basis of experimental data on the distribution of pollutants and geographic information systems (GIS). For the analysis, eighty-one soil samples were collected at the highway crossroads in the selected district. The chemical analysis of samples was performed using the Innov XX 5000 X-ray fluorescence spectrometer. To simplify the LUR analysis, we proposed to use roads as the area sources of polluting emissions. We analyzed the regression models to describe the spatial distribution of pollutants on the basis of experimental data on soil contamination and geographical information of Vtuzgorodok district of Yekaterinburg city.

**Keywords:** *Land Use Regression, regression model, interpolation techniques, modeling techniques scattering, univariate analysis.*