

# ИСПОЛЬЗОВАНИЕ МЕТОДОВ РЕГРЕССИОННОГО АНАЛИЗА ДЛЯ РАСЧЕТА БИОПРОДУКТИВНОСТИ СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР НА ТЕРРИТОРИИ РЕС- ПУБЛИКИ БЕЛАРУСЬ

**Бондаренко Ю.А.**

*Белорусский государственный университет, г. Минск;*

*BondarenYU@bsu.by;*

*науч. рук. – С. А. Лысенко, д-р физ.-мат. наук, проф.*

В статье рассматривается возможность использования регрессионного анализа для прогнозирования урожайности различных сельскохозяйственных культур в Республике Беларусь. Предлагается в качестве аппроксимации использовать нелинейный тренд, а в качестве предикторов использовать метеорологические данные и данные дистанционного зондирования Земли. Результаты опробованы на статистических данных по урожайности за 2017-2019. Получена высокая оценка оправдываемости прогноза, в среднем более 85%.

**Ключевые слова:** прогноз урожайности, регрессионная модель, информационная технология, NDVI.

## **ВВЕДЕНИЕ**

В современном мире одной из важных задач любой страны является обеспечение продовольственной безопасности. В условиях изменения климата вопрос заблаговременного знания объема сельскохозяйственной продукции приобретает еще большую актуальность. Урожайность сельскохозяйственных культур — сложный с точки зрения прогнозирования показатель, что связано с тем, что на формирование урожая влияет не только производственные факторы, но и погодные условия.

В настоящее время в агрометеорологии преобладают методики для прогнозирования урожайности в период вегетации культур. За длительную историю разработки эти прогнозы доказали свою надежность и эффективность и на данный момент времени отличаются высокой точностью, эффективностью и дифференциацией по территории. Данные прогнозы характеризуются значительным разнообразием методик при заблаговременности от 1 до 3 месяцев. В территориальном плане преобладают прогнозы областного уровня, гораздо меньшее распространение в силу сложности доступа к локальным данным и метеорологическим параметрам получили методики ориентированные на административный район или отдельное хозяйство.

## ПОСТРОЕНИЕ РЕГРЕССИОННЫХ МОДЕЛЕЙ

Все многообразие факторов, влияющих на урожайность сельскохозяйственных культур, можно условно разделить на две группы: уровень культуры земледелия и метеорологические факторы. Уровень культуры земледелия оказывает значительное влияние на урожайность сельскохозяйственных культур и связан с материально-технической базой. Данный фактор можно охарактеризовать трендовой компонентой, а на отклонения от тренда будут влиять метеорологические и другие случайные факторы.

Таким образом, урожайность можно аппроксимировать следующим уравнением:

$$Y_i = T_i + dY_i,$$

где  $Y_i$  – урожайность в год,  $T_i$  – трендовый компонент, который связан с сельскохозяйственными технологиями,  $dY_i$  – случайный компонент, который обусловлен климатическими и погодными условиями в текущем году.

В качестве аппроксимации зачастую используется положительный линейный тренд, который характеризует улучшение сельскохозяйственных технологий [1,2]. Но использование линейного тренда не способно полностью отобразить характер и скорость долгосрочных изменений в материально-технической базе и среднесрочных изменений в экономической ситуации. Для учета этих факторов имеет смысл использовать нелинейный тренд. На рисунке 1 приведен пример влияния тренда на предсказание отклонения, связанного с климатической составляющей.

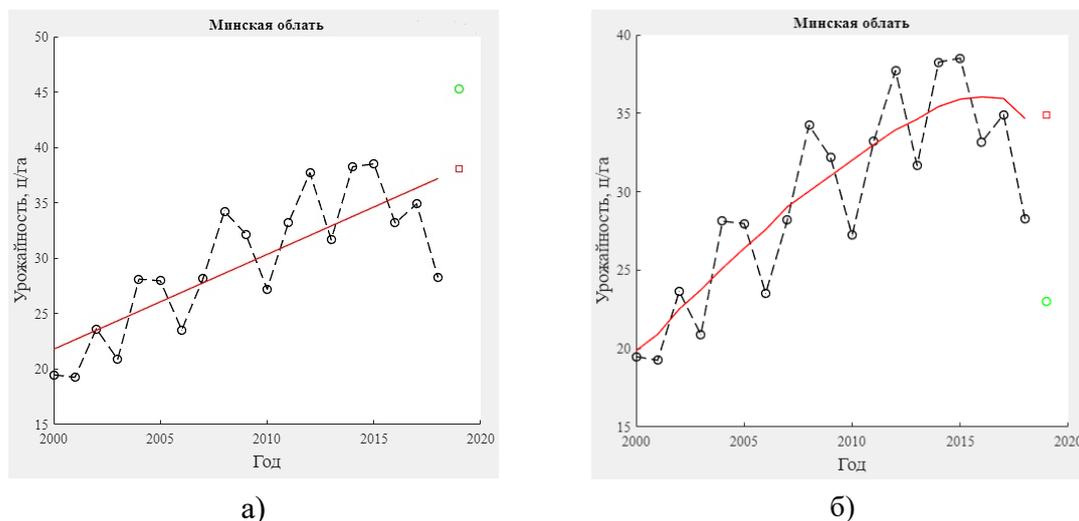


Рис. 1. Использование линейной (а) и нелинейной (б) аппроксимации для линии тренда урожайности зерновых и зернобобовых

Для вычисления климатообусловленной составляющей строится регрессионная модель, связывающая отклонение с множеством факторов,

исчерпывающий перечень которых, как правило, не известен. Общий вид регрессионной модели можно представить как:

$$dY_i=f(x_{1i},x_{2i},\dots,x_{ni}),$$

где  $x_{1i}, x_{2i}, \dots, x_{ni}$  – временные ряды факторных показателей.

В качестве регрессионной функции используем линейную, а в качестве предикторов среднедекадные показатели метеорологических параметров (температура, осадки и др.) и 16-ти дневный компонент вегетационного индекса NDVI [1].

## ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

Таблица

**Результаты прогнозирования урожайности зерновых и зернобобовых для каждой области на 2017-2019 года**

Область	Брест- ская	Витеб- ская	Гомель- ская	Гроднен- ская	Мин- ская	Могилев- ская
2017 год						
R <sup>2</sup>	0,958	0,917	0,782	0,945	0,818	0,852
Предсказанная уро- жайность, ц/га	36,4	32,1	31,4	42,5	40,3	34,9
Фактически получен- ная, ц/га	34,3	27,8	27,9	39,6	34,9	32,7
Оправдываемость, %	93,8	84,4	87,5	92,7	84,5	93,5
2018 год						
R <sup>2</sup>	0,955	0,890	0,783	0,942	0,865	0,844
Предсказанная уро- жайность, ц/га	32,3	23,7	25,8	34,8	35,6	27,1
Фактически получен- ная, ц/га	28,4	22,4	23,2	29,5	28,3	24,3
Оправдываемость, %	86,1	94,0	89,0	82,2	73,9	88,5
2019 год						
R <sup>2</sup>	0,961	0,903	0,825	0,957	0,878	0,883
Предсказанная уро- жайность, ц/га	28,7	24,8	29,7	35,9	35,5	27,0
Фактически получен- ная, ц/га	34,4	26,7	23,2	36,7	33,2	26
Оправдываемость, %	83,6	92,8	71,9	97,7	93,2	96,3

В таблице представлены результаты сопоставления данных официальной статистики по урожайности зерновых и зернобобовых культур в целом с результатами прогнозирования при использовании линейной регрессии, основанной на индексе NDVI и метеорологических параметрах,

для моделей, обученных на данных 2017-2019 гг. Произведена оценка агрометеорологических прогнозов исходя из квалификации количественных значений оправдываемости [3].

Можно заметить, что в среднем оценка оправдываемости прогнозов превышает 85%, но присутствуют неточности, например в Минской области за 2018 год и в Гомельской за 2019. Это связано с аномально низкой урожайностью и требует ввода дополнительных предикторов.

Исследование показало, что величина отклонения от тренда урожайности при прогнозировании с использованием множества предикторов может отличаться от области к области по своей величине, но при этом она совпадает по знаку. Следует отметить, что коэффициент детерминации модели не имеет прямой зависимости на оправдываемость, но влияет на стабильность результата.

## **ВЫВОДЫ**

В работе построена линейная регрессионная модель с нелинейным трендом для прогнозирования биопродуктивности сельскохозяйственных культур в Республике Беларусь на основе спутниковых данных и метеорологических параметров. В качестве предикторов использованы: вегетационный индекс NDVI, среднесуточная температура, осадки. Использование результатов прогнозирования урожайности на основе комбинированных данных наряду с результатами, полученными другими методами позволяет повысить качество прогнозов, что может повысить эффективность управления решениями в сельском хозяйстве.

### **Библиографические ссылки**

1. Куссуль Н.Н., Кравченко А.Н., Скакун С.В., Адаменко Т.И., Шелестов А.Ю., Колотий А.В., Грипич Ю.А. Регрессионные модели оценки урожайности сельскохозяйственных культур по данным MODIS // Сборник научных статей "Современные проблемы дистанционного зондирования Земли из космоса". — 2012. — Том 9, №1. — С. 95–107
2. Kogan F., Salazar L., Roytman L. Forecasting crop production using satellite-based vegetation health indices in Kansas, USA // International Journal of Remote Sensing. — 2012. — 33, N 9. — P. 2798–2814
3. ТКП 17.10-19-2017 (33140). Порядок составления и оценки агрометеорологических прогнозов. – Минск, 2017.