

УДК 004.942

DOI 10.58649/1694-8033-2024-2(118)-4-11

БОГАТЫРЕВ М.М.

Ж. Баласагын атындагы КУУ

БОГАТЫРЕВ М.М.

КНУ имени Ж. Баласагына

BOGATYREV M.M.

KNU named after J. Balasagun

ORCID: 0009-0008-7791-4579

АЙЫЛ ЧАРБА ӨСҮШҮН БОЛЖОЛДООНУН МОДЕЛДЕРИ ЖАНА ЫКМАЛАРЫ, АНЫН
ИЧИНДЕ МАШИНАЛЫК ОКУУНУ КОЛДОНУУ

**МОДЕЛИ И ПОДХОДЫ К ПРОГНОЗИРОВАНИЮ РОСТА СЕЛЬСКОГО
ХОЗЯЙСТВА, В ТОМ ЧИСЛЕ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ**

**MODELS AND APPROACHES TO FORECASTING AGRICULTURAL GROWTH,
INCLUDING MACHINE LEARNING**

Кыскача мүнөздөмө: Айыл чарбасынын өсүш арымы белгилүү бир экономикалык көрсөткүчтөргө, ошондой эле өлкөдөгү макроэкономикалык жана социалдык туруктуулукка түздөн-түз таасир этет. Компьютердик технологияларды иштеп чыгуу менен татаал процесстерди, анын ичинде экономикалык процесстерди моделдөө популярдуулугуна ээ болууда, анткени ал эсептөөлөрдүн убактысын кыскартууга жана натыйжаларды алуу тактыгын жогорулатууга мүмкүндүк берет. Бул макалада айыл чарбасынын өсүш чендеринин динамикасын алдын ала айтууга мүмкүндүк берген кээ бир моделдер жана ыкмалар, анын ичинде машиналык окууну колдонгондор көрсөтүлөт.

Аннотация: Темпы роста сельского хозяйства напрямую отражаются на определенных экономических показателях, а также влияют на макроэкономическую и социальную стабильность в стране. С развитием компьютерных технологий моделирование сложных процессов, в том числе экономических, набирает популярность, поскольку позволяет сократить время на расчеты и повысить точность получения результатов. В данной статье приводятся некоторые модели и подходы, в том числе с применением машинного обучения, позволяющие прогнозировать динамику темпов роста сельского хозяйства.

Abstract: The growth rate of agriculture has a direct impact on certain economic indicators, as well as on macroeconomic and social stability in the country. With the development of computer technologies, modeling of complex processes, including economic ones, is gaining popularity, since it allows you to reduce the time for calculations and increase the accuracy of obtaining results. This article presents some models and approaches, including those using machine learning, that make it possible to predict the dynamics of agricultural growth rates.

Негизги сөздөр: математикалык моделдер; болжолдоо; айыл чарбасы; машиналык окуу; регрессия анализи.

Ключевые слова: математические модели; прогнозирование; сельское хозяйство; машинное обучение; регрессионный анализ.

Keywords: mathematical models; forecasting; agriculture; machine learning; regression analysis.

Введение. Экономические процессы представляют собой сложный механизм взаимодействия хозяйствующих субъектов, развитие которых определяется множеством различных факторов. Сельское хозяйство, как одна из составляющих экономики страны, не является исключением. Такие факторы? как погодные условия, уровень развития химической промышленности, в части производства удобрений и минеральных веществ, стоимость и достаточное количество семян, горюче-смазочных материалов и многое другое напрямую отражаются на объемах производства сельскохозяйственной продукции [1].

Многофакторность процессов осложняет анализ и прогнозирование. Однако с развитием вычислительной техники расширяются возможности для оценки перспективной динамики, в том числе сложных систем.

В данной статье дается описание некоторых наиболее распространенных моделей и подходов, которые могут быть применены для прогнозирования динамики роста сельского хозяйства. Отмечается важность формирования достаточно емкой и достоверной статистической базы данных по рассматриваемому объекту исследования (темпы роста сельского хозяйства) и возможным его факторам. Особое внимание уделяется подготовке информации, которая непосредственно участвует в модельных расчетах и на основе которой строятся прогнозы.

Модели и подходы к прогнозированию экономических показателей. В данном разделе отметим наиболее распространенные направления для оценки перспективной динамики роста сельского хозяйства. Однако, прежде чем перейти к рассмотрению возможных моделей и подходов необходимо отметить значимые факторы, непосредственно влияющие на динамику производства в сельском хозяйстве.

В первую очередь влияние идет со стороны природно-климатических условий. Маловодье, заморозки и другие погодные явления негативно отражаются на общем производстве сельскохозяйственных культур. Состояние почвы также является значимым фактором урожайности в сельскохозяйственном секторе. При этом дотации со стороны государства, научно-технические инновации и достижения, а также здоровая конкуренция в значительной степени определяют перспективы развития сельскохозяйственного сектора. Перечисленные факторы следует напрямую или косвенно отражать в моделях, что позволит улучшить качество результатов прогнозирования.

Перейдем непосредственно к описанию моделей и подходов к получению прогнозов, существующих в настоящее время и получивших наибольшее распространение [2].

Метод средних. Основывается на предположении о цикличности объекта исследования или процесса, который в определенное время в будущем возвращается к своему среднестатистическому значению. Метод предполагает расчет среднего значения переменной за определенный месяц/день/час за ряд лет/месяцев. Полученные значения принимаются в качестве прогнозных.

Факторная модель. В общем понимании представляет процесс дезагрегации объекта исследования на составляющие компоненты. После этого проводится их анализ и прогнозирование в отдельности. Полученные оценки синтезируются в прогнозные значения в целом по исследуемой переменной.

Регрессионное уравнение. Данный подход предполагает использование уравнения зависимой переменной с включением возможных ее факторов в качестве влияющих переменных. Выбор перечня объясняемых переменных основывается на теории, практике, экспертных суждениях и др. С помощью специализированного прикладного программного обеспечения (к примеру: Eviews, Stata и т.п.) ряды данных приводятся к стационарности и рассчитываются коэффициенты влияния. На основе полученных коэффициентов производится оценка будущей динамики исследуемого объекта.

VAR (Vector autoregression/векторная авторегрессия) представляет собой подход, основанный на прогнозировании по VAR-модели. Модель оценивает поведение нескольких переменных во взаимосвязи с другими переменными, входящими в VAR-модель. Подобные модели позволяют строить отклики изменения одной переменной на изменение другой, тем самым проследить теоретическую состоятельность модели. При этом ее структура представляет собой определенный шаблон, в который не представляется возможным вносить корректировки, что является неким ее недостатком. В целом VAR характеризуется как система уравнений, при этом уравнения ни что иное как авторегрессионная модель (модель временного ряда, в которой значения ряда в текущий момент линейно зависят от предыдущих значений этого же ряда, в зависимости от порядка авторегрессии) и распределенного лага (текущее значение ряда линейно зависит от своих прошлых значений и от текущих и прошлых значений других временных рядов).

Преимущества этого метода: позволяет решать сложные системы уравнений и получать приемлемые результаты моделирования. Недостатки: алгоритм расчета сложно скорректировать с учетом реалий исследуемых процессов.

На основе векторной авторегрессии разработаны модернизированные подходы. К ним можно отнести следующие: BVAR, SVAR, FAVAR и др.

BVAR (Bayesian vector autoregression/Байесовская векторная авторегрессия) отличается от VAR-модели тем, что в BVAR-модели параметры перед переменной представляют собой случайно выбранное значение, однако возможно устанавливать предполагаемое для него распределение вероятностей. Моделирование на основе данного метода позволяет получать более сглаженные и обоснованные результаты относительно результатов по VAR-модели.

Преимущества: позволяет решать сложные системы уравнений и получать приемлемые результаты моделирования, также возможно частично скорректировать алгоритм расчета с учетом реалий исследуемых процессов. Недостатки: высокая зависимость от установленного предполагаемого распределения вероятностей и исторической динамики временного ряда.

SVAR (Structure Bayesian vector autoregression/Структурная байесовская векторная авторегрессия) отличается от предыдущих моделей тем, что в SVAR параметры перед переменной возможно изменять, тем самым устанавливая желаемые взаимосвязи между объясняемыми переменными.

Преимущества: позволяет решать сложные системы уравнений на основе частично скорректированного алгоритма расчета с учетом реалий исследуемых процессов и получать приемлемые результаты моделирования. Недостатки: высокая зависимость от исторической динамики временного ряда.

VECM (Vector error correction model/Векторная модель коррекции ошибки) представляет собой вспомогательный инструмент при моделировании и добавляется к мультифакторным моделям, примером может служить VAR модель. В основе настоящего подхода лежит динамическая система, предполагающая использовать отклонения показателя от долгосрочного тренда (состояния равновесия) и тем самым корректировать результаты моделирования. Преимущества: позволяет значительно улучшить результаты моделирования. Недостатки: сложный процесс подбора корректных переменных для включения их в модель.

OLS (Ordinary Least Squares/Метод наименьших квадратов (МНК)) предполагает построение моделей для прогнозирования на основе математического метода, основанного на минимизации суммы квадратов отклонений функций от искомым значений переменных. МНК является одной из основ регрессионного анализа для оценки неизвестных параметров уравнений исследуемого объекта. Данный метод широко используется для решения систем линейных уравнений, а также систем с превышающим числом уравнений над неизвестными и систем нелинейных уравнений. Преимущества подхода: является наиболее доступным математическим методом для решения сложных задач. Важным преимуществом выступает возможность экономической интерпретации полученных коэффициентов при моделировании. Недостатки подхода: алгоритм расчета сложно скорректировать с учетом реалий исследуемого процесса, прослеживается высокая зависимость от исторической динамики временного ряда.

Приведенный выше список моделей и подходов успешно интегрирован и реализован в таких средах программного обеспечения, как: «EViews», «Stata», которые уже содержат часть кода, имеют понятный интерфейс с диалоговыми окнами и получили распространение среди начинающих моделистов. При этом для анализа, моделирования и прогнозирования применяются такие более сложные инструменты/программы, как: «Matlab», «R», «Python»,

которые требуют навыков программирования и знаний применения методов прогнозирования.

Набирают популярность моделирование и прогнозирование с использованием искусственного интеллекта. Данный подход представляет собой область компьютерных наук, исследующую возможности решения определенных задач машинами, требующими человеческого интеллекта. Реализация искусственного интеллекта осуществляется различными методами. В рамках статьи следует отметить такой метод, как машинное обучение, которое позволяет обрабатывать большой набор данных [3]. При помощи такой разновидности машинного обучения как глубокое обучение, обработанный массив данных используется для обучения нейронной сети с последующей генерацией прогнозов, либо группировки составляющих объекта исследования, либо других задач.

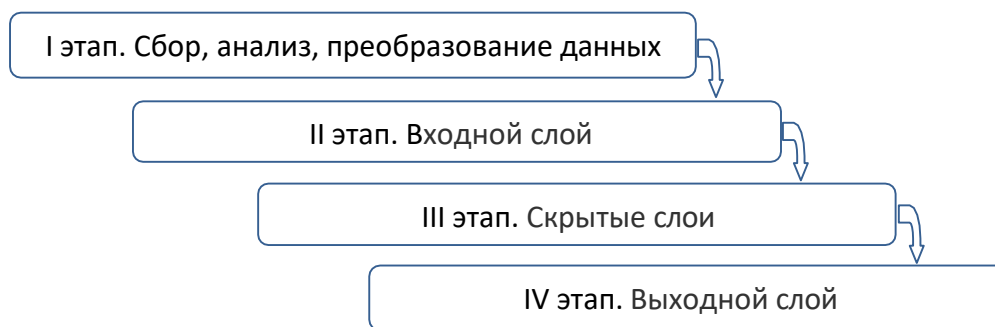
Машинное обучение определяется как класс методов искусственного интеллекта, направленных на решение задач за счет обучения на примерах различных схожих задач, т.е. отсутствует прямое решение. Использование предполагает применение различных математических подходов и численных методов. К таким следует отнести теорию вероятностей и графов, математический анализ и математическую статистику. Сравнивая машинное обучение с распространенными статистическими подходами (регрессионные модели, VAR, BVAR и пр.) следует выделить его особенность в возможности оценки параметров и доверительных интервалов. Таким образом, использование машинного обучения в анализе и прогнозировании различных социально-экономических и бюджетно-налоговых переменных набирает популярность и нашло широкое распространение в различных сферах деятельности человека.

Выделяют различные способы машинного обучения, из которых в данной статье рассматриваются способы, основанные на применении нейронных сетей (методы, основанные на обучающей выборке, в данной статье рассматриваться не будут) [4].

При разделении способов обучения по видам возьмем за основу типы алгоритмов обучения нейронных сетей. Согласно подобной типологизации алгоритмов выделяется обучение с учителем, обучение без учителя, обучение с подкреплением, активное обучение и др. В данной работе предполагается рассмотрение метода «глубокое обучение», который входит в способ обучения с учителем, а именно в понятие «искусственная нейронная сеть» [5]. Отметим, что помимо искусственной нейронной сети способ обучения с учителем включает в себя такие виды как: метод коррекции ошибки, метод обратного распространения ошибки и метод опорных векторов.

Машинное обучение с учителем представляет собой извлечение закономерной информации алгоритмом из предоставленного массива данных, который уже имеет набор заранее отмеченных способов решения той или иной задачи. Другими словами, искусственная нейронная сеть обучается на определенном наборе данных, где отмечены задачи и способы их решения. Анализируя большой объем подобной информации, нейронная сеть учится находить правильное/оптимальное решение поставленных задач.

При этом искусственная нейронная сеть имеет дополнительный подвид – глубокое обучение (Deep Learning (DL)), которое в широком понимании представляет собой набор алгоритмов для машинного обучения как с учителем, так и другими способами [6]. Глубокое обучение реализуется путем обучения представлениям или принципам, а не конкретным решениям определенных задач (см. рисунок).



Схематическое представление процесса глубокого обучения

Глубокое обучение впервые применено в 80-х годах XX века, однако наибольшее распространение получило уже в XXI веке на фоне ускоренного развития компьютерных технологий и значительного увеличения производительных мощностей вычислительной техники. Основой подобного обучения является глубокая нейронная сеть (Deep neural networks – DNN).

Преимуществами выступает возможность обучать больше нейронов и больше слоев нейронной сети, что, в свою очередь, позволяет повысить точность прогнозов. Недостатками являются сложность алгоритма расчета и необходимое наличие навыков программирования.

Базы данных и их обработка. Для построения и реализации отмеченных выше моделей и подходов необходимо создать базу данных основных показателей, которые будут применяться в анализе и прогнозировании объекта исследования. Временной ряд в базе данных должен содержать максимально доступные исторические данные. Немаловажной является категория данных: дневные, месячные, квартальные, годовые. Дневные данные содержат максимальную информацию о показателе, но несут в себе много «шума», информации не полезной для анализа. Выбор категории данных исходит из задач анализа и горизонта прогнозирования. Краткосрочный прогноз целесообразнее строить на месячных данных, а среднесрочный – на квартальных данных [7].

Важным аспектом выступает очистка данных от сезонности. Данный аспект приобретает особое значение при прогнозировании показателей с ярко выраженными циклическими/сезонными колебаниями, которым и является выпуск сельского хозяйства. Очистка данных от сезонности основывается на выделении повторяющейся из года в год циклической компоненты анализируемого показателя. В этих целях применяют сглаживание методом Census X13, Census X12, X11, Tramo/Seats и т.п. [8].

Следующим этапом обработки данных выступают логарифмирование и расчет первой разности логарифмов. Логарифмирование позволяет привести уравнения, основанные на экономической теории, в линейный вид. При этом первая разность логарифмов характеризует темп прироста переменных уравнения, что позволяет привести ряд к стационарному виду. Важным замечанием является необходимость обратного преобразования полученных результатов моделирования в уровни показателей, путем экспонирования (производится на завершающем этапе анализа и прогнозирования).

В силу сложности исследуемого объекта (темпов роста сельскохозяйственной продукции), итоговый прогноз можно определять на основе консенсус-прогноза, рассчитываемого как средневзвешенный результат всех используемых моделей и подходов. При этом наибольший вес следует присваивать той модели или подходу, который за несколько наблюдений давал бы прогнозы, более близкие к фактическим значениям. Такой

подход мог бы «справиться» с неопределенностью, которая связана с «непредсказуемым» влиянием различных факторов.

Заключение.

В статье приведен обзор различных методов и подходов прогнозирования временных рядов, которые могут быть применены для получения будущих значений такого сложного объекта исследования, как темпы роста сельского хозяйства. Показаны их достоинства и недостатки, условия их применения. Предлагается подход, который позволяет получать более достоверный прогноз с использованием комбинации различных методов и моделей.

Список использованной литературы

1. Экономический рост в сельском хозяйстве. – URL: <http://www.imf.org>. – Загл. с экрана. (дата обращения: 01.03.2024).
2. Библиотека МВФ. – URL: <http://www.imf.org>. – Загл. с экрана. (дата обращения: 01.03.2024).
3. Мырзакулова К.М. Технологии искусственного интеллекта, выполняющие математические вычисления // Вестник КНУ им. Ж. Баласагына, 2023, №4 (116), с. 174-178. – URL: <https://vestnik.knu.kg/wp-content/uploads/2024/01/Вестник-КНУ-4-2023....pdf> (дата обращения: 05.03.2024).
4. Amazon Web Services, Inc. Что такое нейронная сеть? – URL: <https://aws.amazon.com/ru/what-is/neural-network/> (дата обращения: 05.03.2024).
5. Машинное обучение. – URL: https://ru.wikipedia.org/wiki/Машинное_обучение (дата обращения: 03.03.2024).
6. Rammer W., Seidl R. Глубокое обучение в экологии. – URL: <https://habr.com/ru/articles/537132/> (дата обращения: 02.03.2024).
7. Nowcasting: The real-time informational content of macroeconomic data [Text] / D. Giannone, L. Reichling, D. Small // J. of Monetary Economics, 2008, Vol. 55, No 4, p. 665-676.
8. Seasonal Adjustment. EViews 12 Users Guide I. ISBN: 978-1-880411-63-6, p. 487.
9. Усенов И.А., Жапарова З.А. Математическое моделирование процессов развития зеленой экономики // Вестник КНУ, 2023, Спецвыпуск, с. 357-362.