

© Рапаков Г.Г., Лебедева Е.А.,
Горбунов В.А., Абдалов К.А., Мельничук О.В., 2018

Рапаков Георгий Германович

Кандидат технических наук, доцент,
Вологодский государственный университет
(Вологда, Россия)
E-mail: grapakov@yandex.ru

Rapakov Georgij Germanovich

PhD (Technology), Associate Professor,
Vologda State University (Vologda, Russia)
E-mail: grapakov@yandex.ru

Лебедева Елена Александровна

Кандидат технических наук, доцент,
Вологодский государственный университет
(Вологда, Россия)
E-mail: lebedeva.elena@inbox.ru

Lebedeva Elena Aleksandrovna

PhD (Technology), Associate Professor,
Vologda State University
(Vologda, Russia)
E-mail: lebedeva.elena@inbox.ru

Горбунов Вячеслав Алексеевич

Доктор физико-математических наук,
профессор, Вологодский государственный
университет
(Вологда, Россия)
E-mail: gorbunov1945@inbox.ru

Gorbunov Vyacheslav Alekseevich

Doctor of Sciences (Physico-mathematical),
professor, Vologda State University
(Vologda, Russia)
E-mail: gorbunov1945@inbox.ru

Абдалов Кенгес Абдалович

Доктор медицинских наук, заведующий
хирургическим отделением № 2 БУЗВО
Вологодской городской больницы № 1
(Вологда, Россия)
E-mail: muzgb1@inbox.ru, muzgb1@visp.ru

Abdalov Kenges Abdalovich

Doctor of Medicine,
M.D. Head of the surgical department no. 2,
Vologda municipal hospital no. 1
(Vologda, Russia)
E-mail: muzgb1@inbox.ru, muzgb1@visp.ru

Мельничук Олеся Валерьевна

Врач-стоматолог, городская
стоматологическая поликлиника №33
(Санкт-Петербург, Россия)
E-mail: dr_move@mail.ru

Mel'nichuk Olesia Valer'evna

Stomatologist, Municipal Stomatological
Polyclinic no. 33
(Saint-Petersburg, Russia).
E-mail: dr_move@mail.ru

**АНАЛИЗ ПРОСТРАНСТВЕННЫХ
КЛАСТЕРОВ И ИДЕНТИФИКАЦИЯ
ВЫБРОСОВ С ПОМОЩЬЮ
ГЕОИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ**

**SPATIAL CLUSTER ANALYSIS AND
EMISSION DETECTION BY USING
GEOINFORMATION TECHNOLOGY**

Аннотация. Работа представляет результаты исследований в области пространственной кластеризации селитебной территории г. Вологды. Анализ кластеров и выбросов выполнен на основе статистического

Abstract. The paper reports the research results in the field of the spatial cluster analysis and emission detection of Vologda's residential area. To handle the data of the geoeological monitoring, Anselin Local Moran's I statistical

показателя Anselin (локальный индекс Морана I) при обработке данных геоэкологического мониторинга. В ходе геоинформационного моделирования вычислено значение среднего расстояния до соседних объектов, реализована пошаговая пространственная автокорреляция, получена сводка индекса Морана по расстоянию, обнаружены кластеры и выявлены выбросы.

Ключевые слова: компьютерное моделирование, пространственная кластеризация, выявление выбросов, геоинформационные системы, геостатистика, геоэкологический мониторинг, поддержка решений

value is used to make the analysis of the clusters and emissions. While geoinformatic modeling, the value of the average distance to the nearest objects is calculated, the step-by-step spatial autocorrelation and the distance report of the Moran's index are made, the clusters and emissions are found.

Keywords: computer modeling, spatial clustering, emission detection, geographic information systems, geostatistics, geoecological monitoring, decision support

Введение

Пространственная кластеризация геоинформационных данных позволяет выполнить обнаружение статистически значимых территориальных неоднородностей, возникающих в случаях, когда критическое число исследуемых признаков встречается в непосредственной близости друг от друга. Кластеры могут быть представлены в виде географического распределения – группировки объектов с малыми или высокими значениями изучаемых свойств. Визуализация результатов пространственной кластеризации наглядно отображает холодные точки, горячие пятна и пространственные выбросы, которые отличаются резко выделяющимися значениями экспериментальных величин [9], [12].

Исследование влияния среды обитания на человеческое здоровье является одним из перспективных научных направлений [6]. Растущая антропогенная нагрузка на биосферу обуславливает возникновение и особенности распространения целого ряда неинфекционных заболеваний. Так, длительное воздействие на организм фтора в зоне влияния выбросов промышленных предприятий способствует развитию дистрофии миокарда и ишемической болезни сердца [11]. Оценки многосредового риска для здоровья детей демонстрируют взаимосвязь между экологическим состоянием урбанизированных территорий и возникновением врожденных пороков развития новорожденных. Техногенный прессинг увеличивает генетическую нагрузку на популяцию [7]. Решению проблемы экологической медицины содействуют смешанные подходы, основанные на привлечении традиционного статистического анализа и методов искусственного интеллекта [13]. С ростом количества и качества территориально пространственных антропогенных факторов риска, как объектов регионального управления, возрастает роль интеллектуальных систем поддержки принятия решений [10].

Актуальной является задача геоинформационного моделирования по идентификации кластеров и определению выбросов при анализе пространственного распределения загрязнения снежного покрова как характеристики техногенного воздействия на окружающую среду. Практическая значимость исследования определяется возможностью использования результатов при информационно-аналитической поддержке задач обеспечения экологической безопасности селитебных территорий. Цель настоящей работы – исследование методов пространственной статистики на основе компьютер-

ного моделирования территориального распределения и анализа структурных закономерностей. Расчет статистически значимых показателей интенсивности кластеризации дает представление о пространственных процессах и визуализации результатов анализа геоэкологических данных, что обуславливает новизну работы.

Основная часть

Методы пространственной статистики позволяют определять ключевые характеристики распределения изучаемой трехмерной переменной $Z(X, Y)$, выполнить оценку общих пространственных структурных закономерностей и идентифицировать статистически значимые пространственные кластеры и выбросы. При концептуализации пространственных отношений каждый объект оценивается в контексте окрестных соседей. Для обеспечения надежности анализа требуется вычислить диапазон расстояний до соседних объектов так, чтобы для каждого элемента геоданных имелся как минимум один сосед. В качестве метода определения расстояния используется евклидова метрика. При применении географических координат пространственный расчет выполняется с помощью хордовых измерений, которые дают хорошие оценки геодезических расстояний, так как область исследования существенно меньше 30° . В противном случае необходимо производить проецирование входных данных. Выбранному масштабу анализа с учетом границ области исследования соответствует, как правило, среднее значение диапазона расстояний до соседнего объекта, которое используется в качестве начального для пошаговой пространственной автокорреляции (ППА). Минимальное расстояние не применяется в случае, если во входном наборе геоданных предполагается наличие выбросов. Использование максимального значения приводит к неоправданному росту числа соседей, когда у каждого объекта их слишком много. При этом для ряда объектов соседями будут все или большая часть остальных элементов геоданных. Метод ППА позволяет определить расстояние, на котором пространственная кластеризация наиболее очевидна [15]. Пространственная автокорреляция измеряется для серии расстояний. Приращение задается как половина от величины среднего значения диапазона расстояний до соседнего объекта.

Интенсивность пространственной кластеризации отражается при помощи Z – оценки. Ее статистически значимая пиковая величина свидетельствует о наличии выраженной кластеризации во входных геоинформационных данных. Статистическую значимость определяют на основе p – значения, которое используется для проверки нулевой гипотезы. Для индекса Морана нулевая гипотеза утверждает, что изучаемые атрибуты геоданных в области исследования распределены случайно и пространственные процессы, являющиеся источником структур, хаотичны. При 95 % доверительном уровне $p < 0,05$, а критические значения Z – оценки представлены $-1,96$ и $+1,96$ стандартными отклонениями. Если $Z < -1,96$, то пространственная закономерность является дисперсионной. Для $Z > 1,96$ структурированность носит кластеризованный характер.

Расстояние, соответствующее пиковому значению ППА, используется в качестве порогового для метода анализа кластеров и выбросов на основе статистического показателя Anselin локального индекса Морана I. Данный метод позволяет определить присутствие статистически значимых кластеров (выбросов) во входном классе объ-

ектов с достоверностью 95 %. Для статистически значимого кластера с высокими схожими значениями (НН) Z – оценка имеет большое положительное значение. Малые положительные величины Z – оценки отвечают статистически значимому кластеру с низкими подобными значениями (LL). Если Z – оценки имеют малые отрицательные значения, то присутствуют статистически значимые пространственные выбросы: HL, если объект с высоким значением окружен малыми элементами, и LH, когда атрибут геоданных с низким значением выявлен среди высоких объектов. При определении пространственных взаимоотношений для точечных данных используется метод фиксированного диапазона расстояний. В этом случае при анализе объекта данные, расположенные за пределами порогового расстояния, игнорируются [1], [8]. Расчет Z – оценки производится следующим образом:

$$Z_{I_i} = \frac{I_i - \mathbf{E}[I_i]}{\sqrt{\mathbf{V}[I_i]}} ,$$

где:

$$\begin{aligned} \mathbf{E}[I_i] &= -\frac{1}{n-1} \sum_{j=1, j \neq i}^n w_{i,j} , & \mathbf{V}[I_i] &= \mathbf{E}[I_i^2] - \mathbf{E}[I_i]^2 , \\ I_i &= \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^n w_{i,j} (x_j - \bar{X}) , & S_i^2 &= \frac{1}{n-1} \sum_{j=1, j \neq i}^n (x_j - \bar{X})^2 \\ \mathbf{E}[I_i^2] &= A - B , & A &= \frac{n - b_{2i}}{n-1} \sum_{j=1, j \neq i}^n w_{i,j}^2 , \\ B &= -\frac{2b_{2i} - n}{(n-1)(n-2)} \sum_{k=1, k \neq i}^n \sum_{h=1, h \neq i}^n w_{i,k} w_{i,h} , & b_{2i} &= \sum_{i=1, i \neq j}^n (x_i - \bar{X})^4 \left/ \left(\sum_{i=1, i \neq j}^n (x_i - \bar{X})^2 \right)^2 \right. , \end{aligned}$$

где локальный индекс Морана I_i статистически оценивает пространственную связь объектов; x_i – числовой атрибут i -го объекта; $\bar{X}$ – среднее значение данного атрибута; $w_{i,j}$ – пространственный вес пары i, j ; n – число объектов [3].

В ходе работы выполнен литературный обзор научных источников по теме исследования. Предыдущие публикации авторского коллектива посвящены компьютерному моделированию методов пространственной интерполяции на основе сопоставления алгоритмов гридинга и сравнительному вариограммному анализу. В результате визуальной оценки гридинга и статистического анализа невязок для пространственной визуализации отобран метод кригинга. Оценка качества пространственного распределения удельной электропроводности снежного покрова получена при помощи перекрестной проверки грида. Выполнена визуализация результатов кросс-валидации для метода кригинга. Рассчитаны геостатистические показатели моделей вариограмм [10], [11]. В работе [5] обсуждается кластеризация территорий как источников чрезвычайных ситуаций по степени их рискованности. В качестве метода определения близости используется расстояние по прямой линии между двумя точками. Для исследования чрезвычайных ситуаций предлагается цветовая карта территории. Автор [2], тестируя гипотезу о зависимости темпов роста городов от их территориальной привязки, рассчитывает пространственный коэффициент ав-

токорреляции Морана для логарифмов темпов роста всех городов России. В публикации [1] приведен анализ пространственной кластеризации торговых точек. При обнаружении автокорреляции были использованы общие статистики пространственной корреляции Морана и Гири (Moran's I, Geary) и пространственная диаграмма Морана. Работа [9] знакомит с результатами выделения 4-х территориальных кластеров муниципальных образований Самарской области на основе глобального индекса Морана и 6 подкластеров с помощью локального индекса Морана (LISA – Local Index Spatial Autocorrelation). Выявлены зоны влияния для ядер – городских округов (рис. 1).

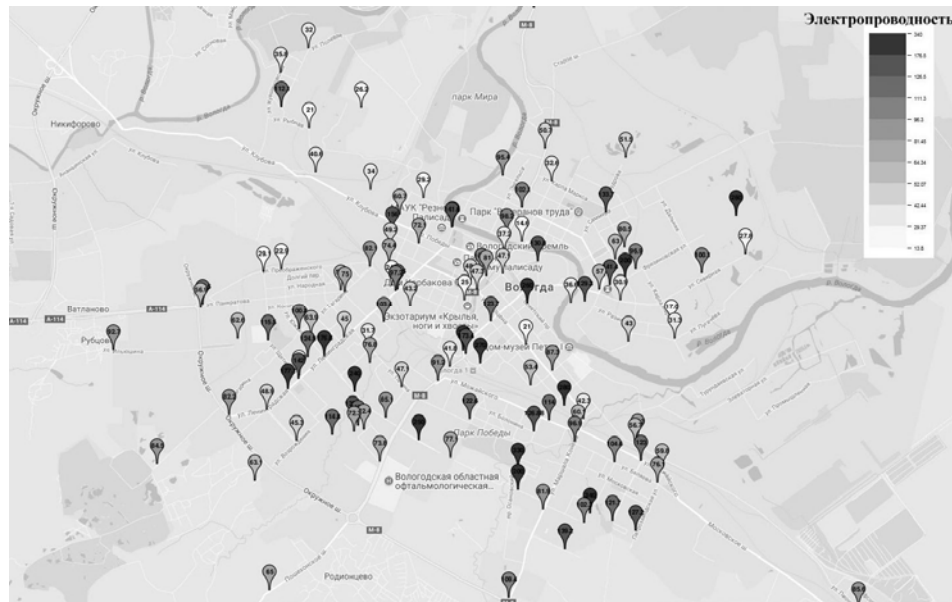


Рис. 1. Карта сбора точечных проб электропроводности

К перечню социально значимых региональных задач относится снижение рисков здоровью от химического загрязнения окружающей среды. Аккумулятором загрязнений атмосферы и индикатором экологического состояния среды обитания в городе является снег, который накапливается в течение холодного периода года. Системный подход к обеспечению экологической безопасности селитебных территорий и обеспечению надлежащего санитарно-гигиенического состояния поверхностных и подземных источников питьевого водоснабжения требует выполнения комплексных мероприятий [8]. Достижение максимальной эффективности капитальных затрат и инвестиций при планировании и реализации мер по снижению воздушного загрязнения городской среды, сбору и утилизации снега, водоотведению, обработке и доочистке сточных вод возможно при наличии статистически достоверной и легко интерпретируемой лицом, принимающим решения, информации. Отбор проб снега был выполнен в ходе полевого этапа исследования в период начала весеннего снеготаяния в марте 2017 года. Географические координаты точек определялись путем моделирования геопространственной сетки. При этом были использованы статистические

данные о состоянии здоровья населения г. Вологды (рис. 1). Объем выборки обеспечивает точность оценки не ниже 5 % с доверительной вероятностью $\alpha = 0,95$. Физико-химические исследования проб расплавленного снега выполнены по показателям: рН, взвешенные вещества и электропроводность.

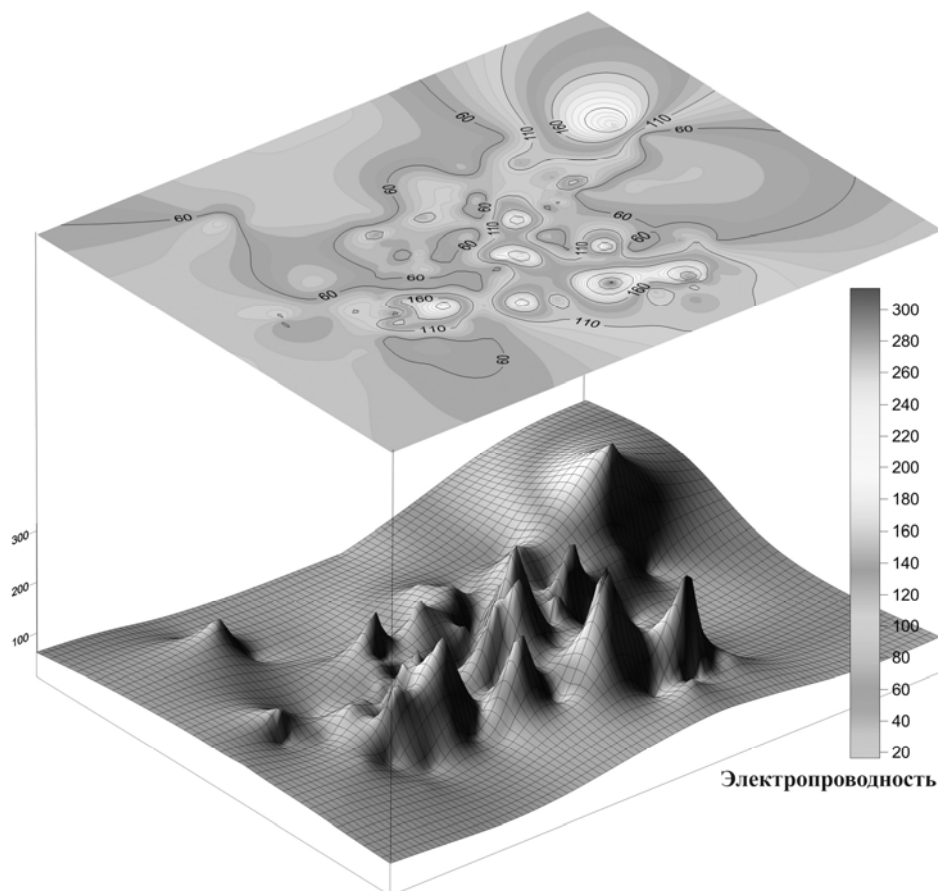


Рис. 2. Совмещенные карты изолиний и 3D поверхности для электропроводности

Наглядность сведений для лица, принимающего решение, обеспечивается за счет зрительного представления пространственно-распределенных данных. Карта пространственного распределения удельной электропроводности снегового фильтрата с интервалом между изолиниями 10 мкСм/см, полученная методом кригинга с применением исследуемых моделей вариограмм, представлена на рис. 2 [10], [11]. В целях удобства анализа она совмещена с соответствующей 3D поверхностью и образует стековую 3D картограмму. Полутоновое изображение карты рельефа наглядно отображает зоны вариации электропроводности и выделяет ее локальные максимумы.

Выбранному масштабу анализа соответствует среднее значение диапазона расстояний до соседнего объекта – 390 м. Это начальная величина расстояния, для которого выполняется ППА. От него задается приращение, равное половине среднего

значения диапазона расстояний до соседнего объекта – 195 м. ППА измеряется для серии из 10 расстояний. Получена сводка индекса Морана по расстоянию, содержащая для серий из последовательности приращений: индекс Морана, ожидаемый индекс, дисперсию, Z – оценку и p – значение (см. таблицу).

Таблица

Пошаговая пространственная автокорреляция по расстоянию

Расстояние	Индекс Морана	Ожидаемый индекс	Дисперсия	Z – оценка	p – значение
390	0,265500	-0,013333	0,013699	2,382341	0,017203
585	0,079670	-0,009346	0,007347	1,038481	0,299046
780	0,060815	-0,008929	0,004230	1,072324	0,283574
975	0,044007	-0,008696	0,002874	0,983087	0,325565
1170	0,048195	-0,008621	0,001780	1,346675	0,178085
1365	0,049014	-0,008547	0,001373	1,553360	0,120337
1560	0,076994	-0,008475	0,001056	2,629691	0,008546
1755	0,112167	-0,008403	0,000848	4,140661	0,000035
1950	0,089197	-0,008403	0,000651	3,825790	0,000130
2145	0,082863	-0,008333	0,000605	3,707863	0,000209

Пиковое значение зарегистрировано для расстояния 1755 м. Максимум Z – оценки равен 4,14, что свидетельствует о статистически значимой кластеризации ($p = 0,000035$). С учетом расстояния пикового значения (как порогового для метода анализа кластеров и выбросов на основе статистического показателя Anselin локального индекса Морана I) в результате геоинформационного моделирования получен выходной класс пространственных объектов, содержащий статистически значимые кластеры (НН и LL) и выбросы (НЛ и LH) (рис. 3). Для геообработки использовалась ГИС ArcGIS (Spatial Statistics Analysis Tools). Учитывался точечный характер сбора проб электропроводности при концептуализации пространственных отношений, поэтому был выбран метод фиксированного диапазона расстояний.



Рис. 3. Карта анализа кластеров и выбросов электропроводности

Выводы

В ходе компьютерного моделирования территориальных структурных закономерностей выполнен анализ пространственных кластеров и идентифицированы выбросы значений удельной электропроводности снежного покрова, полученные в ходе обработки данных геоэкологического мониторинга. На основе методов пространственной статистики вычислено значение среднего расстояния до соседних объектов, реализована пошаговая пространственная автокорреляция и получена сводка индекса Морана по расстоянию. Нулевая гипотеза, утверждающая, что изучаемые атрибуты геоданных в области исследования распределены случайно и пространственные процессы, являющиеся источником структур, хаотичны, отвергнута на уровне значимости $p < 0,001$. Максимум Z – оценки зарегистрированы для расстояния 1755 м, максимум равен 4,14, что свидетельствует о статистически значимой кластеризации. При анализе порогового расстояния на основе статистического показателя Anselin локального индекса Морана I были обнаружены кластеры и выявлены выбросы. Карта анализа кластеров и выбросов электропроводности наглядно воспроизводит статистически значимые пространственные закономерности двух типов, которые носят кластеризованный характер с высокими/низкими схожими значениями (HH/LL). Локальные атмогеохимические неоднородности, отличающиеся резко выделяющимися значениями геоэкологических данных, – статистически значимые пространственные выбросы – представлены в двух вариантах: HL и LH. Объект с высоким/ низким значением находится в окружении элементов с низким/ высоким значением атрибута геоданных. Результаты исследований уже использованы в региональном проекте

улучшения качества водоснабжения (обеспечение экологической безопасности сели-
тебных территорий). Перспективы работы – в изучении пространственной анизотро-
пии.

Литература

1. Балаш О.С. Пространственный анализ размещения торговых предприятий Саратова // Поволжский торгово-экономический журнал. 2010. № 4. С. 13–22.
2. Балаш О.С. Пространственный анализ темпов роста городов России // Известия Сара-
товского университета. Новая серия. Серия: экономика, управление, право. 2013. № 2. Т. 13.
С. 186–191.
3. Бердышев О.В., Долинов А.Л. Кластеризация поверхностно распределенного риска // Вестник Пермского национального исследовательского политехнического университета. Безопасность и управление рисками. 2015. № 3. С. 7–13.
4. Демьянов В.В., Савельева Е.А., Арутюнян Р.В. Геоestatистика: теория и практика. М.: Наука, 2010. 327 с.
5. Ершов Е.В., Бабенко А.И., Пониц Е.С., Хаснулин В.И. Система мониторинга состояния здоровья работников газодобывающего предприятия на Крайнем Севере // Бюллетень Сибир-
ского отделения Российской академии медицинских наук. 2008. № 2. Т. 28. С. 57–62.
6. Касаткина А.Л. Клинико-эпидемиологическая характеристика врожденных пороков челюстно-лицевой области на территории Волгограда и Волгоградской области: автореф. дис. ... канд. мед. наук. Волгоград, 2000. 18 с.
7. Мамадкулов Ю.О., Швецов А.Н. Модель мультиагентного сбора и обработки информа-
ции о состоянии окружающей среды // Информатизация процессов формирования открытых систем на основе СУБД, САПР, АСНИ и систем искусственного интеллекта: материалы V Меж-
дународной научно-технической конференции «ИНФОС–2009». Вологда: ВоГУ, 2009. С. 166–
169.
8. Онищенко Г.Г. Проблемы качества питьевой воды в Российской Федерации и пути их
решения // Водоснабжение и санитарная техника. 2010. № 12. С. 5–8.
9. Павлов Ю.В., Королева Е.Н. Пространственные взаимодействия: оценка на основе
глобального и локального индексов Морана // Пространственная экономика. 2014. № 3. С. 95–
110. DOI 10.14530/se.2014.3.95-110
10. Рапаков Г.Г., Лебедева Е.А., Горбунов В.А., Кузьминов А.Л., Абдалов К.А. Геоинфор-
мационное картографирование на основе машинного моделирования вариограммы // Вестник
Череповецкого государственного университета. 2018. №3 (84). С. 46–54. DOI 10.23859/1994-
0637-2018-3-84-6
11. Рапаков Г.Г., Лебедева Е.А., Горбунов В.А., Кузьминов А.Л., Абдалов К.А. Компью-
терное моделирование и анализ методов сеточной интерполяции при обработке геоинформа-
ционных данных // Вестник Череповецкого государственного университета. 2018. №2 (83).
С. 31–39. DOI 10.23859/1994-0637-2018-1-83-4
12. Филимонов С.Н., Семенова Е.А., Панев Н.И., Коротенко О.Ю., Филимонов Е.С., Ер-
шов Е.В., Никифоров В.В. Факторы риска ишемической болезни сердца у рабочих с хрониче-
ской фтористой интоксикацией // Вестник Кузбасского научного центра. 2014. №19. С. 116–
118.
13. Швецов А.Н., Суконщиков А.А., Кочкин Д.В., Андрианов И.А. Ситуационные интел-
лектуальные системы поддержки принятия решений. Курск: Университетская книга, 2018.
251 с.
14. Anselin L. Local Indicators of Spatial Association – LISA. Geographical Analysis, 1995.
Vol. 27 (2). Pp. 93–115. DOI 10.1111/j.1538-4632.1995.tb00338.x
15. Griffith D. Spatial Autocorrelation and Spatial Filtering: Gaining Understanding Through

Theory and Scientific Visualization. Springer-Verlag Berlin Heidelberg, 2003. 260 p. DOI 10.1007/978-3-540-24806-4

16. Haining R. Spatial Data Analysis: Theory and Practice. Cambridge University Press, 2003. 454 p. DOI 10.1017/CBO9780511754944

17. Hunsaker C., Goodchild M., Friedl M., Case T. Spatial Uncertainty in Ecology: Implications for Remote Sensing and GIS Applications. Springer-Verlag New York, Inc., 2001. 417 p. DOI 10.1007/978-1-4613-0209-4

18. Tiefelsdorf M. Modelling Spatial Processes: The Identification and Analysis of Spatial Relationships in Regression Residuals by Means of Moran's I. Berlin: Springer-Verlag Heidelberg, 2000. 167 p. DOI 10.1007/BFb0048754

References

1. Balash O.S. Prostranstvennyi analiz razmeshcheniia torgovykh predpriatii Saratova [Spatial analysis of the location of trading enterprises in Saratov]. *Povolzhskii torgovo-ekonomicheskii zhurnal* [Povolzhsky trade and economic journal], 2010, no. 4, pp. 13–22.

2. Balash O.S. Prostranstvennyi analiz tempov rosta gorodov Rossii [Spatial analysis of growth rates of Russian cities]. *Izvestiia Saratovskogo universiteta. Novaia seriia. Serii: ekonomika, upravlenie, pravo* [Izvestiya of Saratov university. New series. Series: economics, management, law], 2013, no. 2, vol. 13, pp. 186–191.

3. Berdyshev O.V., Dolinov A.L. Klasterizatsiia poverkhnostno raspredelen'nogo riska [Clustering of surface distributions risk]. *Vestnik Permskogo natsional'nogo issledovatel'skogo politekhnicheskogo universiteta. Bezopasnost' i upravlenie riskami* [Bulletin of the Perm national research polytechnic university. Safety and risk management], 2015, no. 3, pp. 7–13.

4. Dem'ianov V.V., Savel'eva E.A., Arutiunian R.V. *Geostatistika: teoriia i praktika* [Geostatistics: the theory and practice]. Moscow: Science, 2010. 327 p.

5. Ershov E.V., Babenko A.I., Ponich E.S., Khasnulin V.I. Sistema monitoringa sostoiianiia zdorov'ia rabotnikov gazodobyvaiushchego predpriatii na Krainem Severe [The system of health monitoring of workers of the gas enterprise in the far north]. *Biulleten' sibirskogo otdeleniia rossiiskoi akademii meditsinskikh nauk* [The bulletin of Siberian branch of Russian academy of medical sciences], 2008, no. 2, vol. 28, pp. 57–62.

6. Kasatkina A.L. *Kliniko-epidemiologicheskaiia kharakteristika vrozhdennykh porokov cheliustno-litsevoi oblasti na territorii Volgograda i Volgogradskoi oblasti* [Clinical and epidemiological characteristics of congenital malformations of the maxillofacial area in the territory of Volgograd and Volgograd region. PhD dis.] Volgograd, 2000. 18 p.

7. Mamadkulov Iu.O., Shvetsov A.N. Model' mul'tiagentnogo sbora i obrabotki informatsii o sostoianii okruzhaiushchei sredy [Model of multi-agent collection and processing of information on the state of the environment]. *Informatizatsiia protsessov formirovaniia otkrytykh sistem na osnove SUBD, SAPR, ASNI i sistem iskusstvennogo intellekta: materialy 5-i Mezhdunarodnoi nauchno-tekhnicheskoi konferentsii "INFOS-2009"* [Informatization of the processes of formation of open systems based on DBMS, CAD, ASNI and artificial intelligence systems: proceedings of the 5th international scientific and technical conference INFOS-2009]. Vologda: Vologda Technical State University, 2009, pp. 166–169.

8. Onishchenko G.G. Problemy kachestva pit'evoi vody v Rossiiskoi Federatsii i puti ikh resheniia [Problems of potable water quality in the Russian Federation and ways to solve them]. *Vodosnabzhenie i sanitarnaia tekhnika* [Water supply and sanitary technique], 2010, no. 12, pp. 5–8.

9. Pavlov Iu.V., Koroleva E.N. Prostranstvennye vzaimodeistviia: otsenka na osnove global'nogo i lokal'nogo indeksov Morana [Spatial interactions: evaluation with the help of global

- and local Moran's index]. *Prostranstvennaia ekonomika* [Spatial economics], 2014, no. 3, pp. 95–110. DOI 10.14530/se.2014.3.95-110
10. Rapakov G.G., Lebedeva E.A., Gorbunov V.A., Kuz'minov A.L., Abdalov K.A. Geoinformatsionnoe kartografirovaniye na osnove mashinnogo modelirovaniia variogrammy [Geoinformation mapping based on computer-based semivariogram modeling]. *Vestnik Cherepovetskogo gosudarstvennogo universiteta* [Cherepoves State University Bulletin], 2018, no. 3 (84), pp. 46–54. DOI 10.23859/1994-0637-2018-3-84-6
11. Rapakov G.G., Lebedeva E.A., Gorbunov V.A., Kuz'minov A.L., Abdalov K.A. Komp'yuternoe modelirovaniye i analiz metodov setochnoi interpolatsii pri obrabotke geoinformatsionnykh dannykh [Computer modeling and analysis of methods of grid interpolation in the processing of geoinformation data]. *Vestnik Cherepovetskogo gosudarstvennogo universiteta* [Cherepoves State University Bulletin], 2018, no. 2 (83), pp. 31–39. DOI 10.23859/1994-0637-2018-1-83-4
12. Filimonov S.N., Semenova E.A., Panev N.I., Korotenko O.Iu., Filimonov E.S., Ershov E.V., Nikiforov V.V. Faktory riska ishemicheskoi bolezni serdtsa u rabochikh s khronicheskoi fluoridnoi intoksikatsiei [Risk factors for coronary heart disease in workers with chronic fluoride intoxication]. *Vestnik Kuzbasskogo nauchnogo tsentra* [Bulletin of the Kuzbass scientific centre], 2014, no. 19, pp. 116–118.
13. Shvetsov A.N., Sukonshchikov A.A., Kochkin D.V., Andrianov I.A. *Situatsionnye intellektual'nye sistemy podderzhki priniatiia reshenii* [Situational intelligent decision support systems]. Kursk: University Book, 2018. 251 p.
14. Anselin L. Local Indicators of Spatial Association – LISA. *Geographical Analysis*, 1995. vol. 27 (2), pp. 93–115. DOI 10.1111/j.1538-4632.1995.tb00338.x
15. Griffith D. Spatial Autocorrelation and Spatial Filtering: Gaining Understanding Through Theory and Scientific Visualization. *Springer-Verlag Berlin Heidelberg*, 2003. 260 p. DOI 10.1007/978-3-540-24806-4
16. Haining R. *Spatial Data Analysis: Theory and Practice*. Cambridge University Press, 2003. 454 p. DOI 10.1017/CBO9780511754944
17. Hunsaker C., Goodchild M., Friedl M., Case T. *Spatial Uncertainty in Ecology: Implications for Remote Sensing and GIS Applications*. Springer-Verlag New York, Inc., 2001. 417 p. DOI: 10.1007/978-1-4613-0209-4
18. Tiefelsdorf M. *Modelling Spatial Processes: The Identification and Analysis of Spatial Relationships in Regression Residuals by Means of Moran's I*. Springer-Verlag Berlin Heidelberg, 2000. 167p. DOI 10.1007/BFb0048754

Для цитирования: Рапак Г.Г., Лебедева Е.А., Горбунов В.А., Абдалов К.А., Мельничук О.В. Анализ пространственных кластеров и идентификация выбросов с помощью геоинформационных технологий // Вестник Череповецкого государственного университета. 2018. №5(86). С. 25–35. DOI: 10.23859/1994-0637-2018-5-86-3

For citation: Rapakov G.G., Lebedeva E.A., Gorbunov V.A., Abdalov K.A., Mel'nychuk O.V. Spatial cluster analysis and outlier detection using geoinformation technology. *Bulletin of the Cherepovets State University*, 2018, no. 5 (86), pp. 25–35. DOI: 10.23859/1994-0637-2018-5-86-3