

## СОВРЕМЕННЫЕ НАПРАВЛЕНИЯ МОДЕЛИРОВАНИЯ УРОЖАЙНОСТИ ЗЕРНОВЫХ КУЛЬТУР

*Архипова Марина Юрьевна, профессор департамента статистики и анализа данных, НИУ ВШЭ*

**Аннотация.** *Представлена методика статистического моделирования и прогнозирования урожайности зерновых культур с использованием новых цифровых возможностей и методов машинной обработки данных.*

**Ключевые слова:** *моделирование, нейросетевые модели, урезанная выборка, зерновые культуры, прогнозирование.*

Сельское хозяйство является одной из важнейших отраслей народного хозяйства и основным поставщиком продуктов питания и сырья для многих отраслей промышленности. Сельскохозяйственный сектор России в последнее время переживает обновление и подъем благодаря интенсификации и применению современных инновационных технологий, контролю за состоянием полей с помощью космических фотоснимков на основе систем компьютерного зрения. Вместе с тем, остается еще широкий пласт задач, требующий оперативного решения. Одной из таких задач является разработка новых моделей и методов, позволяющих прогнозировать основные результирующие показатели развития сельского хозяйства и обладающих преимуществом по сравнению с существующими моделями. Для повышения точности прогнозных моделей необходимо опираться на широкий спектр доступных статистических показателей и новый современный инструментарий [1].

Цель исследования заключалась в разработке методики статистического моделирования и прогнозирования результативности производственно-финансовой деятельности в сельском хозяйстве на примере растениеводства. Современные исследования в этой сфере охватывают поиск новых источников информации и обоснование методических платформ и статистических моделей, позволяющих повысить точность и надежность традиционных моделей прогнозирования урожайности полей.

Следует отметить, что, не смотря на значимость выделенной проблемы для развития сельского хозяйства и обеспечения продовольственной безопасности страны, в России ощущается недостаток исследований, посвященных выделенной проблематике.

Для решения поставленных задач в работе использовались данные по сельскохозяйственным полям, расположенным в муниципальных образованиях регионов России. Отметим, что регионы России, также, как и количество наблюдений от каждого региона, выбирались пропорционально

## Список независимых переменных

Переменная	Название	Един. измерения	Источник
X3	Чернозем	1 – да, 0 – нет	<a href="https://www.etomesto.ru">https://www.etomesto.ru</a>
X4	NDVI на начало лета	%	<a href="https://eos.com">https://eos.com</a>
X5	NDRE на начало лета	%	<a href="https://eos.com">https://eos.com</a>
X6	NDVI на конец лета	%	<a href="https://eos.com">https://eos.com</a>
X7	NDRE на конец лета	%	<a href="https://eos.com">https://eos.com</a>
X8	Минимальная температура	t°C	<a href="https://eos.com">https://eos.com</a>
X9	Максимальная температура	t°C	<a href="https://eos.com">https://eos.com</a>
X10	Максимальные осадки	мм	<a href="https://eos.com">https://eos.com</a>
X11	Вывоз твердых коммунальных отходов в год	тысяча кубометров на душу населения	<a href="https://www.gks.ru">https://www.gks.ru</a>
X12	Доля растениеводства в сельском хозяйстве	%	<a href="https://www.gks.ru">https://www.gks.ru</a>
X13	Доля сельского населения	%	<a href="https://www.gks.ru">https://www.gks.ru</a>
X14	Наличие города	1 – да, 0 – нет	<a href="https://www.gks.ru">https://www.gks.ru</a>
logX15	Среднемесячная зарплата во всех секторах	ln(руб)	<a href="https://www.gks.ru">https://www.gks.ru</a>
logX16	Среднемесячная зарплата в сельскохозяйственном секторе	ln(руб)	<a href="https://www.gks.ru">https://www.gks.ru</a>
X17	Количество тракторов	шт / га	<a href="https://www.gks.ru">https://www.gks.ru</a>
X18	Минеральные удобрения	шт / га	<a href="https://www.gks.ru">https://www.gks.ru</a>
X19	Доля посевной площади от площади МО	%	<a href="https://www.gks.ru">https://www.gks.ru</a>
logX20	Сельское население на 1 января 2018 г.	ln(количество людней)	<a href="https://www.gks.ru">https://www.gks.ru</a>
logX21	Объем инвестиций в основной капитал (без учета бюджетных средств) на душу населения	ln(руб)	<a href="https://www.gks.ru">https://www.gks.ru</a>
X22	Специализированные продуктовые магазины на душу населения	единицы на душу населения	<a href="https://www.gks.ru">https://www.gks.ru</a>
X23	Загрязнение воздуха (превышение индекса загрязнения атмосферы больше 7)	0 – да, 0 – нет	<a href="https://www.igce.ru">https://www.igce.ru</a>
X24	Загрязнение воздуха (превышение предельно допустимой концентрации)	1 – да, 0 – нет	<a href="https://www.igce.ru">https://www.igce.ru</a>
X25	S на снежном покрытии	г / км <sup>2</sup> ·мес	<a href="https://www.meteorf.ru">https://www.meteorf.ru</a>
X26	N на снежном покрытии	кг / км <sup>2</sup> ·мес	<a href="https://www.meteorf.ru">https://www.meteorf.ru</a>
X27	pH на снежном покрытии	pH	<a href="https://www.meteorf.ru">https://www.meteorf.ru</a>
X28	Качество воды водоемов,% по классам загрязнения 4 и 5	% водных ресурсов	<a href="https://www.gidrohim.com">https://www.gidrohim.com</a>
X29	Инсоляция	кВт/м <sup>2</sup>	<a href="https://solargis.com/">https://solargis.com/</a>

В качестве основных типов моделей для прогнозирования урожайности зерновых культур были рассмотрены следующие модели.

объему продукции растениеводства данного региона в 2018 г. (удельному весу продукции растениеводства в продукции сельского хозяйства по категориям хозяйств по субъектам РФ). Это обеспечило выбор регионов, в которых продукция растениеводства является важной частью сельского хозяйства и экономики региона в целом. Исследование было проведено по данным Росстата<sup>2</sup>. Также в ряд моделей была добавлена информация космических фотоснимков вегетации полей. Космические фотоснимки вегетации полей имеют специфическую структуру: различные элементы снимков соответствуют речкам, пересушенным участкам, постройкам, незасаженным участкам и т. д., что может быть использовано при моделировании.

Для построения модели урожайности сельскохозяйственных культур использовалась одна зависимая (эндогенная) переменная  $Y_1$  - урожайность в рублях на гектар засеянных территорий муниципального округа. Данную переменную исследователи часто включают в анализ в качестве зависимой переменной, так как она отражает продуктивность поля, как в физическом объеме, так и в денежном выражении (см., например, работу Salvati L. et al. [2]. Перечень отобранных для анализа независимых (экзогенных) переменных ( $X_j$ ) представлен в таблице 1.

1. Традиционные модели, основанные на методе наименьших квадратов (OLS). Выбор этой формы модели был определен исходя из ряда статей, посвященных прогнозированию урожайности сельскохозяйственных культур (см., например, работы [3].

2. Нелинейные регрессионные модели.

3. Регрессионные модели по урезанной выборке (*truncated regression*). Преимуществом последней модели является возможность избежать ошибки прогнозирования отрицательных значений, которые могут быть получены с использованием традиционных регрессионных моделей (OLS). Выбор данной модели опирался на результаты, полученные в исследовании Basso F. et al. [4].

4. Нейросетевые модели. Выбор нейросетей в работе не случаен. В последние годы их использование набирает все большую популярность и признание среди исследователей и аналитиков в связи с высокой предсказательной способностью и возможностью интерпретации промежуточных результатов.

Сопоставление качества эконометрических моделей и нейронных сетей проводилось на основе средней квадратичной ошибки (MSE) [5], которая

---

<sup>2</sup> Росстат - gks.ru

является одна из самых распространённых функций потерь для решения задач такого рода.

В результате построения моделей и анализа полученных результатов были получены ответы на основные исследовательские вопросы. Так, было статистически доказано, что спутниковые фотоснимки являются значимым фактором для прогнозирования урожайности поля муниципалитета и повышают точность прогнозных моделей.

Сопоставление эконометрических моделей и моделей нейронных сетей позволило сделать выбор в пользу последних, которые показали лучшие результаты по прогнозированию урожайности сельскохозяйственных культур для муниципальных образований сельскохозяйственных регионов.

### **Библиографический список**

1. Архипова М.Ю., Смирнов А.А. Современные направления прогнозирования урожайности сельскохозяйственных культур на основе использования эконометрических моделей // Вопросы статистики – № 5. – 2020. – С.81-92.

2. Salvati L. et al. Exploring the relationship between agricultural productivity and land degradation in a dry region of Southern Europe //New Medit. – 2010. – Т. 9. – №. 1. – С. 35-40.

3. Мхитарян В.С., Архипова М.Ю., Дуброва Т.А., Миронкина Ю.Н., Сиротин В.П. Анализ данных: учебник для академического бакалавриата. Сер. 58 Бакалавр. Академический курс (1-е изд.) М.: Издательство Юрайт, 2019. – 490 с.

4. Basso F. et al. Evaluating environmental sensitivity at the basin scale through the use of geographic information systems and remotely sensed data: an example covering the Agri basin (Southern Italy) // Catena. – 2000. – Т. 40. – №. 1. – С. 19-35.

5. De la Casa A. et al. Soybean crop coverage estimation from NDVI images with different spatial resolution to evaluate yield variability in a plot // ISPRS journal of photogrammetry and remote sensing. – 2018. – Т. 146. – С. 531-547.

УДК 332.055.2, 311.3/4

### **ИЗМЕНЕНИЕ КЛИМАТА И ИНКЛЮЗИВНОЕ РАЗВИТИЕ СЕЛЬСКОГО ХОЗЯЙСТВА В РЕГИОНАХ РОССИИ**

*Зинченко А.П., член-корр. РАН, д.э.н., профессор ФГБОУ ВО РГАУ-МСХА имени К.А. Тимирязева*

*Демичев В.В., к.э.н., доцент ФГБОУ ВО РГАУ-МСХА имени К.А. Тимирязева*