

УДК 633/635:57.081:004.94

ИСПОЛЬЗОВАНИЕ ПРОЦЕДУРЫ ОБОБЩЕННОЙ ЛИНЕЙНОЙ МОДЕЛИ ДЛЯ АНАЛИЗА РЕЗУЛЬТАТОВ СЕЛЬСКОХОЗЯЙСТВЕННЫХ ИССЛЕДОВАНИЙ

Мельник А.Ф., Шуметов В.Г., Кондрашин Б.С.

ФГБОУ ВО «Орловский государственный аграрный университет им. Н.В. Парухина», Орел,
e-mail: melnik.anat202@yandex.ru, shumetov@list.ru

Рассмотрены основные этапы моделирования результатов сельскохозяйственных исследований с использованием процедуры обобщенной линейной модели, показаны ее преимущества и возможности при анализе данных полевых опытов. Приведены примеры использования процедуры для оценки статистической значимости влияния факторов агротехнических опытов, построения доверительных областей выборочных параметров и проверки гипотез. Показано, что моделью, адекватно отражающей влияние предшественников (кукурузы на зеленую массу, гороха на зерно и ячменя) и нормы удобрений (2 и 4 ц/га азофоски) на урожайность озимой пшеницы сорта Московская 39, является двухфакторная линейная модель дисперсионного анализа, причем влияние предшественника больше влияния нормы минеральных удобрений. Получены мнк-оценки параметров модели, доверительные интервалы эффектов предшественников и норм удобрений. Доказано, что оптимальными для повышения урожайности озимой пшеницы являются предшественник горох и норма азофоски 4 ц/га. Сформированы однородные группы предшественников по критерию множественного сравнения Тьюки, при этом предшественник горох образует самостоятельную подгруппу, обеспечивающую большую урожайность озимой пшеницы, а ячмень и кукуруза входят в общую подгруппу предшественников, обеспечивающих меньшую урожайность. Получены двухфакторные линейные модели дисперсионного анализа, которые также адекватно отражают влияние предшественников и нормы удобрений на показатели качества зерна озимой пшеницы сорта Московская 39 – содержание сырого белка и клейковины. Доказано, что, как и для урожайности, оптимальными являются предшественник горох и норма азофоски 4 ц/га. Предложено визуально оценивать качество моделирования путем сравнения диаграмм зависимости показателей продуктивности от уровней факторов, построенных по фактическим и расчетным данным. Существенным преимуществом использования процедуры обобщенной линейной модели для анализа результатов полевых опытов является возможность моделирования по средним данным, при отсутствии информации по повторностям, что позволяет строить модели по данным публикаций.

Ключевые слова: озимая пшеница, предшественники, норма удобрений, обобщенная линейная модель, дисперсионный анализ, статистическая значимость, критерий Тьюки, различия средних

PROCEDURE OF THE GENERALIZED LINEAR MODEL FOR THE ANALYSIS OF AGRICULTURAL RESEARCH RESULTS

Melnik A.F., Shumetov V.G., Kondrashin B.S.

Orel State Agrarian University named after N.V. Parakhin, Orel, e-mail: melnik.anat202@yandex.ru,
shumetov@list.ru

The main stages of agricultural research results' modeling using the procedure of generalized linear model are considered; its advantages in the analysis of field experiments are shown. Concrete examples of using the procedure for assessing the statistical significance of the factors' influence of agrotechnical experiments, building confidence areas of sample parameters and testing hypotheses are given. It is shown that the model adequately reflecting the effect of precursors (corn for green mass, peas for grain and barley) and fertilizer rates (2 and 4 centners / ha of nitrogen-phosphorus-potassium fertilizer) on the yield of winter wheat of variety Moscovskaya 39 is a two-factor linear model of analysis of variance, and predecessor is more powerful than the norm of mineral fertilizers. LS method of the model parameters, the confidence intervals of the effects of the precursors and fertilizer rates were obtained. It has been proven that the precursor peas and a nitrogen concentration of 4 centners / ha are optimal for increasing the yield of winter wheat. Homogeneous groups of predecessors were formed according to Tukey's multiple comparison criterion, with the predecessor «peas» forming an independent subgroup providing greater yields of winter wheat, while barley and corn are included in the common subgroup of predecessors providing lower yields. Two-factor linear models of analysis of variance were obtained, which also adequately reflect the influence of the precursors and fertilizer norms on the grain quality indicators of winter wheat variety Moscovskaya 39 – the content of crude protein and gluten. It is proved that, as well as for yields, the predecessor peas and the norm of nitrogen-phosphorus-potassium fertilizer are 4 t / ha. It is proposed to visually assess the quality of modeling by comparing the diagrams of dependence of productivity indicators on the factors' levels, built on the basis of actual and calculated data. A significant advantage of using the procedure of generalized linear model for analyzing the results of field experiments is the possibility of modeling according to average data, in the absence of information on replications, which makes it possible to build models according to publications.

Keywords: winter wheat, predecessors, rate of fertilizers, generalized linear model, analysis of variance, statistical significance, Tukey criterion, differences in averages

Важным инструментом при внедрении современных наукоемких технологий в практику растениеводства являются динамические модели продукционного процесса

сельскохозяйственных растений, роль которых возрастает в свете нового перспективного тренда развития сельскохозяйственной отрасли – устойчивого и точного земледелия

лия [1]. Необходимым условием возможности создания таких моделей является наличие полного набора данных как о внешних факторах (погода, агротехника), так и об индикаторах роста и развития растений, при этом отправной точкой является информация о прошлых данных полевых опытов [2]. Научный полевой опыт позволяет изучать вклад факторов в формирование урожая (севооборот, обработка почвы, удобрения и др.), обобщать полученные результаты, создавая модели [3, 4].

Вплоть до настоящего времени исследователи в области агрономии при разработке моделей ориентируются на классические методы однофакторного, и лишь в отдельных случаях многофакторного, дисперсионного анализа, используя рекомендации, изложенные еще в «докомпьютерную» эру в учебнике Б.А. Доспехова [5]. Но в связи с распространением компьютерной техники появились новые возможности. В вышедшем в 2013 г. учебнике для агрономических специальностей [6] отмечается необходимость компьютерной оценки данных научного полевого исследования. При этом, помимо широко распространенного табличного процессора *MS-EXCEL*, рекомендуются к использованию такие пакеты программ анализа данных, как *SPSS*, *STATGRAPHICS Plus for Windows*, *STATISTICA*, предлагающие исследователям широкий спектр многомерных методов моделирования, основанных на алгоритмах регрессионного, дискриминантного, факторного, кластерного анализа. Однако в широкую практику использование указанных пакетов программ для анализа данных пока не вошло, обучение студентов-аграриев в отечественных вузах искусству моделирования по-прежнему ведется по методикам дисперсионного анализа или, в лучшем случае, универсального пакета *MS-EXCEL*. На актуальность «внедрения» новых технологий анализа данных в практику сельскохозяйственных исследований указывали различные авторы, в частности, В.М. Кузнецов [7]. Он отмечал, что «в большинстве работ российских исследователей-животноводов анализ экспериментальных и «полевых» данных... ограничивается расчетом средних значений и, в лучшем случае, их стандартных ошибок. Лишь в небольшом числе работ используется однофакторный дисперсионный анализ и очень редко – *многофакторные обобщенные линейные модели*» [7, с. 27].

В растениеводстве ситуация аналогичная. Лишь в немногих работах использо-

ваны компьютерные методики многофакторного многомерного дисперсионного анализа. Так, в работе [8] многофакторный многомерный дисперсионный анализ реализован с помощью процедуры обобщенной линейной модели пакета *STATISTICA*, что позволило авторам данной работы получить ряд новых выводов по засоренности зерновых агроценозов, при этом, как правило, по всем выделяемым эффектам нулевые гипотезы отвергались с высокой вероятностью [8, с. 5].

Вышесказанное актуализирует наши исследования, направленные на адаптацию новых процедур получения моделей по результатам сельскохозяйственных экспериментов к специфике отрасли.

Цель исследования: исследование возможностей и эффективности применения одной из перспективных процедур дисперсионного анализа – обобщенной линейной модели – для разработки моделей полевых опытов. Особенность этой модели дисперсионного анализа заключается в том, что она ставит в соответствие результирующий показатель значениям воздействующих факторов, которые могут быть как количественными, так и качественными. В этом ее преимущество против регрессионных, в которых все предикторы, как правило, количественные. Еще одним преимуществом процедуры обобщенной линейной модели является возможность оценки статистической значимости эффектов и их доверительных границ, что обеспечивает надежность обоснования достоверности и значимости результатов полевых опытов.

Материалы и методы исследования

В исследовании нами использовалась одна из версий пакета статистических программ анализа данных общественных наук *SPSS* – версия 8.0 [9], а конкретнее, процедура «Обобщенная линейная модель» данного пакета. Прямое назначение данной процедуры – нахождение параметров модели, связывающей результирующую количественную переменную с двумя и более качественными (номинальными) переменными, однако команды процедуры «Общая линейная модель» позволяют выполнять и однофакторный дисперсионный анализ [10, 11]. В качестве примера использовали эмпирические данные полевых исследований по изучению влияния предшественников на продуктивность озимой пшеницы с применением различных норм минеральных удобрений, полученные

в производственных условиях СПК «Русь» Знаменского района Орловской области [12]. Исследования проведены в 2006–2007 гг. по плану полного двухфакторного эксперимента ПФЭ 3×2: фактор А – предшественники: 1) кукуруза на зеленую массу, 2) горох на зерно, 3) ячмень; фактор В – нормы внесения удобрений: 1) 2 ц/га азофоски, 2) 4 ц/га азофоски. Объектом исследования являлся районированный в Орловской области сорт озимой пшеницы Московская 39, способный формировать зерно с высокими технологическими качествами. Опыты проводились в трехкратной повторности, однако обработке подвергали лишь средние данные по повторностям,

тем самым воспроизводя ситуацию отсутствия первичных (сырых) данных полевых опытов.

Результаты исследования и их обсуждение

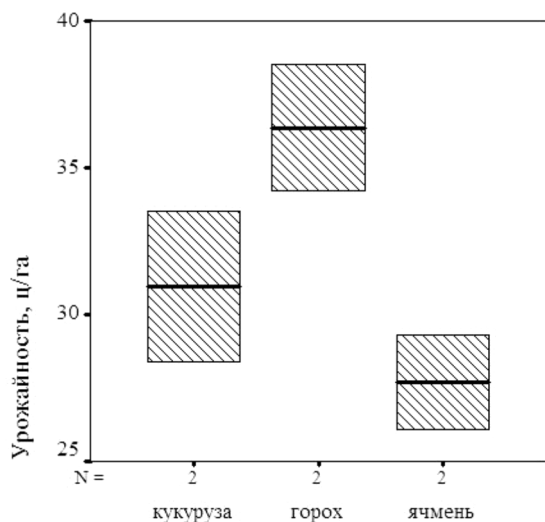
Методику двухфакторного дисперсионного анализа, реализуемую с помощью процедуры обобщенной линейной модели, рассмотрим на основе исходных данных по урожайности и качеству зерна озимой пшеницы сорта Московская 39, приведенных в табл. 1.

На первом этапе методики проводим визуальный анализ диаграмм Тьюки по предшественникам и норме удобрений (рис. 1).

Таблица 1

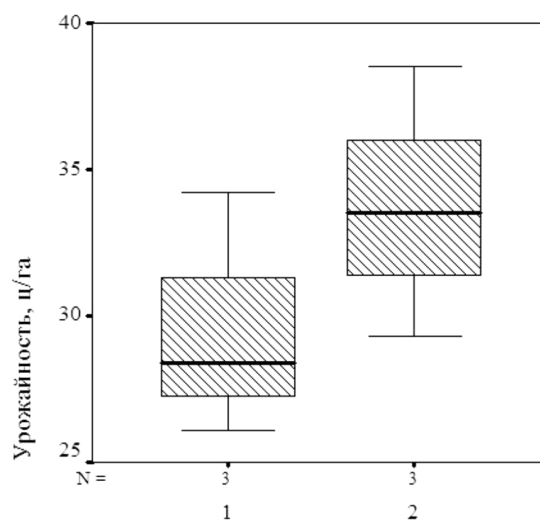
Урожайность и качество зерна озимой пшеницы Московская 39 (в среднем за два года).
Источник: [12, с. 30]

Предшественники	Норма удобрений	Урожайность, ц/га	Содержание, %	
			клейковины	белка
1. Кукуруза	1. 2 ц/га азофоски	28,4	24,7	11,8
	2. 4 ц/га азофоски	33,5	26,1	12,5
2. Горох	1. 2 ц/га азофоски	34,2	26,8	13,2
	2. 4 ц/га азофоски	38,5	27,5	14,7
3. Ячмень	1. 2 ц/га азофоски	26,1	22,0	11,3
	2. 4 ц/га азофоски	29,3	23,6	12,7



Предшественники

а)



Норма удобрения

б)

Рис. 1. Урожайность озимой пшеницы: а) в зависимости от предшественников; б) в зависимости от нормы азофоски (1 – 2 ц/га; 2 – 4 ц/га)

Из рис. 1, а, видно, что урожайность озимой пшеницы по гороху превосходит урожайность по ячменю и кукурузе, однако по кукурузе и ячменю наблюдается перекрытие диапазонов изменения показателя. По норме удобрений также наблюдается перекрытие значений показателя, хотя медианы при этом различаются – рис. 1, б.

На следующем этапе проводим оценку статистической значимости различия урожайности по предшественникам и норме удобрений «в целом» в процедуре «Обобщенная линейная модель», используя модель двухфакторного дисперсионного анализа (табл. 2).

Из табл. 2 следует, что оба фактора – предшественники и норма удобрения – статистически значимы на уровне не хуже 0,05. Это позволяет признать адекватной модель

$$Y_{ij} = \mu_0 + \alpha_i + \beta_j + \varepsilon_{ij}$$

и оценить ее параметры (табл. 3). В этой модели: Y_{ij} – наблюдаемое значение выходной переменной Y (урожайность) на i -м уровне фактора «предшественник» и j -м уровне фактора «норма удобрения»; μ_0 – оценка свободного коэффициента модели; α_i и β_j – оценки главных эффектов; ε_{ij} – случайная ошибка.

Поясним табл. 3. В ней постоянная $\mu_0 = 29,8$ ц/га, эффекты предшественника 3

(ячмень) и нормы удобрения 2 (4 ц/га азотом) приняты за нулевые. Эффекты других предшественников и нормы удобрения отсчитываются уже от этого уровня; так, урожайность озимой пшеницы по предшественнику 2 (гороху) характеризуется добавкой $\alpha_2 = 8,65$ ц/га. Значения 95%-ного доверительного интервала всех эффектов не включают в себя нуль, что свидетельствует о статистической значимости параметров модели.

Дополнительную информацию о значимости разности средних дают табл. 4 множественных сравнений и табл. 5 однородных подгрупп предшественников, полученные при использовании критерия Тьюки.

Согласно табл. 4, статистически значимыми на уровне надежности не хуже 95% могут быть приняты разности урожайности озимой пшеницы по предшественнику гороху с урожайностью по предшественникам ячменю и кукурузе, так как соответствующие значения 95%-ного доверительного интервала не включают в себя нуль. Напротив, разность урожайности озимой пшеницы по предшественникам ячменю и кукурузе статистически значима на уровне 0,073, превышающем нормативную величину 0,05; кроме того, доверительный интервал данной разности включает в себя нуль.

Таблица 2

Тест межсубъектных эффектов

Источник изменчивости	Сумма квадратов	Ст. св.	Средний квадрат	F-критерий	Значимость
Исправленная модель	102,823	3	34,274	75,328	0,013
Постоянная	6016,667	1	6016,667	13223,443	0,000
Предшественник	76,363	2	38,182	83,916	0,012
Удобрение	26,460	1	26,460	58,154	0,017
Ошибка	0,910	2	0,455		
Сумма	6120,400	6			
Исправленная сумма	103,733	5			

Таблица 3

МНК-оценки параметров двухфакторной модели влияния предшественника и нормы удобрения на урожайность озимой пшеницы

Параметр	В	Стд. ошибка	t	Значимость	95% доверительный интервал	
					нижняя граница	нижняя граница
Постоянная	29,800	0,551	54,107	0,000	27,430	32,170
[ПРЕДШЕСТ=1]	3,250	0,675	4,818	0,040	0,348	6,152
[ПРЕДШЕСТ=2]	8,650	0,675	12,824	0,006	5,748	11,552
[ПРЕДШЕСТ=3]	0	,	,	,	,	,
[УДОБРЕНИ=1]	-4,200	0,551	-7,626	0,017	-6,570	-1,830
[УДОБРЕНИ=2]	0	,	,	,	,	,

С этими результатами согласуются данные табл. 5: горох образует самостоятельную подгруппу, тогда как ячмень и кукуруза входят в общую подгруппу 1.

Наглядное представление о качестве прогноза урожайности по двухфакторной модели дает рис. 2: видно, что расчетные данные лишь немного отличаются от опытных.

Таблица 4

Апостериорные парные сравнения средних по критерию Тьюки

(I) Предшественники	(J) Предшественники	Средняя разность (I-J)	Стд. ошибка	Знач. (2-сторон)	95% доверительный интервал	
					нижняя граница	верхняя граница
Кукуруза	Горох	-5,400	0,6745	0,028	-9,374	-1,426
	Ячмень	3,250	0,6745	0,073	-0,724	7,224
Горох	Кукуруза	5,400	0,6745	0,028	1,426	9,374
	Ячмень	8,650	0,6745	0,011	4,676	12,624
Ячмень	Кукуруза	-3,250	0,6745	0,073	-7,224	0,724
	Горох	-8,650	0,6745	0,011	-12,624	-4,676

Таблица 5

Однородные подгруппы предшественников по критерию Тьюки (уровень значимости критерия различия между подгруппами $p = 0,05$)

Предшественники	N	Урожайность озимой пшеницы, ц/га	
		1	2
Ячмень	2	27,700	
Кукуруза	2	30,950	
Горох	2		36,350
Уровень значимости критерия различия в подгруппе		0,073	1,000

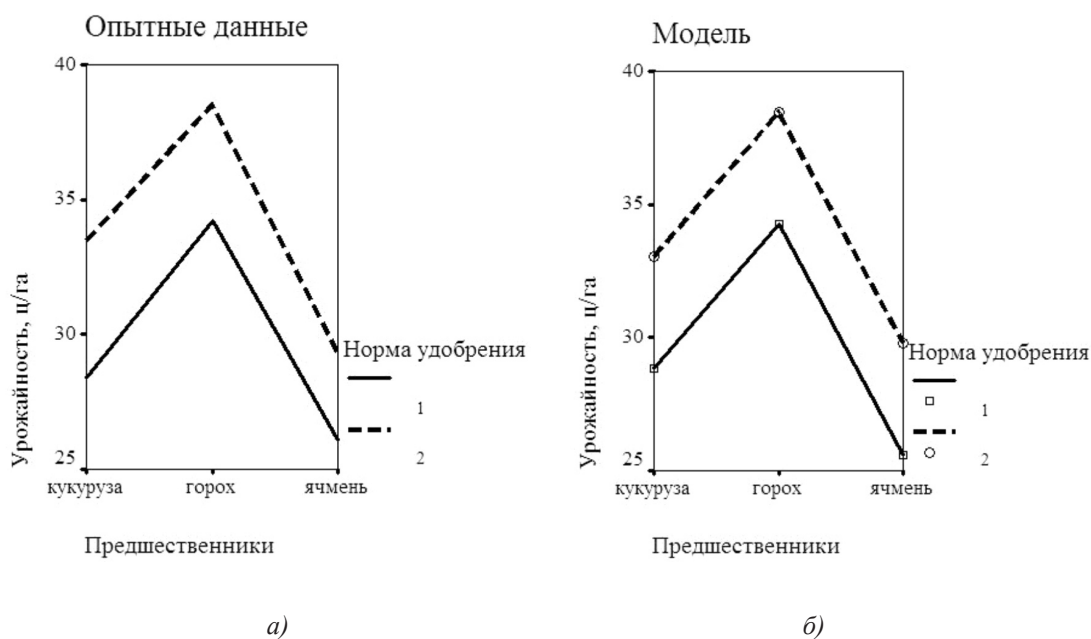


Рис. 2. Урожайность озимой пшеницы в зависимости от предшественников и нормы азотоса: а) опытные данные; б) расчет по двухфакторной модели

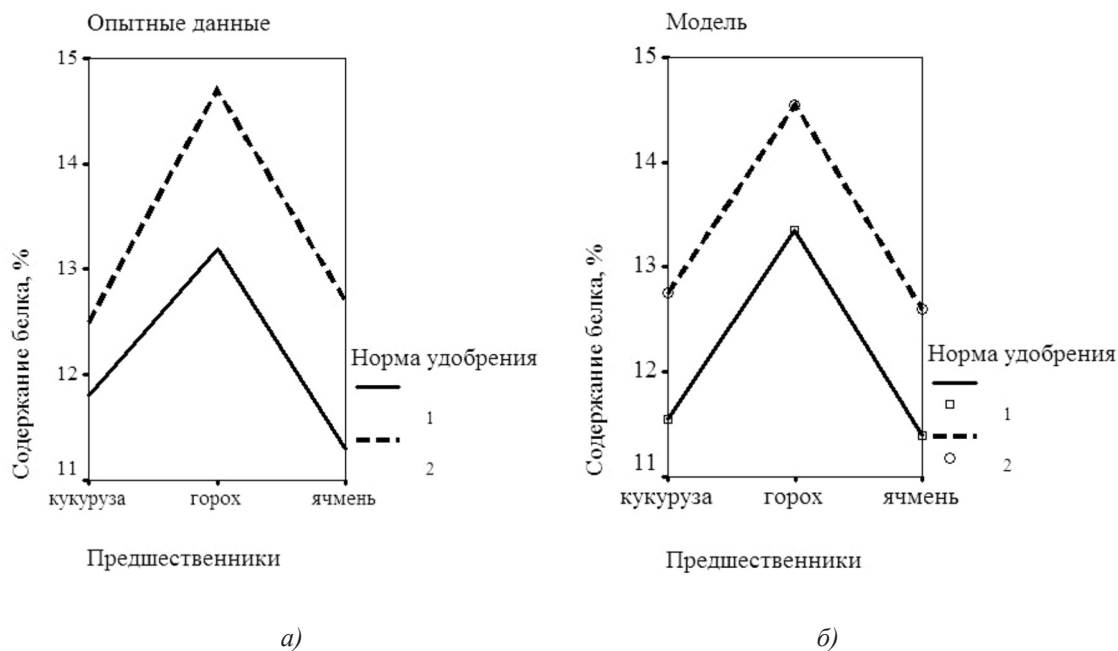


Рис. 3. Содержание в зерне озимой пшеницы сырого белка в зависимости от предшественников и нормы азотоски: а) опытные данные; б) расчет по двухфакторной модели

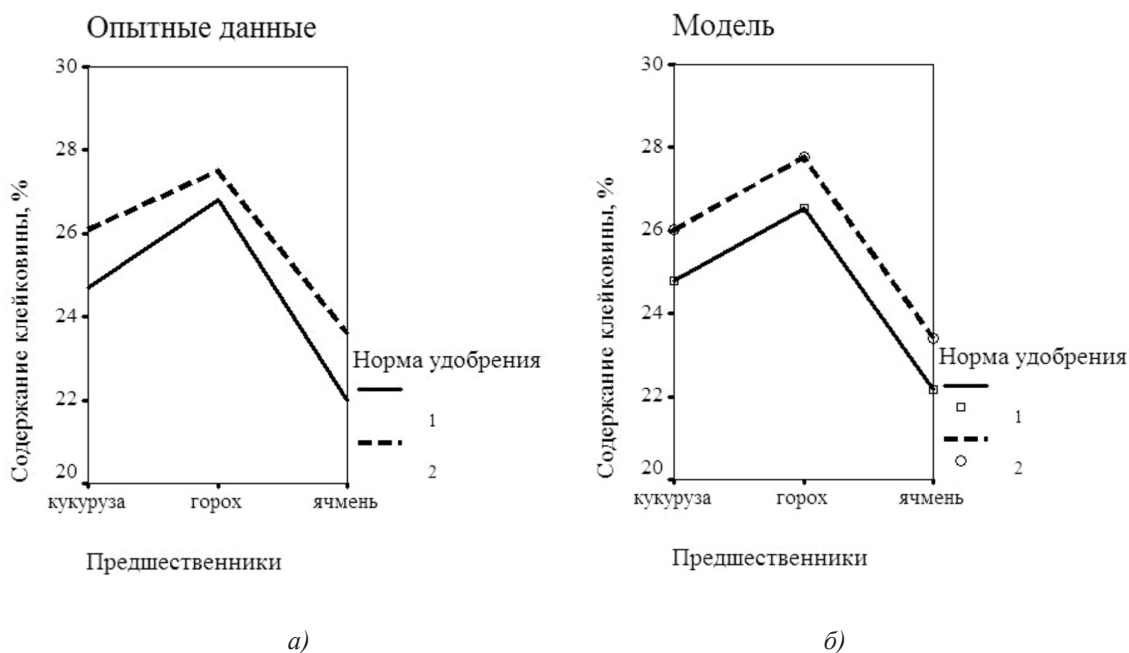


Рис. 4. Содержание в зерне озимой пшеницы клейковины в зависимости от предшественников и нормы азотоски: а) опытные данные; б) расчет по двухфакторной модели

Следует особо отметить, что проведенный выше анализ выполнен для средних, без учета данных по повторностям. Именно этим объясняется довольно высокое значение стандартной ошибки разностей урожайности – 0,67 ц/га. Соответственно этому наблюдается значительная ширина 95%-ных доверительных интервалов эффектов прибавки урожайности от агротехнических факторов. Так, при средней прибавке урожайности

– 0,67 ц/га. Соответственно этому наблюдается значительная ширина 95%-ных доверительных интервалов эффектов прибавки урожайности от агротехнических факторов. Так, при средней прибавке урожайности

сти озимой пшеницы по предшественнику гороху относительно предшественника кукурузы 5,40 ц/га доверительный интервал составляет от 1,43 до 9,37 ц/га, а относительно предшественника ячменя – 8,65 ц/га с 95%-ным доверительным интервалом от 4,68 до 12,62 ц/га. Впрочем, эта ситуация вполне соответствует значительной изменчивости условий возделывания большинства сельскохозяйственных культур.

Аналогично нами были получены двухфакторные модели влияния предшественников и норм удобрений на содержание в зерне озимой пшеницы сорта Московская 39 сырого белка и клейковины, объясняющие соответственно 97,3% и 99,0% общей дисперсии.

Согласно приведенным графикам, отражающим качество двухфакторных моделей (рис. 3, 4), установлено, что, так же как и на урожайность, предшественники оказывают более сильное влияние на оба показателя качества зерна, чем нормы удобрений, причем лучшие результаты достигаются при посеве озимой пшеницы по гороху.

Высокое качество полученных моделей подтверждается также близкими к единице значениями коэффициента детерминации R^2 : для модели урожайности $R^2 = 0,991$, для содержания белка и клейковины 0,973 и 0,990 соответственно.

Закключение

Представленный пример показывает, что преимуществом использования процедуры обобщенной линейной модели для анализа результатов полевых опытов является возможность построения моделей по средним данным, в случае отсутствия информации по повторностям. Это позволяет строить модели также и по материалам публикаций, которые, как правило, содержат лишь средние данные, а результаты измерений по повторностям не приводятся.

Список литературы / References

1. Якушев В.П., Якушев В.В. Информационное обеспечение точного земледелия. СПб.: Издательство ПИЯФ РАН, 2007. 384 с.
 Yakushev V.P., Yakushev V.V. Information support for precision farming. SPb.: Izdatel'stvo PIYAF RAN, 2007. 384 p. (in Russian).
 2. Медведев С.А., Топаж А.Г., Белов А.В., Глядченко-ва Н.А., Лекомтцев П.В. Распределенный измерительно-моделирующий комплекс для оперативного сопровождения полевого опыта // АгроЭкоИнфо. Электронный научно-производственный журнал. 2015. № 2 [Электронный ресурс]. URL: http://agroecoinfo.narod.ru/journal/STATYI/2015/2/st_08.doc (дата обращения: 18.02.2019).
 Medvedev S.A., Topazh A.G., Belov A.V., Glyadchenkova N.A., Lekomtsev P.V. Distributed measuring and modeling

complex for operational support of field experience // AgroEcoInfo. Elektronnyy nauchno-proizvodstvennyy zhurnal. 2015. № 2 [Electronic resource]. URL: http://agroecoinfo.narod.ru/journal/STATYI/2015/2/st_08.doc (date of access: 18.02.2019) (in Russian).

3. Морозов В.И., Тойгильдин А.Л. Полевой опыт как метод познания и практического освоения инновационных технологий // Вестник Ульяновской государственной сельскохозяйственной академии. 2012. № 1. С. 40–44.

Morozov V.I., Toigildin A.L. Field experience as a method of knowledge and practical development of innovative technologies // Vestnik Ul'yanskovy gosudarstvennoy sel'skokhozyaystvennoy akademii. 2012. № 1. P. 40–44 (in Russian).

4. Соколов А.С., Байрамбеков Ш.Б., Соколова Г.Ф. Влияние обработки почвы, удобрений, гербицидов на засоренность и урожайность овощных культур в севообороте // Успехи современного естествознания. 2018. № 8. С. 78–84.

Sokolov A.S., Bayrambekov Sh.B., Sokolova G.F. Influence of Soil Dressing, Fertilizers, Herbicides Upon Pollution and Yields of Vegetable Cultures in Sawing Turnover // Advances in current natural sciences. 2018. № 8. P. 78–84 (in Russian).

5. Доспехов Б.А. Методика полевого опыта. М.: Альянс, 2011. 352 с.

Dospikhov B.A. Methods of field experience. M.: Al'yans, 2011. 351 p. (in Russian).

6. Кирюшин Б.Д., Усманов Б.Д., Васильев И.П. Основы научных исследований в агрономии. М.: Колос, 2009. 398 с.

Kiryushin B.D., Usmanov B.D., Vasilyev I.P. Fundamentals of research in agronomy. M.: Kolos, 2009. 398 p. (in Russian).

7. Кузнецов В.М. Разведение по линиям и голштинизация: методы оценки, состояние и перспективы // Проблемы биологии продуктивных животных. 2013. № 3. С. 25–79.

Kuznetsov V.M. Breeding lines and holsteinization: analytical methods, state of the art and prospects // Problemy biologiy produktivnykh zhivotnykh. 2013. № 3. P. 25–79 (in Russian).

8. Синещев В.Е., Южаков А.И., Красноперов А.Г. Методы регулирования сорных растений в зерновых агроценозах Западной Сибири // АгроЭкоИнфо. 2008. № 2 (3) [Электронный ресурс]. URL: http://agroecoinfo.narod.ru/journal/STATYI/2009/2/st_18.doc (дата обращения: 18.02.2019).

Sineschekov V.E., Yuzhakov A.I., Krasnoperov A.G. Methods of weed control in grain agrocenoses of Western Siberia // AgroEcoInfo. 2008. № 2 [Electronic resource]. URL: http://agroecoinfo.narod.ru/journal/STATYI/2009/2/st_18.doc (date of access: 18.02.2019) (in Russian).

9. SPSS Base 8.0 для Windows. Руководство по применению. М.: СПСС Русь, 1998. 397 с.

SPSS Base 8.0 for Windows. Application Guide. M.: SPSS Rus', 1998. 397 p. (in Russian).

10. Бююль А., Цёфель П. SPSS: Искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей. М.: «ДиаСофт», 2005. 608 с.

Buyuyul' A., Tsofel' P. SPSS: The Art of Information Processing. Analysis of statistical data and restoration of hidden patterns. M.: OOO «DiaSoft», 2005. 608 p. (in Russian).

11. Шуметов В.Г., Моисеенко А.М., Кондрашин Б.С. Применение процедуры общей линейной модели для статистической обработки результатов сортоиспытаний // Успехи современной науки. 2016. № 11. Т. 10. С. 97–101.

Shumetov V.G., Moiseenko A.M., Kondrashin B.S. Application Procedure General Linear Model for the Statistical Processing of the Results Variety Testing // Advances in modern science. 2016. № 11. T. 10. P. 97–101 (in Russian).

12. Мельник А.Ф., Кондрашин Б.С., Митюшкин Н.В. Влияние предшественников на урожайность и качество зерна озимой пшеницы // Вестник Орел ГАУ. 2009. № 4. С. 27–30.

Melnik A.F., Kondrashin B.S., Mityushkin N.V. The influence of predecessors on the yield and grain quality of winter wheat // Vestnik Orel GAU. 2009. № 4. P. 27–30 (in Russian).