

**НЕКОТОРЫЕ НАПРАВЛЕНИЯ ПРИМЕНЕНИЯ
СТАТИСТИЧЕСКИХ МЕТОДОВ В РАМКАХ
ИНФОРМАЦИОННОЙ СИСТЕМЫ «ПОЧВЕННО-
ГЕОГРАФИЧЕСКАЯ БАЗА ДАННЫХ РОССИИ»**

*Алябина Ирина Олеговна¹, Голозубов Олег Модестович¹,
Чернова Ольга Владимировна^{1,2}*

¹ Московский государственный университет им. М.В. Ломоносова,
факультет почвоведения, Москва, Россия

² Институт проблем экологии и эволюции им. А.Н. Северцова РАН,
Москва, Россия

E-mail: oleggolozubov@gmail.com

Аннотация. Приводится несколько примеров применения статистических методов расчетов в Информационной Системе «Почвенно-географическая база данных России». Часть примеров касается проектов ФАО ООН по расчету всемирных карт секвестрации и запасов органического углерода в почвах.

Ключевые слова: почвенные базы данных, статистические методы, распределенные системы, органический углерод.

Введение. Информационная система «Почвенно-географическая база данных России» – ИС ПГБД РФ (<https://soil-db.ru/>) насчитывает уже более 10 лет своего существования. Основанная на фундаментальных работах В.А. Рожкова по Почвенной информатике и по математической статистике в почвоведении Е.А. Дмитриева [1, 2], функционирующая в содружестве с научными и учебными организациями и центрами Агрехимической службы Минсельхоза РФ, ИС ПГБД РФ представляет собой сегодня распределенную сеть почвенных дата-центров, собрание методик, алгоритмов и программ, применяемых как в научно-исследовательской, так и в учебной и практической деятельности. Объекты ИС ПГБД – пространственно распределенные данные, базы данных и знаний, в том числе формализованные классификации как базы знаний в информационных системах [1], данные в качественных и количественных шкалах измерений, то есть практически все категории переменных [2].

Ниже приводится краткая характеристика наборов почвенных данных, аккумулированных в ИС ПГБД:

Профильные данные: 10 500 профилей, из них на территории России – 3 000, в том числе представительных профилей с полным

описанием – 900, и соответственно 22 000, 13 000 и 5 000 описаний горизонтов.

Мелкомасштабные почвенные карты и карты почвенно-экологического районирования – ПЭР (от М 1:1 000 000 и более мелко-масштабных): 36 000 контуров, из них на территории России 27 000;

Мелкомасштабные тематические карты: климатические, растительности, лесов, почвообразующих пород, природно-сельскохозяйственного районирования и другие.

Среднемасштабные почвенные карты: почвообразующих пород, эрозии (М: от 1:200 000 до 1:600 000): 30 000 контуров, из них на территории России 27 000.

Крупномасштабные почвенные карты: 287 000 контуров, из них на территории России 184 000 контуров, более 25 000 000 га.

ИС ПГБД также аккумулирует большой объем сопутствующей информации: данные агрохимических обследований, геоботанические и геоморфологические описания, карты землеустройства, севооборотов, эрозии и негативных факторов, справочники методов измерения, классификаторов и многое другое.

Одновременно с формированием и наполнением баз данных разрабатывались методики расчетов, анализа информации, генерации тематических карт и отчетов. Учитывая такие базовые свойства почвенной информации как разреженность, неполнота, историчность, а также существенные погрешности регистрируемых переменных, статистические методы находят широкое применение в алгоритмах и методах, применяемых в ИС ПГБД, три примера из которых приводится в настоящей работе.

1. Оценка связи распространения почвенных горизонтов с климатическими параметрами [3] была выполнена для равнинной территории России. Исходный массив данных подготовлен с использованием цифровых карт в ГИС MapInfo. Путем пересечения различных карт получено векторное покрытие на равнинную часть РФ, включающее 19729 полигонов. Атрибутивная информация, связанная с каждым полигоном: преобладающая почва (с Почвенной карты РСФСР М 1:2 500 000); горизонты, слагающие профиль; значения 15-ти климатических параметров.

Массив данных обработан с использованием метода ранговой корреляции. Для 12-ти почвенных горизонтов рассчитаны: **r** – выборочный коэффициент корреляции Спирмена; **Z** – статистика; **P** – уровень значимости. По величине Z-статистики получена оценка влияния климатических параметров на распространение почвенных горизонтов.



Рис. 1. Влияние климатических параметров на распространение почвенных горизонтов

2. В 2017 году ФАО ООН инициировал проект по созданию всемирной карты запасов органического углерода в 30-сантиметровом слое почвы (GSOC17) [4], была предложена единая методика расчетов, для реализации которой, наряду с прочими характеристиками, были необходимы сведения о редко определяемой экспериментально объемной массе почвенных горизонтов (bulk density). Кроме того, требовалось составить карту оценки погрешности расчетов. В обоих случаях для расчетов применяли статистические методы.

Для расчета объемной массы минеральных горизонтов использовали пятипараметрическую функцию нелинейной регрессии, которая позволяет прогнозировать объемную массу в зависимости от содержания гумуса и глубины горизонта [5]. На основе анализа массива данных, аккумулированных в ПГБД РФ (<https://soil-db.ru/map/profiles>), было показано, что для 3 крупных групп генетически сходных почв «Таежных», «Луговых» и «Степных» плотности, вычисленные с использованием указанного уравнения, хорошо совпадают с экспериментальными значениями [6]. Верификацию уравнения и коэффициентов проводили на заведомо независимых информационных массивах: для почв группы «Таежные» – на основе данных по 125 горизонтам из 31 разреза дерново-подзолистых почв (средняя относительная ошибка составила 7.5%); для почв группы «Степные» – по 307 горизонтам из 111 разрезов черноземов (относительная ошибка – 7.6%) (рис. 2).

Анализ большого объема разновременной пространственно-атрибутивной информации, аккумулированной в цифровой форме в объединенной сети почвенных дата-центров, позволил проследить динамику содержания и запасов гумуса в пахотных черноземах в пределах двух провинций ПЭР на территории Ростовской области и

наметить подходы к снижению их дегумификации с учетом специфических характеристик территорий [7].

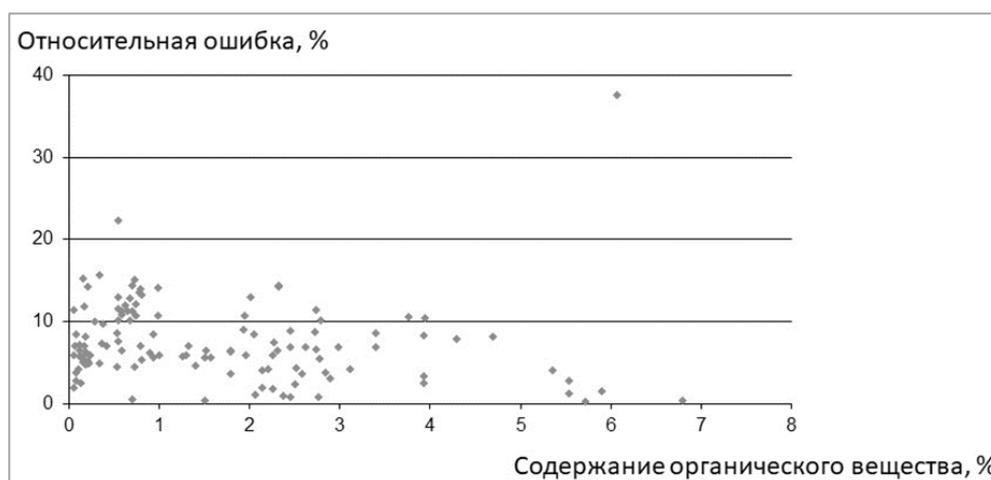


Рис. 2. Зависимость относительных ошибок расчета объемной массы дерново-подзолистых почв от их обогащенности органическим веществом

3. В 2020 году в связи с проблемами изменения климата ФАО предложило сформировать мировую карту секвестрации органического углерода SOCSeq на основе единой методики, в которой одним из параметров расчета является процентное содержание частиц <0.002 мм в слое почвы 0-30 см [8]. Для получения этой информации использован следующий расчетный алгоритм:

- данные представительных профилей, содержащие минимально необходимый набор показателей (гранулометрический состав и содержание гумуса), усреднялись по принадлежности к одинаковым группам почв (рис. 3)

- с помощью регрессионной формулы, предложенной в [9], рассчитывали процентное содержание частиц <0.002 на основе данных о гранулометрическом составе в градациях, принятых в отечественном почвоведении:

$$\Phi_{<0.002} = -1.148 + 0.43 \Phi_{<0.005} + 0.53 \Phi_{<0.001}.$$

В случае отсутствия данных для фракции <0.005, часто встречающегося в практике обследований, использовали другой вариант регрессионной формулы:

$$\Phi_{<0.002} = 2.556 + 2.11 \Phi_{<0.001}.$$

- далее рассчитывали средневзвешенные значения показателей для слоя 0-30 см по формуле (на примере 3 горизонтов):

$$C1 * M1 + C2 * M2 + C3 * (30 - G13) / 30,$$

где $C1$, $C2$, $C3$ – содержание фракции $\Phi_{<0.002}$ в горизонтах 1, 2 и 3; $M1$, $M2$ – мощность горизонтов 1 и 2, $G13$ – верхняя граница горизонта 3.

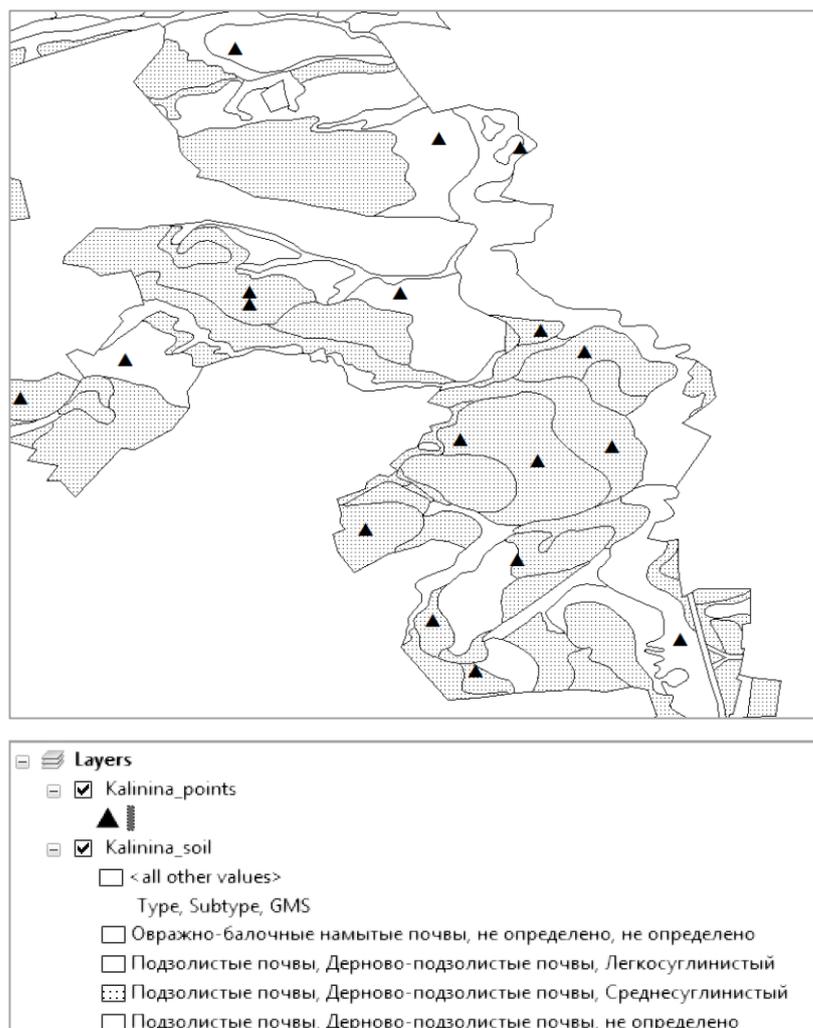


Рис. 3. Фрагмент почвенной карты с разрезами

В работе приведены примеры статистических методов, использованных в проектах, реализованных на базе ИС ПГБД. Для решения научных и практических задач применяются различные технологии отбраковки, фильтрации и гармонизация данных, сочетание эмпирических и статистических методов обобщения информации, имитационные модели и др. Возможности распределенной БД по накоплению, верификации, использованию данных, поддержке их совместимости, многопользовательский режим, а также поддержка программных языков манипуляции пространственными данными и языка R для пространственно-статистического анализа являются основными преимуществами в исследованиях на больших территориях.

Благодарности

Работа выполнена в рамках НИР «Почвенные информационные системы и оптимизация использования почвенных ресурсов» ЦИТИС: 121040800147-0.

Список литературы

1. Рожков В.А. Об информационном подходе в классификации почв // Бюл. Почв. ин-та им. В.В. Докучаева. 2012. Вып. 69. С. 4-23.
2. Дмитриев Е.А. Математическая статистика в почвоведении: учебник / Научн. ред. Ю.Н. Благовещенский. Изд. 4-е, доп. М.: Книжный дом «ЛИБРОКОМ», 2010. 336с.
3. Алябина И.О., Неданчук И.М. Оценка связи распространения почвенных горизонтов с климатическими параметрами // Почвоведение. 2014. № 10. С. 1165-1176.
4. FAO. 2020. A protocol for measurement, monitoring, reporting and verification of soil organic carbon in agricultural landscapes – GSOC-MRV Protocol. Rome. <https://doi.org/10.4060/cb0509en>
5. Честных О.В., Замолодчиков Д.Г. Зависимость плотности почвенных горизонтов от глубины их залегания и содержания гумуса // Почвоведение. 2004. № 8. С. 937-944.
6. Чернова О.В., Голозубов О.М., Алябина И. О., Щепашенко Д. Г. Комплексный подход к картографической оценке запасов органического углерода в почвах России // Почвоведение. 2021. № 3. С. 273-286. DOI:10.31857/S0032180X21030047.
7. Чернова О. В., Алябина И. О., Безуглова О. С., Литвинов Ю. А. Современное состояние гумусированности пахотных черноземов настоящих степей (на примере Ростовской области) // Юг России: экология, развитие. 2020. Т. 15. N 4. С. 99-113. <https://doi.org/10.18470/1992-1098-2020-4-99-113>
8. FAO. 2020. Technical specifications and country guidelines for Global Soil Organic Carbon Sequestration Potential Map (GSOCseq). Rome.
9. Теории и методы физики почв / ред. Е.В. Шеин, Л.О. Карпачевский. М.: Гриф и К, 2007. 616 с.

SOME AREAS OF APPLICATION OF STATISTICAL METHODS WITHIN THE FRAMEWORK OF THE INFORMATION SYSTEM "SOIL-GEOGRAPHICAL DATABASE OF RUSSIA"

Alyabina Irina¹, Golozybov Oleg¹, Chernova Olga^{1,2}

¹ Lomonosov Moscow State University, Soil Science Faculty, Moscow, Russia

² Severtsov Institute of Ecology and Evolution RAS, Moscow, Russia

E-mail: oleggolozubov@gmail.com

Abstract. Several examples of the application of statistical methods of calculations in the Information System "Soil-geographical database of Russia" are given. Some examples relate to FAO projects on the calculation of world maps of organic carbon stocks and sequestration of organic carbon.

Keywords: soil databases, statistical methods, distributed systems, organic carbon.