

**И. В. Белько**, д-р физ.-мат. наук, проф., проф. кафедры высшей математики,  
**О. Л. Сапун**, канд. пед. наук, доц., зав. кафедрой информационных технологий и моделирования экономических процессов,  
**Е. А. Криштапович**, ст. преп. кафедры высшей математики,  
**Е. И. Подашевская**, ст. преп. кафедры информационных технологий и моделирования экономических процессов  
(Белорусский государственный аграрный технический университет,  
Минск, Республика Беларусь)

## **ОСОБЕННОСТИ ЭКОНОМЕТРИЧЕСКОГО АНАЛИЗА ПРИ ПРОГНОЗИРОВАНИИ УРОЖАЙНОСТИ ЗЕРНОВЫХ**

*Аннотация.* Отбор факторов, которые могут оказывать влияние на урожайность зернобобовых, проведен с использованием анализа наличия мультиколлинеарности с пошаговым отбором, что дает возможность создавать модели, пригодные для практической деятельности. Адекватность полученной модели подтверждается найденным интервалом прогноза.

*Ключевые слова:* эконометрическая модель, урожайность, мультиколлинеарность, ошибка прогноза, интервал прогноза.

Взаимосвязи между основными составляющими сельскохозяйственного производства носят специфический характер и требуют индивидуального подхода и учета особенностей конкретного хозяйства или региона [1], а также возможной изменчивости определяющих факторов.

Возможности современных компьютерных технологий обеспечивают исследователей полной технической поддержкой, позволяя сосредоточиться на главных вопросах: постановке задачи, отборе факторов, анализе полученных результатов, построении и оценке прогноза.

Целью работы является анализ существенности факторов, влияющих на урожайность зерновых, их взаимосвязей и возможностей использования при прогнозировании. Для выделенных значений наиболее значимых факторов находим прогнозные значения и строим интервальные оценки прогноза. Изучаем влияние одного из факторов на урожайность при зафиксированном среднем значении другого фактора.

В основе исследования содержится классический метод множественной регрессии, дополненный развернутым анализом возможностей его применения для получения адекватного результата.

В проведенном исследовании в качестве результирующего показателя выступает урожайность зернобобовых по хозяйствам Гродненской области. Экзогенные переменные, отобранные для исследования, следующие (табл. 1):

$x_1$  – балл плодородия почвы (балл);

$x_2$  – число благоприятных дней (погода);

$x_3$  – число колесных тракторов и зерноуборочных комбайнов на 100 га (техника);

$x_4$  – число орудий поверхностной обработки почвы на 100 га (орудия);

$x_5$  – количество химических удобрений, расходуемых на гектар, ц (удобрения).

Эндогенная переменная  $y$  – урожайность, ц.

## 1. Исходные данные

$y_i$	$x_{i1}$	$x_{i2}$	$x_{i3}$	$x_{i4}$	$x_{i5}$
26,9	49	53	5,4	2,8	0,09
25,6	46	56	3,4	1,7	0,23
26,2	49	52	1,2	1,8	0,15
25,8	48	58	3,0	1,7	0,34
22,6	46	62	1,9	0,9	0,18
24,5	44	57	1,4	0,7	0,09
28,5	53	54	4,7	2,1	0,12
26,5	46	55	1,2	1,1	0,61
27,8	51	62	1,7	0,9	0,48
24,9	49	60	3,1	1,6	0,57
23,2	44	57	3,1	2,5	0,26
29,5	52	52	1,2	0,7	0,35
27,0	50	61	1,2	0,6	0,47
25,5	47	64	3,9	1,6	0,41
27,1	49	66	2,0	1,0	0,31
22,6	46	54	1,7	0,8	0,09
29,2	50	56	4,7	3,3	0,48
21,8	44	55	2,9	1,5	0,36
25,1	49	54	1,8	1,0	0,62
33,5	51	66	5,9	3,9	0,64

Анализ матрицы парных коэффициентов корреляции показывает, что результативный показатель наиболее тесно связан с фактором  $x_1$  (баллом) ( $r_{yx_1} = 0,794$ ).

В то же время связь между факторами достаточно тесная: существует практически линейная связь между  $x_3$  и  $x_4$  ( $r_{x_3x_4} = 0,872$ ).

Учитывая тесную взаимосвязь этих факторов, в регрессионную модель урожайности следует включить только один из них, например  $x_4$ .

Чтобы продемонстрировать отрицательное влияние мультиколлинеарности, рассмотрим регрессионную модель урожайности, включив в нее все исходные показатели:

$$\hat{y} = -16,701 + 0,728x_1 + 0,109x_2 - 0,526x_3 + 1,694x_4 + 0,985x_5;$$

$$t_{a_1} = 5,469; t_{a_2} = 1,273; t_{a_3} = -1,068; t_{a_4} = 2,171; t_{a_5} = 0,474; F_{\text{набл}} = 10,805.$$

Для проверки гипотезы о значимости отдельных коэффициентов регрессии сравним их с критическим значением  $t_{\text{кр}} = 2,1448$ . Статистически значимыми являются коэффициенты регрессии при  $x_1$  и  $x_4$ .

Не поддается экономической интерпретации отрицательный знак коэффициента регрессии при  $x_3$ , из чего следует, что повышение насыщенности сельского хозяйства колесными тракторами и зерноуборочными комбайнами отрицательно сказывается на урожайности. Таким образом, полученное уравнение регрессии неприемлемо.

Для получения уравнения регрессии со всеми значимыми коэффициентами используем пошаговый алгоритм регрессионного анализа. Первоначально используем пошаговый алгоритм с исключением переменных [2, 3].

Исключим из модели переменную  $x_3$ , коллинеарную переменной  $x_4$ , имеющую экономически недопустимый знак и незначимый коэффициент регрессии.

Для оставшихся переменных построено уравнение регрессии значимо.

Проведя пошаговое исключение переменных, получим значимое уравнение регрессии со значимыми и интерпретируемыми коэффициентами при регрессорах:

$$\hat{y} = -10,731 + 0,733x_1 + 1,001x_4;$$

$$t_{a_1} = 5,693; \quad t_{a_4} = 2,656;$$

$$F_{\text{набл}} = 24,063.$$

Полученное уравнение является лучшей моделью урожайности для используемых данных среди выделенных факторов.

В условиях мультиколлинеарности применяется также пошаговый алгоритм с включением переменных. На первом шаге в модель урожайности включим объясняющую переменную  $x_1$ , имеющую самый высокий коэффициент корреляции с зависимой переменной  $y$ . На втором шаге включаем в уравнение наряду с  $x_1$  переменную  $x_4$  и получаем модель, которая по экономическим соображениям и статистическим характеристикам превосходит полученную ранее.

Включение в это уравнение дополнительно любой из трех оставшихся переменных ухудшает свойства модели.

Множественный коэффициент детерминации  $\hat{R}^2 = 0,739$  свидетельствует о том, что 73,9% вариации урожайности объясняется вошедшими в модель факторами  $x_1$  и  $x_4$ . Остальная часть вариации обусловлена действием неучтенных факторов.

Для построения прогноза задаем прогнозные значения  $x_1 = 54$  и  $x_4 = 4,5$ . По уравнению регрессии находим прогнозное значение урожайности, равное 33,4. Для вычисления ошибки прогноза составляем матрицу  $X$ , столбцами которой служат значения факторов  $x_1$  и  $x_4$ , первый столбец матрицы для введения свободного члена уравнения заполняем единицами.

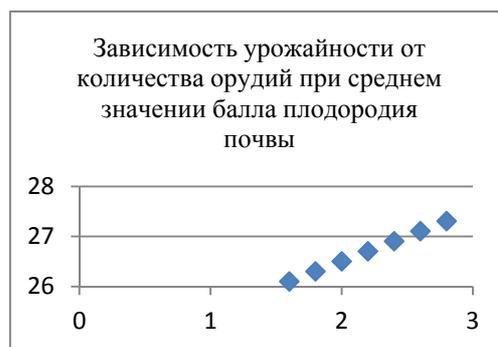
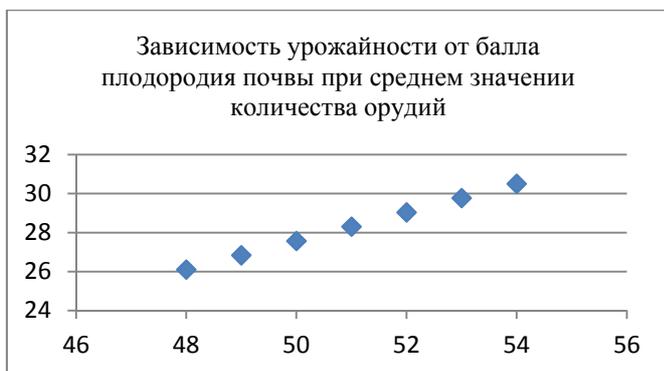
Используя матричные вычисления в Excel, находим ошибку индивидуального значения прогноза, равную 1,93. Таким образом, с вероятностью 0,95 прогнозное значение урожайности зернобобовых находится в интервале (31,5; 35,3).

Полученные результаты показывают возможность использования балла плодородия почвы и числа орудий ее поверхностной обработки для построения адекватной модели.

Найдем средние значения факторов  $x_1$  и  $x_4$  и исследуем зависимость урожайности при изменении фактора  $x_1$  и неизменном среднем значении фактора  $x_4$ . Затем фиксируем среднее значение фактора  $x_1$  и варьируем фактор  $x_4$ . Находим прогнозные значения урожайности для выбранных значений факторов. Полученные зависимости изображены на рис. 1 и 2.

Следует отметить, что среди отобранных нами показателей отсутствует показатель использования гербицидов, хотя он оказывает значительное влияние на урожайность и частично входит в состав показателя  $x_5$  – количество химических удобрений, расходуемых на гектар [4]. Причиной его невключения в данное исследование является отсутствие данных.

Особого внимания требуют показатели использования техники. Исключение из модели показателя  $x_3$  позволяет сделать вывод, что степень использования числа колесных тракторов



**Рис. 1. Зависимость урожайности от исследуемых факторов**



**Рис. 2. Зависимость ошибки прогноза от прогнозных значений**

и зерноуборочных комбайнов на 100 га (техника) должна оцениваться с учетом качества их использования. Это подтверждается мультиколлинеарностью фактора  $x_3$  с фактором  $x_4$  – числом орудий поверхностной обработки почвы.

Подводя итоги проведенного исследования, можно отметить, что главной проблемой моделирования сельскохозяйственных процессов является отсутствие не только качественно структурированных данных, но просто систематических данных.

### Список литературы

1. Производственно-экономический потенциал сельского хозяйства Беларуси: анализ и механизмы управления / Т. А. Тетеринец, В. М. Синельников, Д. А. Чиж, А. И. Попов. – Тамбов : Изд-во ФГБОУ ВО «ТГТУ», 2018. – 160 с.
2. Многомерный статистический анализ в экономических задачах: компьютерное моделирование в SPSS : учеб. пособие / под ред. И. В. Орловой.– М. : Вузовский учебник, 2013. – 310 с.
3. Эконометрика и экономико-математические методы и модели : учеб. пособие / Г. О. Читая и др. ; под ред. Г. О. Читая, С. Ф. Миксюк. – Минск : БГЭУ, 2018. – 511 с.
4. Белько, И. В. Анализ урожайности зерновых на основе метода частных наименьших квадратов / И. В. Белько, Е. А. Криштапович // Материалы 6-й Междунар. практ. конф. «Научно-инновационная деятельность в агропромышленном комплексе». – Минск : БГАТУ, 2014. – С. 45.