

## ГЕНЕЗИС И ГЕОГРАФИЯ ПОЧВ

УДК 631.4

### КОМБИНАТОРНЫЙ АНАЛИЗ И КЛАССИФИКАЦИЯ ПОЧВ

© 2019 г. Н. П. Кириллова<sup>1, \*</sup>, Т. М. Силева<sup>1</sup>, М. И. Макаров<sup>1</sup>, З. С. Артемьева<sup>2</sup>, Е. К. Бурова<sup>1</sup>

<sup>1</sup>МГУ им. М.В. Ломоносова, Россия, 119991, Москва, Ленинские горы, 1

<sup>2</sup>Почвенный институт им. В.В. Докучаева, Россия, 119017, Москва, Пыжевский пер., 7, стр. 2

\*e-mail: npkirilova@yandex.ru

Поступила в редакцию 13.02.2018 г.

После доработки 04.05.2018 г.

Принята к публикации 26.09.2018 г.

На экспериментальном материале о строении 1029 профилей почв, распространенных на территории Учебно-опытного почвенно-экологического центра МГУ им. М.В. Ломоносова “Чашниково” (Московская область), создана классификация-перечисление на уровне групп горизонтов в соответствии с теоретической концепцией генератора (машины) классификаций почв по В.А. Рожкову и концепцией пространства логических возможностей по Г.А. Заварзину. С помощью комбинаторного анализа набора признаков, заданного распределением 17 диагностических горизонтов по восьми группам, создано универсальное пространство возможностей, которое характеризует 255 возможных комбинаций (сочетаний) групп горизонтов. Пространство логических возможностей насчитывает 168 комбинаций, а в реальном пространстве количество комбинаций уменьшилось до 46, то есть в 3.6 и 5.5 раз по сравнению с логическим и универсальным пространствами соответственно. Соотношение между сочетаниями групп горизонтов и подтипами почв по Классификации и диагностике почв СССР, 1977 характеризуется следующим образом: одна комбинация не принадлежит к двум разным подтипам, но в одном подтипе может быть несколько комбинаций. Это позволило связать классификацию-перечисление и классификацию почв и обеспечить переход из одного классификационного поля в другое.

**Ключевые слова:** пространство логических возможностей, базы данных, систематика, цифровая картография почв, классификация-перечисление

**DOI:** 10.1134/S0032180X19020072

#### ВВЕДЕНИЕ

Развитие методов цифровой почвенной картографии, особенно для построения карт таксонов, привело к возрастанию интереса к почвенным классификациям. Алгоритмы создания всех цифровых почвенных карт предусматривают наличие этапа поиска сходства между точками опробования и точками интерполяции [11, 18, 19]. Он может быть основан на принципах ландшафтно-индикационных связей [8, 20] или на геостатистических расчетах [17]. В результате расчетов каждая ячейка раstra принимает значение того номера (индекса) точки опробования, с которой найдено максимальное сходство. Например, при 30 точках опробования для раstra размерностью 200 столбцов × 100 строк количество ячеек будет равно 20000, и значение каждой ячейки может быть в пределах от 1 до 30 в соответствии с количеством номеров (индексов) точек опробования. В любом варианте на следующей стадии необходимо перейти от индекса точки опробования к таксономическим индексам.

Таксономический индекс — это числовой индекс, однозначно характеризующий название таксона в справочной таблице, где собраны все данные по таксону данного уровня для использованной классификации. На основе Классификации и диагностики почв СССР (КиД-77) [10] создано 6 справочных таблиц на следующих таксономических уровнях: тип, подтип, род, вид, разновидность, разряд [8]. Например, таблица типов включает 73 записи — по количеству типов, входящих в классификацию, для их обозначения используется 73 числовых индекса.

Каждая точка опробования характеризуется шестью таксономическими индексами, поэтому можно перейти от индекса точек опробования к таксономическим индексам. Примыкающие друг к другу точки с одинаковым таксономическим индексом формируют области распространения данного таксона.

От того насколько правильно определены таксономические индексы точек опробования в значительной мере зависит точность всей карты. Когда число точек опробования невелико, ввод

таксономических индексов в базу данных не представляет сложности. При наличии схемы распределения профилей по почвенным таксонам [3, 4] этот процесс легко осуществляется вручную, выбором соответствующих позиций из таблиц. С увеличением количества точек опробования процедура усложняется: для нее требуется больше времени, а также возрастает число случаев, когда в классификации нет примера профиля точки опробования по тому или иному таксону.

Проблема конвертации совокупности горизонтов в классы (таксоны) переходит на новый уровень, когда карта таксонов строится только с помощью интерполяционных методов расчета мощности горизонтов [17]. Это будет требовать увеличения списка строк в справочной таблице, добавляемых вручную, чтобы обеспечить автоматическую идентификацию типа почв для всего массива входных данных. Смена методов интерполяции приведет к смене возможных комбинаций горизонтов (профилей), и процесс модификации справочных таблиц придется проводить заново.

Поясним это на конкретном примере. В табл. 1 приведены предсказанные значения мощностей гор. А2, рассчитанные различными методами интерполяции в SAGA GIS на одном и том же множестве точек опробования для одной и той же территории, состоящей из 10820 ячеек раstra. Совпадение предсказаний наличия горизонта для всех трех интерполяций наблюдается только в 51% случаев, для двух интерполяций – еще 45% и для одной – 4%. Таким образом, совокупность комбинаций горизонтов в зависимости от использованных методов интерполяции будет меняться.

В более общем виде проблему можно сформулировать следующим образом: существует ли способ описать совокупность комбинаций горизонтов в пределах любой классификации таким образом, чтобы полученное множество было исчерпывающим и не требовало расширения (смены числа) таксонов в классификации?

С точки зрения математики ответ на этот вопрос является утвердительным, поскольку из комбинаторного анализа известно, что совокупность (сумма) всех сочетаний для  $n$ -элементов по  $k = \{1, 2, \dots, n\}$  (набор  $k$  элементов, выбранных из данного множества, содержащего  $n$  различных элементов) составляет  $2^n - 1$ , то есть это множество является конечным, и все его элементы можно перечислить. Наборы, отличающиеся только порядком следования элементов (но не составом), считаются одинаковыми [16].

Этот подход реализуется в предложенной В.А. Рожковым теоретической концепции генератора (машины) классификаций почв [15]. Цель классификации определяется через фиксированную систему выбранных почвенных по-

**Таблица 1.** Распределение предсказанных значений наличия–отсутствия гор. А2 (1 – наличие (мощность слоя >0), 0 – отсутствие (мощность слоя = 0)) в зависимости от метода интерполяции: NatN – Natural Neighbour; NearN – Nearest Neighbour; IDW – Inverse Distance Weighted

NatN	NearN	IDW	$n$	%
0	0	1	430	4
1	0	1	4909	45
1	1	1	5481	51

казателей (признаков), организованных в виде исчерпывающей (comprehensive) классификации-перечисления. Число объектов в классификации определяется числом выбранных признаков [15]. Концепция В.А. Рожкова базируется на идее классификации-перечисления Ю.А. Воронина, которая была предложена для применения в геологии [1].

На необходимость расширения мысленного пространства признаков указывал А.А. Роде в 1947 г. В работе [14] он подчеркивал, что попытки описать разнообразие почв в природе “страдают существенным недостатком – мы черпаем материал из существующих, наблюдаемых нами объектов. Между тем, несомненно, что возможные пути эволюции почв” (а, следовательно, и почвенные профили) “гораздо разнообразнее, но остаются для нас неизвестными. Приходится при этом признать, что такая ограниченность нашего мышления при данном состоянии науки является неизбежной” [14, с. 131].

Для территории подзоны южной тайги (УОПЭЦ “Чашниково” в Солнечногорском районе Московской области) показано [6], что максимальная совокупность идентифицируемых горизонтов зональных и интразональных почв составляет 17 (табл. 2). Для оценки размера классификации-перечисления, где в качестве признаков (элементов) используются почвенные горизонты, надо подсчитать число всех комбинаций (сочетаний) для  $n = 17$  горизонтов, когда элементы (горизонты) группируются по 1, 2, 3... и так далее до 17 ( $k = \{1, 2, \dots, 17\}$ ). Эта величина будет составлять  $2^{17} - 1 = 131071$ . Она характеризует все возможные комбинации из заданного набора признаков и по определению Г.А. Заварзина относится к универсальному пространству возможностей (УВП). Часть этого пространства будет соответствовать реально существующим объектам, а часть – только теоретически возможным. Бессмысленные комбинации должны быть исключены как логически несовместимые. Это позволит уменьшить универсальное пространство до пространства логических возможностей (ЛВП) [2].

В нашем примере к категории запрещенных почвенных комбинаций относятся все комбина-

**Таблица 2.** Список горизонтов в буквенно-цифровом обозначении (для изученной территории)

Горизонт, подгоризонт, Hor	Индекс группы горизонтов, idHorGrp	Группа горизонтов, HorGrp	Термины, включенные в название подтипа	Название
H	2	H	Перегноино-	Перегноинный
AT	3	T1_	Торфянисто-	Торфянистый
T1	3	T1_	Торфяно-, торфяные	Торфяный
A1	4	A1	Дерново-, дерновые	Гумусово-аккумулятивный
Ap	1	Ap	Освоенные дерново-	Пахотный
A1A2	5	A2_	Подзолистые, оподзоленные	Элювиально-гумусово-аккумулятивный
A2	5	A2_	Подзолистые, оподзоленные	Элювиальный (подзолистый)
A2B	5	A2_	Подзолистые, оподзоленные	Элювиально-иллювиальный (оподзоленный иллювиальный)
AB	6	A_		Переходный от гумусово-аккумулятивного к срединному B
B	7	B		Срединный: иллювиальный или переходный (в аллювиальных почвах)
AC	6	A_		Переходный от гумусово-аккумулятивного к породе
BC	0			Переходный к породе
T	0			Торф как порода
ALG	8	G_	Глеевые	Оглеенный аллювий
AL	0			Неоглеенный аллювий
C	0			Порода, затронутая почвообразованием
G	8	G_	Оглеенные, глееватые глеевые	Глеевый

ции, куда входят сочетания гор. AC или AB с любым горизонтом из группы A1A2, A2, A2B, поскольку наличие AC/AB исключает присутствие какого-либо горизонта между A1 и C или B. Но даже после такого сокращения трудно ожидать, что пространство логических возможностей комбинаций 17 горизонтов будет соответствовать по порядку величины количеству подтипов почв, которое составляет для этой территории всего 18 [6]. Очевидно, что КиД-77 оперирует не отдельными горизонтами, а их группами. Если мы хотим найти классификацию-перечисление, соответствующую по размеру КиД-77, то надо понять алгоритм группировки горизонтов на уровне подтипа и построить по найденному алгоритму классификацию-перечисление, которая будет охватывать все возможные комбинации пространства логических возможностей. Это станет основанием для

создания исчерпывающего списка строк в справочных таблицах и обеспечит полную автоматизацию перехода от профиля точки к ее таксономической характеристике.

Цель работы – на основании экспериментальных данных о профилях почв, расположенных в подзоне южной тайги на территории УОПЭЦ “Чашниково”, изучить соответствие между их комбинаторикой и группами горизонтов, составляющих подтипы по КиД-77, оценить возможности создания классификации-перечисления на этом таксономическом уровне.

## ОБЪЕКТЫ И МЕТОДЫ

База данных почвенных профилей в виде набора горизонтов собрана в период 2012–2017 гг. в рамках летней полевой практики по картографии

почв студентов факультета почвоведения МГУ им. М.В. Ломоносова на территории УОПЭЦ “Чашниково”, расположенной в Солнечногорском районе Московской области.

Характеристика территории площадью 339 га подробно дана при составлении карт почвенных подтипов [6] и почвообразующих пород [7]. Таксономическую диагностику почв проводили в соответствии методикой [4]. Составляли список горизонтов в буквенно-цифровом обозначении (табл. 2). Список основан на КиД-77, но в тех случаях, когда в наименование горизонта входили нижние индексы в буквенном обозначении в кириллице, их заменили на индексы в латинице, чтобы избежать наличия двух кодировок в таблице. Это относится к пахотному (А пах – А<sub>р</sub>), перегнойному (А пер – Н) и торфянистому (А<sub>т</sub> – А<sub>Т</sub>) горизонтам. Вместо СD для аллювия использовали обозначение АL, чтобы из названия сразу было видно, что это аллювиальный горизонт.

Обработку данных (Data processing) проводили в модуле “Программа комбинаторного анализа почвенных горизонтов” [5]. Входными данными являлись два набора: совокупность разрезов с указанием их профилей в буквенно-цифровом обозначении (экспериментальные данные полевых исследований) и справочные данные, основанные на КиД-77. Справочник горизонтов позволяет проводить их группировки по сходству свойств. На основании выбранных группировок по принципам комбинаторики строили таблицу всех возможных сочетаний для групп горизонтов. Для определения подтипа находят соответствие между индексами группы горизонтов и их количеством в профиле. При полном совпадении этих параметров в справочной и экспериментальной таблице профилю приписывали индекс подтипа почв.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

**Создание пространства универсальных возможностей групп горизонтов.** Для определения уровня, на котором можно сгруппировать горизонты с целью создания классификации-перечисления использовали два принципа:

1. В одну группу объединяли горизонты, которые имели однокоренные названия подтипов. Например, группа с индексом 3 (idHorGrp = 3, HorGrp = T1) включала гор. АТ и Т1 (торфяно-сто-, торфяно-, торфяные) а группа с индексом 5 (idHorGrp = 5, HorGrp = А2\_) – гор. А1А2, А2, А2В (подзолистые, оподзоленные).

2. Горизонты объединяли по положению в профиле, если их присутствие никак не обозначено в названии подтипа. Так, гор. АВ, АС вошли в группу А\_, которая характеризует переходные от гумусово-аккумулятивного к срединному В или к породе. Эти горизонты также связаны логически:

их присутствие исключает наличие горизонтов группы подзолистых (HorGrp = А2\_). Наличие срединного гор. В также не отражено в названии подтипа, но он является диагностическим. Например, в типе аллювиальных луговых кислых почв присутствие гор. В различает аллювиальные луговые кислые слоистые (А1–АLГ) и собственно аллювиальные луговые кислые почвы (А1–В–АLГ).

Проведенное объединение позволяет выделить 8 групп горизонтов (ГГ) с индексами от 1 до 8. Группа с индексом 0 (горизонты переходные к почвообразующей породе или почвообразующая порода) из создания комбинаторной таблицы была исключена: ее наличие логически связано с присутствием группы 8: если группа G (G, ALG) отсутствует, то присутствует группа 0 (C, T, AL).

Из 8 элементов мы можем создать  $N_u = 2^8 - 1 = 255$  сочетаний, в которых группы горизонтов будут комбинированы по 1, 2, ... 8. Это число ( $N_u$ ) определяет наше УВП.

**Создание пространства логических возможностей групп горизонтов.** Уменьшение УВП до ЛВП происходит за счет удаления логически несовместимых комбинаций [2]. Число элементов в ЛВП мы обозначим как  $N_{logic}$ . Исключаются 15 комбинаций, в которых отсутствуют верхние горизонты, имеющие idHorGrp < 5: А2\_ ; А\_ ; В ; G\_ ; А\_ А2\_ ; А2\_ В ; А\_ В ; G\_ В ; А2\_ G\_ ; А\_ G\_ ; А\_ А2\_ В ; А2\_ G\_ В ; А\_ А2\_ G\_ ; А\_ G\_ В ; А\_ А2\_ G\_ В. Число составит  $N_{logic} = 255 - 15 = 240$ .

В том случае, когда почвы сильно эродированы или нарушены естественным образом за счет роящей деятельности животных или криогенной турбации и др., верхние горизонты могут отсутствовать, но в настоящей работе мы не рассматриваем такие профили.

Исключаются 60 комбинаций, в которых одновременно присутствуют группы А2\_ и А\_ :  $N_{logic} = 240 - 60 = 180$ .

Исключаются 12 комбинаций, в которых А\_ присутствует без А1 или А<sub>р</sub>:  $N_{logic} = 180 - 12 = 168$ .

Таким образом, ЛВП в 1.5 раза меньше, чем УВП.

Полученные 168 комбинаций ГГ составляют справочную таблицу (Look up table, Lut), с индексами Lut.IdCombHorGrp = {1, 2, ..., 168}. Одному индексу комбинации (Lut. IdCombHorGrp) соответствует набор из ГГ (idHorGrp).

Структура Lut и пример записей представлены в табл. 3. Индекс комбинации ГГ меняется от 1 до 168. В примере даны комбинации с индексами 46 и 47. В каждую комбинацию входят три группы горизонтов. Индекс комбинации ГГ с номером 46 соответствует комбинации групп 2,6,9. Номер 47 соответствует комбинации групп 5, 6, 9.

Общее число строк в справочной таблице составило 644. Это число получается, если умножить значение из столбца 1 табл. 4 (число

**Таблица 3.** Справочные данные (Lut) пространства логических возможностей (фрагмент)

IdCombHorGrp	idHorGrp	CountHorGrp
46	2	3
46	6	3
46	9	3
47	5	3
47	6	3
47	9	3

Примечание. IdCombHorGrp – индекс комбинации ГГ; idHorGrp – индекс ГГ (см. табл. 2); CountHorGrp – число ГГ в комбинации.

**Таблица 4.** Распределение данных в пространствах универсальных возможностей, логических возможностей и реальном в зависимости от количества групп горизонтов в комбинации

Количество групп горизонтов	УВП		ЛВП		РП	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
1	8	3.1	4	2.4	4	8.7
2	28	11.0	20	11.9	13	28.3
3	56	22.0	43	25.6	19	41.3
4	70	27.4	51	30.4	9	19.6
5	56	22.0	35	20.8	1	2.2
6	28	11.0	13	7.7	0	0
7	8	3.1	2	1.2	0	0
8	1	0.4	0	0	0	0
Сумма	255	100	168	100	46	100

Примечание. УВП – пространство универсальных возможностей, ЛВП – пространство логических возможностей, РП – реальное пространство, *n* – число комбинаций.

групп ГГ) на *n* для ЛВП и получить сумму произведений:  $1 \times 4 + 2 \times 20 + 3 \times 43 + 4 \times 51 + 5 \times 35 + 6 \times 13 + 7 \times 2 = 644$ .

Представляется, что использование ЛВП целесообразно, когда мы хотим сопоставить реально существующие объекты в природе с логически возможными. Оно создает фильтр, который убирает комбинаторный шум. В том случае, когда комбинации горизонтов создаются методами интерполяции, то есть вычисляются математическими методами, использование УВП (по Заварзину) или полной классификации-перечисления (по Рожкову) будет предпочтительнее.

**Реальное пространство групп горизонтов.** Реальное пространство (РП) основано на экспериментальных данных, которые включали информацию о 1029 разрезах (точках опробования). Описание каждого разреза включает от 1 до 6 горизонтов. В результате для исследуемой территории РП включает суммарно 3807 записей отдельных гори-

зонтов. По правилам, изложенным выше (табл. 2), каждый горизонт конкретного профиля был отнесен к одной из восьми ГГ. Таким образом, формула каждого профиля была преобразована в последовательность индексов ГГ. Поскольку в некоторых профилях почв несколько горизонтов относились к одной ГГ (например, A1A2–A2–A2B преобразовывались в A2\_), общее число записей индексов ГГ в РП составило 2764. Структура таблицы с экспериментальными данными (DataT) и пример записей для одной точки опробования представлены в табл. 5.

Путем запроса к базе данных можно для каждой точки опробования из пространства РП получить значение уникального индекса комбинации ГГ, который определен в справочной таблице ЛВП. Рассмотрим данные из табл. 3 и 5. Для разр. 1 экспериментальным путем установлено 3 индекса ГГ: 5, 6, 9 (табл. 5). В справочной таблице (табл. 3) указано два индекса комбинации ГГ с номером 46 и 47, в обоих случаях в комбинацию входят три ГГ. Номер 46 соответствует комбинации групп 2, 6, 9. Номер 47 соответствует комбинации групп 5, 6, 9.

Для первого варианта сумма совпадающих индексов с точками опробования равна двум (2, 6, 9 и 5, 6, 9), во втором варианте сумма совпадающих индексов равна трем (5, 6, 9 и 5, 6, 9), и число индексов равно трем, следовательно, этому разрезу соответствует Lut.IdCombHorGrp = 47.

Таким образом, каждый разрез можно охарактеризовать по количеству ГГ и подсчитать, сколько разрезов с числом ГГ от 1 до 8 присутствует в экспериментальных данных. Результаты показывают, что распределение разрезов по количеству ГГ в РП существенно отличается от УВП и ЛВП. Такое распределение в УВП симметрично, с максимальным процентом сочетаний, имеющих 4 и 5 ГГ. ЛВП перестает быть симметричным и максимум смещается к сочетаниям 4 ГГ. В РП асимметрия вырастает еще больше, максимум смещается к разрезам, имеющим 3 ГГ в профиле, а варианты с количеством ГГ более 5 отсутствуют. Количество реализованных комбинаций в РП составило 46 (табл. 4). Общее соотношение между количеством сочетаний в трех пространствах составило УВП : ЛВП : РП = 1 : 0.66 : 0.18. Таким образом, реальное пространство в 5.5 раз меньше универсального и в 3.6 раз меньше логического.

**Комбинаторика групп горизонтов и подтипы почв в классификации.** Все почвенные профили, как описанные в классификации, так входящие в нашу базу данных, были диагностированы по ключам [4] до уровня подтипа по КИД-77. Это дало возможность определить соотношения между сочетаниями групп горизонтов и подтипом почв. Пойменные почвы рассматриваются отдельно от почв вне поймы, поскольку одна и та же комби-

нация групп горизонтов может принадлежать к разным подтипам почв в пойме и не в пойме. В пределах выделенных территорий отношения между комбинацией ГГ и индексом подтипа характеризуются следующим образом: одна комбинация не принадлежит к двум разным подтипам, но в одном подтипе может быть несколько комбинаций. Из 46 комбинаций, реализованных на изученной территории, 23 комбинации общие для пойменных и не пойменных почв, 11 – уникальны для почв вне поймы и принадлежат к 5 подтипам по КиД-77. Еще 12 уникальных комбинаций принадлежат к 5 подтипам аллювиальных типов почв (табл. 6).

Таким образом, с помощью всего 46 комбинаций можно осуществить однозначный переход к 17 подтипам почв. Для сравнения отметим, что в настоящее время количество ключей в программе “Автоматизированный определитель почв”, основанных на комбинации горизонтов, достигает 235 [3]. Очевиден выигрыш в эффективности и простоте применения классификации-перечисления, основанной на комбинаторике групп горизонтов, а не горизонтов.

Другим следствием является то, что мы могли бы заранее провести процедуру отнесения к имеющимся подтипам в КиД-77 всех оставшихся незадействованными комбинаций ЛВП ( $168 - 46 = 122$ ). Рассмотрим примеры комбинаций по 2, 3, 4, 5, 6 и 7 горизонтов, входящих в ЛВП, но не вошедших в рабочее пространство (табл. 7).

Из этого списка комбинация с 4 ГГ с возможным профилем Н–А2(А2В)–В–G подходит под описание подтипа перегнойно-подзолистых грунтово-оглеенных почв из типа болотно-подзолистых. Для комбинаций с количеством ГГ 6 и 7 возможно не существует в природе комплекса факторов, обеспечивающих формирование такого профиля, и эта позиция в РП не будет задействована. Поскольку мы не можем заранее сказать, какие комбинации никогда не будут реализованы, то ЛВП избыточно по отношению к РП. Рожков [15] считает, что избыточность в виде вхождения в классификацию несуществующих в природе сочетаний не мешает организации и анализу данных, но обеспечивает сбалансированность классификации.

Таким образом, часть комбинаций войдет в имеющиеся подтипы, а часть останется не классифицированной. Для нее нужно будет расширить существующую классификацию.

Конкретный пример необходимости такого расширения связан с диагностикой на изученной территории дерново-глеевых почв, которые все имеют грунтовое оглеение. По КиД-77 их можно отнести к двум подтипам: дерново-грунтово-глеевые и перегнойные грунтово-глеевые. Самый распространенный профиль имеет вид А1g–G.

**Таблица 5.** Экспериментальные данные (DataT) реального пространства (фрагмент)

N_TO	idHorGrp	CountHorGrp
1	5	3
1	6	3
1	9	3

Примечание. Обозначение полей таблицы: N\_TO – номер точки опробования; idHorGrp – индекс ГГ (см. табл. 2); CountHorGrp – число ГГ в комбинации.

Единственный подтип, с помощью которого можно описать этот профиль по КиД-77 – это дерново-грунтово-глеевые почвы. Из этого примера следует, что в классификации пропущен подтип дерново-грунтово-глеевые, поэтому мы вынуждены пользоваться наиболее близким подтипом – дерново-грунтово-глеевые.

Проблема встраивания в существующую классификацию новых почвенных образований, отличных по строению и свойствам от установленных, поднималась ранее [12]. Теперь этот вопрос может быть решен: после добавления новых вхождений в классификацию никакого дальнейшего ее расширения не понадобится. На этом этапе использование ЛВП предпочтительнее по сравнению с УВП, потому что позволяет существенно сократить процедуру соотношения классификации-перечисления и КиД-77, за счет отказа от определения логически бессмысленных комбинаций.

Переход, а точнее связывание двух классификаций, это пример того, как, по выражению Рожкова [15], в классификационное поле можно подвесить любую другую классификацию, лишь бы ее объекты (их предметы) содержали показатели, принятые в классификации-перечислении. Это действительно является путем “для решения проблемы смены со временем систем классификации или легенды почвенной карты без существенных исследований корреляций между их вариантами” [15, с. 124].

Часть объектов старой классификации может оказаться за пределами нового классификационного поля. В нашем случае это те подтипы, которые включают в себя малые индексы, например дерново-подзолистые поверхностно-оглеенные почвы из типа болотно-подзолистых почв, которые характеризуются наличием оглеения в подзолистом горизонте (А2g).

Выход из этой ситуации заключается в том, чтобы сознательно исключить такие вхождения в классификацию-перечисление, а информацию о наличии оглеения перенести на более низкий таксономический уровень. Этот подход реализован в Классификации и диагностике почв России 2004 г. [9] и Полевом определителе почв России 2008 г. [13], где на уровне типа присутствуют толь-

Таблица 6. Распределение комбинаций групп горизонтов по подтипам почв

Не пойма		Пойма	
сочетания групп горизонтов	подтип	сочетания групп горизонтов	подтип
A1 B C	дп	A1 B AL	садк
A1 A2_ B C	дп	A1 A2_ B AL	адко
A1 A_ B C	дп	A1 A_ B AL	садк
A1 A2_ C	дп	A1 A2_ AL	адко
Ap A1 A2_ B C	одп		
Ap A2_ C	одп		
Ap A_ C	одп		
Ap B C	одп		
Ap A1 A2_ C	одп		
Ap A2_ B C	одп	Ap A2_ B AL	адко
Ap A_ B C	одп		
T1_ A2_ B G_	тпго		
A1 T1_ A2_ G_	тпго		
A1 A2_ G_	дпго	A1 A2_ G_	адко
A1 A2_ B G_	дпго	A1 A2_ B G_	адко
A1 G_	дгг	A1 G_	алкс
A1 B G_	дгг	A1 B G_	салк
A1 A_ G_	дгг		
A1 A_ B G_	дгг	A1 A_ B G_	салк
A1 T1_ B G_	пгг	A1 T1_ B G_	салб
H B G_	пгг	H B G_	абпг
A1 T1_ G_	пгг	A1 T1_ G_	салб
A1 H G_	пгг	A1 H G_	салб
A1 H B G_	пгг		
A1 H T1_ G_	пгг	A1 H T1_ G_	салб
H G_	бнтг	H G_	абпг
T1_ G_	бнтг	T1_ G_	баитг
T1_ B G_	бнтг	T1_ B G_	баитг
H T1_ G_	бнтг	H T1_ G_	баитг
H T1_ T3	бнт	H T1_ T3	баит
T1_ T3	бнт	T1_ T3	баит
H T3	бнт	H T3	баит
		Ap A1	адкс
		Ap A2_ G_	адко
		Ap A2_ B G_	адко
		Ap H A2_ B G_	адко
		Ap G_	алкс
		Ap A1 G_	алкс
		Ap B G_	салк
		A1 H B G_	салб
		Ap T1_ G_	салб
		A1 H G_	салб
		A1 T1_ G_	салб
		Ap H G_	салб
A1 C	сд	A1 AL	адкс
A1 A_ C	сд		

Обозначения подтипов почв по КиД-77: дп – дерново-подзолистые почвы, одп – освоенные дерново-подзолистые почвы, тпго – торфянисто-подзолистые грунтово-оглеенные, дпго – дерново-грунтово-глееватые, пгг – перегнойные грунтово-глеевые, бнтг – болотные низинные (типичные) торфяно-глеевые почвы, бнт – болотные низинные (типичные) торфяные почвы, сд – собственно дерновые почвы, адкс – аллювиальные дерновые кислые слоистые почвы, садк – собственно аллювиальные дерновые кислые почвы, адко – аллювиальные дерновые кислые оподзоленные почвы, алкс – аллювиальные луговые кислые слоистые почвы, салк – собственно аллювиальные луговые кислые почвы, салб – собственно аллювиальные лугово-болотные почвы, абпг – аллювиальные болотные перегнойно-глеевые почвы, баитг – болотные аллювиальные иловато-торфяно-глеевые почвы, баит – болотные аллювиальные иловато-торфяные почвы.

**Таблица 7.** Примеры комбинаций групп горизонтов пространства логических возможностей, не вошедшие в реальное пространство

Число ГГ в комбинации	Комбинация ГГ	Возможный профиль
2	H A2_	H-A2(A2B)-C
3	Ap A1 T1_	Ap-T1-A1-AL
3	H A2_ B	H-A2(A2B)-B-C
4	H A2_ B G_	H-A2(A2B)-B-G
5	A1 H T1_ A2_ G_	
6	A1 H T1_ A2_ B G_	
7	Ap A1 H T1_ A2_ B G_	

ко диагностические (генетические) горизонты, а на уровне подтипа – генетические признаки.

### ЗАКЛЮЧЕНИЕ

На экспериментальном материале о строении профилей почв подзоны южной тайги, распространенных на территории УОПЭЦ “Чашниково”, создана классификация-перечисление на уровне групп горизонтов, в соответствии с теоретической концепцией генератора (машины) классификаций почв по Рожкову [15]. Охарактеризовано пространство универсальных возможностей, которое включает все возможные комбинации (255) из набора признаков, заданного распределением 17 диагностических горизонтов по восьми группам. В результате анализа проведено исключение бессмысленных комбинаций, как логически несовместимых, и формирование пространства логических возможностей (168 элементов).

В реальном пространстве, основанном на экспериментальных данных, количество элементов (комбинаций групп горизонтов) составило 46, что в 3.6 раз меньше, чем в логическом и в 5.5 раз меньше, чем в универсальном пространстве.

Соотношение между сочетаниями групп горизонтов и подтипами почв характеризуется следующим образом: одна комбинация не принадлежит к двум разным подтипам, но в одном подтипе может быть несколько комбинаций. Это позволило связать две классификации: классификацию-перечисление и КиД-77, и обеспечить переход из одного классификационного поля в другое.

Нам представляется, что данный подход чрезвычайно перспективен в тех областях, где предусматривается автоматизация диагностирования больших массивов данных.

### СПИСОК ЛИТЕРАТУРЫ

1. *Воронин Ю.А.* Геология и математика. Новосибирск: Наука, 1967. 224 с.

2. *Заварзин Г.А.* Фенотипическая систематика бактерий. Пространство логических возможностей. М.: Наука, 1974. С. 1–142.

3. *Кириллова Н.П.* Автоматизированный определитель почв. Св-во о гос. регистрации программы для ЭВМ № 2014617393. 18 июля 2014 г.

4. *Кириллова Н.П., Силева Т.М.* Таксономическая диагностика почв с помощью автоматизированного определителя // Вестник Моск. ун-та. Сер. 17, почвоведение. 2015. № 3. С. 31–36.

5. *Кириллова Н.П., Силева Т.М., Макаров М.И.* Программа комбинаторного анализа почвенных горизонтов. Св-во о гос. регистрации программы для ЭВМ № 2018614417. 05 апреля 2018 г.

6. *Кириллова Н.П., Силева Т.М., Ульянова Т.Ю., Розов С.Ю., Ильашенко М.А., Макаров М.И.* Цифровая почвенная карта УОПЭЦ Чашниково МГУ им. М.В. Ломоносова // Вестник Моск. ун-та. Сер. 17, почвоведение. 2015. № 2. С. 22–29.

7. *Кириллова Н.П., Силева Т.М., Ульянова Т.Ю., Розов С.Ю., Смирнова И.Е.* Цифровая крупномасштабная карта почвообразующих пород и принципы ее составления // Вестник Моск. ун-та. Сер. 17, почвоведение. 2017. № 3. С. 3–10.

8. *Кириллова Н.П., Силева Т.М., Ульянова Т.Ю., Савин И.Ю.* Метод “совпадений” и его применение для построения цифровой крупномасштабной почвенной карты // Почвоведение. 2014. № 10. С. 1193–1203. doi 10.7868/S0032180X14080073

9. Классификация и диагностика почв России. Смоленск: Ойкумена, 2004. 342 с.

10. Классификация и диагностика почв СССР. М.: Колос, 1977. 224 с.

11. *Козлов Д.Н., Сорокина Н.П.* Традиции и инновации в крупномасштабной почвенной картографии // Цифровая почвенная картография: теоретические и экспериментальные исследования. М.: Почв. ин-т им. В.В. Докучаева. 2012. С. 35–57.

12. *Лебедева И.И., Герасимова М.И.* Факторы почвообразования в классификациях почв // Почвоведение. 2009. № 12. С. 1515–1520.

13. Полевой определитель почв России. М.: Почв. ин-т им. В.В. Докучаева, 2008. 182 с.

14. *Роде А.А.* Почвообразовательный процесс и эволюция почв. М.: ОГИЗ, 1947. С. 131.

15. *Рожков В.А.* Концепция генератора (машины) классификаций почв // Бюл. Почв. ин-та им. В.В. Докучаева. 2016. Вып. 85. С. 115–129. doi 10.19047/0136-1694-2016-85-115-130

16. *Стенли Р.* Перечислительная комбинаторика. М.: Мир, 1990. 440 с.

17. *Хитров Н.Б.* Создание детальных почвенных карт на основе интерполяции данных о свойствах почв // Почвоведение. 2012. № 10. С. 1045–1056.

18. *Grunwald S.* Multi-criteria characterization of recent digital soil mapping and modeling approaches // Geoderma. 2009. № 152. P. 195–207.

19. *McBratney A.B., Mendonca-Santos M.L., Minasny B.* On digital soil mapping // Geoderma. 2003. V. 117. № 1–2. P. 3–52.

20. *Zhu A.X., Hudson B., Burt J.E., Lubich K.* Soil mapping using GIS, expert knowledge, and fuzzy logic // Soil Sci. Soc. Am. J. 2001. V. 65. P. 1463–1472.



## Combinatorics and Soil Classification

N. P. Kirillova<sup>a,\*</sup>, T. M. Sileva<sup>a</sup>, M. I. Makarov<sup>a</sup>, Z. S. Artem'eva<sup>b</sup>, and E. K. Burova<sup>a</sup>

<sup>a</sup>Lomonosov Moscow State University, Leninskie Gory 1, Moscow, 119991 Russia

<sup>b</sup>Dokuchaev Soil Science Institute, per. Pyzhevskii 7, Moscow, 119017 Russia

\*e-mail: npkirillova@yandex.ru

Experimental data on 1029 soil profiles examined in the area of the Chashnikovo Experimental and Training Soil-Ecological Center of Moscow State University in Moscow region were used to create an enumerative classification at the level of the groups of soil horizons according to the theoretical concept of a machine generator of soil classifications suggested by V.A. Rozhkov and the concept of logical possibilities suggested by G.A. Zavarzin. Combinatorics of the sets of features of 17 diagnostic horizons separated into eight groups made it possible to characterize 255 possible combinations of the groups of horizons. In turn, the space of logical possibilities consisted of 168 combinations. In the real space, where there were 46 combinations, i.e., their number was 3.6 and 5.5 times smaller in comparison with the logical and universal spaces of combinations, respectively. The relationship between the separated groups of combinations of soil horizons and the subtypes of soils identified in the *Classification and Diagnostics of Soils of the Soviet Union* (1977) was as follows: any given combination does not belong to two different subtypes, but a given subtype may include several combinations. This made it possible to link the suggested enumerative classification and the classification system of 1977 to ensure mutual transition between these classification fields.

*Keywords:* space of logical possibilities, data bases, systematics, digital soil cartography, enumerative classification