

К ВОПРОСУ МНОЖЕСТВЕННОГО ОЦЕНИВАНИЯ ЗАВИСИМОСТИ УРОЖАЙНОСТИ ОТ СОПРЯЖЕННОГО КОМПЛЕКСА ПОКАЗАТЕЛЕЙ ПОЧВЕННОГО ПЛОДОРОДИЯ

Усманов Раиф Рафикович, к.с.-х.н., доцент,

Хохлов Николай Федорович, д.с.-х.н., профессор кафедры земледелия и методики опытного дела

ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», E-mail: usman@rgau-msha.ru

***Аннотация:** Проведена статистическая оценка зависимости урожайности ячменя от агрофизических и агрохимических показателей плодородия почвы. Рекомендовано производить отбор переменных в модель множественной регрессии в программе Stat.Soft Statistica.*

***Ключевые слова:** агрофизические и агрохимические показатели, корреляция, множественная регрессия, мультиколлинеарность.*

Определение количественной зависимости урожайности от базовых показателей почвенного плодородия является распространенной процедурой экспериментальной агрономии. Корректное статистическое построение, анализ и интерпретация результирующей множественной модели дает понимание сложных динамических отношений между свойствами почвы, часть из которых может находиться в тесной взаимосвязи [1,2,5]. При этом задача исследователя заключается в построении модели с полноценным и одновременно минимальным числом показателей, необходимых и достаточных для отражения совокупного воздействия на функциональный признак [3,4].

Вместе с тем, в отдельных публикациях авторы, оценивая множественную зависимость свойств почвы и урожайности, ограничиваются нахождением исключительно коэффициентов парной корреляции, или в окончательную модель между зависимым и независимыми признаками некорректно включают все коэффициенты регрессии исследуемых показателей плодородия без обязательного учета тесноты и значимости взаимосвязей между независимыми переменными.

Если включение в модель тесно коррелированных между собой независимых показателей почвенного плодородия закрывает перспективу сокращения перечня определяемых показателей и соответственно экономику исследований, то игнорирование обеспечения должного уровня значимости ставит под сомнение обоснованность результатов научного исследования.

Поэтому отбор независимых признаков, выступающих в качестве предикторов для оценки множественной зависимости, является одним из ответственных этапов корреляционно-регрессионного анализа. Предикторы,

включаемые во множественную регрессию, должны быть между собой независимыми, или в терминах статистики интеркоррелированными.

Для осмысления базисных положений и полноценной корректной оценки множественных связей нами приводится пример статистического анализа данных модельного полевого многофакторного опыта, в котором изучалась зависимость урожайности ячменя (Y) от агрохимических показателей плодородия дерново-подзолистой почвы: содержание гумуса (X_1), содержание подвижного фосфора (X_2), обменного калия (X_3) и агрофизических: плотность почвы (X_4), порозность (X_5) и водопрочность структуры (X_6).

Множественный корреляционно-регрессионный анализ данных опыта провели в статическом пакете Stat Soft Statistica 10.

На основании проведенных расчетов (табл.1) по величине коэффициента множественной корреляции ($R=0,97$) и коэффициента детерминации ($R^2 = 0,94$) была установлена очень тесная и существенная на 0,0001 уровне значимости связь урожайности ячменя со всем комплексом представленных агрофизических и агрохимических показателей плодородия почвы.

Таблица 1. Результаты множественного корреляционно-регрессионного анализа по влиянию показателей плодородия почвы на урожайность ячменя

Переменная	$R=0,97$ $R^2=0,94$ $F_{\phi}=33,83$ $p<0,0001$					
	Коэффициенты				Критерий Стьюдента $t_{\phi ак}$	р-знач.
	парной корреляции r_{yx}	частной корреляции $r_{yx.z}$	регрессии b_{yx}	бета β		
Гумус	0,88	0,74	0,325	0,42	4,00	0,001
P ₂₀₅	0,88	0,66	0,020	0,38	3,20	0,006
K ₂₀	0,32	0,08	0,002	0,02	0,30	0,768
Плотность	-0,80	0,07	0,147	0,05	0,23	0,817
Порозность	0,82	0,26	0,010	0,20	0,98	0,347
Водопрочность	0,61	0,46	0,018	0,16	1,85	0,087

Коэффициенты парной корреляции показали, что между урожайностью ячменя и содержанием гумуса, подвижного фосфора ($r=0,88$), а также порозностью ($r=0,82$) отмечается положительная тесная связь, в то время как между плотностью почвы и урожайностью она обратная и тесная ($r= -0,80$), влияние водопрочной структуры ($r=0,61$) и содержание обменного калия ($r=0,32$) на урожайность носит средний характер.

Так как представленные в опыте факториальные признаки имеют разные единицы измерения, для определения силы влияния каждого из них на результативный признак помимо традиционных коэффициентов регрессии b_{yx} были рассчитаны β (бета)- коэффициенты регрессии. По значениям бета-коэффициентов можно судить о вкладе независимых признаков на урожайность ячменя. Из данных таблицы 1 видно, что из всех признаков наибольший и значимый ($p<0,01$) вклад в изменение урожайности вносит содержание в почве гумуса и подвижных форм фосфора.

Для учета взаимосвязей между независимыми признаками провели корреляционный анализ между всеми агрохимическими и агрофизическими показателями, результаты анализа представлены на рис. 1.

Переменная	Корреляции (Множ кор -гумус-урож) Отмеченные корреляции значимы на уровне $p < 0,05000$ N=20 (Построчное удаление ПД)						
	Гумус	P205	K20	Плотность	Порозность	Водопрочность	Урожайность
Гумус	1,00	0,71	0,38	-0,63	0,64	0,45	0,88
P205	0,71	1,00	0,22	-0,75	0,77	0,44	0,88
K20	0,38	0,22	1,00	-0,23	0,23	0,12	0,32
Плотность	-0,63	-0,75	-0,23	1,00	-0,93	-0,61	-0,80
Порозность	0,64	0,77	0,23	-0,93	1,00	0,58	0,82
Водопрочность	0,45	0,44	0,12	-0,61	0,58	1,00	0,61
Урожайность	0,88	0,88	0,32	-0,80	0,82	0,61	1,00

Рис. 1. Матрица парных коэффициентов корреляции

Анализ матрицы парных корреляций между всеми переменными показал на весьма сильную, приближающуюся к функциональной связь между плотностью почвы и порозностью ($r = -0,93$), что свидетельствует о *мультиколлинеарности* между этими признаками. При такой сильной зависимости между независимыми признаками, когда коэффициент корреляции между ними $r > \pm 0,8$ рекомендуется один из независимых признаков исключить из рассмотрения.

Для устранения эффекта мультиколлинеарности и нахождения наилучшего регрессионного уравнения в программе Statistica предусмотрены две схемы пошаговой регрессии: пошаговое включение и пошаговое исключение независимых признаков.

Таблица 2. Итоги пошаговой регрессии методом исключения

Показатели $R=0,96$ $R^2=0,92$ $F_{\phi}=64,38$ $p<0,0000$

Показатели	Бета – коэффициент, β	Коэффициент регрессии, b_{yx}	Критерий Стьюдента, $t_{\text{фак}}$	p -знач.
Свободный член		0,52		
Гумус	0,46	0,34	4,57	<0,0003
P ₂₀₅	0,36	0,02	3,02	0,0080
Порозность	0,24	0,01	2,20	0,043

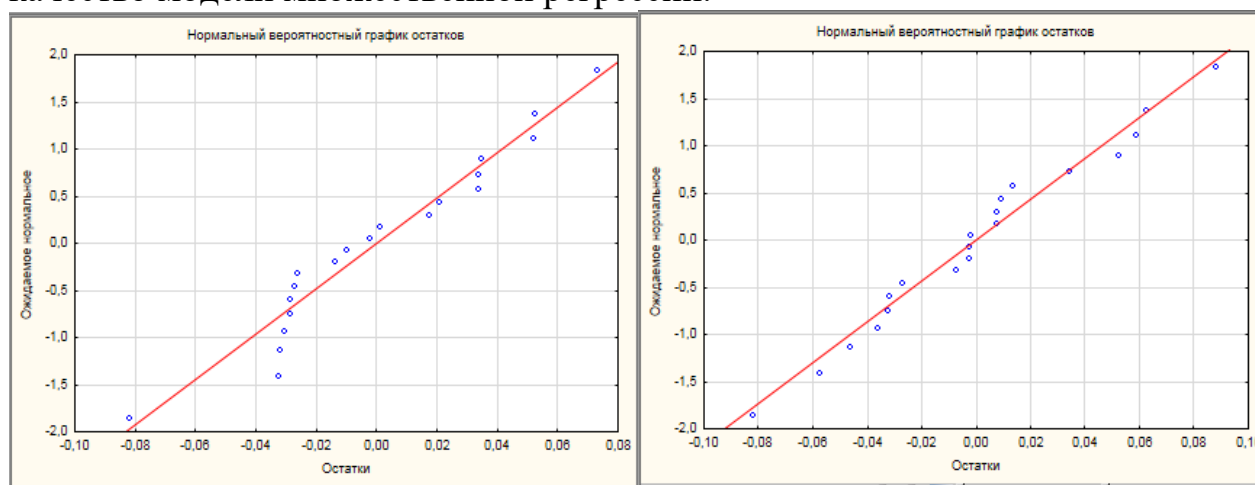
При использовании метода *пошагового включения признаков*, первым в итоговую регрессию включается независимый признак, наиболее тесно коррелирующий с результирующим признаком, который дает максимальное значение коэффициента корреляции и бета-коэффициент регрессии.

При пошаговом *исключении независимых признаков* на последнем шаге программа Statistica автоматически оставляет только те признаки, бета-

коэффициенты которых являются значимыми, что свидетельствует о наличии в регрессионной модели только существенных факторов.

В таблице 2 приведены результаты заключительного, третьего шага гребневой регрессии методом исключения. Из шести независимых признаков в модели оставлены по значениям бета-коэффициентов регрессии ($p < 0,05$) следующие показатели: гумус, подвижный фосфор и порозность почвы. Содержание обменного калия и водопрочность из-за незначимости их бета-коэффициентов, и плотность почвы из-за эффекта мультиколлинеарности исключены из модели множественной регрессии (табл. 1, рис.1).

Оценка адекватности уравнения множественной регрессии осуществляется по нормальному вероятностному графику остатков. Чем меньше разброс значений остатков около линии регрессии тем, очевидно, выше качество модели множественной регрессии.



а)

б)

Рис.2. График остатков: а — до исключения переменных; б — после исключения незначимых переменных

Как видно из двух нормальных вероятностных графиков остатков (рис.2), пошаговая регрессия с исключением незначимых переменных позволила описать множественную регрессию более точно. Так, если на первом графике остатков в верхней и нижней части точки приближены к прямой, а в средней части отмечается значительный разброс точек от теоретической линии регрессии, то после проведения пошаговой (гребневой) регрессии эти точки имеют незначительный разброс от прямой линии.

На основании проведенного регрессионного анализа можно составить следующее уравнение множественной регрессии:

$$Y = 0,52 + 0,34 \cdot X_1 + 0,02 \cdot X_2 + 0,01 X_5$$

Таким образом, после проведения пошагового регрессионного анализа методом исключения переменных с вероятностью 99% значимыми признаками в множественной регрессионной модели, оказывающими существенное влияние на урожайность ячменя, являются содержание гумуса ($t_{\phi} = 4,57$; $p = 0,0003$) и содержание подвижного фосфора в почве ($t_{\phi} = 3,02$; $p = 0,008$) и с 95% – порозность почвы ($t_{\phi} = 2,20$; $p = 0,043$).

Судя по полученным значениям бета-коэффициентов можно сделать вывод о необходимости приоритетов в экспериментальном поиске регулирующего воздействия на гумусное состояние и содержания фосфора в почве и соответственно ослабления акцента на сопутствующие программные вспомогательные исследования.

Библиографический список

1. Жеряков Е.В. Взаимосвязь параметров почвенного при систематическом применении удобрений /Е.В. Жеряков // Международный научно-исследовательский журнал ISSN 2303-9868 Meždunarodnyj naučno-issledovatel'skij žurnal. – 2014. –№10 (29). – С. 58-60.
2. Попок Л.Б. Корреляционно-регрессионный анализ в изучении взаимосвязи содержания тяжелых металлов с агрофизическими и агрохимическими свойствами почв / Л.Б. Попок, И.С. Белюченко // Экологический вестник Северного Кавказа. – 2013. – Т.9. - №4. – С. 45-49.
3. Усманов Р.Р. Методика экспериментальных исследований в агрономии: учебное пособие для вузов / Р. Р. Усманов.– Москва: Издательство Юрайт, 2021. – 197 с. – (Высшее образование). – Текст: непосредственный.
4. Шабаев А.Г. Взаимосвязь показателей состояния плодородия дерново-подзолистых почв с урожайностью сельскохозяйственных культур / А.Г. Шабаев// Электронный научный журнал «Инженерный вестник Дона». – 2014. – № 4.
5. Шевченко В.А. Корреляционная взаимосвязь агрофизических показателей мелиорированных земель Верхневолжья при различных способах их основной обработки / А.В. Шевченко, А.И. Беленков, А.М. Соловьев // Земледелие. – 2018. – №8. – С. 26-29.

ON THE ISSUE OF MULTIPLE EVALUATION OF THE DEPENDENCE OF YIELD ON THE CONJUGATE COMPLEX OF INDICATORS OF SOIL FERTILITY

Usmanov R.R., K.Sc. in Agricultural Sciences, Hohlov N.F., D.Sc. in Agricultural Sciences, Russian Timiryazev State Agrarian University - Moscow Timiryazev Agricultural Academy

Abstract: A statistical assessment of the dependence of barley yield on agrophysical and agrochemical indicators of soil fertility has been carried out. It is recommended to select variables in the multiple regression model in the Stat program. Soft Statistica.

Key words: agrophysical and agrochemical indicators, correlation, multiple regression, multicollinearity.