

ПРЯМЫЕ МЕТОДЫ ЦИФРОВОЙ ПОЧВЕННОЙ КАРТОГРАФИИ И ПРИКЛАДНЫЕ АСПЕКТЫ ЕЕ ИСПОЛЬЗОВАНИЯ

Киндеев А.Л.

*Белорусский государственный университет,
г. Минск, Республика Беларусь
AKindeev@tut.by*

Приводятся основные причины развития прямых методов в цифровой почвенной картографии (ЦПК) и дается краткая характеристика их методологических основ и развития на территории Республики Беларусь. Проводится сравнительный анализ автоматических и ручных способов проведения геостатистического анализа, который показывает погрешности автоматических методов вплоть до 50 %. Демонстрируется один из способов, предлагаемый геостатистикой для создания сбалансированной сети мониторинга, который может быть использован не только для наблюдения за почвенным покровом, но и за другими компонентами природной среды.

Ключевые слова: география почв, картографирование, геостатистика, сеть мониторинга.

Введение. Основные принципы почвенной картографии были заложены еще в конце 19 века В.В. Докучаевым, но их дальнейшее развитие в рамках отечественной и зарубежной географии почв пошло различными путями, продиктованными потребностями общества того времени. Факторное картографирование, основанное на отображении генетического типа почв по косвенным признакам (растительной, рельеф, климат и др.) получило активное развитие в СССР. Развитию данного направления способствовало наличие обширных территорий, которые необходимо было осваивать. Некоторые оппоненты такого метода, например Г.Ф. Нефедов выступал с критикой подобного подхода [1], утверждая, что прикладные задачи должны решаться с помощью крупного и детального картографирования, делая упор не на генетическую принадлежность почвы, а на ее конкретные свойства (кислотность, влажность, гумус и др.), другими словами – прямыми (математическими) методами.

Трудозатратность методов, декларируемых Нефедовым Г.Ф., не позволили им активно развиваться в СССР, но нашли свое место в западных школах географии почв [2, 3]. Для частных фермеров было необходимо знать детальное, вплоть до квадратных метров, распределение почвенных свойств, а малые территории хозяйств, наличие финансовых и технических ресурсов, позволяли оперативно получать необходимую информацию. В свою очередь наличие большего количества данных способствовало применению математико-статистических методов и теории вероятности для прогнозирования почвенных свойств.

Появление аппаратно-технических средств в 80-ых годах XX века позволили решать многомерные уравнения и количественно оценивать неоднородность почвенного покрова, что привело к появлению нового научного направления в западной географии почв – педометрики, науке изучающей математические и статистические методы, с помощью которых можно понять, как в почве распределяются разные свойства, как произошли почвы, какой у них генезис и как они распределены в пространстве [4].

К сожалению, приходится признать, что с прошлого века на территории Республики Беларусь работ в данной области до последних нескольких лет практически не проводилось. В рамках развития данного направления становятся актуальными исследования, направленные на методологические совершенствования основ педометрики, а также демонстрации некоторых практических задач, которые могут быть решены с помощью методов этого направления. Вышесказанное определяет основную цель и актуальность настоящего исследования

Материалы и методы исследований. Методической основой прямых методов, используемых педометрикой является представление почвенного покрова, как набора случайно распределенных свойств в пространстве, что основывается на теории вероятности и континуальности почвенного покрова.

Прямое (математическое) картографирование и геостатистика не противопоставляется классическому подходу, однако из теории хаоса известно, что если на одном иерархическом уровне присутствует упорядоченность и закономерность, то на другом они могут вовсе отсутствовать. Эксперименты, опыты и исследования в области педометрики, доказали, что при рассмотрении почвенных свойств при крупномасштабном и детальном картографировании системный подход перестает работать и, как следствие, на место классификационного почвоведения должна приходиться педометрика.

Об объединении этих двух подходов, базирующихся на противоположных философских началах, Красильниковым П.В. в 2019 г. начал подниматься вопрос о переходе к «новой географии почв», предлагая почвенно-географическому сообществу сформировать теоретическую основу для осмысленного развития познания пространственной структуры педосферы и начать синтез прямого (математического) и косвенного (факторного) картографирования почв [5].

Для этого отечественной школе географии почв, в первую очередь необходимо понимание углубленных методов пространственного статистико-математического анализа, используемых при прямом почвенном картографировании, которые объединила в себе геостатистика. При всем богатстве приемов, предлагаемых геостатистикой для анализа пространственной структуры почвенного покрова основным является вариография.

Вариограммный анализ является ключевым моментом геостатистического оценивания данных. Вариограмма описывает пространственную дисперсию распределения свойств в зависимости от расстояния и определяется следующей формулой (1):

$$\gamma(h) = \frac{1}{2N(h)} \sum [z(x_i) - z(x_i + h)]^2,$$

где z – варьирующая величина; $z(x_i) - z(x_i + h)$ – результаты измерений в точках x_i и $x_i + h$; $N(h)$ – количество пар точек, разделенных расстоянием h (которое обычно называется лагом).

Величина $\gamma(h)$ является средней мерой сходства значений признака в точках, разделенных расстоянием h . Таким образом, дисперсия для каждого h рассчитывается как половина среднего квадрата разности значений показателя, измеренных на расстоянии h друг от друга.

При этом подбор вариограмм, без серьезных навыков довольно трудоемкий и требует временных затрат, в связи с этим ряд исследователей придерживались мнения что автоматизация процесса гораздо эффективнее [6]. При этом одни из основоположников геостатистики в географии почв М.А. Oliver и R. Webster утверждали, что необдуманное «нажатие нескольких кнопок» может привести к ненадежным и даже ошибочным результатам» [7].

Технологический прогресс, машинное обучение с того времени сделали значительные шаги вперед, однако без специалистов в области ЦПК невозможно оценить качество полученных моделей, схем, картограмм, карт и т.д., кроме этого, одной из задач, стоящих перед геостатистикой и педометрикой является определения необходимого числа точек пробоотбора для мониторинга, что также требует экспертной оценки.

Для подтверждения необходимости использования методических аспектов геостатистики в прямом почвенном картографировании, а также демонстрации одного из возможных способ создания математически обоснованных сетей мониторинга почвенного покрова были заложен участок недалеко от южной границы г. Минска на сельскохозяйственных землях площадью 25,6 га, с которого было отобрано 66 образцов и получены данные о содержании фосфора. Весь последующий процесс обработки и анализа данных производился в ГИС-среде программном продукте ArcGIS ArcMap.

Результаты и их обсуждение. При проведении геостатистического анализа значения содержания P_2O_5 , прошли 4 основных этапа обработки – определение вида распределения данных, анализ глобальных трендов (направлений), подбор вариограмм и кросс-валидацию (проверка) итог чего является картограммы изучаемых свойств.

Так, было определено, что значение коэффициента асимметрии составила 1,17, при допустимой ошибке в $\pm 0,44$ говоря о правостороннем отклонении данных, а показатель эксцесса составил 3,20 при ошибке в $\pm 0,74$, что показывает наличие выбросов у ряда данных. Поэтому для уменьшения ошибок было выполнено преобразование данных.

При определении тренда был обнаружен тренд второго порядка, графический соответствующий параболе, что также было включено в подбор вариограммы. Лучше из которых стала экспоненциальная (рисунок 1).

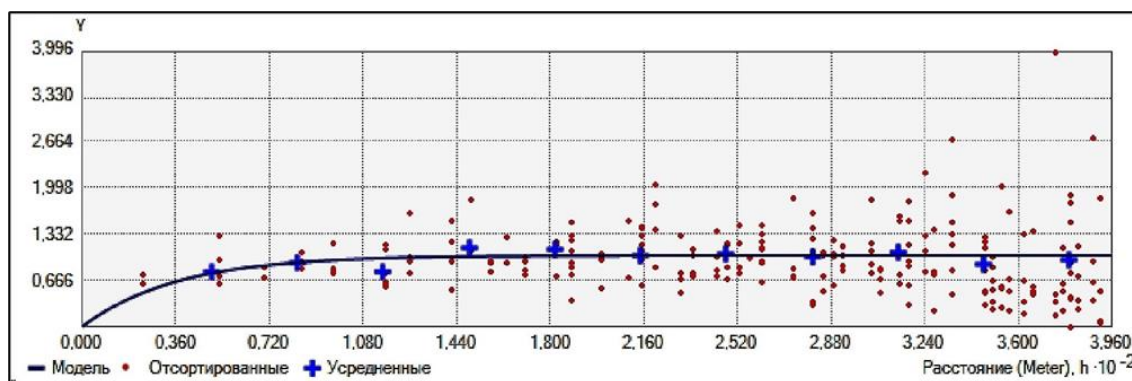


Рисунок 1 – Экспоненциальная вариограмма значений фосфора

Сопоставляя полученные эмпирические и экспериментальные вариограммы, можно констатировать их полное. Также видно отсутствие наггет-эффекта – вариограмма начинается от 0, что говорит об отсутствии «шумов» и наличии сильной пространственной зависимости переменных, а также высоком качестве картограммы, в основе которой будет данная модель (рисунки 2, 3). Также на рисунке представлены 2 картограммы (№ 1 и № 2), смоделированные с помощью автоматических инструментов – кригинг Spatial Analyst (№ 1) и стандартные настройки модуля Geostatistical Analyst (№ 2).

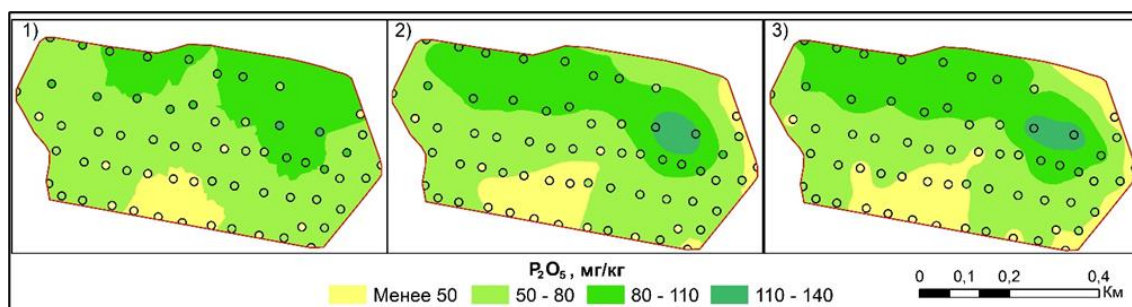


Рисунок 2 – Картограммы фосфора ключевого участка построенные:

1 – автоматическим кригингом (Spatial Analyst); 2 – автоматическим кригингом (Geostatistical Analyst); 3 – кригингом, при соблюдении методики с использованием экспоненциальной вариограммы

Кроме визуального сравнение, достоверным критерием качества картограмм является сопоставления статистических параметров интерполированных значений со значениями полученными лабораторным способом в местах пробоотбора представленных в таблице.

Таблица – Статистические параметры и ошибки интерполяции P_2O_5

Параметр	P_2O_5 , мг/кг	AKSA, мг/кг	Ошибка интерпо- ляции	AKGS, мг/кг	Ошибка интерпо- ляции	КМ, мг/кг	Ошибка интерпо- ляции
Минимум	20,0	43,4	-82,0	39,5	-38,0	20,0	-0,11
Среднее	66,6	66,9	-1,78	66,9	0,37	66,6	0,02
Медиана	65,2	65,6	-0,55	64,9	0,15	65,1	0,00
Максимум	152	100	56,2	115	23,3	152	1,24
Стандартная ошибка	3,03	1,67	2,32	2,15	1,27	3,02	0,02

AKSA – автоматический кригинг (Spatial Analyst); AKGS – автоматический кригинг (Geostatistical Analyst); КМ – кригинг, при соблюдении методики с использованием экспоненциальной вариограммы

Очевидным является значительные расхождения интерполируемых автоматическими кригингами значений. Использование кригинга Spatial Analyst (AKSA) наиболее нецелесообразно – в значительной степени происходит сглаживание ряда данных, игнорируются значения ниже 43,4 и выше 100 мг/кг, что составляет практически 50 %. Среднее значение ошибок равняется -1,78, при этом в ряде случаев инструмент занижал значения до -82,0 мг/кг и завышал на 56,2 мг/кг. В случае с автоматическим кригингом модуля Geostatistical Analyst (AKGS) ситуация

обстоит лучше, но также присутствует игнорирование минимальных и максимальных значений, но уже минимальные и максимальные ошибки в 2 раза стали меньше, среднее уменьшилось практически в 5 раз (0,37). Тем не менее, ручное проведение всего геостатистического анализа (КМ) позволило добиться наилучшего соответствия с измеренными значениями: среднее значение ошибок практически равно нулю (0,02), максимальное занижение составило всего 0,11, а завышение 1,24. В ходе интерполяции весь ряд данных был учтен в картограмме № 3, при этом сохранились значение среднего, медианы и стандартной ошибки, делая данную картограмму и способ интерполяции наилучшим.

Пространственный анализ по картограмме № 3 показал, что выделяются 3 зоны: 1) с очень низким содержанием P_2O_5 (менее 50 мг/кг) в южной части участка, 2) с низкими значениями (50-80 мг/кг) в центре и 3) средними (80-110 мг/га) на севере. Из общей тенденции выклинивается контур повышенных значений (110-140 мг/кг) в восточной части территории. Данный ареал не относится к неучтенной части варьирования, так как вариограммный анализ показал отсутствие нагетт-эффекта у фосфора и тем самым не является «артефактом».

Более подробная оценка по средствам стохастического моделирования позволила установить пространственное распределение среднеквадратических ошибок в каждой из ячеек растра указала на ряд ареалов требующих уточнений при дальнейшем мониторинге содержания фосфора (рисунок 3).

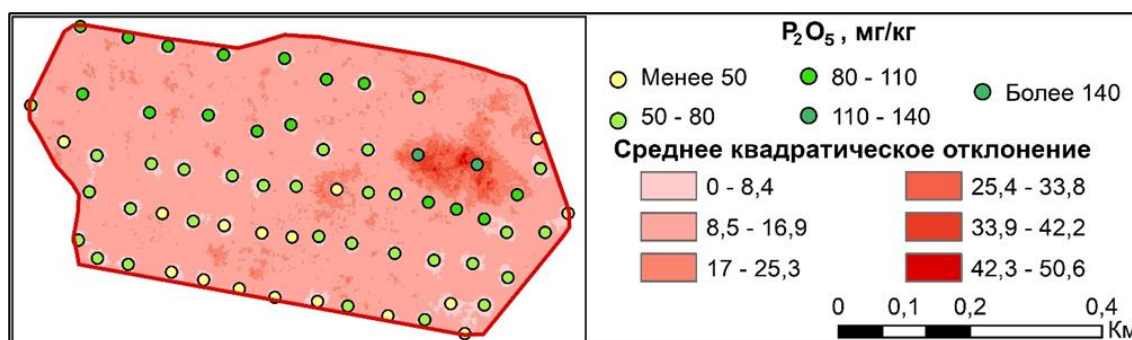


Рисунок 3 – Картограмма среднеквадратических ошибок значений содержания фосфора в почвах ключевого участка

К ним можно отнести территории возле точек с максимальными значениями, где ошибки составляют от 33,9 мг/кг до 50,6 мг/кг. В остальном же отклонения не превышает 16,9 мг/кг, что можно назвать удовлетворительным показателем, т.к. отклонения не превышает интервала диапазона градаций фосфора, установленных методикой агрохимического обследования Республики Беларусь.

Тем не менее для нивелирования территорий с большими ошибками предлагается разработать пространственно-сбалансированные точки, используя соответствующий инструмент в наборе ArcToolbox Geostatistical Analyst на основе картограммы среднеквадратического отклонения значения фосфора (рисунок 4).

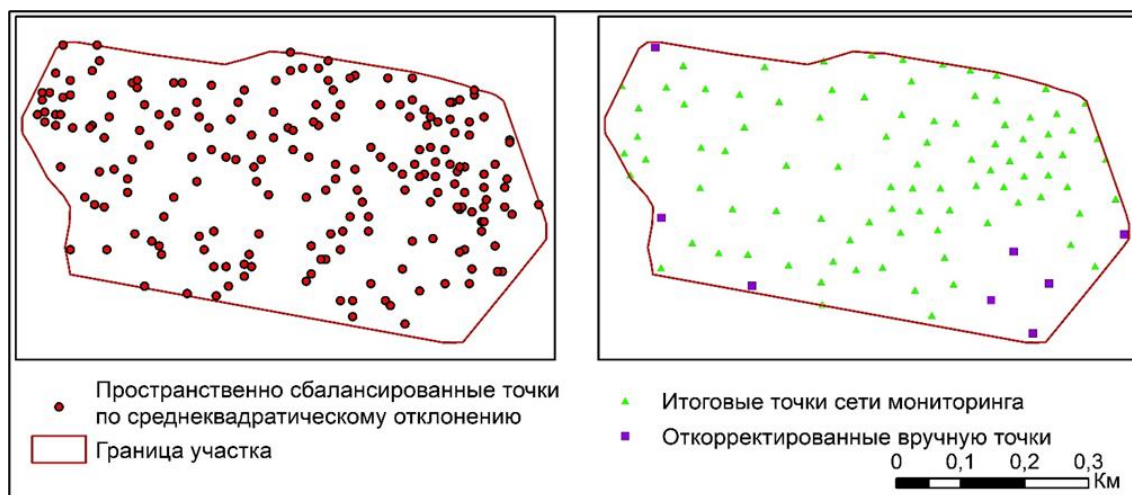


Рисунок 4 – Построение сети мониторинга ключевого участка

Поскольку генерация новой сети мониторинга происходит только с учетом весов местоположений, которыми выступают значения среднеквадратического отклонения, в дальнейшем необходимо будет откорректировать полученную сеть с учетом установленных ранее точек и расстояния между ними, что позволяет сделать инструмент «Уплотнить сеть выборки» в том же ArcToolbox Geostatistical Analyst.

При размере участка 25,2 га предлагается смоделировать сеть мониторинга, состоящую из 100 точек. Для этого было сгенерировано 200 точек (в 2 раза больше необходимого) при помощи первого инструмента с учетом среднеквадратического отклонения.

Как видно из рисунка, на первой картограмме точки приурочены к местам высоких значений среднеквадратического отклонения, однако при этом совершенно не учитывается расстояние между ними. С практической точки зрения это является нерациональным, а с точки зрения геостатистической методики может привести к появлению пространственных кластеров, которые будут нарушать гипотезу стационарности данных.

Для уточнения местонахождений пунктов пробоотбора используется второй упомянутый инструмент, с помощью которого мы генерируем 100 необходимых точек. Полученная итоговая сеть мониторинга учитывает такие параметры, как абсолютные значения показателя (картограмма фосфора), местонахождение исходных опорных точек, установленный в результате подбора вариограммы лаг (33 м), а также среднеквадратическое отклонение.

Новая сеть мониторинга сможет закрыть локальные места повышенной дисперсии значений P_2O_5 и не нарушит пространственной структуры ввиду сильной автокорреляционной взаимосвязи.

Заключение. Проведенное исследование отражает лишь один из многих аспектов и возможностей применения методических основ прямых методов картографирования и доказывает необходимость их изучения в контексте развития отечественной географии почв и почвоведения.

Доказано, что ошибки автоматических способов моделирования могут достигать 50 % от анализируемой выборки, тем самым показывая, что они не в состоянии построить репрезентативные и реалистичные картограммы без участия специалиста.

Использование представленного подхода для разработки сети мониторинга, базирующегося на математических моделях и внутренней пространственной взаимосвязи исследуемого свойства позволяет объективизировать подход к их разработке, как для почвенного покрова, так и для других видов многолетних наблюдений за компонентами окружающей среды.

Библиографические ссылки

1. Нефедов, Г.Ф. К методике почвенной картографии: Ответ на ст. А.П. Левицкого «В защиту докучаевского почвоведения» // Почвоведение. – 1907. – № 4. – 30 с.
2. Webster, R.A. The development of pedometrics // Geoderma. – 1994. – Vol. 62. – P. 3-18.
3. Heuvelink, G.B.M., Webster, R.A. Modelling soil variation: past, present, and future / G.B.M Heuvelink, R.A. Webster // Geoderma – 2001. – № 100. – P. 269-301.
4. McBratney, A.B. Introuction to Pedometrics. A cource of lectures / A.B. MsBratney. – South Australia: CSIRO Division of Soils Technical memorandum, 1986. – 150 p.
5. Красильников, П.В. На пути к «новой географии почв» вызовы и решения (обзор) / П.В. Красильников, В.О. Таргульян // Почвоведение. – 2019. – № 2. – С. 131-139.
6. Russo, D., Jury, W.A. A theoretical study of the estimation of the correlation scale in spatially variable fields: 1. Stationary fields // Water Resour. Res. – 1987. – Vol. 23, № 7. – P. 1257-1268.
7. Webster, R.A. Geostatistics for Environmental Scientist / R.A Webster, M.A. Oliver – John Wiley & Sons, Ltd., 2007. – 318 p.