

ИСПОЛЬЗОВАНИЕ МЕТОДА ТАКСОНОМИЧЕСКОГО АНАЛИЗА Е.С. СМИРНОВА ДЛЯ ОБРАБОТКИ ФЛОРИСТИЧЕСКИХ ДАННЫХ

А.А. Зверев, А.Л. Эбель

Национальный исследовательский Томский государственный университет

USE OF METHOD OF TAXONOMICAL ANALYSIS BY E.S. SMIRNOV FOR PROCESSING OF FLORISTIC DATA

A.A. Zverev, A.L. Ebel

National Research Tomsk State University

Арсенал статистических методов, применяемых при сравнении и группировке данных о растительном покрове, достаточно богат и разнообразен. Используемый в данной работе классификационный метод был изначально предложен Е.С. Смирновым для установления степени сходства таксонов [10, 19]. Он вызвал живой интерес и критику научного сообщества [9], вошел в методические работы по численной таксономии [25, 26], получил математическое обоснование [1, 2, 14], широко и успешно применялся в практике количественной систематики [3, 17, 18].

Основной особенностью метода Е.С. Смирнова, отличающим его от других количественных подходов в классификации, является учет степени выраженности («веса») признака: вклад признака при оценке сходства двух таксонов обратно пропорционален его весу, т.е. большее значение имеет факт одновременного наличия редкого признака у двух таксонов. Редкость признака устанавливается по частоте его встречаемости в матрице сравниваемых объектов, в авторской терминологии – по «фреквенциям» (от англ. frequency – частота). Изначально метод позволял оперировать только бинарными (качественными) признаками таксонов. В более поздних работах [12, 13] Е.С. Смирнов обобщил свои предложения и распространил их на признаки, выраженные в ординируемой номинальной и количественной шкалах, а также на полиморфизмы – одновременное наличие у таксонов двух и более состояний признака.

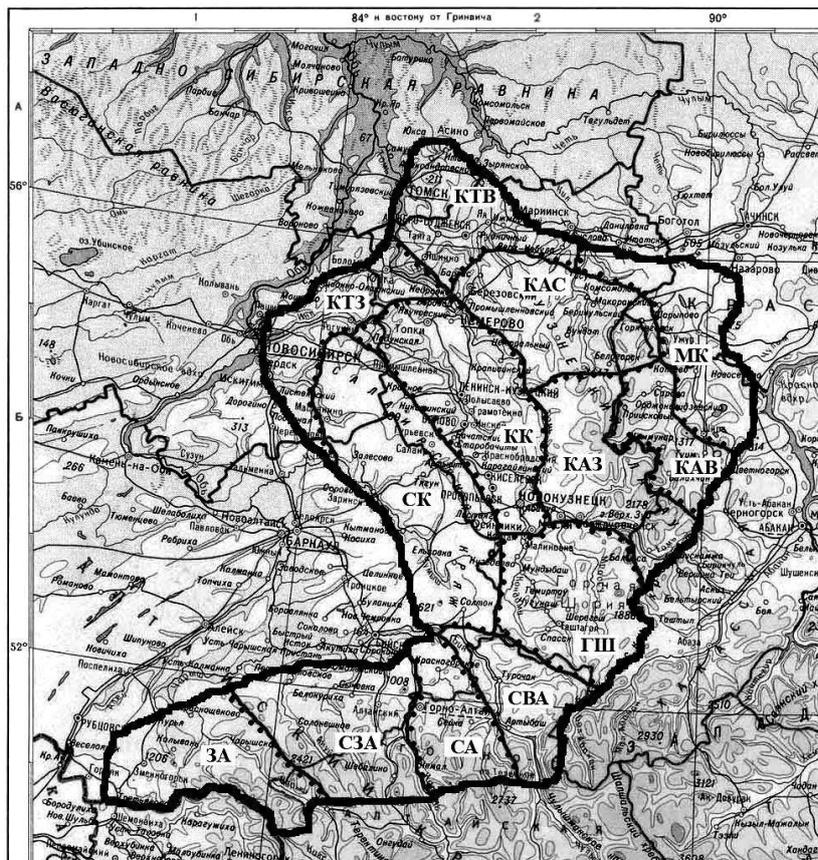
В.М. Шмидт наряду с популяризацией метода Е.С. Смирнова для целей собственно родства степени родства таксонов [16, 18] предложил ввести его в арсенал приемов сравнительно-флористического анализа [15], в частности использовать его для сравнения флор по их историко-географическим связям в целях «объективного» флористического районирования [17–19]. По сути речь идет о возможности использования метода таксономического анализа Е.С. Смирнова как варианта сравнительного Q-анализа – кластерного анализа в геоботанике и флористике, при котором объектами сравнений служат флоры или описания растительности, а таксоны или элементы фактор-множеств (именно последнее и было предложено В.М. Шмидтом) – их признаками с возможным выражением веса в той или иной шкале. Важным было то, что при таком использовании сохранялся весь понятийный аппарат метода Е.С. Смирнова, а ключевым моментом его реализации в новом качестве становился способ кодирования модальностей признаков. Алгоритм такого кодирования В.М. Шмидт подробно рассмотрел на примере учета веса (количества видов) каждого из историко-географических элементов в составе сравниваемых флор.

Целью данного исследования стало сопоставление результатов классификации флор, выполненных с помощью метода таксономического анализа Е.С. Смирнова, и «традиционных» методов сравнительно-флористического. Все иерархические кластерные построения выполнены в программе Statsoft Statistica 8.0 [24], дендрограммы сходства получены методом WPGMA (взвешенное парно-групповое среднеарифметическое связывание). Использованы бинарные и количественные индексы сходства: Czekanowski–Sørensen–Dice [22, 23, 27], Quantitative Ochiai (Cosine) [6, 28] и Relativized Sørensen [6]. Сравнения выполнялись как для полных флористических списков, так и для их полных и головных семейственно-видовых таксономических спектров.

Ввод и хранение модельных флористических списков, вычисление таксономических показателей, построение спектров по таксономическим фактор-множествам [29], расчет матриц сходства по заявленным выше индексам, а также матриц парных таксономических отношений между сравниваемыми флорами по методу Е.С. Смирнова (далее именуется «индексом Смирнова») выполнены в интегрированной ботанической информационной системе IBIS [4].

В качестве модельных флор выступили флоры выделенных и изученных нами [20, 21] 13 ботанико-географических районов, расположенных территориально в северо-западной части Алтае-Саянской гор-

ной области (АСО), а по флористическому районированию Алтае-Саянской провинции (АСП) – в пределах ее Северной подпровинции [8]. АСО возникла на месте синийско-кембрийской геосинклинальной системы в результате складчатости, основные фазы которой имели в разных частях системы разный возраст. На формирование рельефа АСО большое влияние оказали плейстоценовые оледенения, характер которых в различных районах горной страны не был одинаковым. К территории наших исследований относятся западные и северные районы Русского Алтая, Кузнецкий Алатау, Горная Шория, Салаирский кряж, Колывань-Томское плато и межгорные котловины – Кузнецкая и северо-западная часть Назаровско-Минусинской (рис. 1). Территория северо-запада АСП (СЗАСП) относится преимущественно к бассейну Оби; небольшая часть Северо-Минусинской котловины – к области внутреннего стока. Основная часть района наших исследований не покрывалась сплошным ледниковым покровом. В выработке форм мезорельефа, расчленении склонов и вершин хребтов, в создании современных долин огромное значение имели процессы выветривания и эрозионной деятельности водных потоков. Гидрографическая сеть, хорошо развитая в районе исследований, заложилась еще в третичное время на территории древнего пенеплена и за время четвертичной истории претерпела большие превращения.



ЗА – Западный Алтай, СЗА – Северо-Западный Алтай, СА – Северный Алтай, СВА – Северо-Восточный (Прителецкий) Алтай, ГШ – Горная Шория, СК – Салаир, КК – Кузнецкая котловина, КАЗ – Кузнецкий Алатау (западный район), КАВ – Кузнецкий Алатау (восточный район), КАС – Кузнецкий Алатау (северный район), МК – Назаровско-Минусинская котловина, КТЗ – Колывань-Томское плато (западный район), КТВ – Колывань-Томское плато (восточный район).

В формировании климата на территории АСП проявляется важная роль атлантического переноса влаги, а также влияние Азиатского антициклона. Согласно схеме биоклиматических подразделений гор Южной Сибири [7] в СЗАСП представлены 3 группы районов (климатические фации). *Избыточно влажная климатическая фация* – на наиболее влажных наветренных склонах передовых горных хребтов и их отрогов, где среднегодовое количество осадков составляет до 1000 мм и более (СВА, ГШ, КАЗ, среднегорье ЗА, осевая часть СК). *Влажная циклоническая фация* – в низкогорьях и среднегорьях СЗА, в предгорьях и низкогорьях ЗА, на основной части СК и КК, в среднегорьях КАВ и на большей части Колывань-

Томского плато. *Умеренно влажная антициклоническая фация* – в районах влияния эффекта барьерной тени от передовых хребтов горной системы (МК, низкогорья КАВ, западная часть КК).

Для региона в целом характерно большое разнообразие типов почв. В степном поясе распространены различные подтипы черноземов. Преобладающее значение имеют почвы лесного пояса, развитые сплошными массивами на больших площадях. В высокогорьях – горно-луговые и горно-тундровые почвы. В межгорных котловинах (Северо-Минусинская, западная часть Кузнецкой) по понижениям развиты засоленные почвы с сопутствующей им специфической растительностью. Растительность СЗАСП носит весьма пестрый характер. Здесь развиты 5 поясов растительности: степной, лесостепной, лесной, субальпийский, альпийский. В разных районах набор поясов, их характер и мощность вертикального развития различны. Так, только в пределах Горного Алтая представлены 4 типа поясности [5], 3 из которых отмечены на территории, входящей в состав СЗАСП.

Краткая количественная характеристика исследованной флоры в целом и по районам по данным А.Л. Эбеля [20] приведена в таблице.

Таксономические показатели флор районов

Показатель	ЗА	СЗА	СА	СВА	ГШ	СК	КК	КАЗ	КАВ	КАС	МК	КТЗ	КТВ	Вся флора
Общее число видов	1376	1301	1063	1043	948	944	1022	773	1023	834	1160	944	827	2017
Общее число родов	468	460	430	415	402	403	401	348	392	376	427	400	368	565
Общее число семейств	112	109	106	104	110	111	103	97	97	103	107	100	100	125
Среднее число видов в роде	2,94	2,828	2,472	2,513	2,358	2,342	2,549	2,221	2,61	2,218	2,717	2,36	2,247	3,57
Среднее число видов в семействе	12,286	11,936	10,028	10,029	8,618	8,505	9,922	7,969	10,546	8,097	10,841	9,44	8,27	16,136
Число одновидовых родов	230	239	235	233	228	224	213	204	216	219	225	217	210	262
Доля видов в 100 ведущих родах, %	59,08	59,42	58,42	60,69	60,23	59,53	60,86	62,23	63,54	60,07	61,21	59,43	60,58	58,95
Доля видов в 10 ведущих семействах, %	57,56	57,26	54,37	54,65	52,95	54,77	57,63	52,91	59,34	54,08	58,45	56,78	55,14	57,51
Доля сосудистых споровых, %	3,49	4,46	4,8	6,04	5,38	4,34	1,96	6,34	4,2	4,56	2,33	2,86	3,51	3,52
Доля голосеменных, %	0,73	0,69	0,85	0,86	0,84	0,74	0,59	1,16	0,78	0,84	0,6	0,85	0,73	0,69
Доля цветковых, %	95,78	94,85	94,36	93,1	93,78	94,92	97,46	92,5	95,01	94,6	97,07	96,29	95,77	95,79
Доля однодольных среди цветковых, %	24,28	25,04	25,02	24,92	26,1	27,68	26,2	26,43	27,26	26,36	27,44	25,08	26,26	25,67
Доля двудольных среди цветковых, %	75,72	74,96	74,98	75,08	73,9	72,32	73,8	73,57	72,74	73,64	72,56	74,92	73,74	74,33

Примечание. Обозначения районов соответствуют рис. 1.

Сравнение полных видовых списков, выполненных с использованием различных «традиционных» индексов сходства, приводит к сходным результатам (рис. 2, 3). Очень четко выделяется кластер, объединяющий 2 западных алтайских района (ЗА + СЗА). Довольно тесные связи наблюдаются и между 2 другими районами Алтая (СА + СВА), однако в этот кластер входят также (с менее тесными связями) западный р-н Кузнецкого Алатау и Горная Шория. Остальные районы распределяются по 2 кластерам — «Верхне-чулымскому» (МК + КАВ) и «Кузнецкому» (СК + КК + КТЗ + КТВ + КС). Метод Е.С. Смирнова (рис. 4) дает практически ту же картину, что и «традиционные» индексы. Однако при этом обнаруживаются более тесные связи между отдельными районами (Каз + ГШ; МК + КАВ). Особенно наглядно это видно при использовании «количественного» подхода (рис. 5), когда учитывается не только присутствие вида в районе, но и его встречаемость.

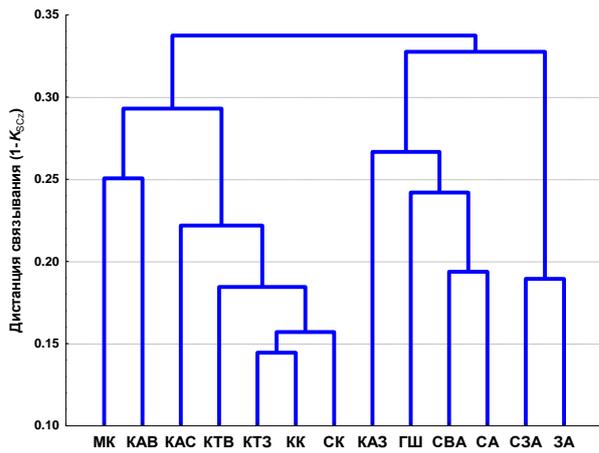


Рис. 2. Дендрограмма сходства флор 13 районов. Индекс Сзекановски–Sørensen–Dice по полным видовым спискам, бинарные признаки (присутствие таксонов)

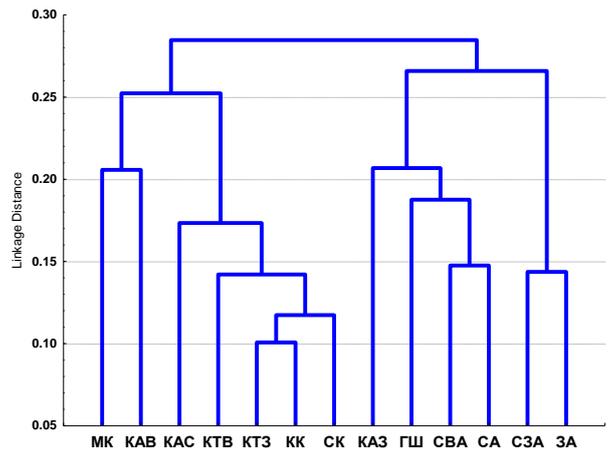


Рис. 3. Дендрограмма сходства флор 13 районов. Количественный вариант индекса Ochiai (Cosine) по полным видовым спискам (встречаемость таксонов в 9-бальной геометрической шкале)

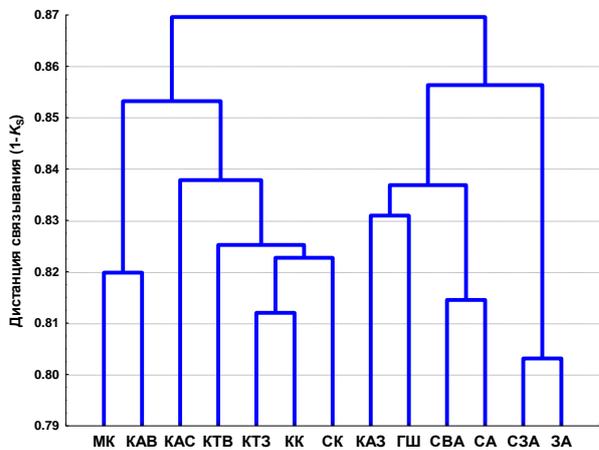


Рис. 4. Дендрограмма сходства флор 13 районов. Индекс Смирнова по полным видовым спискам, бинарные признаки (присутствие таксонов)

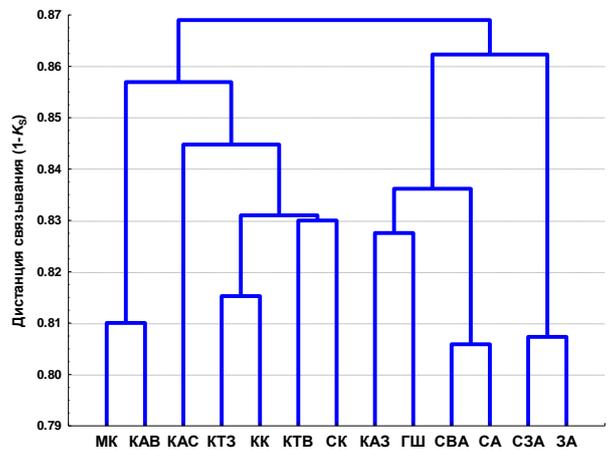


Рис. 5. Дендрограмма сходства флор 13 районов. Индекс Смирнова по полным видовым спискам, количественные признаки (встречаемость таксонов)

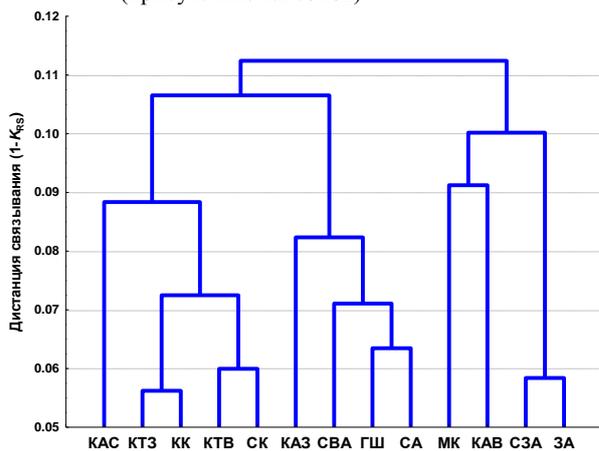


Рис. 6. Дендрограмма сходства флор 13 районов. Индекс Relativized Sørensen по полному семейственно-видовому спектру, взвешенному встречаемостью таксонов

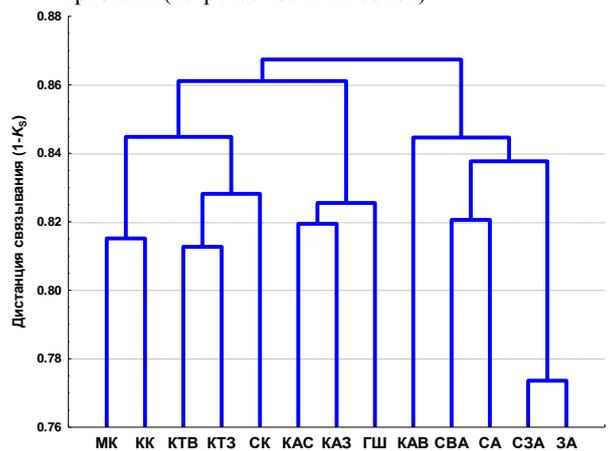


Рис. 7. Дендрограмма сходства флор 13 районов. Количественный вариант индекса Смирнова по полному семейственно-видовому спектру, взвешенному встречаемостью таксонов

Несколько иные результаты дал анализ сходства районов по таксономической структуре флоры (семеjственно-видовой спектр). Использование «релятивизированного индекса Сёренсена» (Relativized Sørensen) (рис. 6) позволило выделить те же самые 4 кластера. Однако географически удаленные кластеры (МК + КВА и ЗА + СЗА) оказались гораздо ближе по таксономической структуре, чем по полному видовому составу.

Существенно иная картина обнаружилась при сравнении таксономической структуры районов с использованием метода Е.С. Смирнова (рис. 7). При этом относительно четко выделились 3 кластера (да и то лишь при учете встречаемости видов): 1) равнинно-низкогорные районы: (КК + МК) + (КТВ + КТЗ + СК); 2) Кузнецкое нагорье (без восточной части): КАС + КАЗ + ГШ; Алтай и восточный район Кузнецкого Алатау: (ЗА + СЗА) + (СВА + СА) + КАВ.

Сравнение районов по головным спектрам (оставлены 24 семейства, аккумулирующие не менее одного процента таксонов в среднем по всем районам) с использованием разных подходов показало не очень устойчивое положение 3 районов: КК, КАВ и МК. Так, при использовании метода Е.С. Смирнова (без учета встречаемости) (рис. 8) МК и КК попадают в «Западноалтайский кластер» (ЗА + СЗА). Вероятно, это можно объяснить обилием видов Fabaceae в данных районах. Однако при этом набор видов в алтайских районах и в МК существенно различен (в частности, за счет «своих» эндемичных видов).

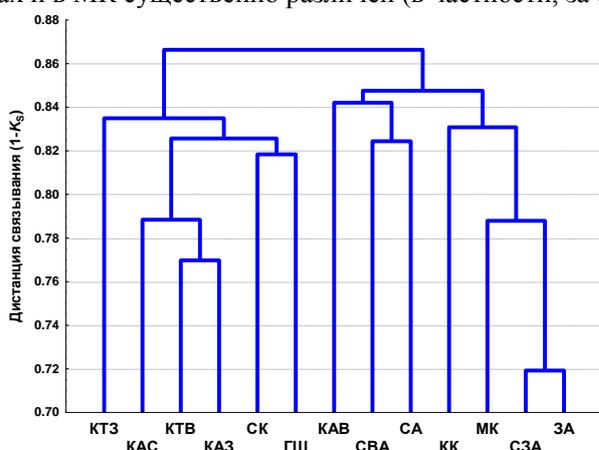


Рис. 8. Дендрограмма сходства флор 13 районов. Количественный вариант индекса Смирнова по головной части семейственно-видового спектра

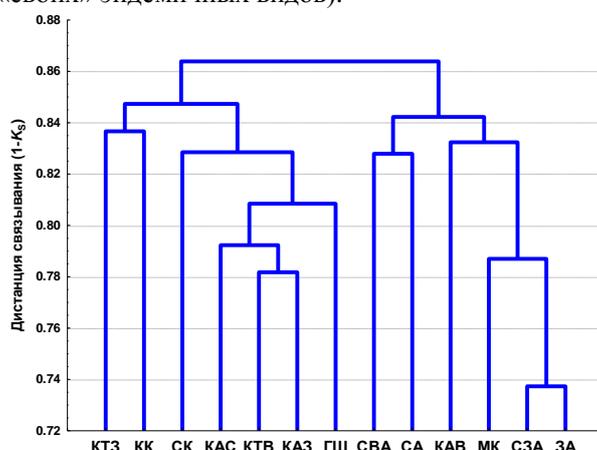


Рис. 9. Дендрограмма сходства флор 13 районов. Количественный вариант индекса Смирнова по головной части семейственно-видового спектра, взвешенного встречаемостью таксонов

Учет встречаемости по этому же методу (рис. 9) показывает более тесное сходство КАВ с МК (вероятно, за счет обилия Роасеae, Сурегасеae и Fabaceae), но при этом МК остается в «Западноалтайском» кластере (опять-таки за счет разных «наборов» многочисленных бобовых), а КК перемещается в группу равнинно-низкогорных районов, расположенных преимущественно на территории Кузбасса.

Проведенное исследование позволяет сделать вывод о перспективности использования метода таксономического анализа Е.С. Смирнова для количественной обработки флористических данных и применения результатов такого анализа для целей флористического и геоботанического районирования.

Литература

1. Дорюфеюк А.А. Алгоритмы автоматической классификации // Автоматика и телемеханика. 1971. № 12. С. 78–113.
2. Дюльдин А.А. Булевы функции в таксономическом анализе Е.С. Смирнова // Журн. общ. биол. 1973. Т. 34, № 5. С. 745–751.
3. Зайцев Г.Н. Алгоритмы таксономического анализа Е.С. Смирнова // Бюллетень Главного ботанического сада. М.: Наука, 1976. Вып. 100. С. 101–107.
4. Зверев А.А. Информационные технологии в исследованиях растительного покрова: учеб. пособие. Томск: ТМЛ-пресс, 2007. 304 с.
5. Огуреева Г.Н. Ботаническая география Алтая. М.: Наука, 1980. 189 с.
6. Песенко Ю.А. Принципы и методы количественного анализа в фаунистических исследованиях. М.: Наука, 1982. 287 с.
7. Поликарпов Н.П., Чебакова Н.М., Назимова Д.И. Климат и горные леса Южной Сибири. Новосибирск: Наука, 1986. 225 с.

8. *Ревушкин А.С.* Материалы к флористическому районированию Алтае-Саянской провинции // Флора, растительность и растительные ресурсы Сибири. Томск: Изд-во Том. ун-та, 1987. С. 32–46.
9. *Сёмкин Б.И., Куликова Л.С., Петухова Е.Л.* Количественный анализ связей энтомофауны сельскохозяйственных посевов и естественной растительности в Приморском крае // Структурная организация компонентов биogeосистем (сравнительный и количественный анализ). Владивосток: ДВО РАН СССР, 1988. С. 128–143.
10. *Смирнов Е.С.* Таксономический анализ рода // Журн. общ. биол. 1960. Т. 21, № 2. С. 89–103.
11. *Смирнов Е.С.* Проблема таксономического сходства в систематике // Журн. общ. биол. 1963. Т. 24, № 3. С. 172–181.
12. *Смирнов Е.С.* Таксономический анализ. М.: Изд-во МГУ, 1969. 187 с.
13. *Смирнов Е.С.* О кодировании признаков для таксономического анализа // Журн. общ. биол. 1971. Т. 32, № 2. С. 224–228.
14. *Тамарин П.В.* Анализ математических методов систематики // Журн. общ. биол. 1971. Т. 32, № 3. С. 277–286.
15. *Шмидт В.М.* О методе таксономического анализа Е.С. Смирнова и некоторых возможностях его применения в ботанике // Бот. журн. 1962. Т. 47, № 11. С. 1648–1654.
16. *Шмидт В.М.* О двух методах таксономического анализа // Бот. журн. 1970. Т. 55, № 3. С. 386–396.
17. *Шмидт В.М.* Статистические методы в сравнительной флористике. Л.: Изд-во ЛГУ, 1980. 176 с.
18. *Шмидт В.М.* Математические методы в ботанике: учеб. пособие. Л.: Изд-во ЛГУ, 1984. 288 с.
19. *Шмидт В.М., Баранова Е.В.* Применение метода таксономического анализа для сравнения флор по их историко-географическим связям // Журн. общ. биол. 1975. Т. 36, № 4. С. 555–562.
20. *Эбель А.Л.* Флора северо-западной части Алтае-Саянской провинции: состав, структура, происхождение, антропогенная трансформация: дис. ... д-ра биол. наук. Томск, 2011. 758 с.
21. *Эбель А.Л.* Конспект флоры северо-западной части Алтае-Саянской провинции. Кемерово: КРЭОО «Ирбис», 2012. 568 с.
22. *Czekanowski J.* «Coefficient of racial likeness» und «durchschnittliche Differenz» // Anthropologischer Anzeiger. 1932. Bd. 9. S. 227–249.
23. *Dice L.R.* Measures of the amount of ecologic association between species // Ecology. 1945. Vol. 26, № 3. P. 297–302.
24. *Hill T., Lewicki P.* STATISTICS Methods and Applications. Tulsa: StatSoft, 2007. 832 p.
25. *Sneath R.H.A., Sokal R.R.* Numerical taxonomy. San Francisco: W.H. Freeman and Company, 1973. 573 p.
26. *Sokal R.R., Sneath P.H.A.* Principles of numerical taxonomy. San Francisco, London: W.H. Freeman and Company, 1963. 359 p.
27. *Sørensen T.* A new method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analysis of the vegetation on Danish commons // Biologiske Skrifter / Kongelige Danske Videnskabernes Selskab. 1948. Bd. 5, № 4. S. 1–34.
28. *Torgerson W.S.* Theory and methods of scaling. New York: John Wiley and Sons, Inc., 1958. 460 p.
29. *Zverev A.A.* Use of equivalence classes and factor sets in the analysis of botanical data // Contemporary Problems of Ecology. 2012. Vol. 5, № 2. P. 165–173.