

Посвящается Ю.М. Возняковской

НЕЙРОСЕТЕВОЕ РАНЖИРОВАНИЕ АГРОТЕХНОЛОГИЙ ПО ИНДЕКСАМ МИКРОБИОЛОГИЧЕСКОЙ АКТИВНОСТИ ПОЧВЫ И ПОЧВЕННОГО ПЛОДОРОДИЯ: НОВЫЕ ВОЗМОЖНОСТИ СТАТИСТИЧЕСКОГО АНАЛИЗА*

Н.И. ВОРОБЬЕВ^{✉ 1, 2}

Нейросети позволяют извлекать ранее недоступную информацию из физиологических и молекулярно-генетических данных и визуализировать неявные зависимости. В представленной работе мы применили нейросетевой подход для анализа данных Ю.М. Возняковской с соавт. (Сельскохозяйственная биология, 1994), полученных в длительном опыте (1962-1991 годы, Северо-Западный регион Российской Федерации) по изучению влияния севооборотов и минеральных удобрений на микробиологические и биохимические характеристики дерново-подзолистых почв под картофелем и ячменем. В опыте, цель которого заключалась в выявлении микробиологических показателей, наиболее четко характеризующих уровень плодородия почвы, был определен состав почвенных микроорганизмов, измерена интенсивность выделения CO_2 , разложения целлюлозы, активности почвенных полифенолоксидаз и пероксидаз как показателей интенсивности процессов гумификации растительных остатков, а также инвертаз и уреаз для оценки накопления нитратов и аммония в почве. На этом основании к наиболее информативным микробиологическим показателям уровня эффективного плодородия почв авторы отнесли общую биогенность почвы, видовое разнообразие почвенных микроорганизмов, соотношение трофических групп микроорганизмов, граф трофических связей. Проведенный в настоящей работе анализ фрактальных профилей физиологических групп микроорганизмов, выделяемых на селективных средах, продемонстрировал уникальные возможности статистического анализа с привлечением вычислительных нейросетей. Их применение для обработки эмпирических микробиологических и физико-биохимических почвенных данных опыта Ю.М. Возняковской с соавторами позволило определить специфичность воздействия элементов использованных агротехнологий на плодородие почв под разными культурами. Микробиологическая активность почвы и почвенное плодородие представляют собой данные, характеризующие разнородные биологические объекты — почвенное микробное сообщество (индекс $\text{CSI}_{\text{micro}}$) и растения ($\text{CSI}_{\text{yield}}$), где CSI — Cognitive Saliency Index. Поэтому для визуализации зависимости $\text{CSI}_{\text{yield}} = f(\text{CSI}_{\text{micro}})$ нами была использована нейросеть с построением матрицы Scale, ячейки которой с координатами $\text{CSI}_{\text{yield}}; \text{CSI}_{\text{micro}}$ заполнялись соответствующими номерами варианта опыта. Фрактальные профили почвенных физиологических групп микроорганизмов, полученные в опыте, были подвергнуты нейросетевому анализу. В результате по виду матрицы Scale удалось оценить зависимость $\text{CSI}_{\text{yield}} = f(\text{CSI}_{\text{micro}})$ и влияние удобрительных и севооборотных агротехнологий на интенсивность накопления гумуса в почве и плодородие почв. В результате обнаружено, что максимальное значение индекса почвенного плодородия $\text{CSI}_{\text{yield}}$ не совпадает с максимальным значением индекса микробиологической активности почвы $\text{CSI}_{\text{micro}}$. Предложенные нами оригинальные индексы позволяют оптимизировать агротехнологии применения органоминеральных удобрений и бактериальных препаратов для получения стабильно высоких урожаев сельскохозяйственных культур.

Ключевые слова: корреляционный анализ, кластерный анализ, дискриминантный анализ, фрактальный анализ микробиоты, вычислительная нейронная сеть, микробиологическая активность почвы, почвенное плодородие.

Микроорганизмы, обитающие в почве, непрерывно ее преобразуют. Почвенные микроорганизмы трансформируют растительные остатки в гумусовые и питательные вещества, поддерживая этим почвенное плодородие и, как следствие, высокую продуктивность растений (1-3). При объединении микроорганизмов в биосистемы интенсивность преобразовательной деятельности микроорганизмов значительно возрастает, так как в биосистемах микроорганизмы совместно трансформируют органические субстраты растительных остатков, используя ускоренные и малозатратные биохимические схемы (4-6). Микроорганизмы в биосистемах организованно синтезируют

* Подготовлено при финансовой поддержке Российского научного фонда (проект № 23-26-00234).

ферменты для разложения органических молекул растительных клеток и останавливают преобразования при появлении простых органических мономеров, а не конечных продуктов деструкции — воды и парникового газа CO_2 , как это происходит вне биосистемных отношений. Кроме этого, в настроенных микробных биосистемах устанавливаются фиксированные (фрактальные) соотношения численности микроорганизмов (7, 8), что исключает непроизводительные затраты энергетических и материальных ресурсов из-за отклонения от оптимального соотношения микроорганизмов в биосистеме или при неконтролируемых хаотических биохимических процессах в почве. В совокупности биосистемные организационные процессы в почвенных микробных сообществах обеспечивают ускоренную и эффективную трансформацию органических субстратов в питательные и гумусовые вещества, а также снижают выделение парникового газа CO_2 из почвы в атмосферу.

В почве микробиологические процессы происходят в условиях регулярной видовой ротации растительных остатков и при чередовании легко разлагаемых субстратов с трудно разлагаемыми субстратами — лигноцеллюлозными растительными композитами (9). Это вынуждает микроорганизмы оперативно менять схему биохимических преобразований органических субстратов для эффективного непрерывного накопления гумусовых и питательных веществ в почве, определяющих почвенное плодородие (10, 11). Вынужденные реорганизационные перестройки микробных биосистем сопровождаются количественным и генотипическим варьированием компонентного состава биосистем, сукцессией микроорганизмов (12) и колебательной временной динамикой численности групп микроорганизмов в почве (13). Таким образом, именно в составе биосистем преобразовательная деятельность микроорганизмов гарантирует непрерывное накопление в почве питательных и гумусовых ресурсов для микробных сообществ и растений (14).

Частотно-физиологические характеристики почвенных микробных сообществ можно определить с использованием селективных сред (15-17), а частотно-таксономические характеристики — молекулярно-генетическим методом (18-20). Молекулярно-генетический метод имеет высокую генотипическую разрешающую способность, позволяет выделить большое число геномных таксонов (до нескольких сотен тысяч) и обнаружить даже те микробные генотипы, которые не удастся выделить селективным методом. На ограниченном наборе селективных сред удастся выделить и определить численность физиологических групп микроорганизмов только по физиолого-функциональному признаку и в динамическом диапазоне, составляющем 6-7 порядков. Такого широкого динамического диапазона не удастся достичь молекулярно-генетическим методом (максимум 3-4 порядка). Поэтому для контроля численности редких целлюлозолитических бактерий приходится применять оба метода. Целлюлозолитические бактерии осуществляют первичную рестрикцию лигноцеллюлозных молекул при многоэтапном процессе деструкции—гумификации растительных остатков (21). Поэтому эти бактерии входят в основной состав большинства микробиологических препаратов, применяемых для интенсификации преобразовательной и обогачительной деятельности почвенной микрофлоры.

Показатели микробиологической активности почвы (22, 23) и почвенного плодородия представляют собой неоднородный набор цифровых данных, характеризующих разнородные биологические объекты, — почвенное микробное сообщество и сельскохозяйственные растения. Измерения характеристик этих объектов могут происходить в разное время, в разных условиях, разными методами, а их результаты выражаться в несовместимых

единицах. Например, численность физиологических групп почвенных микроорганизмов может представляться в КОЕ/г почвы, а урожайность растений — в ц/га. Подобная несовместимость и неконтролируемое влияние внешних факторов на растения в период их вегетации снизили качество регрессионных уравнений, описывающих зависимость продуктивности растений от удобрительных технологий, что в итоге сказалось на низкой надежности прогнозов урожая сельскохозяйственных растений (24, 25).

Нейросети позволяют извлекать ранее недоступную информацию из физиологических и молекулярно-генетических данных и визуализировать неявные зависимости. В настоящей работе нами разработан алгоритм нейросетевого ранжирования агротехнологий по индексам микробиологической активности почвы и почвенного плодородия. Для этого решены следующие задачи. В Excel-среде создана вычислительная нейросеть, выполняющая преобразование фрактальных профилей физиологических групп почвенных микроорганизмов в значения индексов CSI_{micro} и CSI_{yield} (CSI — Cognitive Saliense Index). В режиме настройки-обучения нейросети осуществлен поиск корректного и обоснованного алгоритма вычисления индексов CSI_{yield} и CSI_{micro} на основе сопоставления значений индексов с почвенными физико-биохимическими данными и урожайностью растений. С использованием вычисленных значений индексов CSI_{yield} и CSI_{micro} и матрицы Scale для ячменя и картофеля визуализированы зависимости CSI_{yield} = f (CSI_{micro}) и проведена оценка их качественных и количественных характеристик.

Целью исследования была визуализация неявной неформализуемой зависимости плодородия почв от микробиологической активности почвы с помощью созданной вычислительной нейросети. Для этого мы предлагаем трансформировать исходные эмпирические микробиологические данные (Ю.М. Возняковская с соавт., 1994) в индексы CSI_{micro} и CSI_{yield}. После вычисления индексов CSI_{yield} и CSI_{micro}, загрузки их значений в матрицу Scale и визуализации зависимости CSI_{yield} = f (CSI_{micro}) появляется возможность оценить влияния микробиологической активности почвы, режимов применения минеральных удобрений и севооборотов на почвенное плодородие.

Методика. Проблему визуализации неформализуемых зависимостей для неоднородных данных, принадлежащих разным компонентам аграрных экосистем, удалось решить, основываясь на преобразовании таких данных в безразмерные (б/р) индексы когнитивной значимости CSI (Cognitive Saliense Index) (26). При этом было предложено делегировать каждому индексу свои права количественного представления: CSI_{micro} отражает уровень микробиологической активности почв, CSI_{yield} — уровень плодородия почвы. Далее было предложено, используя вычисленные значения индексов CSI_{yield} и CSI_{micro}, заполнить двумерную матрицу шкалирования Scale (27, 28) и визуализировать искомую зависимость CSI_{yield} = f (CSI_{micro}).

Для индексации микробиологических данных можно воспользоваться суммарной численностью физиологических групп почвенных микроорганизмов (29). Мы полагаем, что суммарная численность физиологических групп микроорганизмов менее информативна, чем соотношение численностей физиологических групп почвенных микроорганизмов. Для анализа соотношения по численности групп микроорганизмов микробиологические данные были представлены в форме фрактального профиля микроорганизмов (7, 8). Фрактальный профиль микроорганизмов — это убывающий числовой ряд, в котором значения частот физиологических групп микроорганизмов располагаются в порядке уменьшения подобно убывающему степенному числовому ряду (30).

Плохая обусловленность задачи вычисления индексов CSI_{yield} и CSI_{micro} и визуализации зависимости $CSI_{yield} = f(CSI_{micro})$ вынуждает использовать для ее поиска вычислительные нейронные сети (31-33). Вычислительные нейронные сети отличаются тем, что в их структуру могут быть включены сервисы, позволяющие видоизменять алгоритм нейросетевых вычислений (34) и проводить экспертизу моделируемых процессов (35, 36).

В структуре нейросети (37, 38), использованной для определения CSI_{yield} и CSI_{micro} , алгоритм их расчета был найден с помощью вычислительных процедур корреляционного, кластерного и дискриминантного статистических анализов (39-41), примененных для оценки количественных соотношений физиологических групп микроорганизмов в их фрактальных профилях.

Корреляционный анализ в предложенном нами алгоритме вычислений формирует матрицу $Corr$, заполненную коэффициентами корреляции признаков микробных биосистем (42). С помощью матрицы $Corr$ впервые удалось построить граф максимальных корреляций и продемонстрировать этим ведущую роль тандема из аммонификаторов и бактерий, использующих минеральный азот, в круговороте азота в аграрных экосистемах (43).

Кластерный анализ формирует матрицы $FDist$ и $CDist$, содержащие евклидовы дистанции между фрактальными профилями почвенных микроорганизмов по вариантам опыта. По данным этих матриц удается построить дендрограммы кластеризации вариантов опыта (44). Построение филогенетических дистанционных дендрограмм позволило, например, проанализировать эволюцию бактериальных генотипов (45).

Дискриминантный анализ формирует матрицы $FCom$, $CCom$, представляющие ортогональные и ранжированные системы координат (системы главных компонент) в многомерном метрическом пространстве фрактальных профилей микроорганизмов. Визуализация эмпирических данных в ортогональных системах координат позволяет, например, изучить влияние агротехнологий на генотипическую конфигурацию почвенных микробных сообществ (46).

Использованные для вычислений эмпирические микробиологические и физико-биохимические почвенные данные были получены в опыте 1962-1991 годов (СЗНИИСХ, Белогорка, Ленинградская обл.) (29). В вариантах опыта применяли минеральные удобрения ($N_{90}P_{90}K_{90}$ и торфяно-микробиологическое азотное удобрение $TMAU_{30}$) и севооборот (однолетние травы; рожь; картофель; овес; картофель; ячмень с подсевом трав). В образцах почвы с полей под растениями ячменя и картофеля на селективных средах была определена численность физиологических групп почвенных микроорганизмов. Биохимические характеристики почв определяли по стандартным методикам: выделение CO_2 — газохроматографическим методом; интенсивность разложения целлюлозы — методом Кристенсена; активность почвенных полифенолоксидаз, пероксидаз (ПФО, ПО), инвертаз и уреаз — ФЭК-методом. ПФО и ПО были выбраны в качестве индикаторов интенсивности процессов гумификации растительных остатков (47), а инвертазы и уреазы в качестве индикаторов накопления нитратов и аммония в почве (48, 49).

Результаты. Алгоритм нейросетевых вычислений включал следующие этапы.

1) Слой нейронов $L1$ выполняет логарифмирование частот групп микроорганизмов во фрактальных профилях (матрица $FPro$) и преобразует логарифмированные частоты разностной WaveLet-процедурой (50) с последующей записью полученных данных в матрицу FWL .

$$FWL_k = (FProk_{k+1} - FProk)/2, \quad [1]$$

где $FProk = \ln(10^4 f_k)$ — логарифмированные значения частот (f_k) групп микроорганизмов во фрактальном профиле; $k = 1..N$ — порядковые номера позиции микроорганизмов во фрактальном профиле; N — число физиологических групп микроорганизмов во фрактальном профиле.

2) Слой нейронов L1 с помощью стандартной функции $Normalization()$ (51) выполняет построчную нормализацию данных матрицы FWL и записывает результаты в матрицу $FNorm$:

$$FNorm = Normalization(FWL). \quad [2]$$

3) Слой нейронов L1 вычисляет матрицу $Corr$, используя данные матрицы $FNorm$ и стандартную процедуру $Correlation()$ (52, 53):

$$Corr = Correlation(FNorm). \quad [3]$$

4) Слой нейронов L1 выполняет построчную нормализацию процедурой $Normalization()$ данных произведения матриц $FNorm \times Corr$, и записывает результаты нормализации в матрицу $CNorm$:

$$CNorm = Normalization(FNorm \times Corr). \quad [4]$$

5) Слой нейронов L2 вычисляет матрицы евклидовых расстояний $FDist$ и $CDist$, используя цифровые данные, расположенные в столбцах матриц $FNorm$ и $CNorm$, и стандартную вычислительную процедуру $EuclidDistance()$ (54):

$$FDist_{kl} = EuclidDistance(FNorm1k, \dots, FNormNk; FNorm1l, \dots, FNormNl), \quad [5]$$

$$CDist_{kl} = EuclidDistance(CNorm1k, \dots, CNormNk; CNorm1l, \dots, CNormNl), \quad [6]$$

где $k, l = 1, \dots, N$ — порядковые номера позиции микроорганизмов во фрактальном профиле; N — число физиологических групп микроорганизмов во фрактальном профиле.

6) Слой нейронов L2 вычисляет матрицы $FCom$ и $CCom$, используя данные диагональных симметричных матриц $FDist$, $CDist$ и стандартную процедуру $EigenVectors()$ (55):

$$FCom = EigenVectors(FDist), \quad [7]$$

$$CCom = EigenVectors(CDist). \quad [8]$$

7) Слой нейронов L3 вычисляет значения индексов $CSImicro$ и $CSYield$, используя данные матриц $FNorm$, $CNorm$, $FDist$, $CDist$, $FCom$, $CCom$ и выбранный алгоритм из базы алгоритмов $Comby$.

8) Слой нейронов L4 анализирует на каждом цикле обучения-поиска $Learning$ корреляцию (r) индексов $CSImicro$ и $CSYield$ с данными матрицы $Fbio$. Если $Abs(r) < 0,7$ то выбор алгоритма из базы $Comby$ продолжается, а при $Abs(r) > 0,7$ обучение нейросети заканчивается и вычисленные значения индексов $CSImicro$ и $CSYield$ принимаются как корректные.

9) После выхода из цикла обучения-поиска $Learning$ ячейки матриц $Scale$, визуализирующих зависимости $CSYield = f(CSImicro)$, были заполнены вычисленными значениями индексов $CSImicro$ и $CSYield$ для вариантов опыта.

Для вычисления индекса $CSImicro$, характеризующего микробиологическую активность почвы, использовали фрактальные профили физиологических групп почвенных микроорганизмов (рис. 1). Алгоритм преобразования фрактальных профилей в индекс

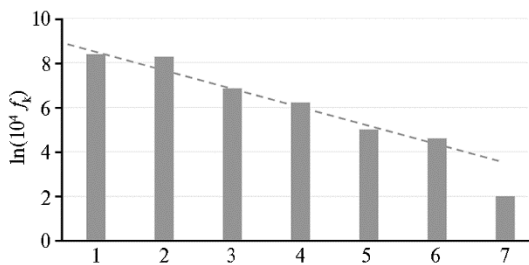


Рис. 1. Фрактальный профиль частот физиологических групп почвенных микроорганизмов (f_k) для варианта опыта 1 (29): 1 — использующие минеральный азот, 2 — эпифиты, 3 — актиномицеты, 4 — аммонификаторы, 5 — споровые, 6 — маслянокислые, 7 — грибы. Пунктирная линия обозначает степенной числовой ряд.

CSImicro априори не известен. Нейросеть (рис. 2) позволяет найти такой алгоритм в процессе ее обучения.

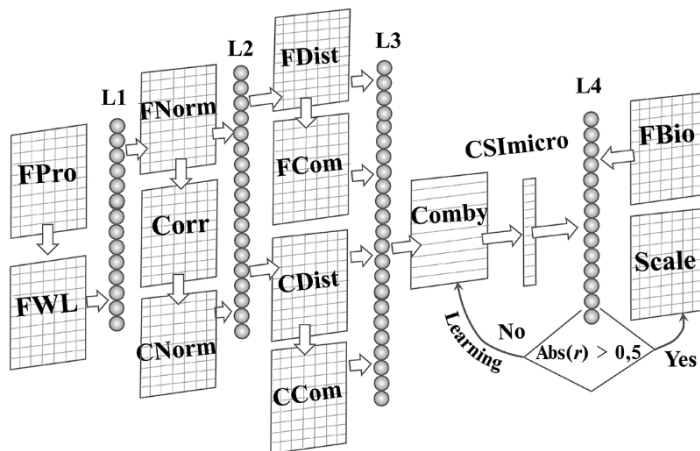


Рис. 2. Нейросеть, вычисляющая индекс CSImicro: L1-L4 — слои искусственных нейронов; FPro, FWL — матрицы исходных и WaveLet-преобразованных фрактальных профилей физиологических групп почвенных микроорганизмов; FNorm, CNorm — матрицы, содержащие нормализованные данные матрицы FWL и матрицы FNorm × Corr; Corr — матрица коэффициентов корреляции между строками матрицы FNorm; FDist, CDist — матрицы евклидовых дистанций между столбцами матриц FNorm, CNorm; FCom, CCom — матрицы собственных векторов матриц FDist, CDist; Scale — матрица шкалирования цифровых данных; Learning — цикл обучения-поиска алгоритма нейросети; Comby — база алгоритмов нейросети; FBio — матрица, включающая физико-биохимические почвенные данные и индекс CSIyield.

Схема полевого эксперимента, положенного нами в основу нейросетевого анализа, представлена в таблице 1 (29).

1. Схема длительного полевого эксперимента с выращиванием ячменя и картофеля (29), результаты которого использовались в качестве эмпирических данных при нейросетевом анализе

Изменяемые условия	Ячмень				Картофель			
	бесменная	— NPK	севооборот	— NPK	бесменная	— NPK+ TMAУ	севооборот	— NPK+ TMAУ
Агротехнология	—	—	—	—	—	—	—	—
Удобрения	—	—	—	—	—	—	—	—
Номер варианта опыта	1	2	3	4	5	6	7	8

Примечание. TMAУ — торфяно-микробиологическое азотное удобрение. Прочерки означают, что удобрения не использовались.

Значения численности физиологических групп почвенных микроорганизмов (29), использованные при нейросетевом анализе, приведены в таблице 2, которая дополнена индексами CSImicro и CSIyield и коэффициентами их корреляции с микробиологическими показателями.

2. Частота встречаемости физиологических групп почвенных микроорганизмов (29) и коэффициенты корреляций (r) с индексами CSImicro (M) и CSIyield (У)

Группа, ×10 ⁶ КОЕ/г почвы	Вариант опыта								r c	
	1	2	3	4	5	6	7	8	M	У
Аммонификаторы на МПА	0,050	0,166	0,167	0,232	0,198	0,419	0,324	0,354	0,60	—
Маслянокислые бактерии на картофеле	0,010	0,004	0,003	0,004	0,003	0,002	0,002	0,001	—	-0,60
Использующие минеральный азот на КАА	0,437	0,436	0,379	0,424	0,370	0,227	0,274	0,277	—	—
Актиномицеты на КАА	0,095	0,073	0,072	0,030	0,076	0,031	0,065	0,057	-0,73	—
Грибы на среде Чапека, ×10 ⁻³	0,75	0,33	0,43	0,30	0,50	0,35	0,21	0,47	—	—
Эпифиты на капустной среде	0,392	0,316	0,368	0,302	0,343	0,316	0,327	0,303	-0,58	—
Споровые на МПА	0,015	0,006	0,011	0,008	0,008	0,005	0,009	0,007	—	-0,50

Индекс CSImicro, ±0,1	3,9	5,0	3,9	8,8	4,3	5,3	4,0	4,7	–
Индекс CSYield, ±0,1	1,0	8,2*	3,2*	11,2*	1,0	8,5**	7,1**	33,6**	–

Примечание. МПА — мясопептонный агар, КАА — крахмалоаммиачный агар. Прочерки означают, что статистически значимые корреляции не обнаружены.

* Отношение урожая ячменя в опытном варианте к контрольному урожаю (3,1 ц/га, без применения удобрений и севооборотов).

** Отношение урожая картофеля в опытном варианте к контрольному урожаю (9,5 ц/га, без применения удобрений и севооборотов).

Почвенные физико-биохимические характеристики (29) были дополнены коэффициентами их корреляции с индексами CSImicro и CSYield (табл. 3).

3. Физико-биохимические характеристики почв (29) и коэффициенты корреляций (r) с индексами CSImicro (M) и CSYield (Y)

Параметр	Вариант опыта								r c	
	1	2	3	4	5	6	7	8	M	Y
Общая биогенность, млн/г	98,5	149,7	86,2	141,6	89,5	156,1	94,6	125,9	0,60	–
Содержание гумуса, %	1,72	2,51	2,01	3,77	1,97	2,15	1,29	2,19	0,94	–
Выделение CO ₂ , мг/кг	2,7	3,95	3,34	4,94	3,33	8,84	2,12	3,83	–	–
Разложение целлюлозы, %	8,0	17,7	6,3	12,1	10,9	78,9	10,6	63,1	–	0,57
Нитрифицирующая способность, мг NO ₃ /100 г	40,7	97,7	50,1	138,0	43,6	275,4	57,5	177,2	–	–
Полифенолоксидазы, мг/г	31,5	38,7	51,0	41,4	24,5	33,0	35,5	32,7	–	–
Пероксидазы, мг/г	35,4	36,0	42,0	31,8	28,5	29,1	37,5	37,8	–	–
Инвертазы, мг/г	10,6	11,9	9,7	12,1	10,9	46,0	15,2	34,0	–	–
Уреазы, мг/100 г	15,6	22,0	15,6	20,0	13,5	35,6	16,4	23,5	–	–

Примечание. Прочерки означают, что статистически значимые корреляции не обнаружены. Значения индексов CSImicro и CSYield представлены в таблице 2.

Матрицы Scale (рис. 3) демонстрируют немонотонный экстремальный тренд зависимости CSYield = f(CSImicro).

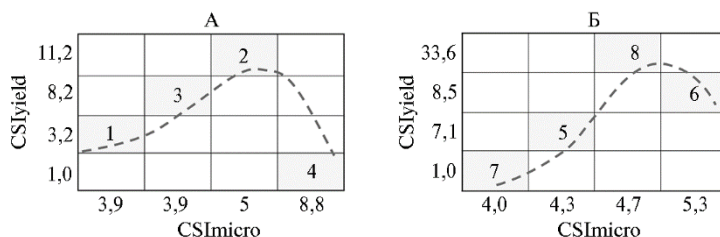


Рис. 3. Матрицы Scale, визуализирующие зависимость CSYield = f(CSImicro) для вариантов 1-4 (А) и вариантов 5-8 (Б). Порядковые номера вариантов опыта (см. табл. 1) помещены в ячейки матриц Scale в соответствии с координатами ячеек CSYield;CSImicro. Пунктирными линиями обозначен немонотонный экстремальный тренд зависимостей.

В работе Ю.М. Возняковской с соавт. (29) в качестве наиболее информативных показателей плодородия почвы были выбраны общая биогенность почвы, видовое разнообразие почвенных микроорганизмов, соотношение трофических групп микроорганизмов, граф трофических связей. Мы для описания сообщества почвенных микроорганизмов под растениями предложили индекс CSImicro, который был вычислен по полученным ранее микробиологическим показателям (29) с помощью нейросети, разработанной нами. Это позволило ранжировать агротехнологии на основе значений CSImicro. В вариантах опыта 1-4 (29) максимальный индекс почвенного плодородия CSYield = 11,2 достигается при CSImicro = 5,0. В вариантах 5-8 максимальный индекс почвенного плодородия CSYield = 33,6 достигается при CSImicro = 4,7. Максимальный индекс микробиологической активности почвы CSImicro = 8,8 в вариантах 1-4 достигается при CSYield = 1,0 < 11,2 (максимальный CSYield), а в вариантах 5-8 максимальный индекс CSImicro = 5,3 достигается при CSYield = 8,5 < 33,6 (максимальный

CSIyield). В рассмотренном эксперименте (29) для зерновой культуры индекс CSIyield был максимальным при бессменной агротехнологии, а для корнеплодной — в севообороте. Возможно, у зерновых культур из-за менее интенсивного выделения корневых минеральных и органических экссудатов — сахаров, аминокислот и органических кислот (56, 57) для формирования сообщества почвенных микроорганизмов, эффективно взаимодействующих с растениями, требуется период, превышающий один сезон вегетации. У корнеплодных культур при бессменной агротехнологии повышается вероятность поражения фитопатогенами, что приводит к снижению урожайности (58-61).

В таблицах 2, 3 приведены статистически значимые ($p < 0,05$) коэффициенты корреляции вычисленных значений индексов CSImicro и CSIyield с микробиологическими и физико-биохимическими характеристиками почв. Обращает на себя внимание значимая корреляция индекса CSImicro с частотой встречаемости в почве аммонификаторов ($r = 0,60$; см. табл. 2), актиномицетов ($r = -0,73$; см. табл. 2) и эпифитов ($r = -0,58$, см. табл. 2). Причем повышение микробиологической активности почвы (увеличение индекса CSImicro) сопровождается уменьшением численности актиномицетов и эпифитов и увеличением численности аммонификаторов, что подтверждает ключевую роль аммонификаторов в круговороте азота в аграрных экосистемах и питании минеральным азотом растений (43). Высокая положительная корреляция индекса CSImicro с содержанием гумуса в почве ($r = 0,94$; см. табл. 3) свидетельствует о биоконсолидации почвенных микроорганизмов в направлении гумификации растительных остатков и повышения почвенного плодородия (62-64).

Итак, в Excel-среде создана нейросеть, которая преобразованием фрактальных профилей физиологических групп почвенных микроорганизмов вычисляет безразмерный индекс CSImicro, представляющий микробиологическую активность почвы. В нейросетевых вычислениях матрицу Corr коэффициентов корреляции признаков микробных биосистем формирует корреляционный анализ, матрицы FDist и CDist, содержащие евклидовы дистанции между фрактальными профилями почвенных микроорганизмов по вариантам наблюдения, — кластерный анализ, матрицы FCom и CCom, представляющие ортогональные и ранжированные системы координат (системы главных компонент) в многомерном метрическом пространстве фрактальных профилей микроорганизмов, — дискриминантный анализ. Выходные данные этих анализов преобразуются в индекс CSImicro по алгоритму, найденному в процессе обучения нейросети. Зависимость CSIyield = f (CSImicro) была визуализирована с помощью матрицы ранжирования Scale, где индекс CSIyield получали, как отношение урожая культуры в опыте к урожаю в контроле. Разработанный нами метод нейросетевой анализ позволил извлечь ранее недоступную информацию о микробиологической активности почвы (CSImicro) и зависимости урожая растений от CSImicro. Нейросетевая обработка почвенных микробиологических данных, полученных в длительном опыте (1962-1991 годы, Северо-Западный регион Российской Федерации), показала, что максимальное значение индекса почвенного плодородия CSIyield не совпадает с максимальным значением индекса микробиологической активности почвы CSImicro. Предложенные оригинальные индексы позволяют оптимизировать агротехнологии применения органоминеральных удобрений и бактериальных препаратов для получения стабильно высоких урожаев сельскохозяйственных культур.

¹ФГБНУ Всероссийский НИИ сельскохозяйственной микробиологии,

196608 Россия, г. Санкт-Петербург, ш. Подбельского, 3,
e-mail: Nik.IvanVorobyov@yandex.ru ✉;

²ФГБОУ ВО Московская государственная академия ветеринарной медицины и биотехнологии им. К.И. Скрябина,

109472 Россия, г. Москва, ул. Академика Скрябина, 23

Поступила в редакцию

28 января 2024 года

Принята к публикации

16 октября 2024 года

Sel'skokhozyaistvennaya biologiya [Agricultural Biology], 2025, V. 60, № 1, pp. 70-81

Dedicated to the memory of Yu.M. Voznyakovskaya

THE NEURAL NETWORK RANKING OF AGRO-TECHNOLOGIES BY INDICES OF SOIL MICROBIOLOGICAL ACTIVITY AND SOIL FERTILITY: NEW POSSIBILITIES OF STATISTICAL ANALYSIS

N.I. Vorobyov^{1, 2} ✉

¹*All-Russian Research Institute for Agricultural Microbiology*, 3, sh. Podbel'skogo. St. Petersburg, 196608 Russia, e-mail Nik.IvanVorobyov@yandex.ru (✉ corresponding author);

²*Skryabin Moscow State Academy of Veterinary Medicine and Biotechnology*, 23, Akademika Skryabina, Moscow, 109472 Russia

ORCID:

Vorobyov N.I. orcid.org/0000-0001-8300-2287

The author declares no conflict of interests

Acknowledgements:

Supported financially by Russian Science Foundation (project № 23-26-00234)

Final revision received January 28, 2024

doi: 10.15389/agrobiology.2024.5.70eng

Accepted February October 16, 2024

Abstract

Neural networks make it possible to extract previously inaccessible information from physiological and molecular genetic data and visualize implicit relationships. In the presented work we applied the neural network approach to analyze the data of Yu.M. Vozniakovskaya et al. (*Sel'skokhozyaistvennaya Biologiya/Agricultural Biology*, 1994) obtained in a long-term experiment (1962-1991, North-West region of the Russian Federation) on the influence of crop rotations and mineral fertilizers on microbiological and biochemical characteristics of sod-podzolic soils under potato and barley. The experiment, the aim of which was to identify microbiological indicators that most clearly characterize the level of soil fertility, determined the composition of soil microorganisms, measured the intensity of CO₂ release, cellulose decomposition, activity of soil polyphenol oxidases and peroxidases as indicators of the intensity of plant residue humification, as well as invertases and ureases to assess the accumulation of nitrates and ammonium in the soil. On this basis, the authors attributed the most informative microbiological indicators of the level of effective soil fertility to the general soil biogenicity, species diversity of soil microorganisms, the ratio of trophic groups of microorganisms, and the graph of trophic relationships. In the present work, the analysis of fractal profiles of microbial physiological groups isolated on selective media demonstrated unique possibilities of statistical analysis using computational neural networks. Their application for processing empirical microbiological and physico-biochemical soil data reported by Yu.M. Vozniakovskaya et al., 1994 allowed determination of the specificity of the impact of the used agrotechnology elements on soil fertility under different crops. Soil microbiological activity and soil fertility are data characterizing heterogeneous biological objects? Soil microbial community (CSI_{micro} index) and plants (CSI_{yield} index), where CSI means Cognitive Saliency Index. Therefore, to visualize the function $CSI_{yield} = f(CSI_{micro})$ we used neural network with the construction of Scale matrix the cells of which with coordinates CSI_{yield};CSI_{micro} were filled with corresponding numbers of treatment variants. For this purpose, fractal profiles of soil microbial physiological groups of microorganisms obtained in the experiment were subjected to neural network analysis. As a result, from the Scale matrix, it was possible to estimate the function $CSI_{yield} = f(CSI_{micro})$ and the influence of fertilizer and crop rotation agrotechnologies on the intensity of humus accumulation in soil and soil fertility. It was found that the maximum soil fertility index CSI_{yield} did not coincide with the maximum value of soil microbiological activity index CSI_{micro}. The original indices proposed by us allow optimization of agrotechnologies for organ mineral fertilizers and bacterial preparations to obtain stably high yields of agricultural crops.

Keywords: soil microbiological activity, soil fertility, fractal analysis of microbiota, correlation analysis, cluster analysis, discriminant analysis, computational neural network.

REFERENCES

1. Aristovskaya T.V. *Mikrobiologiya protsessov pochvoobrazovaniya* [Microbiology of soil formation processes]. Leningrad, 1980 (in Russ.).
2. Fedotov G.N., Lysak L.V., Shalaev V.S. *Lesnoy vestnik*, 2013, 7: 111-115 (in Russ.).
3. Shlegel' G.G. *Istoriya mikrobiologii* [History of microbiology]. Moscow, 2020 (in Russ.).
4. Ben-Jacob E. Bacterial self-organization: coenhancement of complexation and adaptability in a dynamic environment. *Phil. Trans. R. Soc. Lond. A*, 2003, 361: 1283-1312.
5. Young I.M., Crawford J.W. Interactions and self-organization in the soil-microbe complex. *Science*, 2004, 304(5677): 1634-1637 (doi: 10.1126/science.1097394).
6. Crawford J.W., Deacon L., Grinev D., Harris J.A., Ritz K., Singh B.K., Young I. Microbial diversity affects self-organization of the soil-microbe system with consequences for function. *Journal of the Royal Society Interface*, 2012, 9(71): 1302-1310 (doi: 10.1098/rsif.2011.0679).
7. Vorobyov N.I., Egorov I.A., Kochish I.I., Nikonov I.N., Lenkova T.N. Fractal analysis of frequency-taxonomic profile of broiler's gut microbiota for studying the influence of probiotics on bird development. *Sel'skokhozyaistvennaya biologiya* [Agricultural Biology], 2021, 56(2): 400-410 (doi: 10.15389/agrobiologiya.2021.2.400rus).
8. Kochish I.I., Vorob'ev N.I., Nikonov I.N., Selina M.V. *Veterinariya i Zootekhniya*, 2022, 12: 57-65 (in Russ.).
9. Petrov V.B., Chebotar' V.K. Management of destruction and humification of the post-harvest rests of cereal crops using microbiological preparation Extrasol. *Sel'skokhozyaistvennaya biologiya* [Agricultural Biology], 2012, 3: 103-108 (in Russ.).
10. Matyuk N.S., Shevchenko V.A., Mazirov M.A., Polin V.D., Nikolaev V.A., Savos'kina O.A., Chebanenko S.I. *Monitoring i upravlenie plodorodiem pochv v agroekosistemakh: elektronnoe uchebnoe posobie* [Monitoring and management of soil fertility in agroecosystems: electronic textbook]. Moscow, 2022 (in Russ.).
11. Berestetskiy O.A., Voznyakovskaya Yu.M., Dorosinskiy L.M. *Biologicheskie osnovy plodorodiya pochvy* [Biological bases of soil fertility]. Moscow, 1984 (in Russ.).
12. Kruglov Yu.V. Microbial community of soil: physiological diversity patterns and assessment (review). *Sel'skokhozyaistvennaya biologiya* [Agricultural Biology], 2016, 51(1): 46-59 (in Russ.).
13. Semenov A.M., Shatalov A.A., Semenova E.V. *Uspekhi sovremennoy biologii*, 2022, 142(6): 591-602 (in Russ.).
14. Emtsev V.T., Mishustin E.N. *Mikrobiologiya* [Microbiology]. Moscow, 2005 (in Russ.).
15. Ovchinnikova T.A., Pankratov T.A. *Metody ekologii pochvennykh mikroorganizmov: uchebnoe posobie* [Methods of ecology of soil microorganisms: a tutorial]. Samara, 2009 (in Russ.).
16. *Osnovnye mikrobiologicheskie i biokhimicheskie metody issledovaniya pochv* /Pod redaktsey Yu.M. Voznyakovskoy [Basic microbiological and biochemical methods of soil research. Yu.M. Voznyakovskaya (ed.)]. Leningrad, 1987 (in Russ.).
17. Tepper E.Z., Shil'nikova V.K., Pereverzeva G.I. *Praktikum po mikrobiologii* [Microbiology workshop]. Moscow, 2004 (in Russ.).
18. Netrusov A.I., Bonch-Osmolovskaya E.A., Gorlenko V.M. *Ekologiya mikroorganizmov* [Ecology of microorganisms]. Moscow, 2004 (in Russ.).
19. Manucharova N.A. *Molekulyarno-biologicheskie metody v pochvovedenii i ekologii* [Molecular biological methods in soil science and ecology]. Moscow, 2014 (in Russ.).
20. Chubenko G.I. *Metody identifikatsii bakteriy* [Methods of bacterial identification]. Blagoveshchensk, 2018 (in Russ.).
21. Shcherbakov A.V. *Endofitnye soobshchestva sfagnovykh mkhov kak istochnik bakteriy – effektivnykh assotsiantov sel'skokhozyaystvennykh kul'tur. Kandidatskaya dissertatsiya* [Endophytic communities of sphagnum mosses as a source of bacteria – effective associates of agricultural crops. PhD Thesis]. St. Petersburg, 2013 (in Russ.).
22. Korzhov S.I., Maslov V.A., Orekhova E.S. *AgroXXI*, 2009, 1-3: 20090103022.
23. Voznyakovskaya Yu.M., Popova Zh.P. *Sel'skokhozyaistvennaya biologiya* [Agricultural Biology], 1989, 5: 86-89 (in Russ.).
24. Dreyper N., Smit G. *Prikladnoy regressionnyy analiz. Mnozhestvennaya regressiya* [Applied regression analysis. Multiple regression]. Moscow, 2007 (in Russ.).
25. Shafran S.A. *Pochva, udobreniya, urozhay* [Soil, fertilizers, harvest]. Moscow, 2019 (in Russ.).
26. Sutrop U. List Task and a Cognitive Salience Index. *Field methods*, 2001, 13(3): 263-276 (in Russ.).
27. Tolstova Yu.N. *Osnovy mnogomernogo shkalirovaniya* [Basics of multidimensional scaling]. Moscow, 2006 (in Russ.).
28. Ermolaev O.Yu. *Matematicheskaya statistika dlya psikhologov* [Mathematical statistics for psychologists]. Moscow, 2003 (in Russ.).
29. Voznyakovskaya Yu.M., Popova Zh.P., Vorob'ev N.I. *Sel'skokhozyaystvennaya biologiya* [Agricultural Biology], 1994, 5: 84-90 (in Russ.).
30. Shreder M. *Fraktaly, khaos, stepennyye ryady* [Fractals, chaos, power series]. Izhevsk, 2001 (in Russ.).
31. Mitsel' A.A. *Vychislitel'nye metody* [Computational methods]. Tomsk, 2013 (in Russ.).

32. Vakulenko S.A., Zhikhareva A.A. *Prakticheskiy kurs po neyronnym setyam* [Practical course on neural networks]. St. Petersburg, 2018 (in Russ.).
33. Dyrmann M., Karstoft H., Midtby H.S. Plant species classification using deep convolutional neural network. *Biosystems Engineering*, 2016, 151: 72-80.
34. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw.*, 2015, 61(1): 85-117 (doi: 10.1016/j.neunet.2014.09.003).
35. Pogodaev A.K., Khabibullina E.L., Inyutin D.M. *Prikladnaya matematika i voprosy upravleniya*, 2021, 2: 73-92 (in Russ.).
36. Kulakov K.A., Dimitrov V.M. *Osnovy testirovaniya programmnoho obespecheniya* [Software testing basics]. Petrozavodsk, 2018 (in Russ.).
37. Gafarov F.M., Galimyanov A.F. *Iskusstvennye neyronnye seti i ikh prilozheniya* [Artificial neural networks and their applications]. Kazan', 2018 (in Russ.).
38. Kruglov V.V., Borisov V.V. *Iskusstvennye neyronnye seti. Teoriya i praktika* [Artificial neural networks. Theory and practice]. Moscow, 2002 (in Russ.).
39. Kim Dzh.-O., Myuller Ch.U., Klekka U.R., Oldenderfer M.S., Bleshfild R.K. *Faktornyy, diskriminantnyy i klasternyy analiz* [Factor, discriminant and cluster analysis]. Moscow, 1989 (in Russ.).
40. Krivtsova N.I., Moyzes O.E. *Dopolnitel'nye glavy matematiki. Statisticheskiy analiz* [Additional chapters in mathematics. Statistical analysis]. Tomsk, 2015 (in Russ.).
41. Tyurin V.V., Shcheglov S.N. *Diskriminantnyy analiz v biologii: monografiya* [Discriminant analysis in biology: monograph]. Krasnodar, 2015 (in Russ.).
42. Vorob'ev N.I., Pishchik V.N., Provorov N.A., Sviridova O.V. *Agrofizika*, 2014, 4(16): 35-41 (in Russ.).
43. Kutuzova R.S., Sirota L.B., Orlova O.V., Vorob'ev N.I. *Agrokimiya*, 2001, 1: 19-33 (in Russ.).
44. Ivanishchev V.V. *Izvestiya TulGU*, 2018, 1: 69-77 (in Russ.).
45. Karasev E.S., Chizhevskaya E.P., Simarov B.V., Provorov N.A., Andronov E.E. Comparative phylogenetic analysis of symbiotic genes of different nodule bacteria groups using the metatrees method. *Sel'skokhozyaistvennaya biologiya [Agricultural Biology]*, 2017, 52(5): 995-1003 (doi: 10.15389/agrobio.2017.5.995rus).
46. Pershina E.V., Andronov E.E., Pinaev A.G., Akhtemova G.A., Doumova V.A., Provorov N.A. The use of T-RFLP method for studying the dynamics of soil microbial communities under xenobiotics treatment. *Sel'skokhozyaistvennaya biologiya [Agricultural Biology]*, 2011, 3: 81-87 (in Russ.).
47. Samusik E.A., Marchik T.P., Golovaty S.E. *Zhurnal Belorusskogo gosudarstvennogo universiteta. Ekologiya*, 2019, 3: 65-79 (in Russ.).
48. Khaziev F.Kh. *Ekobiotekh*, 2018, 1(2): 80-92 (doi: 10.31163/2618-964X-2018-1-2-80-92) (in Russ.).
49. Sal'nikova E.V. *Khimicheskiy analiz ob'ektov okruzhayushchey sredy* [Chemical analysis of environmental objects]. Orenburg, 2021 (in Russ.).
50. D'yakonov V.P. *Veyvlety. Ot teorii k praktike* [Wavelets. From theory to practice]. Moscow, 2004 (in Russ.).
51. Mascarenhas, WF. Fast and accurate normalization of vectors and quaternions. *Comp. Appl. Math.*, 2018, 37: 4649-4660.
52. Nikoli D., Muresan R.C., Feng W., Singer W. Scaled correlation analysis: a better way to compute a cross-correlogram. *European Journal of Neuroscience*, 2012, 35(5): 1-21 (doi: 10.1111/j.1460-9568.2011.07987.x).
53. Batushansky A., Toubiana D., Fait A. Correlation-based network generation, visualization, and analysis as a powerful tool in biological studies: a case study in cancer cell metabolism. *Biomed. Res. Int.*, 2016, 2016: ID8313272 (doi: 10.1155/2016/8313272).
54. Everitt B.S., Landau S., Leese M., Stahl D. *Cluster analysis*. John Wiley & Sons, 2011.
55. Markova L.V., Korchevskaya E.A. *Chislennyye metody nakhozhdeniya sobstvennykh vektorov i sobstvennykh znacheniy matrits* [Numerical methods for finding eigenvectors and eigenvalues of matrices]. Vitebsk, 2011 (in Russ.).
56. Okubo A., Matsusaka M., Sugiyama S. *Rasteniya i pochva*, 2016, 399 (1-2): 345-356 (in Russ.).
57. Kravchenko L.V. *Rol' kornevykh ekzometabolitov v integratsii mikroorganizmov s rasteniyami. Avtoreferat doktorskoy dissertatsii* [The role of root exometabolites in the integration of microorganisms with plants. PhD Thesis]. Moscow, 2000 (in Russ.).
58. Belosokhov A.F., Yarmeeva M.M., Mislavskiy S.M., Kurchaev M.L., Dolgov A.M., Albantov G.P., Skokov D.N., Tsindeliani A.A., Kokaeva L.Yu., Chudinova E.M., Elanskiy S.N. *Mikologiya i fitopatologiya*, 2023, 57(2): 123-133 (in Russ.).
59. Bogoutdinov D.Z., Fominykh T.S., Kastal'eva T.B., Girsova N.V., Pavlovskaya N.E., Gagarina I.N., Mishurov N.P., Nemenushchaya L.A., Piskunova N.A. *Metody diagnostiki vozбудiteley zabolevaniy ovoshchnykh kul'tur: analiticheskiy obzor* [Methods of diagnostics of pathogens of vegetable crops: an analytical review]. Moscow, 2020 (in Russ.).
60. Pavlyusin V.A., Novikova I.I., Boikova I.V. Microbiological control in phytosanitary optimization technologies for agroecosystems: research and practice (review). *Sel'skokhozyaistvennaya biologiya [Agricultural Biology]*, 2020, 55(3): 421-438 (doi: 10.15389/agrobio.2020.3.421rus).

61. Kolesnikov L.E., Balagurova E.D., Popova E.V., Novikova I.I., Kolesnikova Y.R. Application of chitosan to protect wheat from diseases and boost yields. *Appl. Biochem. Microbiol.*, 2022, 58(3): 329-335.
62. Zhemyakin S.V., Vorobyov N.I., Popov A.A., Orlova O.V., Rusakova I.V. Bioconsolidation of soil microflora and subsequent increase in the efficiency of humification of cereals plant residues. *BIO Web of Conferences 'AgriScience2024'*, 2024, 139 (doi: 10.1051/bioconf/202413904009).
63. Moskovkin V.V. *Ekologo-agrokhimicheskaya otsenka vliyaniya mikrobiologicheskikh preparatov-destruktorov rastitel'nykh ostatkov zernovykh kul'tur v agrotsenozakh na dernovo-podzolistykh supeschanykh pochvakh. Kandidatskaya dissertatsiya* [Ecological and agrochemical assessment of the influence of microbiological preparations-destructors of plant residues of grain crops in agrocenoses on sod-podzolic sandy loam soils. PhD Thesis]. Vladimir, 2018 (in Russ.).
64. Masyutenko N.P., Pankova T.I., Kuznetsov A.V., Masyutenko M.N., Breskina G.M., Chuyan N.A. *Yug Rossii: ekologiya, razvitie*, 2021, 16(2): 108-118 (doi: 10.18470/1992-1098-2021-2-108-118) (in Russ.).