

## МЕТОДИ ДОСЛІДЖЕНЬ

УДК 51-78:57.087:577.1:581.1

### ОБ ИСПОЛЬЗОВАНИИ СТАТИСТИЧЕСКИХ МЕТОДОВ В СТРЕССОВОЙ ФИЗИОЛОГИИ И СЕЛЕКЦИИ РАСТЕНИЙ

© 2018 г. В. В. Иванищев

*Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Тульский государственный педагогический университет им. Л.Н. Толстого»  
(Тула, Россия)*

Рассмотрены результаты применения статистических методов (главных компонент и кластерного анализа) при изучении физиолого-биохимических показателей растений в связи с проблемами устойчивости к различным факторам среды и селекции. Показано, что использование таких методов в области физиологии и биохимии растений все еще недостаточно. Одной из причин является сложность и не всегда понятная биологическая интерпретация математических преобразований результатов исследования. При этом также следует учитывать, что, во-первых, использование таких методов может приводить к частичной потере данных (могут учитываться не все возможные варианты связей между изученными параметрами), а, во-вторых, статистические методы основаны на том, что изучаемые признаки подчиняются так называемому нормальному распределению, что не всегда соответствует биологическим процессам, явлениям и объектам. Тем не менее, такие подходы к обработке результатов могут быть полезны для обнаружения новых взаимосвязей между известными физиолого-биохимическими характеристиками растений. Это важно для преодоления субъективности в толковании результатов исследования и выявления вклада (и/или специфичности) отдельных показателей в устойчивость растений к стрессорам разной природы. В работе приведены примеры применения метода главных компонент и кластерного анализа в конкретных исследованиях. Показано, что первый метод используется в меньшей степени, так как его применение требует определенного навыка при анализе результатов. Метод кластерного анализа применяют более часто, поскольку он позволяет построить достаточно ясную иерархическую структуру между исследованными параметрами, а также обнаружить наибольшую близость (взаимосвязь) между изученными показателями или объектами. Показано, что метод кластерного анализа может быть успешно использован при решении задач, связанных с выбором наиболее подходящих форм растений для дальнейшей селекционной работы, а также при оценке взаимосвязей и потенциальных взаимозависимостей между физиолого-биохимическими показателями растений в условиях стресса.

**Ключевые слова:** *статистические методы, метод главных компонент, кластерный анализ, физиолого-биохимические показатели, стресс, засуха, солеустойчивость, селекция*

Изучение стресса у растений находится в фокусе внимания ученых на протяжении многих лет. Многообразие условий, в которых выращивают растения, приводит к накоплению огромного количества экспериментального ма-

териала. При этом одни и те же факты могут иметь не всегда однозначную интерпретацию. В то же время, похоже, что разнообразные стрессы часто активируют одни и те же гены с аналогичными изменениями физиолого-биохимических реакций (Иванищев, Жуков, 2016). Такая картина существенно затрудняет понимание специфики физиолого-биохимических изменений в растениях. В результате представления о механизмах противо-

*Адрес для корреспонденции:* Иванищев Виктор Васильевич, Тульский государственный педагогический университет им. Л.Н. Толстого, пр. Ленина, 125, Тула, 300026, Россия;  
e-mail: avdey\_VV@mail.ru

действия растительного организма стрессу оказываются сильно размытыми. Это затрудняет поиски методов и технологий управления жизнедеятельностью растений и преодоления ими последствий стрессового состояния. В то же время особую ценность подобные исследования приобретают в случаях их практического приложения, связанного, например, с оценкой генотипов селекционного материала, биологической или хозяйственной продуктивности, устойчивости к специфическим стрессовым условиям среды.

Результаты экспериментов в физиологии растений, как правило, подвергают статистической обработке с представлением соответствующих данных. При этом исследователи часто ограничиваются приведением сведений о достоверности полученных результатов в виде коэффициента Стьюдента или  $F$ -критерия и степени значимости/вероятности события:  $P < 0,05$ . Более глубокий математический анализ как правило не проводится. Одной из причин этого явления может служить, с одной стороны, недостаточная подготовка физиологов растений в области использования математических методов статистики, а с другой – сложность интерпретации результатов в соответствии со статистическими математическими моделями. Это связано с тем, что математические выкладки часто представляют собой только цифры/графики/диаграммы, за которыми необходимо обнаружить адекватный биологический смысл. Недаром в относительно недавней работе подчеркивалась необходимость расширения в использовании компьютерных технологий и программ в области физиологии растений (Fernie, 2011).

Сложившаяся ситуация приводит к тому, что каждый автор интерпретирует результаты в соответствии с собственными идеями и сложившимися в данной области науки представлениями. В результате вместо вероятностно-статистической модели, которая, как полагают современные исследователи, в наибольшей степени отражает реальное протекание процессов (Бурева, 2007), хотя и не всегда (Nath, 2017), мы имеем частное мнение автора, которое может быть весьма далёким от сущности явления.

Всё это показывает необходимость расширения арсенала исследователя в области физиологии и биохимии растений для обработки экспериментальных данных. В работе приведены примеры использования методов главных

компонент и кластерного анализа с соответствующей интерпретацией результатов.

### *Приложение метода главных компонент (РСА)*

Вначале следует отметить, что до применения подобных методов анализа часто оказывается важным провести нормирование полученных результатов, поскольку представляемые величины отдельных параметров (характеристик) могут различаться на несколько порядков (например, активность ферментов и содержание пигментов и т.п.). Такая нормировка может проводиться разными способами (Бурева, 2007), наиболее простым из которых является деление каждой величины на среднюю для этой серии наблюдений (например, изменение содержания хлорофилла для нескольких объектов или нескольких точек эксперимента). Тогда сильно различающиеся величины характеристик существенно сблизятся и будут удобны для анализа.

Метод главных компонент (Principal component analysis – PCA) является одним из методов многомерной статистики, с помощью которого рассматривают изменение самых разных свойств объекта. При этом каждая точка включает в себя сразу несколько характеристик. Например, для каждой концентрации вещества в среде, на которой выращивали растение, определена величина того или иного параметра (активность фермента, содержание метаболита и т.п.). В этом случае обобщенные данные невозможно вывести на один график. Поэтому можно попытаться использовать метод главных компонент (Шабанов, Кравченко, 2011). Он позволяет выявить в многообразии характеристик наиболее важные (с точки зрения математической статистики), которые имеют наибольшее значение в рамках рассматриваемой исследователем проблемы. При этом также не следует забывать, что такое выделение главных характеристик может приводить к потере части информации, что выражается в снижении внимания к другим свойствам объекта.

Использование программы STATISTICA и подобных ей вначале позволяет разместить экспериментальные точки в прямоугольных координатах, далее провести прямую – линию регрессии, на которую проецируется максимальная доля общей дисперсии (сигмы от всех точек). Ее и называют первой главной компонентой. Перпендикулярная ей – это вторая главная компонента (Шабанов, Кравченко, 2011). Даль-

нейший анализ таких данных позволяет решить поставленные исследователем задачи.

В работе (Liu et al., 2010) использовали метод главных компонент при анализе результатов исследования спектров отражения рисовых метелок в контроле и при повреждении насекомыми и микроорганизмами. Авторы отметили, что в условиях эксперимента для разных типов спектров были обнаружены всего одна или две главные компоненты, которые вносили наибольший вклад в изменение спектров. Однако более обстоятельного обсуждения в работе не обнаружено, как и в расширенном варианте подобных исследований (Liu et al., 2012).

Большой коллектив авторов из разных стран провел многофакторный анализ физиологических параметров для моделирования ответных реакций листьев взрослых растений европейского бука (*Fagus sylvatica*) на присутствие озона в среде (Löw et al., 2012). В ходе обработки данных изучали изменение величин главных компонент в зависимости от того, какие параметры добавляли или исключали из анализа. Таким образом обнаруживали наиболее значимые характеристики, которые оказывали наибольшее влияние на величины главных компонент, вследствие чего авторы делали вывод о значимости этих параметров растений для их устойчивости к присутствию озона в среде. Например, было показано, что первая компонента PC1 реагировала на наличие озона в среде и форму кроны деревьев и была связана с запуском окислительного стресса и изменениями в обмене углеводов. Поэтому, физиологическая интерпретация PC1 может быть получена из нагрузок, внесенных исходными переменными (набора использованных для расчета характеристик). Хотя в специализированных (более совершенных) программных продуктах, где наблюдается сходство в ответных реакциях объекта, аналогичный вывод можно делать уже в ходе проводимых расчетов.

В работе исследователей Республики Корея показано, что метод главных компонент позволяет полностью различать разные сорта риса (*Oryza sativa* L.) на основании полученных экспериментальных результатов о присутствии ряда метаболитов в пигментированных семенах (Kim et al., 2013). При этом большое значение придается веществам фенольной природы, которым приписывают ряд физиологически важных свойств, в т. ч. противоопухолевую активность. В связи с этим авторы попытались обнаружить корреляции между накоплением ве-

ществ указанной группы с образованием первичных метаболитов: аминокислот, сахаров, органических кислот. Показано, что на долю двух основных главных компонент приходится 56,8% от общей дисперсии в наборе данных. При этом авторы также последовательно исключали по одному из показателей, следя за изменениями величин главных компонент. Было обнаружено, что гликолевая кислота, глицерин,  $\gamma$ -аминомасляная кислота (ГАМК), фруктоза и галактоза оказывали наибольшее влияние на первую компоненту, в связи с чем был сделан вывод об их наибольшей значимости для синтеза веществ фенольной природы.

В работе (Chunthaburee et al., 2016) исследователи из Таиланда использовали метод главных компонент для оценки солеустойчивости разных сортов риса (*Oryza sativa* L.). Анализ физиолого-биохимических показателей растений позволил обнаружить коэффициенты корреляции между показателями, оценить индексы солеустойчивости по изученным характеристикам и с помощью метода главных компонент выявить характеристики растений, которые имели положительную и отрицательную корреляции в отношении устойчивости к соли. Дальнейший кластерный анализ позволил авторам выделить сорта, наиболее близкие по изменчивости изученных признаков в отношении к избытку соли в среде.

В работе других авторов (Nejat, Sadeghi, 2016) проведено изучение морфологических и биохимических показателей растений зизифуса (*Ziziphus spina-christi*) в условиях двух видов стресса: засухи и засоления. Показаны различия в ответных реакциях растений на два вида стресса, например, по накоплению общего хлорофилла, общего белка, ионов натрия и др. В статье авторы также ссылаются на подобные работы иных исследователей.

Можно ли аналогичную процедуру провести при анализе, например, изменчивости физиолого-биохимических признаков в разных условиях эксперимента (например, при разных концентрациях ионов металла в среде, разных температурных условиях, при разной временной экспозиции) – неизвестно, поскольку в литературе нами таких попыток пока не обнаружено. В упомянутой работе (Nejat, Sadeghi, 2016), например, отдельные условия отмечены лишь точками на координатах, что затрудняет понимание их интерпретации. В то же время в работе Chunthaburee и соавт. (2016) предложен коэффициент устойчивости к засолению и включен в общий анализ. В полученных коор-

динах вектор этого фактора совпадал или был противоположен векторам изменчивости других физиолого-биохимических признаков растений. Поэтому и вывод о связи отдельных физиолого-биохимических свойств с солеустойчивостью растений был очевиден. Возможно, что аналогичные коэффициенты необходимо рассчитывать и в иных исследованиях. В противном случае, по-видимому, отдельный вопрос здесь может состоять в том, на каком из этапов следует остановиться с тем, чтобы не искать биологический смысл там, где его может не быть.

### **Использование метода кластерного анализа**

Метод кластерного анализа исследователи применяют шире в силу более понятного подхода и интерпретаций. Его рекомендуют использовать в случаях, когда есть основания предполагать, что объекты распределены по иерархически соподчиненным группам. При этом основные этапы, которые необходимо пройти в этом случае (выбор объектов и признаков, метрики – меры сходства или различия объектов, процедура объединения), могут существенно влиять на конечный результат. Также следует помнить, что любые математические упрощения (допущения) могут приводить к частичной потере или некоторому искажению информации (Шабанов, Кравченко, 2011). Тем не менее, истина, как известно, постепенно проявляется по мере накопления экспериментальных данных. Поэтому интерес представляют все исследования, в которых такой метод использован разными авторами.

Так, в работе Naseem и Dutta (2007) данные о морфологических, химических, физиологических и биохимических свойствах растений гречихи (*Fagopyrum esculentum* и *F. tataricum*) были подвергнуты кластерному анализу с целью оценки сходства и изменчивости параметров их генотипов. Результаты работы позволили авторам разделить 12 сортов (генотипов) на две основные группы, которые были близки по ряду исследованных показателей. На основании этих данных авторы сделали заключение о том, какие параметры в большей степени отражают специфику генотипа, вследствие чего их можно использовать при проведении селекционной работы в выбранном направлении.

В работе Sharma и Bakshi (2014) проведена оценка клонов растений дальбергии *Dalbergia sissoo* по ряду характеристик, существенных для накопления качественной древесины растением. Корреляционный анализ дал

авторам возможность выделить наиболее важные из изученных показателей, а кластерный анализ – разбить исследованные клоны на группы и выбрать для дальнейшей селекционной работы наиболее ценные.

Исследование устойчивости линий арахиса (*Arachis hypogaea* L.) к засухе (de Lima Pereira et al., 2015) позволило группе авторов выделить три основных кластера генотипов, охарактеризованных по свойствам при использовании праймеров. Далее были изучены физиологические и биохимические показатели растений, подвергнутых водному стрессу. Отмечено, что в наибольшей степени генетическое разнообразие растений арахиса отражали такие признаки, как содержание свободного пролина, активность гваяколпероксидазы и содержание углеводов в корнях, содержание пролина в листьях, активность аскорбатпероксидазы в листьях и корнях, активность гваяколпероксидазы и содержание углеводов в листьях, содержание белка в корнях и активность каталазы в листьях.

В исследовании Ни и соавт. (2016) сравнительный анализ показателей фотосинтеза и содержания метаболитов позволил раскрыть механизм улучшения устойчивости к холоду у растений бермудской травы – свиного пальчатого (*Cynodon dactylon* (L), Pers.), который используют для покрытия полей для гольфа, футбола и т.п., путем обработки экзогенным мелатонином. Применение иерархического кластерного анализа дало возможность авторам представить модель, в которой отмечены наиболее характерные реакции чувствительных и устойчивых к холоду генотипов по таким показателям как уровень работы антиоксидантной системы, фотосистемы II, изменчивости метаболома, стабильности мембранных структур после обработки растений мелатонином.

Использование метода кластерного анализа наряду с методом главных компонент было проведено в выше упоминавшейся работе Chunthabugee и соавт. (2016) по изучению солеустойчивости сортов риса на стадии проростков. После обнаруженных методом главных компонент корреляций между солеустойчивостью и рядом физиологических и биохимических показателей авторы с помощью кластерного анализа выделили среди изученных сортов растений риса четыре основные группы, различающиеся по солеустойчивости. При этом отмечено, что ряд изученных физиологических параметров и величина активности антиоксидантных ферментов могут быть первичными

инструментами в отборе устойчивых форм растений риса для их дальнейшего использования в селекционных программах.

Кластерный анализ накопления разных метаболитов в окрашенных семенах риса был проведен, в также уже упоминавшейся выше работе, наряду с методом главных компонент (Kim et al., 2013). В результате исследования авторам удалось обнаружить наиболее взаимосвязанные между собой (наиболее близкие по изменчивости) группы метаболитов, важные для накопления веществ вторичной природы.

Попытки применения метода кластерного анализа для обнаружения наиболее тесно связанных между собой физиологических и биохимических характеристик проростков тритикале при кратковременном солевом (NaCl) стрессе предпринимались в ряде работ (Иванищев, Жуков, 2017; 2018; Иванищев, 2018а). Показано, что при изучении показателей водного обмена и фотосинтеза была обнаружена наиболее тесная связь между величиной транспирации и содержанием пролина, причем такие изменения коррелировали, в первую очередь, с накоплением ионов натрия и хлора и только потом – это отражалось на величине осмотического давления и содержания воды в проростках тритикале (Иванищев, Жуков, 2018). Изучение изменений показателей окислительного стресса и активности компонентов антиоксидантной защиты методом кластерного анализа позволило сделать вывод о том, что взаимопревращение восстановленной и окисленной форм глутатиона при солевом стрессе имеет более важное значение в первичных ответных реакциях организма, чем содержание аскорбиновой кислоты. Также была отмечена тесная взаимосвязь между активностью каталазы и величиной пероксидного окисления липидов в побегах тритикале (Иванищев, Жуков, 2017). Использование всех полученных результатов для анализа позволило выявить наиболее близкие, но отчасти неожиданные связи (корреляции) между такими показателями как (1) содержание пероксида водорода, скорость транспорта электронов, содержание воды и активность глутатионредуктазы; (2) образование супероксидного анион-радикала и активность аскорбатпероксидазы; (3) содержание основных пигментов фотосинтеза и активность гваяколпероксидазы; (4) величина пероксидного окисления липидов, активность каталазы и карбоангидразы (Иванищев, 2018а).

Последний пункт о близкой связи между активностью каталазы и карбоангидразы пред-

ставляет интерес с точки зрения изучения механизма выделения кислорода при фотосинтезе. Так, в недавнем обзоре отмечалось, что для решения вопроса могут быть полезны модельные эксперименты, в которых в качестве дополнительных участников реакции следует использовать карбоангидразу, каталазу, компоненты мембран (Иванищев, 2018б). Таким образом, результаты приложения метода кластерного анализа могут быть важными для постановки новых вопросов в области разных разделов физиологии и биохимии, а не только стресса.

Несколько иное приложение метод кластерного анализа нашел в попытке разбить широкий диапазон концентраций ионов тяжелого металла – никеля на недостаточные, оптимальные и токсические уровни концентраций в среде для прорастания семян вики (Абрамова, Иванищев, 2016) в соответствии с представлениями о двойной природе (эссенциальной – при низких концентрациях и токсической – при избытке) тяжелых металлов для жизнедеятельности организмов (Prasad, 2010). Показано, что концентрации хлорида никеля в области 5-50 мкМ можно отнести к зоне недостаточного и оптимального количества (зоны физиологической необходимости и толерантности) для прорастания семян вики. Концентрации в 100-1000 мкМ отнесены к зоне, где происходит нарушение гомеостаза, которое далее проявлялось в усилении окислительного стресса в побегах. Концентрации 5000-10000 мкМ были отнесены к токсическим.

### **Заключение**

Использование методов математики в области физиологии и биохимии растений в последние годы существенно расширяется. При этом некоторые авторы идут дальше: используют для анализа данных метод опорных векторов (SVM) (Liu et al., 2010), и далее применяют его вместе с методом главных компонент посредством коэффициентов (Liu et al., 2012). Необходимо признать, что использование таких методов математической статистики как PCA, кластерного анализа и других все еще достаточно ограничено. Это весьма затрудняет развитие представлений в отдельных областях науки о растениях и не способствует преодолению сложившихся взглядов и стереотипов о механизмах отдельных процессов и явлений. Более того, такая ситуация очень ограничивает возможности для моделирования процессов в растении, улучшения имеющихся моделей, а также не способствует совершенствованию и разработке новых способов управления жизне-

деятельностью растений и процессом селекции. Поэтому, можно констатировать, что в области физиологии и биохимии растений мы часто оперируем логическими схемами, а не адекватными (математически обоснованными) моделями, которые позволяли бы не только правильно оценивать полученный экспериментальный материал, но и предсказывать поведение растительного организма в разных условиях среды или эксперимента.

## ЛИТЕРАТУРА

- Абрамова Э.А., Иванищев В.В. 2016. Особенности прорастания семян вики в присутствии ионов никеля в среде. Известия ТулГУ. Естественные науки. 2-3 : 70-78. (Abramova E.A., Ivanishchev V.V. 2016. Characteristics of vetch seed germination in the presence of nickel ions in the medium. Izv. Tul'skogo Gos. Univer. Yestestvenniye Nauki. 2-3 : 70-78.).
- Буреева Н.Н. 2007. Многомерный статистический анализ с использованием ППП "STATISTICA". Нижний Новгород : 112 с. (Bureeva N.N. 2007. Multidimensional statistical analysis with an application of "STATISTICA" software package. (Mnogomernyi ststisticheski analiz s ispolzovaniem PPP "STATISTICA"). Nyzhnii Novgorod : 112 p.)
- Иванищев В.В. 2018а. Исследование влияния кратковременного солевого стресса методом кластерного анализа. Бутлеровские сообщения. 54 (4) : 134-139. (Ivanishchev V.V. 2018. Investigation of the effect of short-term salt stress with the method of cluster analysis. Butlerov Communications (Butlerovskie soobscheniya). 54 (4) : 134-139.)
- Иванищев В.В. 2018б. Проблемные вопросы в биохимии фотосинтеза. Вісн. Харків. нац. аграрн. ун-ту. Сер. Біологія. 1 (43) : 76-92. (Ivanishchev V.V. 2018b. Problematic questions in biochemistry of photosynthesis. Bull. Kharkiv. Natl. Agrar. Univ. Ser. Biology. (Visnyk Kharkiv. Natsional. Agram. Univer. Ser. Biologiya). 1 (43) : 76-92.)
- Иванищев В.В., Жуков Н.Н. 2016. Основы физиологии устойчивости растений. Тула : 84 с. (Ivanishchev V.V., Zhukov N.N. 2016. Basics of plant resistance physiology. (Osnovy fiziologii ustojchivosti rastenii). Tula : 84 p.)
- Иванищев В.В., Жуков Н.Н. 2017. Проявления окислительного стресса в проростках тритикале при кратковременном действии хлорида натрия. Бутлеровские сообщения. 52 (11) : 123-130. (Ivanishchev V.V., Zhukov N.N. 2017. Manifestations of oxidative stress in sprouts of triticale under condition of short-term exposure of sodium chloride. Butlerov Communications (Butlerovskie soobscheniya). 52 (11) : 123-130.)
- Иванищев В.В., Жуков Н.Н. 2018. О взаимосвязи показателей водного обмена и фотосинтеза проростков тритикале при кратковременном действии хлорида натрия. Бутлеровские сообщения. 53 (3) : 35-42. (Ivanishchev V.V., Zhukov N.N. 2018. On the interrelation of water exchange and photosynthesis in triticale sprouts with short-term action of sodium chloride. Butlerov Communications (Butlerovskie soobscheniya). 53 (3) : 35-42.)
- Орлов А. 2016. Как работает метод главных компонент (PCA) на простом примере. <https://habr.com/post/304214/>. (Orlov A. 2016. How does the principal component analysis (PCA) using a simple example)
- Шабанов Д.А., Кравченко М.А. 2011. Статистический анализ данных в зоологии и экологии. (Shabanov D.A., Kravchenko M.A. 2011. Statistical analysis of data in zoology and ecology). <https://batrachos.com/biostatistica>.
- Chunthaburee S., Dongsansuk A., Sanitchon J., Pattanagul W., Theerakulpisut P. 2016. Physiological and biochemical parameters for evaluation and clustering of rice cultivars differing in salt tolerance at seedling stage. Saudi J. Biol. Sci. 23 : 467-477.
- de Lima Pereira J.W., Alves da Silva E.C., Nunes da Luz L., Nogueira R.J.M.C., de Albuquerque Melo Filho P., de Lima L.M., Cavalcanti dos Santos R. 2015. Cluster analysis to select peanut drought tolerance lines. Austr. J. Crop Sci. 9 (11) : 1095-1105.
- Fernie A.R. 2011. Editorial overview – computational approaches in aid of advancing understanding in plant physiology. Front. Plant Sci. 2 : 5-7.
- Hu Z., Fan J., Xie Y., Amombo E., Liu A., Gitau M.M., Khaldun A.B.M., Chen L., Fu J. 2016. Comparative photosynthetic and metabolic analyses reveal mechanism of improved cold stress tolerance in bermudagrass by exogenous melatonin. Plant Physiol. Biochem. 100 : 94-104.
- Kim J.K., Park S.-Y., Lim S.-H., Yeo Y., Cho H.S., Ha S.-H. 2013. Comparative metabolic profiling of pigmented rice (*Oryza sativa* L.) cultivars reveals primary metabolites are correlated with secondary metabolites. J. Cereal Sci. 57 : 14-20.
- Liu Z., Li C., Wang Y., Huang W., Ding X., Zhou B., Wu H., Wang D., Shi J. 2012. Comparison of spectral indices and principal component analysis for differentiating lodged rice crop from normal ones. In: 5<sup>th</sup> Computer and Computing Technologies in Agriculture (CCTA), Oct 2011, Beijing, China. Springer, IFIP Advances in Information and Communication Technology, AICT-369 : 84-92.
- Liu Z., Shi J., Zhang L.W., Huang J.F. 2010. Discrimination of rice panicles by hyperspectral reflectance data based on principal component analysis and support vector classification. Zhejiang Univ-Sci B (Biomed & Biotechnol). 11 (1) : 71-78.
- Löw M., Deckmyn G., Op de Beeck M., Blumenröther M.C., Oßwald W., Alexou M., Jehnes S., Haberer K., Rennenberg H., Herbinger K., Häberle K.-H., Bahnweg G., Hanke D., Wieser G., Ceulemans R., R. Matyssek, Tausz M. 2012. Multi-

- variate analysis of physiological parameters reveals a consistent O<sub>3</sub> response pattern in leaves of adult European beech (*Fagus sylvatica*). *New Phytol.* 196 : 162-172.
- Naseem M., Dutta M., Shah S., Kumar P. 2007. Characterization of Buckwheat Cultivars Using Morphological, Chemical, Physiological and Biochemical Parameters. In: *Advances in Buckwheat Research: Proceedings of the 10<sup>th</sup> International Symposium on Buckwheat.* Chai Yan, Zhang Zongwen. Northwest A & F University Press:162-168.
- Nath S. 2017. Two-ion theory of energy coupling in ATP synthesis rectifies a fundamental flaw in the governing equations of the chemiosmotic theory. *Biophys. Chem.* 230 : 45-52.
- Nejat N., Sadeghi H. 2016. Finding out relationships among some morpho-biochemical parameters of Christ's thorn (*Ziziphus spina-christi*) under drought and salinity stresses. *Planta Daninha, Viçosa-MG.* 34 (4) : 667-674.
- Prasad M.N.V. 2010. *Heavy Metal Stress in Plants: From Biomolecules to Ecosystems.* Kindle Edition : 462 p.
- Sharma A., Bakshi M. 2014. Variability in Growth, Physiological, and Biochemical Characteristics among Various Clones of *Dalbergia sissoo* in a Clonal Seed Orchard. *Int. J. Forestry Res.* 2014 : 829368. <http://dx.doi.org/10.1155/2014/829368>.

*Поступила в редакцию  
30.08.2018 г.*

## **ON APPLICATION OF STATISTICAL METHODS IN STRESS PHYSIOLOGY AND BREEDING OF PLANTS**

V. V. Ivanishchev

*Lev Tolstoy Tula State Pedagogical University  
(Tula, Russia)*

*E-mail: avdey\_VV@mail.ru*

The review considers the results of the application of statistical methods (the principal component method and the cluster analysis method) during the study of physiological and biochemical indicators of plants in connection with the problems of resistance to various environmental factors and selection. It is shown that the application of such methods in the field of plant physiology and biochemistry is still insufficient. One of the reasons is the complexity and not always understandable biological interpretation of the mathematical transformations of the research results. It should also be borne in mind that, firstly, the application of such methods can lead to a partial loss of data (more precisely, not all possible variants of the relationships between the parameters studied are taken into account), and, secondly, statistical methods are based on the fact, that the studied features obey the so-called normal distribution, which does not always correspond to biological processes, phenomena and objects. Nevertheless, such approaches to the processing of results can benefit for the detection of new relationships between the known physiological and biochemical characteristics of plants. This is important for overcoming subjectivity in interpreting the results of the study and in identifying the importance (and / or specificity) of individual indicators in resistance of plants to stresses of different nature. The paper gives examples of the application of the principal component method and cluster analysis in specific studies. It is shown that the first method is used to a lesser degree, since its application requires a certain skill in analyzing the results. The method of cluster analysis is applied more often, because it allows to build a fairly clear hierarchical structure between the parameters studied, and also to find the closest relationship between the studied parameters or objects. It is shown that the cluster analysis method can be successfully used in solving problems related to the selection of the most suitable plant forms for further breeding work, as well as in assessing the relationships and potential interdependencies between the physiological and biochemical parameters of plants in stress.

**Key words:** *statistical methods, principal component analysis, cluster analysis, physiological and biochemical parameters, stress, drought, salt tolerance, selection*

## **ОБ ИСПОЛЬЗОВАНИИ СТАТИСТИЧЕСКИХ МЕТОДОВ**

### **ПРО ВИКОРИСТАННЯ СТАТИСТИЧНИХ МЕТОДІВ У СТРЕСОВІЙ ФІЗІОЛОГІЇ І СЕЛЕКЦІЇ РОСЛИН**

В. В. Іванищев

*Федеральна державна бюджетна освітня установа вищої освіти  
«Тульський державний педагогічний університет ім. Л.М. Толстого»  
(Тула, Росія)  
E-mail: avdey\_VV@mail.ru*

Розглянуто результати застосування статистичних методів (головних компонент і кластерного аналізу) при вивченні фізіолого-біохімічних показників рослин у зв'язку з проблемами стійкості до різних чинників середовища і селекції. Показано, що використання таких методів в галузі фізіології і біохімії рослин є все ще недостатнім. Однією з причин є складність і не завжди зрозуміла біологічна інтерпретація математичних перетворень результатів дослідження. При цьому також слід враховувати, що, по-перше, використання таких методів може призводити до часткової втрати даних (можуть враховуватися не всі можливі варіанти зв'язків між вивченими параметрами), а, по-друге, статистичні методи ґрунтуються на тому, що досліджувані ознаки підпорядковуються так званому нормальному розподілу, що не завжди відповідає біологічним процесам, явищам і об'єктам. Проте, такі підходи до обробки результатів можуть бути корисними для виявлення нових взаємозв'язків між відомими фізіолого-біохімічними характеристиками рослин. Це важливо для подолання суб'єктивності у тлумаченні результатів дослідження і виявлення вкладу (та/або специфічності) окремих показників в стійкість рослин до стресорів різної природи. У роботі наведені приклади застосування методу головних компонент і кластерного аналізу в конкретних дослідженнях. Показано, що перший метод використовується меншою мірою, оскільки його застосування потребує певних навичок при аналізі результатів. Метод кластерного аналізу застосовують частіше, оскільки він дозволяє побудувати досить ясну ієрархічну структуру між дослідженими показниками, а також виявити найбільшу близькість (взаємозв'язок) між вивченими параметрами або об'єктами. Показано, що метод кластерного аналізу може бути успішно використаний при вирішенні завдань, пов'язаних з вибором найбільш відповідних форм рослин для подальшої селекційної роботи, а також при оцінці взаємозв'язків і потенційних взаємозалежностей між фізіолого-біохімічними показниками рослин в умовах стресу.

**Ключові слова:** *статистичні методи, метод головних компонент, кластерний аналіз, фізіолого-біохімічні показники, стрес, посуха, солестійкість, селекція*