

УДК 641.3

КЛАССИФИКАЦИЯ ПОЧВ – НЕ МЕСТО ДЛЯ ДИСКУССИЙ*

© 2013 г. В. А. Рожков

*Почвенный институт им. В.В. Докучаева Россельхозакадемии,
119017, Москва, Пыжевский пер., 7*

Из-за недостаточных темпов внедрения теории и математических методов классификации почв делается попытка наглядной иллюстрации имеющихся подходов. Известный афоризм в названии использован не для того, чтобы развенчать «аристократизм» значимой проблемы и перманентно обсуждаемой в нашей науке, но чтобы «в картинках» показать, как современные средства информатики позволяют решать разнообразные задачи построения, исследования и использования классификационных структур на описаниях почвенных объектов. Если не само построение, то исследование готовой классификации должно осуществляется представленным здесь образом, чтобы считать ее завершенной и открытой для обсуждения, использования и развития.

Ключевые слова: классификация, численная классификация, дендрограмма, метод главных компонент, информативность признаков, качество классификаций, классифицирование, распознавание почв.

ВВЕДЕНИЕ

Дискуссии вокруг любых классификаций почв возникают, главным образом, из-за того, что каждая из них является произведением коллектива или отдельного ученого, и отражает, прежде всего, их собственный опыт и мнение. Но всегда есть другие мнения, порой несовместимые с первыми. Противоречия возникают еще в разрезе и будут продолжаться от представлений об общей структуре классификации до границ интервалов значений каждого почвенного показателя.

*Работа выполнена при финансовой поддержке РФФИ, грант № 11-04-01123а.

Возможно, сказывается дефект обучения. В учебниках от классификации требуется учет сразу нескольких несовместимых аспектов: происхождения, строения, состава и плодородия почв, направление эволюции, «роль общественного производства», «пути к открытию новых типов почв»; она должна ориентировать практику в методах рационального использования почв в земледелии, мелиорации, лесоводстве, инженерном деле и др. Здесь смешаны одновременно неосуществимые цели, которые не могут задаваться через совокупность несовместимых почвенных показателей.

Над авторами довлеет собственный опыт, и новая по замыслу конструкция всегда несет в себе уже существующие представления о классификации (Дунаев, Поляков, 1987). Не напрасно в инженерии знаний для экспертных систем не привлекают глубоко эрудированных исследователей – в формировании декларативного или процедурного знания им мешают тонкости и детали.

В качестве примера идеальной естественной классификации приводят периодическую систему элементов Д.Н. Менделеева. В отличие от классификации почв от нее не требуют отражения ни генезиса, ни производственной значимости элементов, ни даже свойств, хотя все они предполагаются уже на основании других представлений и опыта. Поэтому она служит скорее образцом корректной постановки проблемы, чем универсальной классификации.

Классификациям важно признание, поэтому берутся за них наиболее авторитетные ученые или коллективы. Но в настоящей статье невольно понижается уровень «аристократизма» их работы, превращая ее в рутинную процедуру обработки данных. Речь пойдет о вычислениях структур, обладающих всеми свойствами, необходимыми для полноценной классификации. Это безусловная воспроизводимость результата разными авторами, обеспеченная выбранным алгоритмом обработки, геометрическая визуализация результата вычислений и простота распознавания новых объектов. В статье, избегая сложных формул и терминов, на рисунках показана процедура построения классификации. Можно утверждать, что любая авторская классификация, если не строится изначально, то непременно должна пройти подобное исследование, чтобы стать «прозрачной» для понимания, предметного обсуждения и использования.

НЕКОТОРЫЕ ТЕРМИНЫ И ПОНЯТИЯ

Термин «классификация» имеет три значения: законченную систему, которая может быть представлена некоторой структурой; процесс ее создания и распознавание новых *почвенных объектов*.

Речь пойдет о нахождении классов эквивалентности, т.е. таких классов (групп, кластеров), которые не пересекаются в *пространстве признаков*, т.е. не имеют общих объектов. Не обсуждаются формы и характер границ классов, хотя это имеет свое значение. Многомерное *пространство признаков* образуется в воображаемых координатах, оси которого составляют почвенные показатели.

Объект классификации – описание почв, профилей, горизонтов, образцов, наконец, любой другой *массив* почвенных объектов, представленный значениями его признаков. Почвенные свойства, состав, морфометрические показатели, условия почвообразования – все показатели, которые составляют описание почвы и называются *признаками*. Их значения могут быть качественными (ранжируемыми или неранжируемыми) или количественными. На самом деле требуется более детальное деление *шкалы* их значений, поскольку они определяют допустимые арифметические операции, а значит и методы статистической обработки¹ (Рожков, 2011). *Массив* по определению включает данные одной шкалы, и компьютерные программы, как правило, именно на это и рассчитаны. Пренебрежение шкалами чревато ошибками и заблуждениями при интерпретации полученных расчетов, поскольку кодовые обозначения качественных признаков, порядковые номера ранжируемых и даже отдельные количественные (шкалы отношений, разностей) показатели имеют свои ограничения в методах обработки. Коммерческие пакеты программ (SSPS, STAISTICA) не полностью учитывают это обстоятельство.

Многомерные статистические методы и кластер-анализ – математические методы обработки наблюдений за объектами,

¹ Например, нельзя считать средний бонитет (порядковая шкала), абсурден «средний цвет» (номинальная шкала), некорректна статистика кодовых обозначений.

описанными многими признаками. Они позволяют выявлять отдельные скопления объектов (классы, кластеры, группы) в пространстве признаков, оценивать их компактность и существенность отличий от других классов.

Визуализация – графическое или другое представление результатов обработки данных, делающее их наглядным или очевидным.

ВАРИАНТЫ РЕШЕНИЙ ПРОБЛЕМЫ

Как писал В.Г. Зольников (1955), разработка принципов создания классификации более важна, чем сама классификация, т.к. по этим принципам можно построить новые классификации. Действительно, по накопленным данным математическими методами может быть построена классификация или, точнее, результат (структура), обладающий всеми свойствами, необходимыми для классификации как таковой. К таким свойствам относятся: оценка информативности признаков и качества полученной группировки объектов, просто распознавание новых почв. Эти вопросы решаются сложными методами, однако достаточно длительный опыт их решения позволяет показать ход рассуждений и результаты решений на рисунках.

Для простоты первый пример заимствуем из недавней публикации (Рожков, 2011).

Описание почвенного профиля подзолистой почвы включило по три образца из горизонтов: $A_{\text{пах}}$, A_2 , A_2B и B , описанных семью признаками: 1) рН, 2) содержание гумуса (С, %), 3) гидролитическая кислотность (ГК), 4) содержание ила (Ил) и 5) глины (Гл), 6) вынос ила (Ви) и 7) вынос $Ca + Mg$ (CaMg), %. Данные были обработаны методом главных компонент с тем, чтобы извлечь имеющуюся в них информацию о структуре совокупности представленных образцов и ее характеристике (рис. 1).

Образцы из разных горизонтов в поле главных компонент разделились довольно четко. Особенно это касается горизонтов A_1 (образцы 1–3) и A_2 (4–6). Вполне естественно, что различимыми, но близкими оказались и образцы горизонтов B_1 (7–9) и A_2B (10–12).

Образцы горизонтов в пространстве главных компонент:

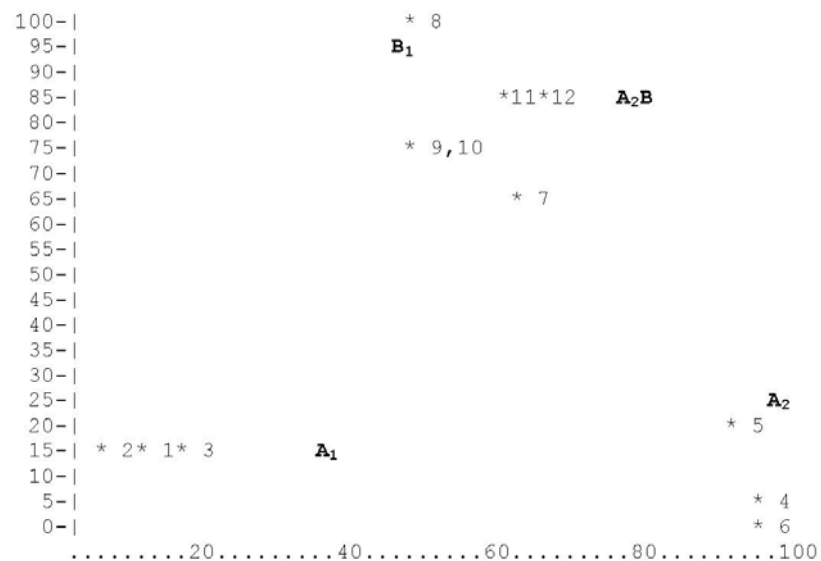


Рис. 1. Ординатная структура совокупности 12-ти образцов.

Критерием качества разделения групп образцов служит дисперсия главных компонент (распечатка расчетов не приводится). Первая главная компонента описывает 67% варьирования признаков в многомерной выборке (распечатка не приводится). Опыт показывает, что если первые две компоненты описывают около 70% варьирования (здесь 85%), разделение объектов на группы будет достаточно выраженным, как это видно в данном примере.

Собственные векторы корреляционной матрицы показывают информационную нагрузку почвенных признаков в полученном разделении образцов (ГК-1 и ГК-2 – главные компоненты):

	pH	C	ГК	Ил	Гл	Ви	СаМг
ГК-1	-0,34	-0,24	0,26	0,45	0,43	0,45	0,40
ГК-2	-0,09	-0,73	-0,67	-0,01	-0,04	0,02	-0,04

Наибольшие значения в первом собственном векторе соответствуют четырем последним признакам почвы, т.е.

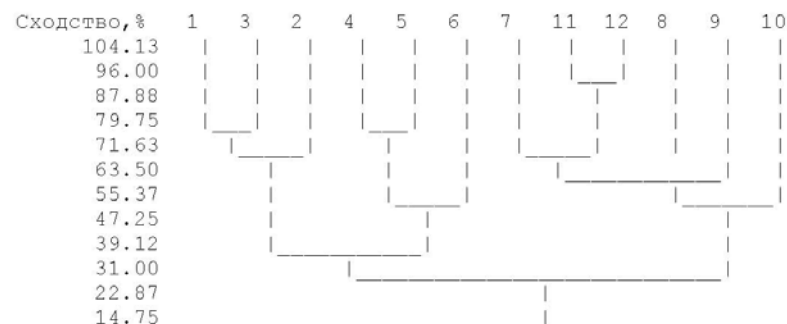


Рис. 2. Иерархическая структура совокупности образцов.

гранулометрическому составу горизонтов. Следовательно, именно они определяют разброс горизонтов по оси ординат (проекции горизонтов на первую главную компоненту). Наибольший вес по второй компоненте (0.73 и 0.67) имеют содержание гумуса и гидролитическая кислотность, они характеризуют разброс горизонтов по оси абсцисс.

Для дальнейшего анализа можно оставить только наиболее весомые показатели, что неоднократно проверено во многих задачах, хотя при этом происходит определенная потеря информации.

Более наглядно разделение видно на дендрограмме этих объектов (рис. 2). Представленная на рис. 2 группировка горизонтов по сходству аналогична таковой в координатах главных компонент. Агломеративная группировка объединяет все представленные образцы в единую систему. Системообразующим фактором выступает принятая мера сходства образцов, а эмерджентным свойством является иерархическая структура с количественными показателями отношений между образцами и их группами.

Дендрограмма является одним из наиболее наглядных средств визуализации результатов классификации. Рис. 1, 2 демонстрируют ординатное и иерархическое представление классификации. Причем метод главных компонент применим для арифметических (количественных) данных, а дендрограммы могут вычисляться для признаков в любой шкале.

На рис. 3 приведена дендрограмма семи выше названных почвенных признаков, отражающая сходство между ними и их группами.

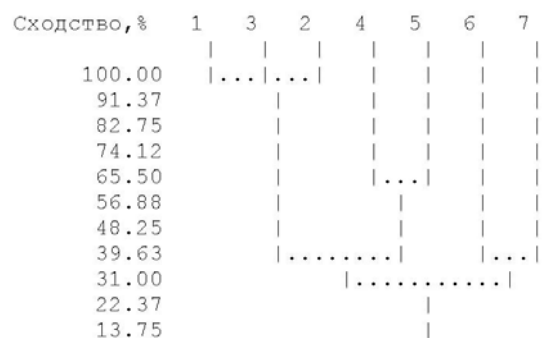


Рис. 3. Дендрограмма семи почвенных показателей.

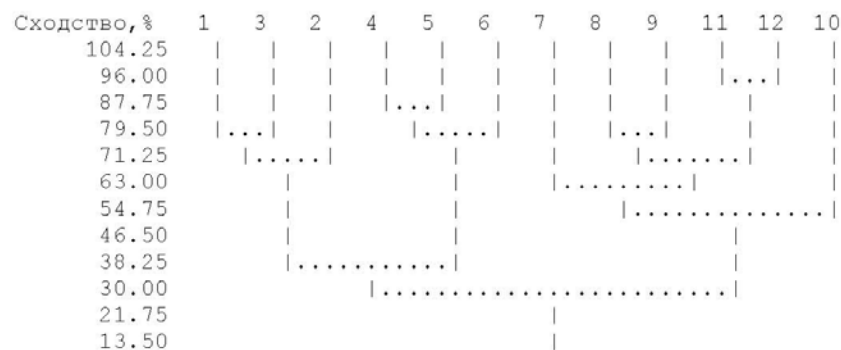


Рис. 4. Дендрограмма образцов по шести признакам (без рН).

Как оказалось, рН, гумус и гидролитическая кислотность проявляются на особенностях образцов одинаково – их сходство в этом проявлении равно 100%. Это означает, что какие-то из них здесь лишние, т.е. не несут информации о разделении образцов, а, следовательно, могут быть исключены. Решение подсказывает ранее рассмотренный результат компонентного анализа. По обоим главным компонентам наименее информативным оказался рН. После его исключения из массива получается новая дендрограмма образцов (рис. 4).

Разделение образцов получилось в некотором смысле даже лучше, чем по семи признакам – более четко разделились образцы

из горизонтов А₂В и В₁. Таким образом, избавляемся от мешающего признака и сокращаем число анализов в дальнейшей работе.

Не загружая статью иллюстрациями, отметим, что дальнейшее сокращение каких-либо признаков ухудшает картину разделения образцов по горизонтам.

В качестве другого объекта для демонстрации методов анализа данных выбрана предлагаемая ФАО система кодирования соотношений площадей почв в контуре (ФАО, 1979). Выделяется 18 вариантов соотношения площадей основной почвы, трех сопутствующих и четырех включенных (таблица). Намеренно для иллюстрации выбран отвлеченный материал, чтобы подчеркнуть возможности некоторых кластерных и информационных алгоритмов.

Система кодировки соотношений почв по ФАО

I	Основная	Сопутствующая			Включенная				
	1	1	2	3	1	2	3	4	
% площади почв в контуре									
	1	2	3	4	5	6	7	8	9
1	24	20	20	20	4	4	4	4	
2	25	20	20	20	5	5	5	0	
3	30	20	20	20	5	5	0	0	
4	30	20	20	0	10	10	10	0	
5	30	20	20	20	10	0	0	0	
6	40	20	20	0	10	10	0	0	
7	40	30	0	0	10	10	10	0	
8	40	20	20	0	5	5	5	5	
9	50	30	0	0	5	5	5	5	
10	50	30	0	0	10	10	0	0	
11	50	20	20	0	10	10	0	0	
12	60	20	20	0	0	0	0	0	
13	60	30	0	0	10	0	0	0	
14	70	30	0	0	0	0	0	0	
15	70	0	0	10	10	10	0	0	
16	80	0	0	0	10	10	0	0	
17	90	0	0	0	10	0	0	0	
18	100	0	0	0	0	0	0	0	

Абстрагируясь от описания структуры почвенного покрова, каждый из 18 описаний принимается за типовое описание конкретной почвы. На этом примере моделируется решение задачи создания классификации, хотя точнее здесь просто осуществляется визуализация структуры представленной совокупности почв в пространстве значений их площадей. Возможно, подобная задача имеет большее отношение к задаче *оптимального кодирования* из теории связи, но в данном случае этот вопрос уходит на второй план.

Рассуждения приводятся в терминах структуры почвенного покрова, однако понимать их нужно в более широком смысле. Это могут быть компоненты ландшафта, категории земель или составляющих их показателей и другие представленные образы на карте. Для общности все они заменены словом «объекты». На рис. 5 приведены результаты применения метода главных компонент.

Примем обозначения признаков: O1 – основная почва; C1, C2, C3 – сопутствующие; B1, B2, B3, B4 – соответственно включенные. Две первые главные компоненты (ГК) описывают всего 52% варьирования площадей, что означает невысокую компактность возможных классов почв. Вес признаков (значений долей площади) на главных компонентах

	O1	C1	C2	C3	B1	B2	B3г	B4
ГК-1	-0.40	-0.52	-0.15	0.29	-0.28	-0.43	-0.34	00.28
ГК-2	-0.06	0.27	0.27	0.05	-0.55	-0.38	0.00	-0.63

Наибольший вклад в варьирование площадей почв по ГК-1 (ось ординат) вносит первая сопутствующая почва, по ГК-2 варьирование связано с первой включенной почвой. Возможно, следует обратить внимание на таком подборе площадей, чтобы основное влияние на общее варьирование оказывала основная почва, т.е. требуется оптимизация описаний объектов. Однако в настоящей статье такая задача не ставилась.

Общая структура рассматриваемых объектов, в качестве которых выступают строки приведенной выше таблицы, наглядно визуализируется дендрограммой (рис. 6).

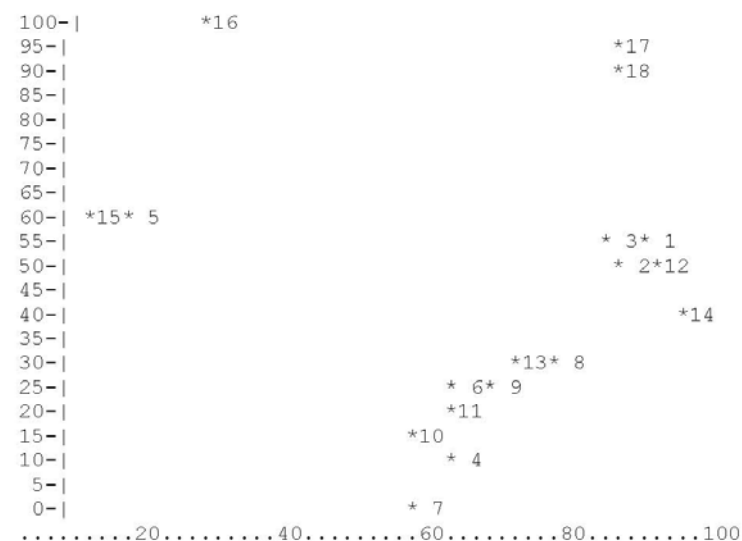


Рис. 5. Ординация объектов методом главных компонент.

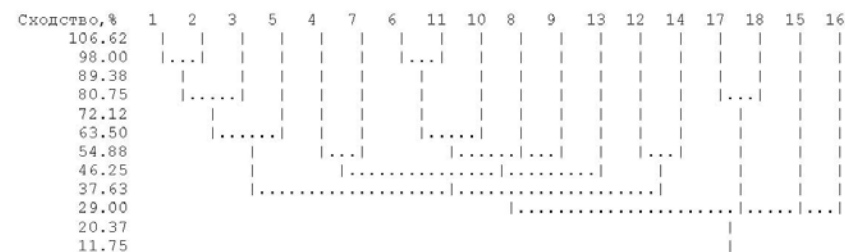


Рис. 6. Дендрограмма строк таблицы.

Если провести сечение дендрограммы на уровне 46% сходства, то выделяются 8 классов объектов, а сечение на уровне сходства 37% дает 6 классов:

S=46%				S=37%			
1)1 2 3 5	4	1 2 3 5	4				
2)4 7	2	4 6 – 11 13	8				
3)6 8 9 10 11 5							
4)13	1						
5)12 14	2	12 14	2				
6)17 18	2	17 18	2				
7)15	1	15	1				
8)16	1	16	1				
56							

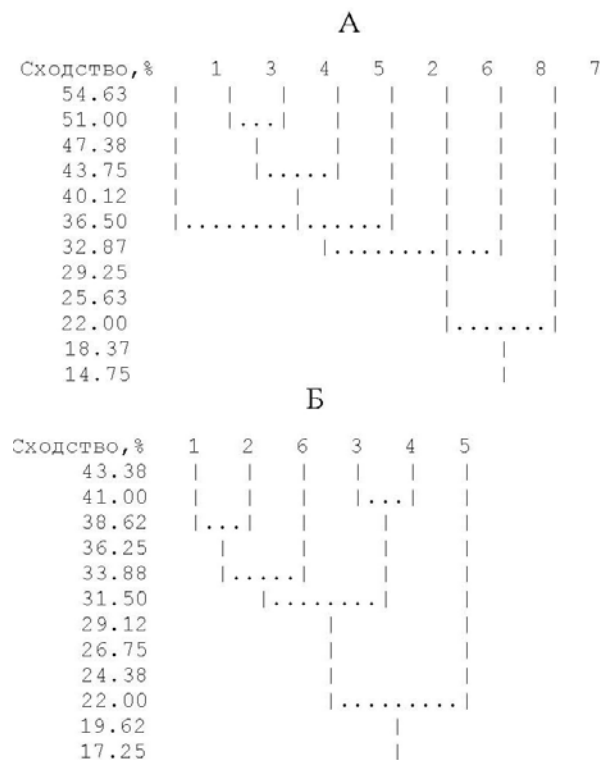


Рис. 7. Дендрографы: А – для группировки на 8 классов, Б – группировка на 6 классов.

Наглядно соотношения классов представлены дендрографами на рис. 7. Для обоих были рассчитаны критерии качества разделения, за которые принята разность среднего внутри- и межклассового сходства объектов (Рожков, 2011). При группировке почв в 8 классов критерий качества составил 45%, при группировке на 6 классов практически не отличался (46%).

В качестве оценки качества разделения объектов была апробирована информационная мера энтропии. Она имеет следующий вид:

$$I = -\sum_i^k p_i \log p_i / \log k,$$

где p_i – доля площади i -той почвы, $i = 1...k$ и k – их число на участке.

При выделении 8-ми классов энтропия составила $I = 0.95$ при 6-ти – $I = 0.68$. Эти показатели свидетельствуют о том, что при более равномерном распределении площадей (что происходит при увеличении числа классов) энтропия растет. Можно условно считать, что классификации с близкими по объему классами являются более оптимальными, чем с классами существенно разного объема. Поэтому показатели энтропии, видимо, могут служить индикаторами, но не критериями качества разделения объектов на классы.

В качестве примера крупной классификационной задачи на рис. 8 представлена дендрограмма почв из классификации 1977 г. и Программы Почвенной карты в масштабе 1 : 2 500 000, а также из экспертных описаний, собранных в свое время Ю.Л. Мешалкиной. Каждый тип почвы представлен большим набором признаков, не имеющим целевой ориентации (архив, к сожалению, не сохранился). Тем не менее, очевидны главные достоинства метода.

На дендрограмме отдельные описания почв объединены в обозримую систему. Системообразующим элементом послужила мера сходства почв и скрытый в алгоритме способ их группировки («критерий ближайшего соседа»). Эмерджентным свойством системы является иерархическая структура, визуализирующая отношения между почвами в системе. На разных уровнях сходства обнаруживаются группы сходных типов почв, требующих своей интерпретации. Если бы не случайный выбор почвенных показателей, полученную систему можно было бы назвать классификацией с заранее назначенной номенклатурой почв.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

На приведенных примерах показаны простейшие подходы для анализа данных на предмет разбиения почвенных объектов на классы и извлечения из данных другой полезной информации.

Подобные представления необходимы для любой классификации. Регулированием состава почвенных признаков достигается соответствие выбранной целевой ориентации и нужное качество. Таким образом, определяется «прозрачность» принципов строящейся классификации, ее обозримость и количественное обоснование выделения таксонов.

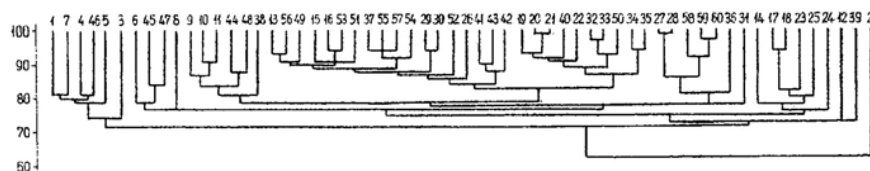


Рис. 8. Дендрограмма почв в пространстве признаков. Список почв на дендрограмме: 1 – подзолистые, 2 – болотно-подзолистые, 3 – дерново-карбонатные, 4 – серые лесные, 5 – бурые лесные, 6 – то же глеевые, 7 – подзолисто-бурые лесные, 8 – то же глеевые, 9 – луговые подбелы, 10 – лугово-черноземные (амурских прерий), 11 – луговые темные, 12 – черноземы, 13 – лугово-черноземные, 14 – каштановые, 15 – лугово-каштановые, 16 – луговые, 17 – бурые полупустынные, 18 – лугово-бурые, 19 – серо-бурые, 20 – такыровидные, 21 – такыры, 22 – песчаные пустынные, 23 – сероземы, 24 – луговые пустынь и полупустынь, 25 – серо-коричневые, 26 – коричневые, 27 – желтоземы, 28 – красноземы, 29 – торфяные болотные верховые, 30 – то же низинные, 31 – солоды, 32 – солонцы автоморфные, 33 – то же полугидроморфные, 34 – солончаки автоморфные, 35 – то же гидроморфные, 36 – аллювиальные дерновые кислые, 37 – то же насыщенные, 38 – дерново-глеевые, 39 – серые лесные глеевые, 40 – лугово-пустынные, 41 – лугово-сероземные, 42 – лугово-серо-коричневые, 43 – лугово-коричневые, 44 – лугово-лесные серые, 45 – желтоземы глеевые, 46 – подзолисто-желтоземные, 47 – то же глеевые, 48 – лугово-болотные, 49 – болотные пустынь и полупустынь, 50 – солонцы гидроморфные, 51 – аллювиальные дерновые насыщенные, 52 – аллювиальные дерновые опустынивающиеся карбонатные, 53 – аллювиальные луговые насыщенные, 54 – то же карбонатные, 55 – аллювиальные лугово-болотные, 56 – аллювиальные болотные, иловато-перегноино глеевые, 57 – аллювиальные болотные иловато-торфяные, 58 – горно-луговые, 59 – горно-луговые черноземовидные, 60 – горно-луговые степные.

При создании классификации могут возникнуть следующие ситуации, в любой из которых полезны описанные методы.

Классификация может строиться, исходя из априорных представлений автора². Тогда необходимо составление тестового при-

² Условно говоря, оперируя с известными ему образцами, автор уже опирается на свой опыт и складывает для себя представление о будущей классификации.

мера (выборки или типовых образцов) для количественного и визуального подтверждения обоснованности предлагаемой структуры.

В другом случае могут классифицироваться накопленные в базе данных описания почв. Непосредственная обработка их показанными методами дает прямой путь к решению задачи.

Распознавание вновь представленных почв проводится по сходству их описаний с почвами, учтенными в системе. Естественно некоторые из новых почв могут оказаться сходны сразу с несколькими другими в системе. Это с одной стороны дает свободу выбора решения, но с другой стороны ставит задачу копить статистику для последующего уточнения состава признаков и градаций их значений. Последующее использование классификации собственно и будет состоять в распознавании вновь предъявленных почв и фиксации результатов для набора статистики ошибок разного рода.

Через определенное время они послужат основанием для коррекции классификации и ее совершенствования или замены.

Ситуация нечеткого разделения множества почв на классы и неоднозначного распознавания почв не уникальная, а скорее является нормой. Реальный пример такой «размытости» представлен на рис. 9, где показано распределение в координатах главных компонент подзолистых почв, имеющих свойства, приведенные в первом примере.

Метод главных компонент – называют методом генерации гипотез. Порой он позволяет выявить в обрабатываемых данных скрытые закономерности (Рожков, 1989). Однако в том случае, если разделение объектов на группы отсутствует, и они распределены более или менее равномерно в пространстве координат, нагрузки на компоненты будут размыты, т.е. собственные числа корреляционной матрицы будут примерно равными по значениям. В таком случае разделение на группы может быть только «силовым» – проводить разделение на заданное число классов, руководствуясь своим опытом или использовать критерии качества классификации (Рожков, 1989, 2011). Иными словами, полностью избежать субъективных решений нельзя. В задачах, подобных рассмотренным, используются эмпирические наблюдения и сложившиеся

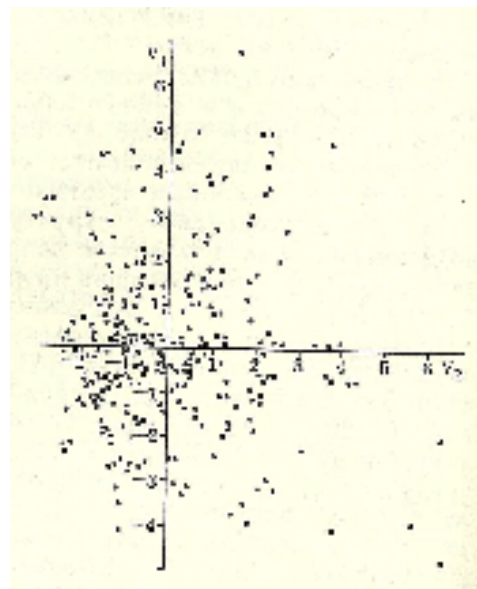


Рис. 9. Распределение подзолистых почв в координатах главных компонент.

представления исследователя. Демонстрируемые методы необходимы для более полного извлечения информации из данных, большей доказательности и визуализации результатов анализа данных. Изложенные подходы и методы применимы в любой дисциплине, поскольку классификация и классификационная деятельность свойственна всем сферам науки и жизни человека в целом. В общем случае любую статистическую обработку данных следует начинать с кластер-анализа – метода главных компонент, дендрограммы с тем, чтобы не смешивать разнородные данные. Особенно важно это при регрессионном анализе – его следует проводить по разным классам отдельно, иначе общее уравнение регрессии может оказаться неадекватным или его коэффициенты слабо достоверны.

Об универсальной применимости методов численной классификации на всех уровнях организации почвенной мегасистемы обсуждалось в специальной публикации (Рожков, Скворцова, 2009).

ЗАКЛЮЧЕНИЕ

Приведенные примеры применения методов из области численной классификации не исчерпывают их разнообразия. Сложные вычисления в алгоритмах многомерной статистики накладывают определенные требования на объемы и свойства многомерных наблюдений. В почвенных исследованиях бывает трудно сформировать выборку описаний почв одного таксономического уровня, чтобы число почв (объектов) было существенно больше числа описывающих их признаков. Без этого невозможно корректное решение задачи. Для выхода из положения могут применяться разные приемы: исключение малоинформативных признаков, разного рода свертки данных (аппроксимация, кодирование, линейные преобразования и др.). В любом случае применение показанных методов необходимо для предметного суждения о результатах любых исследований, для обсуждения их доказательности. Только тогда работа над классификациями станет вполне «прозрачной» свободной от необъяснимой интуиции автора, а выражена в четких формализмах и подкреплена количественными показателями.

Что касается обсуждаемой проблемы построения классификаций почв, нужно признать справедливым следующее высказывание А.Л. Субботина (2001), что требование создать «окончательную», «абсолютно совершенную» классификацию, за исключением, быть может, простейших случаев, выглядит несерьезно; оно способно лишь дезориентировать в оценке реальных шагов в классификационной работе и породить неконструктивный скептицизм. Слишком многогранен объект почвоведения и неограниченны его предназначения для человека. Это требует в каждом конкретном случае формулирования конкретного взгляда на почву, ее свойства и функции. Представленные методы анализа данных могут быть полезны в любом исследовании, их компьютерные программы доступны в коммерческих пакетах SSPS и STATISTICA (есть ограничения по цене и по шкалам признаков) и бесплатно выдаются в Почвенном институте (но в DOC!).

Публикация настоящей статьи – своего рода ответ на статью в «Бюллетене Почвенного института» недавно ушедшего от нас Л.О. Карпачевского (2008) и продолжением статьи по некоторым аспектам теории классификации (Рожков, 2012).

СПИСОК ЛИТЕРАТУРЫ

1. Дунаев В.В., Поляков О.М. Методологические аспекты реляционной теории классификации // Информационный анализ. НТИ. Сер. 2. № 4. 1987. С. 2.
2. Зольников В.Г. Об основных методологических принципах генетической классификации почв // Почвоведение. 1955. № 11. С. 70–79.
3. Карпачевский Л.О. Что есть истина в почвоведении? // Бюл. Почв. ин-та им. В.В. Докучаева. 2008. Вып. 62. С. 108–114.
4. Рожков В.А. Почвенная информатика. М.: Агропромиздат, 1989. 222 с.
5. Рожков В.А. Формальный аппарат классификации почв // Почвоведение. 2011. № 12. С. 1411–1424.
6. Рожков В.А. Об информационном подходе в классификации почв // Бюл. Почв. ин-та им. В.В. Докучаева. 2012. Вып. 69. С. 4–23.
7. Рожков В.А., Скворцова Е.Б. Тектология почвенной мегасистемы (общность организации и анализа данных) // Почвоведение. 2009. № 10. С. 1155–1164.
8. Субботин А.Л. Классификация. М.: ИФ РАН, 2001. 97 с.
9. FAO Soil Bulletin. 1979. № 42. Rome: FAO. 188 p.

SOIL CLASSIFICATION IS NOT QUITE IN PLACE FOR DISCUSSION

© 2013 г. V. A. Rozhkov

V.V. Dokuchaev Soil Science Institute, Moscow, Russia

As the theory and mathematical methods have been so far introduced insufficiently into elaborating the soil classification nowadays, an attempt is made to demonstrate some available approaches. An aphoristic title of this paper is used not to dethrone an “aristocratic” pattern of this significant problem that has been constantly discussed in soil science but to show how the up-to-date means of informatics allow solving numerous tasks related to the elaboration, investigation and utilization of classification structures based upon a description of soil objects. The

research in preparing any classification should be carried out in the way described in this paper in order to consider it as complicated and open for discussion, investigation and further development.

Keywords: classification, digital classification, dendrogram, method of basic components, informative features, classification quality, soil identification.