

Методологические особенности долгосрочного прогнозирования урожайности зерновых культур

Игорь Леонидович Воротников, Александр Владимирович Розанов, Сергей Аркадьевич Богатырев, Аркадий Викторович Ключиков

Саратовский государственный университет генетики, биотехнологии и инженерии имени Н.И. Вавилова, г. Саратов, Россия
e-mail: tettet@inbox.ru

Аннотация. В статье представлены результаты анализа, моделирования и формирования долгосрочных прогнозов урожайности зерновых культур на материале наблюдений динамики временных рядов урожайности в России и США за длительный период, более полутора века. Для повышения надежности и достоверности результатов прогнозирования предложено использовать обобщенную логистическую функцию, что позволяет уменьшить «коридор ошибок» при больших (десять и более лет) горизонтах прогнозирования. Это особенно важно в условиях малой предсказуемости климатических факторов среды, экологических кризисов, санкций и сопутствующей им высокой волатильности урожайности на значительных периодах наблюдений.

Ключевые слова: временные ряды урожайности; прогнозирование; обобщенная логистическая функция; динамика урожайности.

Для цитирования: Воротников И. Л., Розанов А. В., Богатырев С. А., Ключиков А. В. Методологические особенности долгосрочного прогнозирования урожайности зерновых культур // Аграрный научный журнал. 2022. № 11. С. 34–37. <http://10.28983/asj.y2022i11pp34-37>.

AGRONOMY

Original article

Methodological features of long-term forecasting of grain crop yield

Igor L. Vorotnikov, Alexander V. Rozanov, Sergey A. Bogatyrev, Arkady V. Klyuchikov

Saratov State University of Genetics, Biotech and Engineering named after N.I. Vavilov, Saratov, Russia
e-mail: tettet@inbox.ru

Abstract. The article presents the results of analysis, modeling and formation of long-term forecasts of grain crop yields based on observations of the dynamics of time series of yields in the Russia and USA for a period of more than a century and a half. To increase the reliability and reliability of the forecasting results, it is proposed to use a generalized logistic function, which allows reducing the "error corridor" for large (ten or more years) forecast horizon. This is especially important in conditions of low predictability of climatic environmental factors, environmental crises, sanctions and the accompanying high volatility of yields over significant observation periods.

Keywords: yield time series; forecasting; generalized logistic function; yield dynamics.

For citation: Vorotnikov I. L., Rozanov A. V., Bogatyrev S. A., Klyuchikov A. V. Methodological features of long-term forecasting of grain crop yield. Agrarnyy nauchnyy zhurnal = Agrarian Scientific Journal. 2022;(11):34–37. (In Russ.). <http://10.28983/asj.y2022i11pp34-37>.

Введение. Производство зерновых культур является одним из приоритетных и социально-значимых направлений в сфере сельского хозяйства Российской Федерации. За последние 7 лет производство зерна в России составило более 100 млн тонн в год. По данным Росстата, в 2017 г. было собрано рекордное в истории России количество зерновых культур – 135,5 млн тонн при урожайности 29,1 ц/га, а в 2020 г. – 133 млн тонн [1]. В настоящий момент посевные площади зерновых в РФ составляют порядка 60 % от общей площади всех возделываемых сельскохозяйственных культур. Только за 2020 г. посевная площадь для производства зерновых возросла на 1 млн 321 тыс. га. Наиболее рентабельными зерновыми культурами являются пшеница (озимая и яровая), ячмень, кукуруза и овес, которые занимают значительный удельный вес в структуре посевных площадей зерновых культур РФ.

Основными экспортными культурами считаются пшеница, ячмень и кукуруза, по которым в последние годы наблюдается устойчивая тенденция роста урожайности. Следует отметить, что в настоящее время Россия, опередив США, превратилась в одного из ведущих экспортеров зерновых культур.

Экспорт зерна из России осуществляется более чем в 100 стран мира и составляет порядка 53 млн т, тогда как в США – 28,6 млн т, в основном кукуруза. В случае снижения поставок кукурузы в КНР данная сельскохозяйственная культура может стать менее востребованной и рентабельной.





Целью исследования является разработка на основе современных цифровых технологий интеллектуального анализа данных методики формирования надежных долгосрочных прогнозов урожайности зерновых культур в Российской Федерации и США – ведущих экспортеров зерна на мировом рынке.

Методика исследований. Проблемы прогнозирования временных рядов урожайности. Сведения об урожайности зерновых культур в историческом плане принято рассматривать в форме упорядоченных наборов данных, фиксирующих величину урожайности с определенной периодичностью во времени (обычно равной одному году), называемых временными рядами (англ. time-series) вида:

$$\text{Урожайность} = \text{Урожайность}(t_1, \dots, t_N),$$

где t_i – последовательные временные периоды, $t_i \in R$; N – общее число наблюдений [2].

В настоящее время для прогнозирования урожайности применяют несколько различных методов, опирающихся на корреляционно-регрессионный анализ [2], статистику степенных функций насыщения [3], пошаговые рекурсивные процедуры [4] и др. Практика показывает, что указанные методы в силу своих специфических особенностей обеспечивают сравнительно невысокую точность, надежность и достоверность прогноза. Так, например, известные пошаговые рекурсивные процедуры оказываются эффективными только для кратко- и среднесрочных прогнозов, поскольку их «коридор ошибок» очень быстро расширяется при увеличении горизонта прогнозирования.

Обобщенная логистическая функция и ее свойств. В данной статье анализ и прогнозирование урожайности предлагается проводить на основе построения статистических эквивалентов временных рядов с использованием свойств обобщенной логистической функции. В настоящее время логистические функции находят широкое применение в сфере искусственного интеллекта, экономической теории, биологии, физики, теории вероятностей и математической статистики.

Обобщенная логистическая функция $P(t)$ является решением нелинейного дифференциального уравнения первого порядка для вероятности некоторого события при изменениях управляющих параметров, называемых иногда факторами риска [5, 6].

$$P(t) = \frac{P_{\max}}{(1 + A \cdot e^{Bt})^C}, \quad (1)$$

где $P_{\max}$ – асимптотический уровень $P(t)$ при $t \in \mu$; A , B и C – константы, определяющие кривизну, скорость нарастания и асимметрию логистической кривой.

Задача прогнозирования состоит в идентификации числовых значений управляющих параметров с использованием всех имеющихся данных временного ряда урожайности.

В отличие от обычной линейной регрессии подгонка управляющих параметров проводится не на основе минимизации суммы квадратов отклонений, а путем определения вероятности того, что реальные и модельные значения урожайности максимально близки друг к другу. Это особенно важно в условиях слабой предсказуемости на больших (десятки и более лет) интервалах наблюдения климатических факторов природной среды, политических и экономических катаклизмов и сопутствующей им высокой волатильности урожайности. С принципиальной точки зрения наилучшими будут те значения A , B и C , которые максимизируют величину логарифмической функции правдоподобия для всех зарегистрированных величин урожайности [7].

Результаты исследований. На рис. 1 представлены временные ряды урожайности на исторически традиционных территориях выращивания зерновых в Российской империи, СССР и Российской Федерации (нижняя кривая) и США (верхняя кривая) [1].

Рис. 1 показывает, что динамика роста урожайности в рассматриваемых странах характеризуется, во-первых, значительной нелинейностью, а, во-вторых, высокой волатильностью, обусловленной совместным действием стохастических, квазипериодических (сезонных) и трендовых компонент различной природы, сложно связанных друг с другом. Это серьезно усложняет составление достоверных прогнозов, сокращает горизонт планирования производства товарной продукции и снижает качество основанных на них стратегических решений.

На рис. 2 и 3 представлены результаты применения авторской методики аппроксимации, анализа и прогнозирования динамики рядов урожайности зерновых в США и России на историческом материале, охватывающем 165-летний диапазон наблюдений [1].

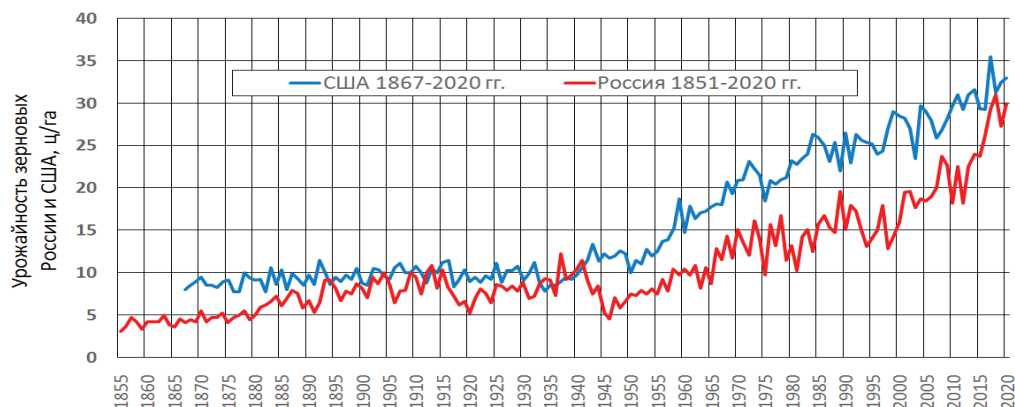


Рис. 1. Динамика урожайности зерновых в России (нижняя кривая) и США (верхняя кривая), 1855–2020 гг.

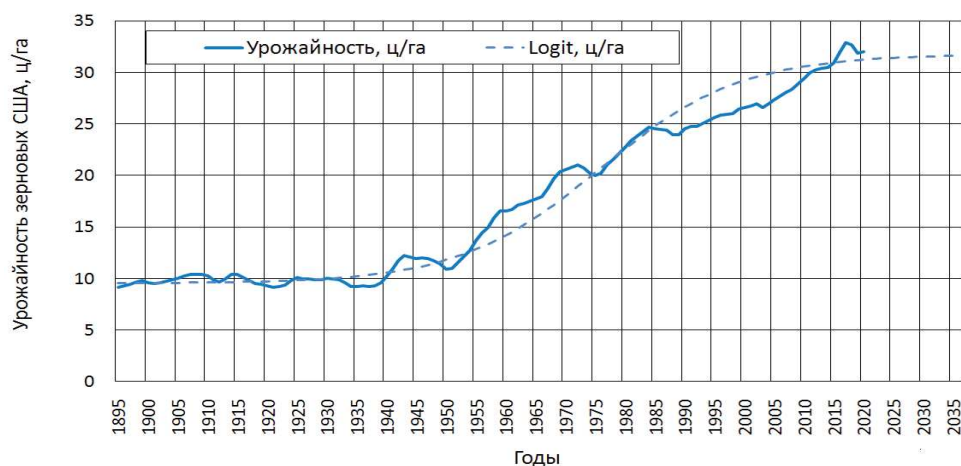


Рис. 2. Урожайность зерновых в США в 1895–2020 гг. (сплошная линия) и прогноз до 2035 г. (пунктир)

Прогнозирование выполнено с использованием надстроек «Анализ данных» и «Поиск решения» таблично-го процессора Microsoft Excel [6]. Горизонт прогнозирования был выбран равным 15 годам с целью обеспечения возможности сопоставления с прогнозом International Grains Council (Международный совет по зерну) для сезона 2025/26 года [10].

Величины факторов риска A , B и C обычно выбирают, руководствуясь опытом, интуицией или на основе экспертных оценок. Это, как правило, трудоемкий процесс, слабо поддающийся цифровизации. В настоящей работе для автоматизации процесса обработки данных и повышения точности и надежности прогнозирования значения коэффициентов A , B и C вычислялись численно средствами надстройки «Поиск решения» табличного процессора MS Excel с использованием эволюционного алгоритма нелинейной оптимизации [9].

Прогнозирование на основе логистической регрессии показало, что урожайность зерновых в США в ближайшее десятилетие будет приближаться к уровню, близкому к 32 ц/га (см. рис. 2). В долгосрочной перспективе на производство зерна в США все более негативное влияние будет оказывать глобальное потепление, следствием которого будет возрастание дефицита влаги из-за уменьшения количества осадков и вынужденное сокращение территорий сельскохозяйственного назначения.

Для Российской Федерации прогноз оказывается заметно более благоприятным. Урожайность зерновых имеет ярко выраженную тенденцию к росту (см. рис. 3). Если эта тенденция сохранится, то в предстоящее десятилетие урожайность зерновых должна приблизиться к 37–39 ц/га.

В противоположность США, глобальное потепление климата для агропромышленного комплекса Российской Федерации не является фактором, принципиально ограничивающим производство зерна. С потеплением климата зона устойчивого земледелия в России будет только расширяться, способствуя развитию зернопродуктового подкомплекса российского АПК.

Сформированные в настоящей работе прогнозы урожайности зерновых в США и Российской Федерации согласуются с результатами, представленными в материалах международных организаций. В докладе «Five-year baseline projections of supply and demand» («Пятилетние базовые прогнозы спроса и предложения») Международного совета по зерну (International Grains Council, IGC) от 11.01.2021 года урожайность зерновых в США в 2025/26 году прогнозируется на уровне 33,4 ц/га, а для РФ – на уровне 42,8 ц/га, т.е. даже выше прогноза на рис. 3 [10]. По прогнозу Продовольственной и сельскохозяйственной организации ООН (ФАО) Россия сохранит лидерство в сфере производства и экспорта зерна как минимум в течение десяти лет.

Следует отметить, что прогноз IGC построен на основе метода экспоненциального сглаживания Хольта-Винтерса, горизонт надежного прогнозирования которого составляет не более 5–6 лет.

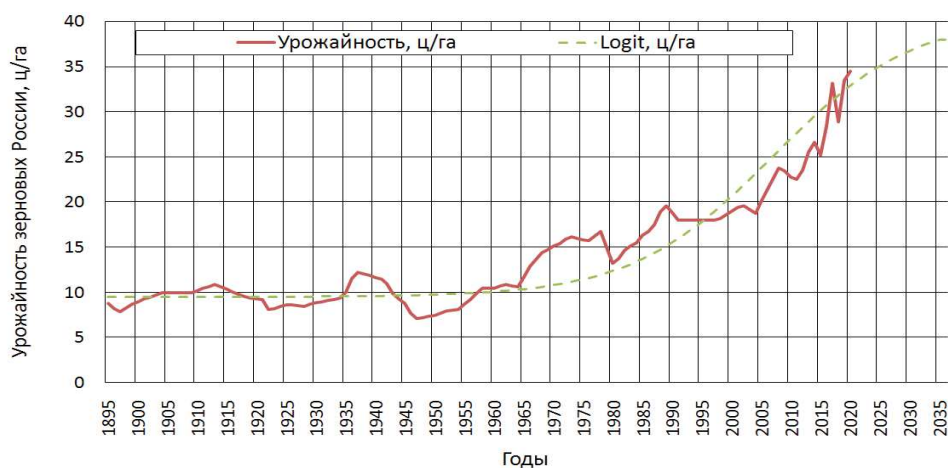


Рис. 3. Урожайность зерновых в России в 1895–2020 гг. (сплошная линия) и прогноз до 2035 г. (пунктир)





Заключение. Опираясь на представленные выше прогнозы, формируемые средствами современных технологий Big data и Data mining, можно с уверенностью констатировать, что зернопродуктовый комплекс Российской Федерации в ближайшие 10–15 лет будет удерживать лидирующие позиции в конкурентной борьбе за мировые рынки сбыта по основным видам продукции.

Дополнительным фактором увеличения урожайности и валового сбора зерна в Российской Федерации в ближайшие десятилетия должна стать масштабная программа вовлечения в оборот неиспользуемых сельскохозяйственных земель, утвержденная Правительством [11]. По данным Росстата, общая площадь земель сельскохозяйственного назначения в РФ превышает 380 млн га, из которых около 44 млн га относятся к неиспользуемым и залежным землям, причем последние составляют примерно 10 %. Большая часть неиспользуемых земель находится в северных районах Нечерноземья и Западной Сибири.

Глобальное потепление, способное негативно повлиять на урожайность и качественные показатели зерновых в США, в Российской Федерации позволит ввести в оборот обширные площади залежных земель в северных районах Нечерноземья и Западной Сибири. Это, по мнению авторитетного соруководителя крупного Гонконгского инвестиционного фонда Value Partners Group Чэа Ченг Хе (Cheah Cheng Hye), даст возможность Российской Федерации в ближайшие три десятилетия достичь статуса продовольственной сверхдержавы [12].

Переходу на более качественный уровень ведения сельского хозяйства в России будут способствовать, прежде всего, стабильно высокие сборы урожая зерновых, потребительские характеристики зерна и адекватная ценовая стратегия на мировых рынках, а также внедрение современных технологий и увеличение мощностей для хранения и переработки зерна в осенне-зимний и весенний периоды.

Разработанная авторами методика долгосрочного прогнозирования динамики временных рядов урожайности может быть адаптирована для формирования надежных прогнозов в других сферах современного сельскохозяйственного производства и агроэкономики.

СПИСОК ЛИТЕРАТУРЫ

1. АгроНовости. [Электронный ресурс]. Режим доступа: <https://agro-bursa.ru> (дата обращения: 2.02.2022).
2. Анализ и модели временных рядов [Электронный ресурс]. Режим доступа: <https://www.statmethods.ru/statistics-metody/modeli-vremennykh-ryadov> (дата обращения: 8.02.2022).
3. Милевский А.С. Эконометрика. Продвинутый уровень. М.: МИИТ, 2017. 207 с.
4. Rajesh S. B. Brief on Regression analysis // Logistic Regression Assumptions. Режим доступа: <https://medium.com/greyatom/logistic-regression-89e496433063> (дата обращения: 04.08.2022).
5. Дроздук А. Логистическая кривая. Торонто: Choven, 2019. Vol.1. 270 с.
6. Много цифр: Анализ больших данных при помощи Excel. / Джон Форман ; пер. с англ. А. Соколовой. 2-е изд. М.: Альпина Паблишер, 2017. 461 с.
7. David W., Hosmer Jr., Stanley Lemeshow, Rodney X. Sturdivant. Applied Logistic Regression. John Wiley & Sons, 2nd ed., 2020.
8. Карлсберг К. Регрессионный анализ в Microsoft Excel. М.: Диалектика, 2019. 400 с.
9. Пселтис Эндрю Дж. Потокковая обработка данных. Конвейер реального времени; пер. с англ. М.: ДМК-Пресс, 2018. 218 с.
10. Five-year baseline projections of supply and demand. [Электронный ресурс]. Режим доступа: <https://www.igc.int/en/markets/marketinfo-forecasts.aspx> (дата обращения: 2.02.2022).
11. Непашаное поле: как вернуть в оборот залежные земли [Электронный ресурс]. Режим доступа: <https://asm-agro.ru/articles/napashanoe-pole-kak-vernut-v-oborot-zaleznye-zemli> (дата обращения: 04.08.2022).
12. Siberia will be full of investment opportunities in the next 3 decades [Электронный ресурс]. Режим доступа: <https://finance.yahoo.com/news> (дата обращения: 28.09.2022).

REFERENCES

1. AgroNews. [Electronic resource]. Access mode: <https://agro-bursa.ru> (date of access: 2.02.2022). (In Russ.).
2. Analysis and models of time series [Electronic resource]. Access mode: <https://www.statmethods.ru/statistics-metody/modeli-vremennykh-ryadov> (date of access: 02.08.2022). (In Russ.).
3. Milevsky A.S. Econometrics. Advanced level. M.: MIIT; 2017. 207 p. (In Russ.).
4. Rajesh S.B. Briefly about regression analysis // Assumptions of logistic regression. Access mode: <https://medium.com/greyatom/logistic-regression-89e496433063> (accessed 08/04/2022).
5. Drozdjuk A. Logistic curve. Toronto: Chauvin, 2019. Vol. 1. 270 p. (In Russ.).
6. Many numbers: Analyzing big data with Excel / John Foreman; per. from English. A. Sokolova. 2nd ed. Moscow: Alpina Publisher; 2017. 461 p. (In Russ.).
7. David W., Hosmer Jr., Stanley Lemeshow, Rodney H. Sturdivant. Applied logistic regression. John Wiley & Sons 2nd ed. 2020.
8. Carlsberg K. Regression analysis in Microsoft Excel. M.: Dialectika, 2019. 400 p. (In Russ.).
9. Pseltis Andrew J. Streaming data processing. Real time pipeline; per. from English. M.: DMC-Press; 2018. 218 p. (In Russ.).
10. Five-year baseline demand and supply forecasts [Electronic resource]. Access mode: <https://www.igc.int/en/markets/marketinfo-forecasts.aspx> (date of access: 02.02.2022).
11. Unplowed field: how to return the turnover of fallow lands [Electronic resource]. Access mode: <https://asm-agro.ru/articles/napashanoe-pole-kak-vernut-v-oborot-zaleznye-zemli> (date of access: 08/04/2022). (In Russ.).
12. Siberia will be full of investment opportunities in the next 3 decades [Electronic resource]. Access mode: <https://finance.yahoo.com/news> (date of access: 09/28/2022).

Статья поступила в редакцию 15.09.2022; одобрена после рецензирования 19.09.2022; принята к публикации 23.09.2022.

The article was submitted 15.09.2022; approved after reviewing 19.09.2022; accepted for publication 23.09.2022.