

УДК 004: 631.1

**ПРИМЕНЕНИЕ МЕТОДОВ КОРРЕЛЯЦИОННО-
РЕГРЕССИОННОГО АНАЛИЗА С ИСПОЛЬЗОВАНИЕМ
ИНСТРУМЕНТОВ PYTHON ДЛЯ ПРОГНОЗИРОВАНИЯ
УРОЖАЙНОСТИ В УСЛОВИЯХ МЕНЯЮЩЕГОСЯ КЛИМАТА**

Кирилл Валерьевич Акиндинов

студент

bokser6831@mail.ru

Станислав Сергеевич Архипов

студент

Олег Валерьевич Пшеничный

студент

Наталья Викторовна Картечина

кандидат сельскохозяйственных наук, доцент

kartechnatali@mail.ru

Мичуринский государственный аграрный университет

г. Мичуринск, Россия

Аннотация. В статье рассматривается применение методов корреляционно-регрессионного анализа для прогнозирования урожайности в условиях изменяющегося климата. Обоснована необходимость использования современных вычислительных инструментов, в частности языка программирования Python, для обработки больших массивов агрометеорологических данных.

Ключевые слова: урожайность, прогнозирование, корреляционный анализ, регрессионный анализ, Python, климатические факторы, озимая пшеница, точное земледелие, мультиколлинеарность, автокорреляция.

В условиях глобального изменения климата точное прогнозирование урожайности стратегически важных культур, таких как озимая пшеница, приобретает ключевое значение для продовольственной безопасности.

Существующие традиционные методы оценки часто оказываются недостаточно эффективными. Они не всегда способны учесть нелинейность климатических процессов и сложность анализа больших объёмов накопленных агрометеорологических данных.

В этой связи цифровая трансформация аграрного сектора требует внедрения новых инструментов анализа [4-5]. Использование современных вычислительных методов, например, на платформе Python, позволяет:

- Автоматизировать обработку значительных массивов статистических данных;
- Выявлять сложные и скрытые взаимосвязи между метеофакторами и конечной продуктивностью посевов.

Реализация такого подхода создаёт основу для перехода к точному земледелию, основанному на данных. Однако для практического применения собранной информации необходимы математические методы, способные количественно оценить вклад каждого фактора в формирование урожая. Именно корреляционно-регрессионный анализ позволяет не только обнаружить статистически значимые связи, но и построить прогностические модели, которые могут быть адаптированы под конкретные региональные условия [1].

Теоретической основой исследования выступает корреляционный анализ, позволяющий установить наличие и тесноту статистической связи между агрометеорологическими факторами и итоговой урожайностью [2]. В агрономических системах зависимости редко носят строго функциональный характер; чаще они являются стохастическими, так как на конечный результат влияет множество случайных переменных. Для количественного описания этих связей применяется линейная и множественная регрессия.

Если парная линейная регрессия позволяет оценить вклад одного доминирующего фактора, то множественная регрессия дает возможность

построить комплексную модель, учитывающую совокупное влияние температуры, осадков и других параметров [3]. Математически такая модель представляется уравнением вида:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$$

где, y - результативный признак (урожайность), x_n - независимые факторы (метеоданные), β_n - коэффициенты регрессии, показывающие вклад каждого фактора в общий результат, а ε - случайная ошибка модели. Построение такой модели в среде Python позволяет не только найти оптимальные веса для каждого параметра [6]. Также оценивается коэффициент детерминации R^2 , который характеризует долю объясненной дисперсии урожайности выбранными погодными условиями.

Построение регрессионной модели требует обязательной проверки статистических гипотез, позволяющей подтвердить, что выявленные зависимости не являются случайными. Для оценки адекватности уравнения в целом применяется F-критерий Фишера, а значимость влияния каждого отдельного агрометеорологического фактора проверяется с помощью t-критерия Стьюдента. При уровне значимости (p-value) менее 0,05 фактор признается статистически достоверным и оставляется в модели для дальнейшего прогнозирования.

Однако при работе с климатическими данными исследователь часто сталкивается с проблемой мультиколлинеарности. Это явление возникает, когда независимые переменные (например, температура воздуха и интенсивность испарения) сильно коррелируют между собой, что приводит к неустойчивости коэффициентов регрессии и затрудняет определение истинного вклада каждого фактора.

Не менее важной проблемой является автокорреляция в рядах динамики урожайности, когда значения текущего года зависят от показателей прошлых лет (например, из-за накопленного инфекционного фона или истощения почвы). Использование инструментов Python позволяет диагностировать данные проблемы через расчет матрицы корреляций и критерия Дарбина-

Уотсона, обеспечивая тем самым высокую надежность и математическую чистоту итогового прогноза.

Несмотря на высокую точность статистических методов, существуют объективные ограничения применимости регрессионных моделей.

Во-первых, такие модели эффективны преимущественно в пределах того диапазона данных, на которых проводилось их обучение. Экстраполяция результатов на экстремальные годы с аномальными погодными условиями, не встречавшимися в выборке ранее, может привести к значительным погрешностям.

Во-вторых, статистическая модель в чистом виде ориентирована на климатические факторы и по умолчанию не учитывает динамические изменения в агротехнике, внедрение новых высокопродуктивных сортов или резкие всплески эпифитотий (массовых заболеваний растений).

Тем не менее, использование современного инструментария на языке Python позволяет частично нивелировать эти ограничения. Автоматизация процесса переобучения модели при поступлении новых данных и возможность интеграции дополнительных переменных (например, индексов вегетации или данных о внесении удобрений) делают такие системы гибкими.

Таким образом, при соблюдении статистических допущений и регулярной актуализации базы данных, регрессионный аппарат остается одним из наиболее эффективных инструментов для оперативного прогнозирования урожайности в условиях изменяющегося климата.

Понимание принципов работы коэффициентов регрессии, методов проверки гипотез и способов устранения мультиколлинеарности позволяет перейти от теоретического описания к технической реализации.

Изученные алгоритмы и выявленные ограничения определяют выбор библиотек и методов обработки данных, которые будут применены в практической части исследования на языке Python.

Список литературы:

1. Акиндинов В. В., Лосева А. С., Мягкова Е. А. Особенности эконометрического анализа при работе со статистическими данными // Инновационное развитие региона: проблемы, перспективы (IV Шаляпинские чтения): материалы Всероссийской (национальной) научно-практической конференции, Мичуринск-научоград РФ, 26 ноября 2021 года. Мичуринск-научоград РФ: Мичуринский государственный аграрный университет, 2021. С. 6-9. EDN QXRERS.
2. Акиндинов В. В. Эконометрический анализ в успешном управлении сельскохозяйственным предприятием // Устойчивое развитие экономики региона (II Шаляпинские чтения): Материалы Всероссийской научно-практической конференции, Мичуринск-научоград РФ, 18–19 декабря 2019 года / под ред. Н.В. Карамновой. Мичуринск-научоград РФ: Мичуринский государственный аграрный университет, 2019. С. 9-12. EDN GPOMBU.
3. Акиндинов В. В., Лосева А. С., Попова В. Б. Статистические методы как средство поддержки принятия решений по развитию сельского хозяйства в условиях цифровизации отрасли // Актуальные проблемы и перспективы развития государственной статистики в современных условиях - 2024: Сборник материалов Всероссийской научно-практической интернет-конференции, посвященной памяти Александра Павловича Дрючина - начальника Саратовского областного статистического управления(1958-1968 гг.), начальника ЦСУ РСФСР (1970-1985 гг.), Саратов, 18–21 ноября 2024 года. Саратов: Территориальный орган Федеральной службы государственной статистики по Саратовской области, 2025. С. 164-167. EDN FRINLS.
4. Использование искусственного интеллекта для прогнозирования урожайности / С. О. Чиркин, Н. В. Картечина, Н. П. Брозгунова, М. Е. Филиппчик // Наука и Образование. 2024. Т. 7, № 4. EDN GJSXLL.
5. Оценка эффективности методов машинного обучения для прогнозирования производства продукции растениеводства / А. А. Тусков, К. С. Советкина, С. Н. Байбекова, А. И. Долотин // Экономика и управление:

проблемы, решения. 2025. Т. 4, № 11(164). С. 202-212. DOI 10.36871/ek.up.p.r.2025.11.04.024. EDN CUDMEQ.

6. Python, корреляция и регрессия: часть 2 // Хабр – URL: <https://habr.com/ru/articles/558084/>

UDC 004: 631.1

**APPLICATION OF CORRELATION AND REGRESSION ANALYSIS
METHODS USING PYTHON TOOLS TO PREDICT YIELDS IN A
CHANGING CLIMATE**

Kirill V. Akindinov

student

bokser6831@mail.ru

Stanislav S. Arkhipov

student

Oleg V. Pshenichny

student

Natalia V. Kartechina

candidate of agricultural sciences, associate professor

kartechnatali@mail.ru

Michurinsk State Agrarian University

Michurinsk, Russia

Annotation. The article discusses the application of correlation and regression analysis methods to predict crop yields in a changing climate. The necessity of using modern computing tools, in particular the Python programming language, for processing large amounts of agrometeorological data is substantiated.

Keywords: yield, forecasting, correlation analysis, regression analysis, Python, climatic factors, winter wheat, precision farming, multicollinearity, autocorrelation.

Статья поступила в редакцию 25.02.2026; одобрена после рецензирования 20.03.2026; принята к публикации 31.03.2026.

The article was submitted 25.02.2026; approved after reviewing 20.03.2026; accepted for publication 31.03.2026.