

DOI 10.58880/DKU.2023.02.016

МНРТ 06.35.51

УДК 332.05

## ВЛИЯНИЕ ЭКОНОМИЧЕСКИХ ФАКТОРОВ НА УРОЖАЙНОСТЬ ЗЕРНОВЫХ КУЛЬТУР В РАСТЕНИЕВОДСТВЕ РК

А.М. Тлеппаев, Д.Т. Омаров

Казахстанско-немецкий университет

Алматы, Республика Казахстан

### Аннотация

В данной статье представлена модель урожайности зерновых культур в растениеводстве РК в виде множественной линейной регрессии, проведена оценка ее качества. В качестве независимых переменных были выбраны: посевная площадь зерновых и бобовых культур, валовой сбор зерновых и бобовых культур, употребление удобрений, численность населения, валовой внутренний продукт.

Для построения множественной линейной регрессии проанализированы данные за период 1996 по 2020 годы, на основе которых построена многофакторная модель урожайности зерновых культур.

Данная модель проверена на мультиколлинеарность, гетероскедастичность. Проведена проверка на автокорреляцию.

Ключевые слова: продовольственная безопасность урожайность, множественная регрессионная модель, мультиколлинеарность, гетероскедастичность.

### Введение

Согласно определению Комитета ООН по всемирной продовольственной безопасности понятие «продовольственная безопасность» означает, что все люди в любое время имеют физический, социальный и экономический доступ к достаточному, безопасному и питательному продовольствию, которое отвечает их предпочтениям в еде и диетическим потребностям в течение жизни, позволяет вести активный и здоровый образ жизни. В ближайшие десятилетия изменение климата, рост населения планеты, рост цен на продовольствие и факторы стресса, связанные с окружающей средой, окажут значительное, но неопределенное воздействие на продовольственную безопасность. В связи с чем необходимы стратегии адаптации и политические меры реагирования на глобальные изменения, включая варианты распределения воды, модели

землепользования, торговлю продуктами питания, послеуборочную переработку продуктов питания, а также цены и безопасность продуктов питания.

В настоящее время обсуждается необходимость дальнейшего повышения потенциала урожайности пшеницы для решения текущих и предстоящих задач, включая рост потребления и потребности в зерне как в качестве топлива, так и в качестве продовольствия [1].

Технологичность производства пшеницы неодинакова между континентами и регионами. Урожайность пшеницы составляет в среднем около 3,5 т/га, но существуют большие региональные различия. Средняя урожайность выше в Восточной Азии и Европейском Союзе (4,3–5,3 т/га). В Южной Азии урожайность составляет 3 т/га, а в Африке 2,6 т/га. Эти различия можно частично объяснить различиями в управлении растениеводством (например, использование удобрений, орошение) и агроэкологии (например, температура, количество осадков, качество почвы). Для преимущественно орошаемого производства пшеницы в Египте средняя урожайность составляет 6,5 т/га. [2] В частности, в Азии высокие урожаи, прибыльность и институциональная поддержка побудили фермеров увеличить площадь выращивания пшеницы за счет других культур. [3] Пшеница является основной культурой примерно для 35% населения мира. [4] Более двух третей мировой пшеницы используется в пищу, а одна пятая — на корм скоту. Но годовое потребление пшеницы на душу населения колеблется в широких пределах — от 170 кг в Центральной Азии до 27 кг в Восточной и Южной Африке. [5] Китай и Индия потребляют по 17–18% мировой пшеницы. [6] Однако их большая и разнообразная диета предполагает, что пшеница обеспечивает в среднем по стране около 500 ккал пищевой энергии на душу населения в день. [2] Использование кормов распространено в развитых странах, но не в Африке. [6]

ФАО прогнозирует, что основные зерновые будут продолжать играть решающую роль в обеспечении продовольственной безопасности до 2050 года, обеспечивая почти половину ежедневного потребления калорий и белка в странах с низким и средним уровнем дохода.[7] Вместе с ростом населения это указывает на необходимость дальнейшего роста производительности зерновых в ближайшие десятилетия. Таким образом, основные злаки будут продолжать играть решающую роль в обеспечении достаточного и доступного потребления калорий и белков в рационе. Тем не менее, основные продукты питания не являются панацеей для обеспечения продовольственной безопасности, а разнообразие рационов питания требует дополнительных инвестиций в другие группы продуктов питания, преобразования продовольственных систем и общего развития. Например, существуют устойчивые к изменению климата и богатые питательными веществами злаки, такие как сорго и просо.[8]

Хотя исторически фермеры приспосабливались к сезонным колебаниям погоды, теперь они сталкиваются с беспрецедентными изменениями климатических условий.[9] Ожидается, что повышение температуры, а также изменение характера осадков негативно повлияют на глобальное сельскохозяйственное производство.[10]

Казахстан, будучи одним из крупнейших поставщиков зерна и зернопродуктов, является важным связующим звеном для более крупного региона Центральной Азии. Казахстан планирует увеличить поголовье скота, чтобы обеспечить 100% внутреннего спроса на мясо, но сталкивается с проблемой ограниченного качества и наличия пастбищ, на которые также влияет изменение климата. По оценке UNDP, что если существующая агротехника не изменится, то к 2030 году урожайность яровой пшеницы снизится на 13-37%, а посевная площадь уменьшится на 23-81%, а прямые экономические потери в отрасли составят 456,93 млрд тенге. А к 2050 году потери урожайности пшеницы составят 20-49%, сократив валовой сбор в отрасли до 608,19 млрд тенге в ценах 2019 года. Учитывая тот факт, что Казахстан занимает 9-е место в мире по производству и 7-е место по экспорту пшеницы и является единственным экспортером в Центральной Азии, отсутствие мер по адаптации к изменению климата в Казахстане может создать угрозу продовольственной безопасности всего региона.[11]

Гипотеза исследования заключается в предположении, что на урожайность зерновых культур в условиях Казахстана влияют такие факторы как площадь посевов, валовой сбор, количество использованных удобрений, численность населения, внутренний валовой продукт на душу населения.

#### Цель и задачи

Цель работы: построить модель урожайности зерновых культур в растениеводстве РК в виде множественной линейной регрессии и оценить ее качество.

Задачи:

1. Выбрать независимые переменные для уравнения множественной регрессии и собрать данные за период с 1996 по 2020 годы.

2. Построить уравнение регрессии для анализа урожайности зерновых культур в РК на основе динамических рядов.

3. Проверить модель на мультиколлинеарность, определить гетероскедастичность, установить наличие автокорреляции.

#### Материалы и методы

Методы исследования: множественная регрессионная модель, методы проверки факторов на наличие мультиколлинеарности, методы проверки модели на гетероскедастичность (тест Голдфельда-Квандта), методы тестирования ошибок регрессии на автокорреляцию остатков (Дарбина-Уотсона).

Для построения модели урожайности были проанализированы динамические ряды и статистические сборники с официального сайта Бюро национальной статистики РК по отрасли сельское хозяйство с 1996 года по 2020 год.

Для анализа урожайности зерновых культур рассмотрена модель множественной регрессии, где в качестве зависимой переменной рассмотрена переменная – урожайность зерновых культур.

В настоящее время разработаны и применяются на практике различные подходы к прогнозированию урожайности:

1. анализ тренда и цикличности в динамичности урожайности,
2. выявление года-аналога,
3. построение регрессионных зависимостей между различными статистическими данными и различными данными,

В настоящее время множественная регрессия – один из наиболее распространенных методов в эконометрике. Основная цель множественной регрессии – построить модель с большим числом факторов, определив при этом влияние каждого из них в отдельности, а также совокупное их воздействие на моделируемый показатель.

Построение уравнения множественной регрессии начинается с решения вопроса о спецификации модели. Он включает в себя два круга вопросов: отбор факторов и выбор вида уравнения регрессии.

Включение в уравнение множественной регрессии того или иного набора факторов связано прежде всего с представлением исследователя о природе взаимосвязи моделируемого показателя с другими экономическими явлениями. Факторы, включаемые во множественную регрессию, должны отвечать следующим требованиям.

1. Они должны быть количественно измеримы. Если необходимо включить в модель качественный фактор, не имеющий количественного измерения, то ему нужно придать количественную определенность.

2. Факторы не должны быть интеркоррелированы и тем более находиться в точной функциональной связи.[12]

Включение в модель факторов с высокой интеркорреляцией, может привести к нежелательным последствиям – система нормальных уравнений может оказаться плохо обусловленной и повлечь за собой неустойчивость и ненадежность оценок коэффициентов регрессии.

Если между факторами существует высокая корреляция, то нельзя определить их изолированное влияние на результативный показатель и параметры уравнения регрессии оказываются неинтерпретируемыми.

## РЕЗУЛЬТАТЫ

### ***1. Построение уравнения множественной регрессии***

В качестве независимых переменных были выбраны:

- посевная площадь зерновых и бобовых культур, тыс. га.;
- валовой сбор зерновых и бобовых культур (в весе после доработки), тыс. т.;
- употребление удобрений, тыс. тонн;
- численность населения на конец периода (года), тыс. чел.;

– валовой внутренний продукт методом производства на душу населения, тыс. тенге.

В таблице 1 представлены исходные данные для построения модели множественной регрессии.

Таблица 1 - Исходные данные для построения модели множественной регрессии урожайности зерновых культур[13,14]

Годы	Урожайность зерновых и бобовых культур (в весе после доработки), ц с 1 га., ут	Посевная площадь зерновых и бобовых культур, тыс. га., X1	Валовой сбор зерновых и бобовых культур (в весе после доработки), тыс. т., X2	Потребление удобрений, тыс. т., X3	Численность населения на конец периода (года), тыс. чел., X4	Валовой внутренний продукт методом производства на душу населения, тенге, X5
1996	6,5	17 187,6	11 237,3	6,3	15480,6	90 880,0
1997	8,7	15 651,4	12 378,0	7,5	15188,2	109 045,2
1998	5,6	13 526,7	6 395,5	8,2	14955,1	115 001,7
1999	13,0	11 392,5	14 264,3	7,3	14901,6	135 075,4
2000	9,4	12 438,2	11 565,0	8,0	14865,6	174 682,0
2001	12,2	13 208,7	15 896,9	11,7	14851,1	218 772,4
2002	11,5	14 022,7	15 959,9	20,9	14866,8	254 141,6
2003	10,8	13 872,6	14 777,4	29,4	14951,2	309 341,3
2004	8,8	14 278,0	12 374,2	36,0	15074,8	391 003,8
2005	10,0	14 841,9	13 781,4	34,4	15219,3	501 127,5
2006	11,7	14 839,8	16 511,5	38,7	15396,9	667 211,6
2007	13,3	15 427,9	20 137,8	54,8	15571,5	829 865,3
2008	10,1	16 190,1	15 578,2	27,9	15982,4	1 024 175,0
2009	12,6	17 206,9	20 830,5	53,0	16203,3	1 056 854,7
2010	8,0	16 619,1	12 185,2	37,5	16440,5	1 336 605,6
2011	16,9	16 219,4	26 960,5	58,2	16673,9	1 705 848,6
2012	8,6	16 256,7	12 864,8	61,1	16910,2	1 847 084,8
2013	11,6	15 877,6	18 231,1	52,3	17160,9	2 113 204,8
2014	11,7	15 291,5	17 162,2	62,6	17415,7	2 294 830,2
2015	12,7	14 982,2	18 672,8	60,6	17669,9	2 330 360,2
2016	13,5	15 403,5	20 634,4	62,8	17918,2	2 639 710,3
2017	13,4	15 405,4	20 585,1	80,8	18157,3	3014720,8
2018	13,5	15 150,0	20 273,7	53,7	18395,6	3 382 469,2
2019	12,3	15 396,6	17 428,6	51,6	18631,8	3 755 744,6
2020	12,8	15 878,4	20 065,3	55,3	18877,1	3 739 597,2

Для вычисления описательных характеристик воспользуемся функцией в Excel «Сервис. Анализ данных. Описательные статистики». В таблице 2 представлены описательные характеристики модели.

Таблица 2 – Описательные характеристики модели

	у	x1	x2	x3	x4	x5
Среднее	11,167	15062,618	16270,064	39,220	16310,377	1361494,152

Стандартная ошибка	0,506	280,350	862,842	4,426	273,122	246680,9147
Медиана	11,7	15396,629	15959,9	38,66	15982,37	1024175
Мода	11,7	#Н/Д	#Н/Д	#Н/Д	#Н/Д	#Н/Д
Стандартное отклонение	2,530	1401,752	4314,208	22,129	1365,612	1233404,574
Дисперсия выборки	6,403	1964907,913	18612394,22	489,714	1864897,287	1,52129E+12
Экссесс	0,339	0,795	0,6815	-1,1304	-1,1677	-0,83340808
Асимметричность	-0,293	-0,846	0,1314	-0,199	0,5404	0,6943
Интервал	11,3	5814,4	20565	74,48	4026,041	3664864,6
Минимум	5,6	11392,5	6395,5	6,3	14851,059	90880
Максимум	16,9	17206,9	26960,5	80,78	18877,1	3755744,6
Сумма	279,165	376565,445	406751,603	980,503	407759,43	34037353,8
Счет	25	25	25	25	25	25

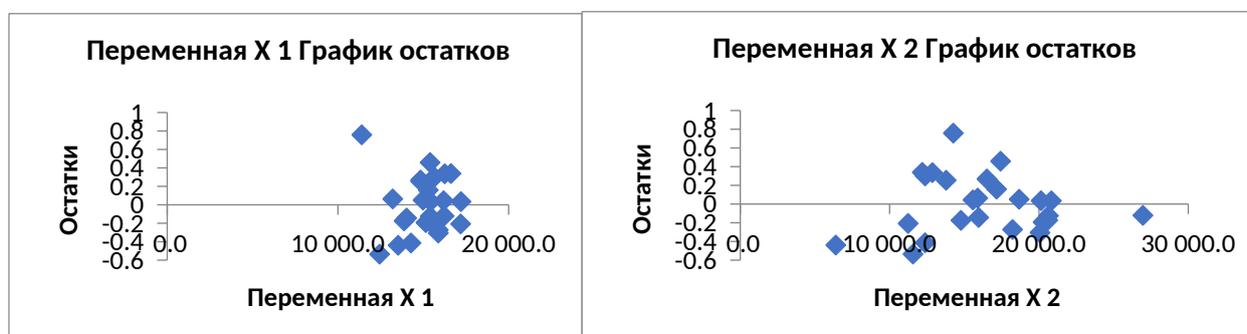
Для проведения множественного регрессионного анализа используем функцию «Сервис. Анализ данных. Регрессия».

Полученные в результате данные представлены в таблице 3.

Таблица 3 – Регрессионно-корреляционный анализ

<i>Регрессионная статистика</i>	
Множественный R	0,992534
R-квадрат	0,985124
Нормированный R-квадрат	0,98121
Стандартная ошибка	0,346852
Наблюдения	25
<i>F</i>	251,6526
<i>Значимость F</i>	1,12941E-16

На рисунках 1 представлены графики остатков.



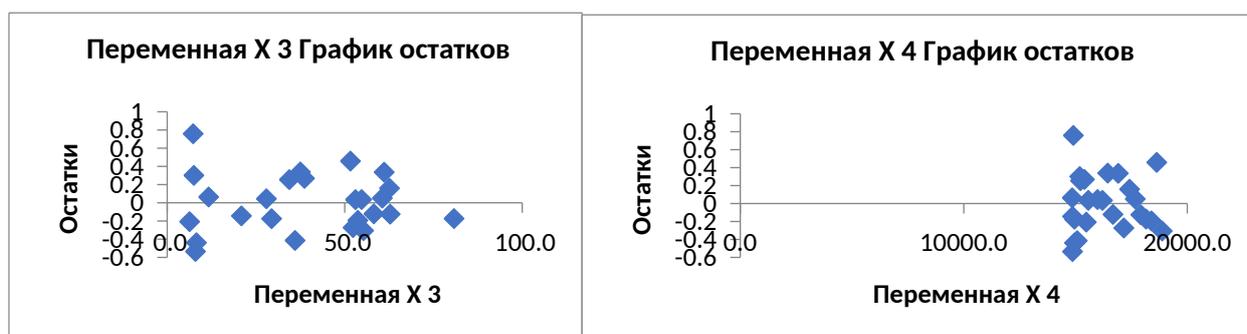


Рисунок 1 – Графики остатков.

В первую очередь обратим внимание на R-квадрат (коэффициент детерминации), значение которого равно 0,985124, или 98,5%. Это означает, что расчетные параметры модели на 98,5% объясняют зависимость между изучаемыми параметрами. Так как коэффициент детерминации выше 80%, можно считать полученную модель достаточно качественной.

В таблице 4 представлены коэффициенты уравнения множественной регрессии.

Таблица 4 – Коэффициенты уравнения множественной регрессии

Наименование коэффициентов	Значение
У-пересечение (b0)	12,33487
Переменная X1	-0,00062
Переменная X2	0,000619
Переменная X3	-0,00361
Переменная X4	-0,00013
Переменная X5	0,0000002

## **2. Проверка уравнения регрессии на мультиколлинеарность**

Мультиколлинеарность – линейная взаимосвязь двух или нескольких объясняющих переменных. Причем, если объясняющие переменные связаны строгой функциональной зависимостью, то говорят о совершенной мультиколлинеарности. На практике можно столкнуться с очень высокой (или близкой к ней) мультиколлинеарностью – сильной корреляционной зависимостью между объясняющими переменными. Мультиколлинеарность затрудняет разделение влияния объясняющих факторов на поведение зависимой переменной и делает оценки коэффициентов регрессии ненадежными. При наличии мультиколлинеарности оценки, формально полученные методом наименьших квадратов (МНК), обладают рядом недостатков и выделяются следующие последствия: увеличиваются стандартные ошибки оценок; уменьшаются t-статистики МНК-оценок регрессии; МНК-оценки чувствительны к изменениям данных; возможность неверного знака МНК-оценок;

трудность в определении вклада независимых переменных в дисперсию зависимой переменной.

При наличии корреляции один из пары связанных между собой факторов исключается. Если статистически незначим лишь один фактор, то он должен быть исключен или заменен другим показателем. В модель регрессии включаются те факторы, которые более сильно связаны с зависимой переменной, но слабо связаны с другими факторами. Другие методы устранения мультиколлинеарности: сбор дополнительных данных или новая выборка; изменение спецификации модели; использование предварительной информации о параметрах; преобразование переменных.

Для анализа мультиколлинеарности используем функцию «Анализ данных. Корреляция». В таблице 5 представлен анализ коэффициентов корреляции.

Таблица 5 – Коэффициенты корреляции

	X1	X2	X3	X4	X5
X1	1				
X2	0,337578	1			
X3	0,492694	0,716211	1		
X4	0,465737	0,578221	0,794532	1	
X5	0,388002	0,594102	0,798576	<b>0,989052</b>	1

Анализ коэффициентов корреляции показывает, что между фактором X4 (численность населения) и X5 (валовый внутренний продукт на душу населения) имеется высокое значение коэффициента корреляции 0,989, что указывает на высокую взаимосвязь между этими факторами.

В связи с этим из регрессионной модели предлагается убрать переменную X5.

Таким образом, уравнение множественной регрессии будет следующим:

$$Y=12,33 - 0,00062X1 + 0,000619X2 - 0,00361X3 - 0,00013X4$$

### 3. Определение гетероскедастичности

Одной из ключевых предпосылок МНК является условие постоянства дисперсий случайных отклонений для любых наблюдений  $D[e] = s^2 = const$ . Выполнимость данной предпосылки называется гомоскедастичностью; невыполнимость данной предпосылки называется гетероскедастичностью.

В настоящее время для определения гетероскедастичности разработаны специальные тесты и критерии для них.

Проведем тест Голдфелда–Квандта для формальной проверки остатков модели на гетероскедастичность. Нулевая гипотеза теста – остатки модели гомоскедастичны, альтернативная – остатки гетероскедастичны.

Тест состоит в следующем:

1. Все  $n$  наблюдений упорядочиваются по величине  $x_j$ .

2. Вся упорядоченная выборка разбивается на три подвыборки размерностей  $k$ ,  $n-2k$  и  $k$  соответственно.

3. Оцениваются отдельные регрессии для первой подвыборки ( $k$  первых наблюдений) и для третьей подвыборки ( $k$  последних наблюдений). Если предположение о пропорциональности дисперсий отклонений значениям  $x_j$  верно, то сумма квадратов отклонений по первой регрессии

$$S_1 = \sum_{i=1}^k e_i^2$$

будет существенно меньше суммы квадратов отклонений по третьей регрессии

$$S_3 = \sum_{i=1}^k e_i^2$$

4. Для сравнения соответствующих дисперсий выдвигается нуль – гипотеза в виде

$$H_0 : S_1^2 = S_2^2 = \dots = S_n^2,$$

которая предполагает отсутствие гетероскедастичности. Для проверки нуль – гипотезы строится следующая статистика

$$F_{набл} = \frac{S_3 / (k - m - 1)}{S_1 / (k - m - 1)} = \frac{S_3}{S_1} \quad (9.3)$$

которая при справедливости нуль – гипотезы имеет распределение Фишера с  $(k-m-1, k-m-1)$  степенями свободы.

5. Если

$$F_{набл} > F_{кр} = F_{табл}(a; k - p - 1; k - p - 1),$$

то гипотеза об отсутствии гетероскедастичности отклоняется на уровне значимости  $\alpha$ .

Разбиваем ряд данных на три группы после сортировки: первые десять, последние десять и в середине пять периодов. В таблице 6 представлены данные после сортировки.[12]

Таблица 6 – Данные после сортировки

Y	X1
13,1058	-0,3058
11,87225	-0,27225
6,707201	-0,2072
13,67997	-0,19327
10,97457	-0,17457
13,57333	-0,17333
11,64373	-0,14373
13,62495	-0,12495
17,0212	-0,1212
12,56535	0,034654
13,2645	0,035504

10,0558	0,044199
12,64889	0,051109
12,13632	0,063676
11,54049	0,159506
9,745516	0,254484
11,43172	0,26828
8,399059	0,300941
8,263838	0,336162
7,661319	0,338681
11,81957	0,458235
12,24168	0,758324

Проводим регрессионный анализ по первой и третьей группе данных.

Таблица 7 - Данные по первой группе

Наименование показателя	Значение
Множественный R	0,529821
R-квадрат	0,28071
Нормированный R-квадрат	0,177955
Стандартная ошибка	2,586664
Наблюдения	9
F	2,731824
Значимость F	0,142346
Остаток	46,83581

Таблица 8 - Данные по третьей группе

Наименование показателя	Значение
Множественный R	0,071159
R-квадрат	0,005064
Нормированный R-квадрат	-0,1193
Стандартная ошибка	1,994416
Наблюдения	10
F	0,040715
Значимость F	0,845125
Остаток	31,82156

По первой группе сумма остатков составила - 46,83581

По третьей группе - 31,82156

По формуле Голдфелда–Квандта определяем расчетную статистику, равную 1,47.

Сравним полученное значение F-статистики с критическим значением  $F_{crit}$  (для заданного уровня значимости  $\alpha = 5\% = 0,05$ ) и степеней свободы.

Расчетная F-статистика меньше Fтабл, поэтому отклоняем нулевую гипотезу о гомоскедастичности и заключаем, что остатки модели являются гетероскедастичными.

#### 4. Определение автокорреляция

Важной предпосылкой построения качественной регрессионной модели по МНК является независимость значений случайных отклонений  $e_i$  от значений отклонений во всех других наблюдениях. Отсутствие зависимости гарантирует отсутствие коррелированности между любыми отклонениями, т.е.  $\text{cov}(e_i, e_j) = 0$  и, в частности, между соседними отклонениями ( $\text{cov}(e_{i-1}, e_i) = 0$ ).

Автокорреляция (последовательная корреляция) определяется как корреляция между соседними значениями случайных отклонений во времени (временные ряды) или в пространстве (перекрестные данные). Она обычно встречается во временных рядах и очень редко – в пространственных данных. В эконометрических исследованиях часто возникают и такие ситуации, когда дисперсия остатков постоянная, но наблюдается их ковариация. Это явление называют автокорреляцией остатков. В экономических задачах значительно чаще встречается положительная автокорреляция остатков ( $\text{cov}(e_{i-1}, e_i) > 0$ ), чем отрицательная автокорреляция ( $\text{cov}(e_{i-1}, e_i) < 0$ ).

Сначала по построенному эмпирическому уравнению регрессии определяются значения отклонений  $e_t = y_t - \hat{y}_t, t = 1, n$ . Рассчитывается статистика

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad (8.3)$$

Далее по таблице критических точек Дарбина – Уотсона определяются два числа  $d_l$  и  $d_u$  и осуществляются выводы по правилу:

- $0 \leq DW \leq d_l$  - положительная автокорреляция;
- $d_l < DW < d_u$  - зона неопределенности;
- $d_u \leq DW \leq 4 - d_u$  - автокорреляция отсутствует;
- $4 - d_u < DW < 4 - d_l$  - зона неопределенности;
- $4 - d_l \leq DW \leq 4$  - отрицательная автокорреляция.

Чтобы вычислить значение DW - статистики Дарбина-Уотсона, необходимо в подтаблице «Вывод остатка» рядом со столбцом «Остатки» сформировать

дополнительные расчетные графы  $(E_i E_{i-1})^2$ ,  $E_i^2$  и определить по ним суммы (таблица 9).

Таблица 9 – Расчет DW - статистики Дарбина-Уотсона

Наблюдение	Предсказанное Y	Остатки	$e_t^2$	$(e_{t-1} - e_t)^2$
1	6,707200969	-0,20720097	0,042932241	
2	8,399058565	0,300941435	0,090565748	0,258208703
3	6,038760013	-0,43876001	0,192510349	0,547158233
4	12,24167565	0,758324353	0,575055824	1,43301098
5	9,934905094	-0,53490509	0,286123459	1,672442401
6	12,13632351	0,063676486	0,004054695	0,358299907
7	11,64373434	-0,14373434	0,020659561	0,043019252
8	10,97456519	-0,17456519	0,030473007	0,000950541
9	9,213778362	-0,41377836	0,171212533	0,057222939
10	9,745516208	0,254483792	0,064762	0,446574307
11	11,43171994	0,268280062	0,071974192	0,000190337
12	13,26449597	0,035504025	0,001260536	0,054184683
13	10,05580124	0,044198763	0,001953531	7,55985E-05
14	12,5653462	0,034653805	0,001200886	9,11062E-05
15	7,661319307	0,338680693	0,114704612	0,092432349
16	17,02120202	-0,12120202	0,014689929	0,21149211
17	8,263837608	0,336162392	0,113005154	0,209182205
18	11,87225226	-0,27225226	0,074121294	0,37016839
19	11,54049395	0,159506051	0,02544218	0,18641524
20	12,64889147	0,051108533	0,002612082	0,011750022
21	13,62494915	-0,12494915	0,015612291	0,030996308
22	13,57333496	-0,17333496	0,030045007	0,002341186
23	13,67996917	-0,19326917	0,037352974	0,000397373
24	11,81956524	0,458234764	0,209979099	0,424457381
25	13,10580362	-0,30580362	0,093515852	0,583754648
сумма			2,285819036	6,994816201
DW				3,060091849
dl				0,953
du				1,886

Согласно таблице 7 значение DW - статистики Дарбина-Уотсона равно 3,06, отсюда верно неравенство  $4 - d_l \leq DW \leq 4$ , то есть  $3,047 \leq DW \leq 4$ . Значит в регрессионной модели присутствует отрицательная автокорреляция.

## Выводы

Таким образом, можно сделать вывод, что сельское хозяйство является одним из наиболее уязвимых секторов экономики, и оценка уязвимости различных культур является важнейшей задачей для обеспечения продовольственной безопасности государства. Одной из важнейших сельскохозяйственных культур являются зерновые культуры.

В работе построена регрессионная модель урожайности зерновых культур, представлено влияние данных факторов на урожайность зерновых и бобовых культур в Республике Казахстан.

Данная модель может быть использована для определения факторов, имеющих наибольшее влияние на снижение урожайности зерновых культур и построения стратегии по обеспечению продовольственной безопасности в современных условиях.

Например, меры по увеличению производства пшеницы могут включать: выведение новых засухоустойчивых сортов пшеницы; выведение новых адаптивных сортов пшеницы, способных приносить гарантированный урожай вне зависимости от погодных условий вегетационного периода; внедрение влагосберегающих технологий, например, no-till и др.

Отметим, что перспективным направлением исследования является исследование влияния изменения климата в Республике Казахстан на урожайность не только зерновых культур, но и в целом культур, имеющих наибольшее значение для экономики страны. Это обусловлено тем, что сады, поля, пастбища, заливные луга, консервирование фруктов, выращивание и переработка зерновых, бахчевых культур, логистика имеют свой предел устойчивости к изменению климата. В связи с вышесказанным, Казахстану нужна соответствующая стратегия адаптации к изменению климата. Риски потери продуктивности сельского хозяйства в Казахстане, связанные с изменением климата (в производстве пшеницы и пастбищах), могут быть смягчены за счет внедрения адаптационных мер.

### Обсуждение

Таким образом, в результате проведенного исследования построена регрессионная модель урожайности зерновых и бобовых культур, в которой предлагается рассмотрение пяти факторов: посевная площадь (тыс.га), валовой сбор (ц), потребление удобрений (тыс.тонн), численность населения (тыс.чел.).

Уравнение регрессии имеет следующий вид:

$$Y=12,33 - 0,00062X_1 + 0,000619X_2 - 0,00361X_3 - 0,00013X_4 + 0,0000002X_5$$

Оценка значимости по F-критерию показывает, что уравнение в целом значимое. Согласно тесту Фишера регрессия адекватна, выбранный набор независимых переменных оказывает линейное влияние на зависимую переменную Y – урожайность зерновых культур.

Анализ с помощью теста Голдфелда–Квандта показал наличие гетероскедастичности в множественной регрессионной модели.

В регрессионной модели присутствует отрицательная автокорреляция.

Коэффициент 12,33 показывает, каким будет  $Y$  (урожайность), если все переменные в рассматриваемой модели будут равны 0. То есть на значение анализируемого параметра влияют и другие факторы, не описанные в модели.

В данной модели переменные  $X_1$  (посевная площадь зерновых и бобовых культур),  $X_3$  (потребление удобрений),  $X_4$  (численность населения на конец периода) имеют отрицательное значение. Чем больше посевные площади при одном и том же значении валового сбора, тем меньше урожайность.

Три переменные – посевные площади, потребление удобрений и численность населения – также имеют отрицательное влияние на зависимую переменную.

Две переменные  $X_2$  (валовой сбор) и  $X_5$  (валовой внутренний продукт на душу населения) имеют положительное влияние. Чем выше их значение, тем выше урожайность.

В нашем случае  $p$ -value для статистики Фишера составило  $1,12941E-16$ , что меньше, чем  $\alpha=0,01$ . Это означает, что с надежностью 99% отвергается нулевая гипотеза о неадекватности уравнения регрессии. Согласно тесту Фишера регрессия адекватна, выбранный набор независимых переменных оказывает линейное влияние на зависимую переменную  $Y$  – урожайность зерновых культур.

Однако в результате проделанного исследования пятый фактор – валовой внутренний продукт на душу населения – исключен в результате анализа мультиколлинеарности, так как коэффициент корреляции составил 0,989, что указывает на высокую взаимосвязь между факторами.

Отсюда можно сделать вывод, что на урожайность зерновых культур влияет четыре фактора: посевная площадь (тыс.га), валовой сбор (ц), потребление удобрений (тыс.тонн), численность населения (тыс.чел.). Уравнение регрессии будет иметь следующий вид:

$$Y=12,33 - 0,00062X_1 + 0,000619X_2 - 0,00361X_3 - 0,00013X_4$$

То есть, посевная площадь, потребление удобрений и численность населения имеют отрицательное влияние на зависимую переменную, это означает, что при увеличении площади посевов на 1 единицу урожайность будет снижаться на 0,00062 ц/га, также при увеличении потреблении удобрений на 1 единицу урожайность снизится на 0,00361ц/га, увеличение численности населения на 1 единицу урожайность будет снижаться на 0,00013ц/га.

При увеличении валового сбора на 1 единицу произойдет увеличение урожайности на 0,000619 ц/га.

Коэффициент детерминации составил 98%, это означает, что расчетные параметры модели на 98,5% объясняют зависимость между изучаемыми параметрами. На долю прочих неучтенных факторов приходится 1,5%. Так как коэффициент детерминации выше 80%, можно считать полученную модель достаточно качественной.

Исходя из результатов проведенного исследования, можно заключить, что гипотеза исследования подтвердилась в том, что на урожайность зерновых культур в Республике Казахстан влияют такие факторы как посевная площадь, валовой сбор, потребление удобрений, численность населения.

#### Список литературы

1. Grote Ulrike, Fasse Anja, Nguyen Trung Thanh, Erenstein Olaf Food Security and the Dynamics of Wheat and Maize Value Chains in Africa and Asia // *Frontiers in Sustainable Food Systems*, vol 4, 2021, 10.3389/fsufs.2020.617009 1.Тлеппаев А.М. Эконометрика. – Алматы, 2017. – 92с.
2. Dixon, J., Braun, H. J., Kosina, P., and Crouch, J., (eds.). (2009). *Wheat Facts and Futures 2009*. Mexico, D.F.: CIMMYT. Available online at: <https://repository.cimmyt.org/handle/10883/1265> (accessed December 06, 2020).
3. Hazell, P. B. R. (2009). The Asian green revolution. *IFPRI Discussion Paper 00911* (Washington, DC: International Food Policy Research Institute).
- 4.IDRC (2010). *Facts and Figures on Food and Biodiversity*. Canada: IDRC Communications, International Development Research Centre. Available online at: <https://www.idrc.ca/en/research-in-action/facts-figures-food-and-biodiversity>
5. Shiferaw, B., Smale, M., Braun, H. J., Duveiller, E., Reynolds, M., and Muricho, G. (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Secur.* 5, 291–317. doi: 10.1007/s12571-013-0263-y
6. RaboResearch (2017). *Global Wheat Consumption*. Utrecht: Rabobank. Available online at: [https://research.rabobank.com/far/en/sectors/grains-oilseeds/global\\_wheat\\_demand\\_article\\_1.html](https://research.rabobank.com/far/en/sectors/grains-oilseeds/global_wheat_demand_article_1.html)
7. FAO (2017). *The Future of Food and Agriculture: Trends and Challenges*. Rome: FAO. Available online at: <http://www.fao.org/3/a-i6583e.pdf> (accessed December 06, 2020).
8. Rodríguez, J. P., Rahman, H., Thushar, S., and Singh, R. K. (2020). Healthy and resilient cereals and pseudo-cereals for marginal agriculture: molecular advances for improving nutrient bioavailability. *Front. Genet.* 11:49. doi: 10.3389/fgene.2020.00049
9. Morton J F 2007 The impact of climate change on smallholder and subsistence agriculture Proc. Natl Acad. Sci. USA 104 19680–5
10. Funk C C and Brown M E 2009 Declining global per capita agricultural production and warming oceans threaten food security Food Security 1 271–89
11. UNDP (2020), <https://www.kz.undp.org/content/kazakhstan/en/home/stories/2020/kazakhstan-may-suffer-economic-losses-in-wheat-production-due-to.html>
12. Тлеппаев А.М. Эконометрика. Продвинутый курс: Учебное пособие. – Алматы: КНУ, 2018. – 210 с.

13.Официальный сайт Бюро национальной статистики Агентства по стратегическому планированию и реформам Республики Казахстан РК // <https://stat.gov.kz> (был доступен 26 апреля 2021г.)

14.Статистические сборники по отрасли сельское хозяйство РК за 2010-2014 гг., 2015-2019 гг.

## REFERENCES

- 1 Grote Ulrike, Fasse Anja, Nguyen Trung Thanh, Erenstein Olaf Food Security and the Dynamics of Wheat and Maize Value Chains in Africa and Asia // *Frontiers in Sustainable Food Systems*, vol 4, 2021, 10.3389/fsufs.2020.617009
- 1.Тлеппаев А.М. Эконометрика. – Алматы,
- 2 Dixon, J., Braun, H. J., Kosina, P., and Crouch, J., (eds.). (2009). *Wheat Facts and Futures 2009 Mexico*, D.F.: CIMMYT. Available online at: <https://repository.cimmyt.org/handle/10883/1265> (accessed December 06,
- 3 Hazell, P. B. R. (2009). *The Asian green revolution*. IFPRI Discussion Paper 00911 (Washington, DC: International Food Policy Research Institute).
- 4.IDRC (2010). *Facts and Figures on Food and Biodiversity*. Canada: IDRC Communications, International Development Research Centre. Available <https://www.idrc.ca/en/research-in-action/facts-figures-food-and->
- 5 Shiferaw, B., Smale, M., Braun, H. J., Duveiller, E., Reynolds, M., and Muricho, G. (2013). Crops that feed the world 10 Past successes and future challenges to the role played by wheat in global food security. *Food Secur.* 5, 291–317. doi: 10.1007/s12571-013-0263-y
- 6 RaboResearch (2017). *Global Wheat Consumption*. Utrecht: Rabobank. [https://research.rabobank.com/far/en/sectors/grains-oilseeds/global\\_wheat\\_demand\\_article\\_1.html](https://research.rabobank.com/far/en/sectors/grains-oilseeds/global_wheat_demand_article_1.html)
- 7 FAO (2017). *The Future of Food and Agriculture: Trends and Challenges*. Rome: FAO. Available online at: <http://www.fao.org/3/a-i6583e.pdf> (accessed December 06, 2020).
- 8 Rodríguez, J. P., Rahman, H., Thushar, S., and Singh, R. K. (2020). *Healthy and resilient cereals and pseudo-cereals for marginal agriculture:*

- molecular advances for improving nutrient bioavailability. Front. Genet. 11:49.
- 9 Morton J F 2007 The impact of climate change on smallholder and subsistence agriculture Proc. Natl Acad. Sci. USA 104 19680–5
- 10 Funk C C and Brown M E 2009 Declining global per capita agricultural production and warming oceans threaten food security Food UNDP (2020), <https://www.kz.undp.org/content/kazakhstan/en/home/stories/2020/kazakhstan-may-suffer-economic-losses-in-wheat-production-due-to.html>
- 12 Tleppaev A.M. Ekonometrika. Prodvintuy kurs: Uchebnoe posobie.
13. Oficial'nyj sajt Byuro nacional'noj statistiki Agentstva po strategicheskomu planirovaniyu i reformam Respubliki Kazahstan RK <https://stat.gov.kz> (byl dostupen 26 aprelya 2021g.)
14. Statisticheskie sborniki po otrasli sel'skoe hozyajstvo RK za 2010-2014gg., 2015-2019gg.

### Түйіндеме

#### ҚР ӨСІМДІК ШАРУАШЫЛЫғыНДАғы ДӘНді ДАҚЫЛДАРдың ӨНІМдіЛІГІНЕ ЭКОНОМИКАЛЫҚ ФАКТОРЛАРдың ӘСЕРІ

ТЛЕППАЕВ А.М.1, ОМАРОВ Д.Т.2

1.Қазақстан-Неміс Университеті, Алматы, Қазақстан Республикасы

Бұл мақалада ҚР өсімдік шаруашылығындағы дәнді дақылдардың өнімділік моделі бірнеше сызықтық регрессия түрінде ұсынылған, оның сапасына баға берілген. Тәуелсіз айнымалылар ретінде таңдалды: дәнді және бұршақты дақылдардың егістік алқабы, дәнді және бұршақты дақылдардың жалпы жиналуы, тыңайтқыштарды қолдану, халық саны, жалпы ішкі өнім.

Бірнеше сызықтық регрессияны құру үшін 1996-2020 жылдар аралығындағы мәліметтер талданды, олардың негізінде дәнді дақылдар өнімділігінің көп факторлы моделі құрылды.

Бұл модель мультиколлинеарлыққа, гетероскедастикаға тексерілген. Автокорреляцияға тексеру жүргізілді.

Көптік регрессия теңдеуін құру және модельді тексеру нәтижесінде ұсынылған факторлардың әсері расталды. Бұл модельде айнымалылар X1 (дәнді және бұршақты дақылдардың егіс алаңы), X3(тыңайтқыштарды тұтыну), X4 (кезең соңындағы халық саны) теріс мәнге ие. Жалпы жинаудың бірдей мәні бар

егістік алқаптары неғұрлым көп болса, өнімділік соғұрлым аз болады.

Үш айнымалы – егістік алқаптары, тыңайтқыштарды тұтыну және халық саны – тәуелді айнымалыға теріс әсер етеді.

Екі айнымалы X2 (жалпы алым) және X5 (жан басына шаққандағы жалпы ішкі өнім) оң әсер етеді. Олардың мәні неғұрлым жоғары болса, өнімділік соғұрлым жоғары болады.

Түйін сөздер: азық-түлік қауіпсіздігі, өнімділік, бірнеше регрессиялық модель, мультиколлинеарлық, гетероскедастикалық.

#### Summary

### INFLUENCE OF ECONOMIC FACTORS ON THE PRODUCTIVITY OF GRAIN CROPS IN CROP PRODUCTION IN THE REPUBLIC OF KAZAKHSTAN

A.M. Tleppeev, D.T. Omarov

Kazakh-German University

Almaty, Republic of Kazakhstan

This article presents a model of grain crop yield in the crop production of the

Republic of Kazakhstan in the form of multiple linear regression, an assessment of its quality is carried out. The following were chosen as independent variables: sown area of grain and legume crops, gross yield of grain and legume crops, fertilizer use, population, gross domestic product.

In order to build a multiple linear regression we analyzed the data for the period 1996 to 2020, based on which we built a multifactor model of grain crop yield.

This model was tested for multicollinearity, heteroscedasticity. The test for autocorrelation was carried out.

As a result of the construction of the equation of multiple regression and verification of the model the influence of the proposed factors is confirmed. In this model the variables X1 (sown area of cereals and legumes), X3 (fertilizer consumption), X4 (population at the end of the period) have a negative value. The larger the acreage at the same value of gross harvest, the lower the yield.

Three variables – acreage, fertilizer consumption and population – also have a negative impact on the dependent variable.

Two variables X2 (gross revenue) and X5 (gross domestic product per capita) have a positive impact. The higher their value, the higher the yield.

Keywords: food security, productivity, multiple regression model, multicollinearity, heteroscedasticity.