

С.Сейфуллин атындағы Қазақ агротехникалық университетінің Ғылым жаршысы (пәнаралық) = Вестник науки Казахского агротехнического университета им. С.Сейфуллина (междисциплинарный). - 2018. - №1 (96). - С.33-40

Статистические свойства и прогнозирование урожайности зерновых в северном зерносеющем регионе Казахстана

Кусаинов Т., Жакупова Ж.О.

Аннотация

В прогнозировании экономических процессов одним из наиболее часто используемых инструментов анализа являются модели временных рядов. При этом часто будущие значения рассматриваемой переменной являются функцией прошлых значений этой же переменной. Другими словами, процесс является авторегрессионным. Динамика урожайности зерновых культур в Северо-Казахстанской и Костанайской областях демонстрируют весьма схожие стохастические свойства. В обоих случаях имеет место положительный линейный тренд, отчетливо прослеживается циклическое развитие процесса. Серьезного внимания заслуживает также наличие цикла в динамике уровня дисперсии урожайности культур. Указанные стохастические особенности показателя необходимо учитывать в сельскохозяйственном прогнозировании.

Ключевые слова: урожайность зерновых, временной ряд, стохастические характеристики, дисперсия, корреляция, автокорреляция, авторегрессия, прогнозирование.

Введение

Ценность прогноза, в большей степени, определяется его точностью, которая зависит от степени схожести будущего значения процесса с оценкой этого значения, сделанного заранее. Поскольку наблюдение за реальным процессом происходит в условиях различных помех, а протекание самих процессов подвержено влиянию различных факторов, то уверенность в точности прогноза будущего значения процесса не может быть абсолютной.

Практика разработки прогнозов опирается на систему методов, среди которых статистические методы прогнозирования занимают особое место. При статистическом подходе к прогнозированию решающую роль играет выбор соответствующей модели. Располагая различными моделями, можно получить варианты прогноза, отвечающие определенным условиям и гипотезам [1-4]. Динамические процессы, происходящие в экономических

системах, обычно представляются в виде ряда значений некоторого экономического показателя, расположенных в хронологическом порядке. В прогнозировании экономических процессов одним из наиболее часто используемых инструментов анализа являются модели временных рядов. Данные типа временных рядов широко распространены в самых разных областях человеческой деятельности. Предсказание будущих значений на основе прошлых наблюдений, а также управление процессом, порождающим временной ряд, представляет собой содержание анализа временных рядов. При

построении модели того или иного ряда во многих случаях исходят из того, что будущие значения рассматриваемой переменной являются функцией прошлых значений этой же переменной. Иначе говоря, предполагается, что тот или иной уровень ряда находится в линейной зависимости от его предыдущих значений. Основная гипотеза в исследовании состоит в том, что динамика урожайности зерновых Северного региона Казахстана представляет собой авторегрессионный процесс, особенности которого необходимо учитывать при прогнозировании уровня урожайности.

Материалы и методы исследования

Объектом исследования выступает динамика урожайности зерновых в Костанайской и Северо-Казахстанской областях Северного зерносеющего региона Казахстана. Материалы – данные об урожайности культур за 1970-2017

г.г. - взяты из официальных источников. В качестве метода исследования взяты приемы и процедуры статистического анализа, в частности корреляционно-регрессионные методы, авторегрессия.

Основные результаты и обсуждение

При анализе временных рядов принято выделять 4 компоненты: 1) тренд– плавно изменяющаяся компонента, описывающая чистое влияние долговременных факторов; 2) циклическая компонента– плавно изменяющаяся компонента, описывающая длительные периоды относительного подъема и спада, состоит из циклов, меняющихся по

амплитуде и протяженности (в экономике бывает связана со взаимодействием спроса и предложения, ростом и истощением ресурсов, изменением в финансовой и налоговой политике и т.п.); 3) сезонная компонента– состоит из последовательности почти повторяющихся циклов (объемы

производства молока, цены на овощи по периодам года); 4) случайная компонента – остается после полного вычленения закономерных компонент.

Волатильность является неотъемлемым свойством результативных показателей в любой сфере экономики, тем более - в сельском хозяйстве и его важнейшей отрасли - растениеводстве. При тестировании гипотез теорий и моделей тех или иных экономических процессов и их использовании для прогнозирования исследователи обычно используют данные в виде временных рядов, то есть хронологических последовательностей результатов наблюдений. В экономике сельского хозяйства в силу особенностей отрасли волатильность результатов экономической деятельности во времени весьма значительна. Динамика урожайности сельскохозяйственных культур демонстрирует периоды существенных всплесков, за которыми следуют более спокойные периоды с меньшей турбулентностью (см., например, рис. 1. График построен на основе данных из [5]). И эта особенность в динамике экономических переменных неслучайна. Например, вполне справедливо утверждать, что в годы большой урожайности из почвы выносятся питательные элементы в значительных объемах, и если при этом процесс

восстановления продуктивных свойств почвы затягивается, то при прочих равных условиях урожайность в последующие годы объективно будет ниже [6-8]. Переход уровня цен на ресурсы и продукцию из экстремального в равновесное состояние также происходит не моментально, а с течением времени. Другими словами, сама волатильность может быть непостоянной и характеризоваться определенными закономерностями в своем поведении. В статистике такую особенность временных рядов принято называть гетероскедастичностью.

Несмотря на непостоянный, подверженный изменению во времени характер волатильности, исследователи в области сельского хозяйства все еще используют методы статистического моделирования, которые предполагают постоянство характера изменчивости временного ряда. Данное обстоятельство не может не приводить к недостаточно корректным выводам и ненадежным прогнозам. Более точные и надежные прогнозы можно получать, если удастся воспользоваться методами и моделями, позволяющими уловить и учесть в расчетах закономерности в развитии волатильности временных рядов (в этом направлении особенно продвинулись исследователи финансовых рынков, см., например, [9-12]). Следовательно, модели стохастических процессов в растениеводстве должны учитывать

возможность присутствия гетероскедастичности урожайности

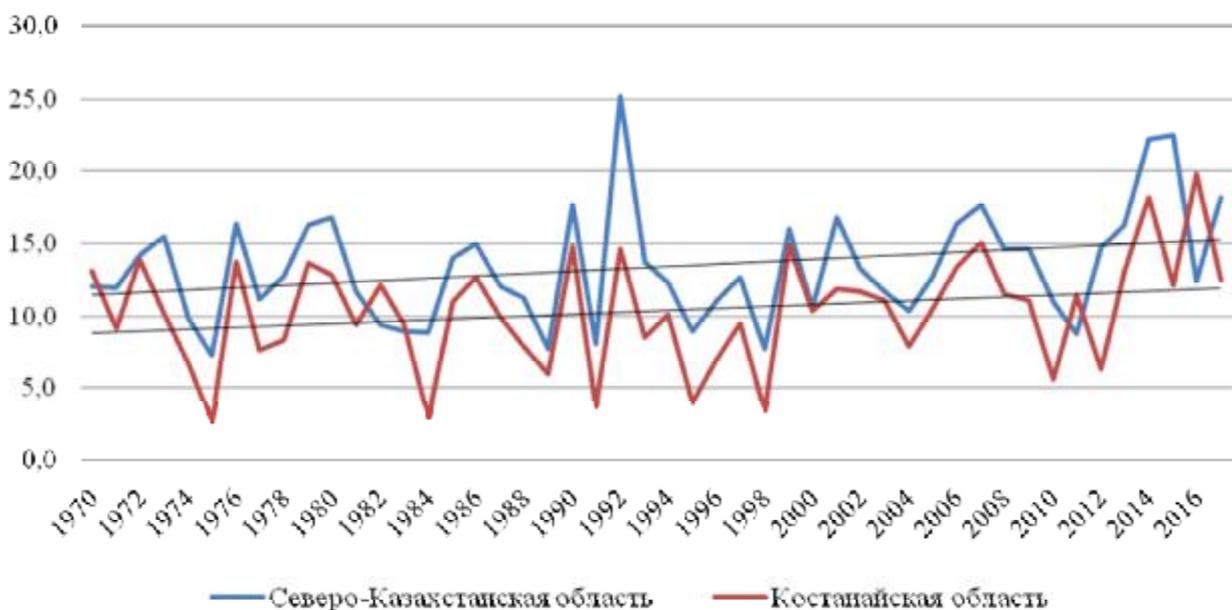
Для того, чтобы измерить силу связи между текущими и предыдущими значениями изучаемого ряда, необходимо рассчитать коэффициенты автокорреляции, которые в совокупности образуют автокорреляционную функцию. Коэффициент автокорреляции первого порядка рассчитывается с лагом в один период, коэффициент автокорреляции второго порядка показывает тесноту связи между значениями, отстоящими на два временных момента и т.д. Рассчитываются коэффициенты автокорреляции всех порядков и далее, проводится оценка значимости каждого из них. В модели оставляются только те лаги, которые являются статистически значимыми. Необходимо иметь в виду, что с увеличением лага k на

возделываемых культур, то есть непостоянства ее дисперсии.

единицу число пар значений, по которым рассчитывается коэффициент автокорреляции, уменьшается на единицу. На основе расчета нескольких коэффициентов автокорреляции определяем лаг k , при котором автокорреляция наиболее высокая. Это позволяет выявить структуру временного ряда.

Для правильной оценки ковариативных свойств временного ряда требуется выполнение условия его стационарности. Напомним, что стационарным называется временной ряд, обладающий постоянной средней и дисперсией, а ковариация внутри ряда может быть обусловлена лишь временным интервалом между отдельными наблюдениями. Говоря более строгим языком, переменная является ковариационно

Рис. 1 - Урожайность зерновых культур в Северо-Казахстанской и Костанайской областях в 1970-2017 гг.

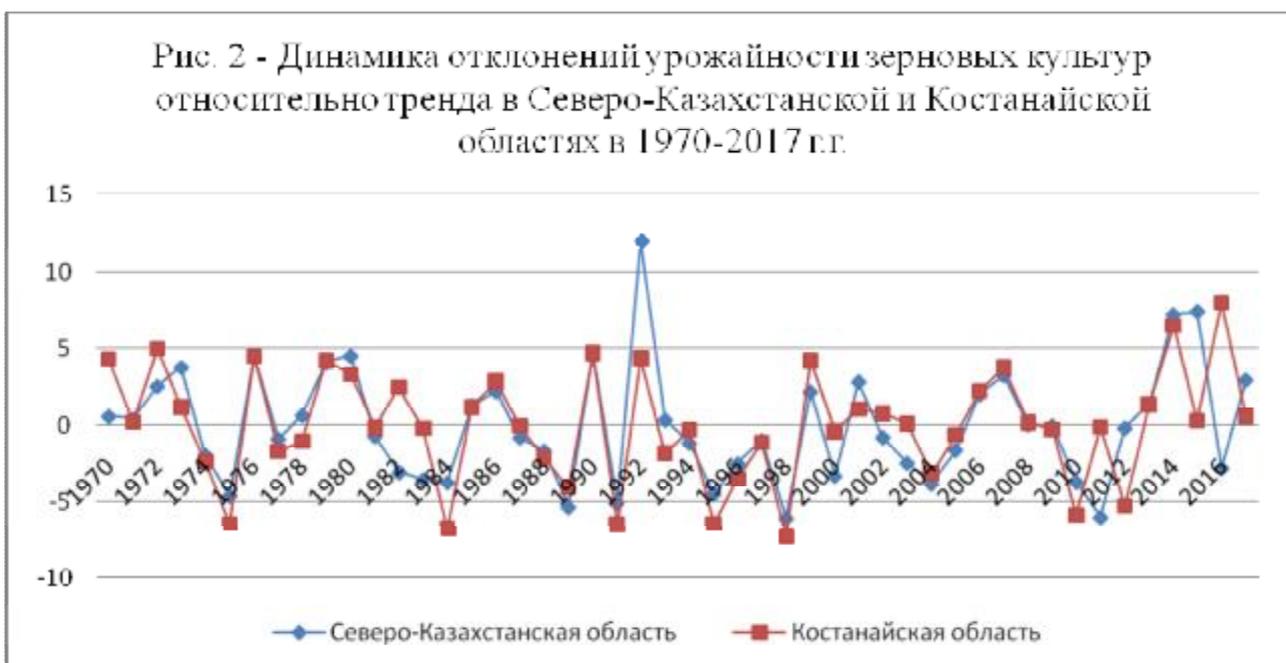


стационарной, если и $E(X_t)$ и $s^2(X_t)$ - конечные константы для корреляции между X_t и X_{t-n} является неизменным для всех t , и, следовательно, ковариация двух наблюдений X обусловлена только временем между наблюдениями.

В сельском хозяйстве немногие временные ряды являются стационарными, поскольку имеются тренды в изменении продуктивности, цены на ресурсы и продукцию имеют свойство расти со временем. К тому же, моделирование на основе нестационарных рядов оказывается проблематичным, поскольку приводит к ложной корреляции. Например, графики изменения урожайности зерновых культур в Северо-Казахстанской и Костанайской областях в период с 1970 по 2017 г.г. позволяют предположить, что имеет место

всех значений t , коэффициент определенная синхронность в скачкообразном изменении показателя в рассматриваемых областях, вызванная влиянием случайных факторов. Однако, также очевидно, что присутствует положительный тренд в обоих случаях, причиной которого является воздействие неслучайных управляемых факторов. Между тем в стохастическом анализе имеют значение случайные, неконтролируемые изменения.

Для анализа стохастических свойств временного ряда необходимо изучать колебания значений ряда относительно тренда. На рис. 2 иллюстрируется динамика отклонений урожайности зерновых по Северо-Казахстанской и Костанайской областям в 1970-2017 г.г. (по данным из [5]).



Коэффициент корреляции между временными рядами составил 0,68.

Наличие или отсутствие циклической составляющей ряда можно проверить на основе анализа коэффициентов автокорреляции при разных лагах. Графическое отображение коэффициентов автокорреляции при разных лагах

называют коррелограммой. На рисунках 3 и 4 приведены коррелограммы, построенные по первым двенадцати коэффициентам автокорреляции урожайности зерновых культур соответственно Северо-Казахстанской и Костанайской областей.

Приведенные графики позволяют выдвинуть гипотезу о том, что в динамике урожайности зерновых в рассматриваемых областях имеется цикл продолжительностью 7 лет. Коэффициент автокорреляции по динамике урожайности зерновых в Северо-Казахстанской области при лаге 7 составляет 0,40, что превышает величину удвоенной стандартной ошибки (0,31). Следовательно, полученный

коэффициент является статистически значимым на уровне доверия $\alpha=0,05$. Иначе говоря, гипотеза о циклическом развитии урожайности в области вполне жизнеспособна.

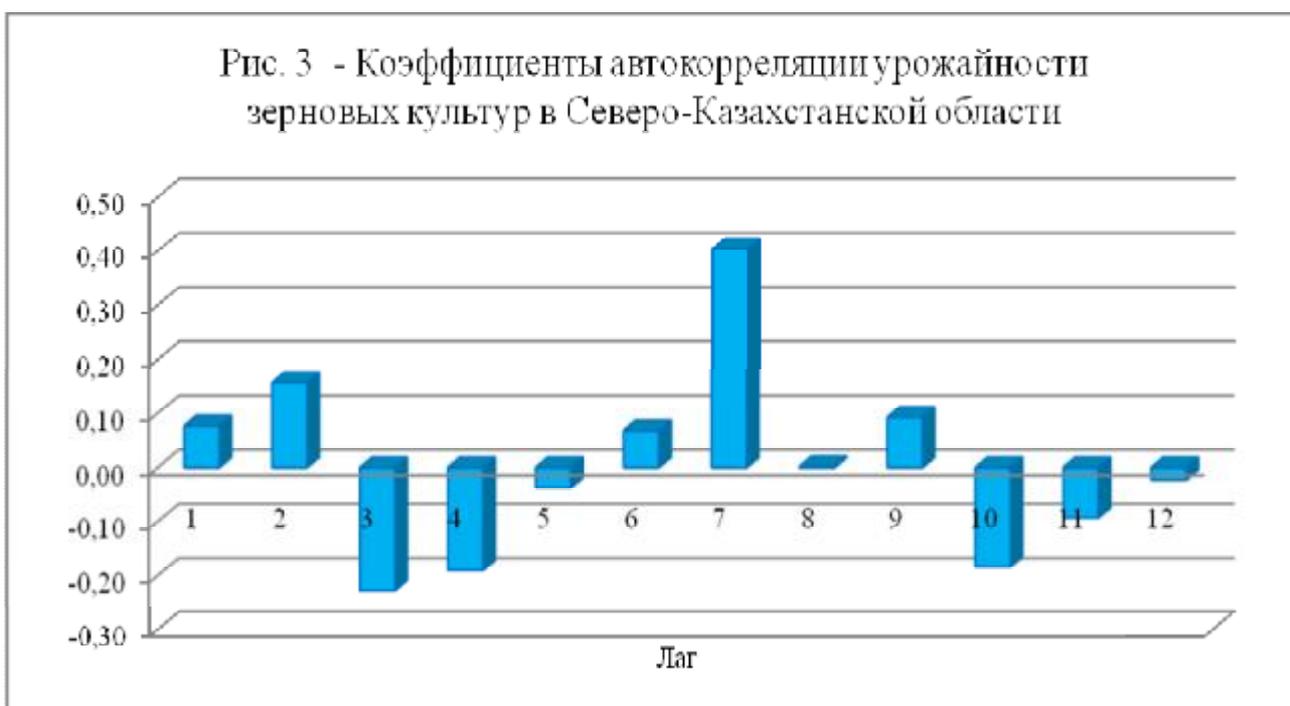
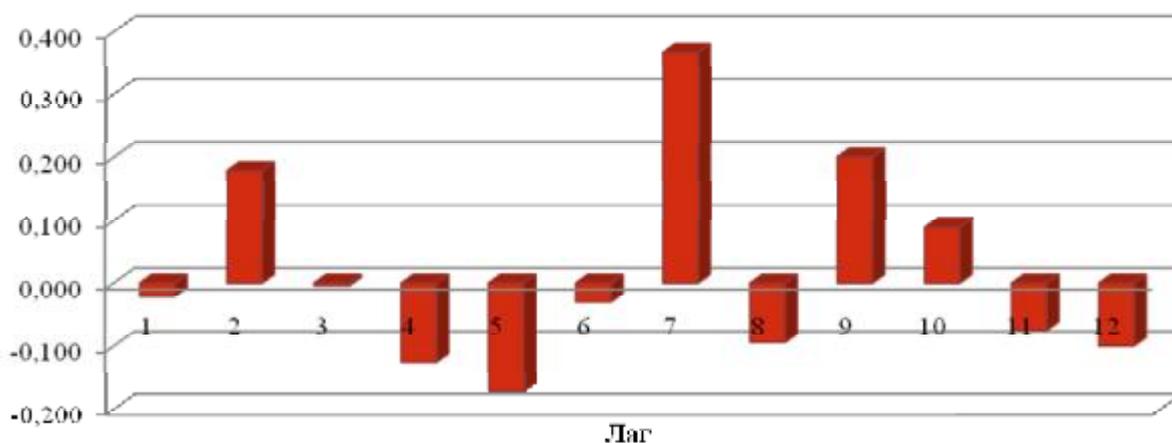


Рис. 4 - Коэффициенты автокорреляции урожайности зерновых культур в Костанайской области



К совершенно аналогичному выводу мы приходим при анализе временного ряда по Костанайской области. Здесь также мы наблюдаем цикл с периодичностью 7 лет (коэффициент автокорреляции при соответствующем лаге наибольший и составляет 0,37, в то время как его удвоенная стандартная ошибка оказывается равной лишь 0,31; другими словами, мы можем утверждать о значимости полученного коэффициента автокорреляции). Указанные свойства свидетельствуют о том, что урожайность в каждом последующем году в существенной мере предопределяется ее величиной, имевшей место 7 лет назад.

Приведенная особенность в поведении урожайности зерновых

$$Y_t = a + bt + gY_{t-7} \quad (1)$$

Таблица 1 - Урожайность зерновых по Северо-Казахстанской и Костанайской областям за 1970-2017 г.г., ц/га

| № | Год | По Северо-Казахстанской | По Костанайской | № | Год | По Северо-Казахстанской | По Костанайской |
|---|-----|-------------------------|-----------------|---|-----|-------------------------|-----------------|
| | | | | | | | |

может играть одну из ключевых ролей в прогнозировании ее уровня. Идея состоит в том, чтобы построить регрессионную модель, в которой в качестве факторной переменной служит урожайность тех же зерновых, взятых с лагом 7. В таблице 1 приведены данные наблюдений по урожайности зерновых в Северо-Казахстанской и Костанайской областях за период с 1970 по 2017 г.г.

При построении уравнения регрессии следует также учесть возможное наличие тренда в динамике урожайности культур. Графики на рисунках 1 и 2 отчетливо свидетельствуют о присутствии линейного тренда в изменении урожайности зерновых Северо-казахстанской и Костанайской областях. Поэтому уравнение регрессии в общем виде записывается следующим образом:

| п/п | | области | области | п/п | | области | области |
|-----|------|---------|---------|-----|------|---------|---------|
| 1 | 1970 | 12,1 | 13,1 | 25 | 1994 | 12,3 | 10,1 |
| 2 | 1971 | 12,0 | 9,1 | 26 | 1995 | 9 | 4 |
| 3 | 1972 | 14,2 | 13,9 | 27 | 1996 | 11,1 | 7 |
| 4 | 1973 | 15,5 | 10,2 | 28 | 1997 | 12,7 | 9,5 |
| 5 | 1974 | 9,9 | 6,7 | 29 | 1998 | 7,7 | 3,4 |
| 6 | 1975 | 7,3 | 2,7 | 30 | 1999 | 16 | 14,9 |
| 7 | 1976 | 16,4 | 13,7 | 31 | 2000 | 10,6 | 10,4 |
| 8 | 1977 | 11,2 | 7,6 | 32 | 2001 | 16,8 | 11,9 |
| 9 | 1978 | 12,8 | 8,3 | 33 | 2002 | 13,3 | 11,7 |
| 10 | 1979 | 16,3 | 13,6 | 34 | 2003 | 11,7 | 11,1 |
| 11 | 1980 | 16,8 | 12,8 | 35 | 2004 | 10,4 | 7,9 |
| 12 | 1981 | 11,7 | 9,4 | 36 | 2005 | 12,7 | 10,5 |
| 13 | 1982 | 9,4 | 12,1 | 37 | 2006 | 16,4 | 13,4 |
| 14 | 1983 | 9,0 | 9,5 | 38 | 2007 | 17,7 | 15 |
| 15 | 1984 | 8,9 | 3 | 39 | 2008 | 14,6 | 11,5 |
| 16 | 1985 | 14,0 | 11 | 40 | 2009 | 14,6 | 11,1 |
| 17 | 1986 | 15,0 | 12,7 | 41 | 2010 | 11 | 5,6 |
| 18 | 1987 | 12,1 | 9,9 | 42 | 2011 | 8,8 | 11,4 |
| 19 | 1988 | 11,3 | 7,9 | 43 | 2012 | 14,7 | 6,3 |
| 20 | 1989 | 7,7 | 6 | 44 | 2013 | 16,3 | 13 |
| 21 | 1990 | 17,7 | 14,8 | 45 | 2014 | 22,2 | 18,2 |
| 22 | 1991 | 8,1 | 3,7 | 46 | 2015 | 22,5 | 12,1 |
| 23 | 1992 | 25,2 | 14,6 | 47 | 2016 | 12,4 | 19,8 |
| 24 | 1993 | 13,7 | 8,5 | 48 | 2017 | 18,2 | 12,5 |

В уравнении (1) текущая урожайность y_t представляет собой

функцию от времени t и ее уровня семилетней давности, то есть y_{t-7} .

В нашем случае, урожайность 2017 года в существенной мере предопределилась урожайностью 2010 года, 2016 года - урожайностью 2009 года, и т.д. Уравнение по существу является авторегрессионным.

Уравнение регрессии относительно урожайности зерновых Северо-Казахстанской области принимает следующий вид:

$$Y_t = 6,17 + 0,09t + 0,43Y_{t-7} \quad (2)$$

уравнений (2) и (3) дает прогнозы, приведенные в таблице 2

Таблица 2 - Ожидаемая урожайность зерновых по Северо-Казахстанской и Костанайской областям в 2018-2024 г.г.

| № п/п | Год | Урожайность зерновых, ц/га | |
|-------|------|---------------------------------|-------------------------|
| | | По Северо-Казахстанской области | По Костанайской области |
| 1 | 2018 | 13,6 | 12,9 |
| 2 | 2019 | 16,3 | 11,1 |
| 3 | 2020 | 17,0 | 13,7 |
| 4 | 2021 | 19,6 | 15,8 |
| 5 | 2022 | 19,9 | 13,6 |
| 6 | 2023 | 15,6 | 16,6 |
| 7 | 2024 | 18,2 | 13,9 |

Анализ стохастических свойств временного ряда предполагает также оценку особенностей поведения его дисперсии. На рис. 5 приведены графики изменения дисперсии урожайности зерновых культур в Северо-Казахстанской и Костанайской областях. Ряды дисперсии построены на основе

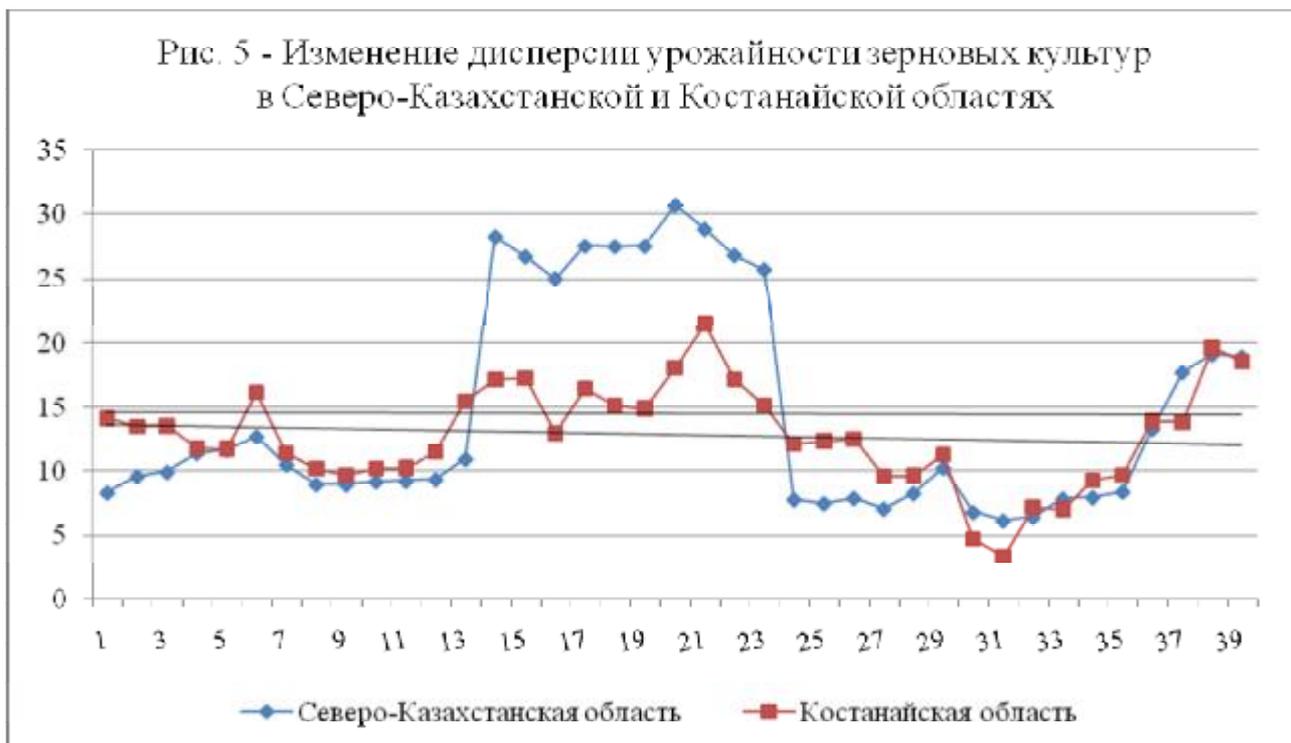
По динамике урожайности зерновых Костанайской области уравнение регрессии приняло вид:

$$Y_t = 4,74 + 0,09t + 0,38Y_{t-7} \quad (3)$$

Оба уравнения, (2) и (3), оказываются значимыми на уровне доверия 0,05. В обоих случаях, коэффициент при основной переменной Y_{t-7} также показал свою значимость на уровне 0,05; в то время как коэффициент при переменной t оказался значимым лишь на уровне 0,1. Использование

скользящих отрезков данных по урожайности зерновых (длина отрезков условно взята равной десяти). То есть, были рассчитаны дисперсии урожайности за 1970 – 1979 г.г., 1971 – 1980 г.г., 1972 – 1981 г.г., ..., 2008 – 2017 г.г. На рисунке 5 первое значение дисперсии соответствует дисперсии

за 1970 – 1979 г.г., второе – 1971 – и т.д.
1980 г.г., третье – за 1972 –1981 г.г.



Графики на рисунке 5 свидетельствуют о том, что имеет место определенная особенность в динамике дисперсии урожайности рассматриваемых культур: с циклическостью примерно 10 лет дисперсия находится выше и ниже своего среднего уровня. Данное

обстоятельство может иметь существенное значение при интервальном прогнозировании урожайности зерновых в Северо-Казахстанской и Костанайской областях.

Заключение

По результатам анализа статистических свойств динамики урожайности зерновых в Северо-Казахстанской и Костанайской областях в период с 1970 по 2017 г.г. можно сделать следующие выводы:

1) в Северо-Казахстанской и Костанайской областях имела место четко выраженная синхронность в изменении

урожайности зерновых, вызванная влиянием неконтролируемых факторов;

2) в обеих рассматриваемых областях имеет место положительный линейный тренд в урожайности зерновых;

3) динамика урожайности зерновых в исследуемых регионах демонстрирует циклические колебания с периодичностью 7 лет;

- 4) дисперсия урожайности зерновых также подвержена циклическим колебаниям: на протяжении приблизительно каждых 10 последовательных лет дисперсия находится выше или ниже своего среднего уровня;
- 5) указанные свойства урожайности зерновых культур в Северо-Казахстанской и Костанайской областях следует учитывать при

разработке сельскохозяйственных прогнозов. При этом авторегрессионные уравнения могут служить инструментом, позволяющим существенно повысить точность и обоснованность прогнозов урожайности сельскохозяйственных культур.

Список литературы

1. Анализ данных, прогнозирование и оптимизация решений: учебное пособие / Т.А. Кусаинов. – Астана, 2011. - 229 с.
2. Дрейпер Н., Смит Г. Прикладной регрессионный анализ: В 2-х книгах - Кн.1. – М.: Финансы и статистика, 1986.
3. Дрейпер Н., Смит Г. Прикладной регрессионный анализ: В 2-х книгах - Кн.2. – М.: Финансы и статистика, 1986.
4. Афанасьев В.Н., М.М. Юзбашев. Анализ временных рядов и прогнозирование. – М.: Финансы и статистика, 2001. - 228 с.
5. Официальная статистическая информация (по отраслям) / Официальный интернет-ресурс Комитета по статистике Министерства национальной экономики Республики Казахстан www.stat.gov.kz
6. Сулейменов М.К. Основы ресурсосберегающей системы земледелия в Северном Казахстане – плодосмен и нулевая или минимальная обработка почвы/Диверсификация растениеводства и No-Till как основа сберегающего земледелия и продовольственной безопасности. – Астана-Шортанды. - 2011. – С. 16-26.
7. Акшалов К., Сагимбаев М., Носачева Н., Эйхвальд Ю. Опыт диверсификации растениеводства и No-till технологии на уровне хозяйств: результаты, проблемы/Диверсификация растениеводства и No-Till как основа сберегающего земледелия и продовольственной безопасности. – Астана-Шортанды. - 2011. – С. 108-114.
8. Кияс А. Зернобобовые культуры в плодосменных севооборотах в условиях Северного Казахстана/Диверсификация растениеводства и No-Till как основа сберегающего земледелия и продовольственной безопасности. – Астана-Шортанды. - 2011. – С. 119-122.
9. Уотшем Т. Дж., К. Паррамор. Количественные методы в финансах. – М.: Финансы, ЮНИТИ, 1999. - 527 с.
10. Engle R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica*, 50, 987-1008.
11. Nelson, D.B. (1991) Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 2, 347-70.

12. Engle R.F. and Rothchild, M. (1992) ARCH models in finance. Supplement to the Journal of Econometrics, 19, 13-29.

Түйін

Зерттеуде негізгі гипотеза ретінде Қазақстанның Солтүстік өңірінде дәнді дақылдар өнімділігінің динамикасы авторегрессиалық қасиеттерге ие екендігі және астық өнімділігі бойынша болжау жасағанда сол қасиеттерді ескеру қажет деген ой алынған. Зерттеуде мәлімет ретінде 1970 - 2017 жылдар арасында Солтүстік Қазақстан және Қостанай облыстарында байқалған дәнді дақылдар өнімділігі алынған, ал зерттеу әдістері болып корреляциялық және регрессиялық талдау әдістері қолданған.

Солтүстік Қазақстан және Қостанай облыстарында дәнді дақылдар өнімділігінің динамикасы өте ұқсас стохастикалық қасиеттер көрсетеді. Екі жағдайда да оң сызықтық тренд және циклдік процесс анық көрінеді. Оның үстіне, дәнді дақылдар өнімділігі дисперсиясының динамикасында циклдық процесстердің орын алуы өзіне назар аударуды талап етеді. Ауыл шаруашылығында болжау жасағанда аталмыш көрсеткіштің стохастикалық қасиеттерін ескеру қажет. Авторегрессиялық теңдеулер ауылшаруашылық болжамдардың дәлдігі мен негізділігін жақсарту үшін құрал ретінде қызмет ете алады.

Summary

The main hypothesis in the study is that the dynamics of grain yield in the northern region of Kazakhstan is an autoregressive process, the peculiarities of which should be taken into account when predicting the yield levels.

The following conclusions can be drawn from the analysis of statistical properties of grain yield dynamics in North-Kazakhstan and Kostanay oblasts in the period from 1970 to 2017: 1) in the North-Kazakhstan and Kostanay regions there was a clearly expressed synchronicity in changing the yield of grain, caused by the influence of uncontrolled factors; 2) in both considered areas there is a positive linear trend in grain yield; 3) dynamics of grain yield in the studied regions demonstrates cyclic fluctuations with a periodicity of 7 years; 4) the dispersion of grain yields is also subject to cyclic fluctuations: for approximately every 10 consecutive years, the dispersion is above or below its average level.

In short, dynamics of grain crops yield in the northern grain producing region of Kazakhstan show a very similar stochastic properties. Noteworthy are the presence of a positive linear trend, clearly apparent cyclical traits of the yield dynamics, as well as noticeable cycles in the dynamics of the level of crop yields dispersion. These stochastic properties of the yields should be taken into account in agricultural forecasting. Autoregressive equations can serve as instruments which greatly improve the accuracy and validity of the crop yields forecasts.

Key words: grain yield, time series, stochastic characteristics, dispersion, correlation, autocorrelation, autoregression, forecasting.