

УДК 004.8:[663.918.2+339.564]

161. СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ПРОГНОЗИРОВАНИЯ ЭКСПОРТА КАКАО-ПРОДУКТОВ

Татарчук В.А., студент гр.372303

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Ладкина М.И., Солодилова М.К. – ассистенты каф. ЭИ

Аннотация. В условиях волатильности мировых цен на какао-бобы прогнозирование экспортной выручки становится критически важным для адаптации стратегий предприятий. В работе проведен сравнительный анализ двух подходов: модели машинного обучения с L1-регуляризацией (LASSO) и эконометрической модели коррекции ошибок (ECM) для прогнозирования стоимости экспорта какао-продуктов Республики Беларусь. Выявлен структурный сдвиг в динамике ряда с 2022 года. Показано, что ECM, учитывающая долгосрочные равновесные связи между показателями, обеспечивает более интерпретируемый и устойчивый прогноз в условиях нарушения стационарности, тогда как LASSO демонстрирует более высокую точность в краткосрочном периоде.

Ключевые слова. Машинное обучение, эконометрическое моделирование, модель коррекции ошибок, LASSO-регрессия, прогнозирование экспорта, структурный сдвиг, какао-продукты.

Резкий рост мировых цен на какао-бобы в последние годы оказывает прямое влияние на стоимость шоколадной продукции и других какао-содержащих товаров. Весной 2025 года биржевые котировки какао-бобов превысили 10 тыс. долл. США за тонну, что стало многолетним максимумом [1]. Несмотря на отсутствие собственного производства какао-бобов, Республика Беларусь имеет место в международной торговле переработанными какао-продуктами. Благодаря развитой кондитерской отрасли, высокому качеству продукции и выгодному географическому положению, страна успешно экспортирует шоколад, какао-порошок и другие изделия в десятки стран мира.

Цель работы – сравнительный анализ двух подходов (LASSO и ECM) для прогнозирования стоимости экспорта какао-продуктов Республики Беларусь и выявление факторов, определяющих точность прогноза.

В качестве исходных данных использованы месячные показатели стоимости экспорта товарной позиции «Шоколад и прочие готовые пищевые продукты, содержащие какао» (код ТН ВЭД 1806) за период с января 2010 по декабрь 2024 года (180 наблюдений). Данные были взяты с сайта Национального статистического комитета Республики Беларусь [2].

Визуальный анализ (см. рис. 1) позволяет выделить устойчивый восходящий тренд: с уровня около 1 855 тыс. долл. США в начале 2010 года показатель достиг пика в 2022 году (свыше 10 827 тыс. долл. США). Начиная с 2022 года наблюдается снижение экспортной выручки. На графике четко прослеживаются регулярные сезонные пики в декабре каждого года, обусловленные предновогодним ростом спроса.

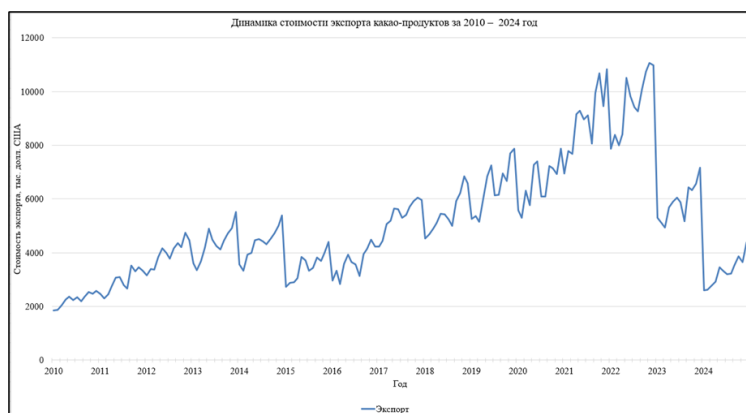


Рисунок 1 – Диаграмма динамики стоимости экспорта какао-продуктов

Для формальной проверки наличия структурного сдвига во временном ряду использован тест Чоу (Chow test), позволяющий оценить стабильность коэффициентов регрессионной модели до и после

предполагаемой точки разрыва. Нулевая гипотеза об отсутствии структурного сдвига проверялась для точки разрыва, соответствующей 2022 году. Результаты теста (F -статистика = 22,87, p -value = 0,00) позволяют отвергнуть нулевую гипотезу на любом уровне значимости, что подтверждает наличие структурного сдвига в динамике экспорта какао-продуктов с 2022 года.

Для выявления структуры зависимостей во временном ряду выполнен анализ автокорреляционной (ACF) и частной автокорреляционной (PACF) функций. Автокорреляционная функция характеризует тесноту линейной связи между членами временного ряда, а её медленное затухание свидетельствует о наличии тренда; значимый пик на первом лаге частной автокорреляционной функции указывает на авторегрессионную структуру порядка $p = 1$ [3]. ACF исходного ряда медленно затухает, оставаясь значимой на всех лагах, что характерно для нестационарного процесса с устойчивым трендом. PACF имеет значимый пик на первом лаге (0,915). Для лаговых переменных коэффициенты автокорреляции на первом лаге составляют 0,905–0,921, подтверждая инерционность экспортного процесса. Наибольшее значение получила переменная lag_{12} , что свидетельствует о годовой сезонной компоненте.

Рассмотрим модель LASSO. LASSO-персепсия (Least Absolute Shrinkage and Selection Operator) представляет собой линейный метод машинного обучения с L_1 -регуляризацией. В отличие от классического метода наименьших квадратов, который минимизирует только сумму квадратов остатков, LASSO добавляет к целевой функции штрафное слагаемое, равное сумме абсолютных значений коэффициентов модели, умноженной на параметр регуляризации λ . Такой подход приводит к тому, что коэффициенты при незначимых или избыточных признаках обнуляются, в то время как значимые переменные сохраняются. Таким образом, LASSO одновременно решает две задачи: снижает риск переобучения за счёт регуляризации и выполняет автоматический отбор наиболее информативных признаков, что существенно повышает интерпретируемость модели.

Для применения метода машинного обучения исходный временной ряд преобразован в формат «признак-цель». Сформированы три группы признаков:

- лаговые значения (lag_1 , lag_3 , lag_6 , lag_{12}), отражающие инерционность процесса;
- скользящие средние (ma_3 , ma_6 , ma_{12}), сглаживающие краткосрочные колебания;
- временные признаки (month, quarter, year, trend) и фиктивные переменные для учета сезонности и структурного сдвига (shift, shift_trend).

Целевой переменной является стоимость экспорта какао-продуктов в прогнозируемом периоде. Горизонт прогнозирования составляет один месяц. Принцип обучения заключается в следующем: для каждого месяца t модель предсказывает значение экспорта в месяце $t+1$, используя данные, известные на момент t , а именно значения экспорта за предыдущие периоды, скользящие средние и временные признаки.

Оптимизация гиперпараметра λ , контролирующего силу штрафа, выполнена с помощью кросс-валидации на скользящем окне (Time Series Split). Данный метод кросс-валидации сохраняет хронологический порядок наблюдений, что исключает утечку данных из будущего и имитирует реальные условия прогнозирования. Оптимальное значение параметра регуляризации составило $\lambda = 10$.

Для учета долгосрочных равновесных связей и краткосрочной динамики использована модель коррекции ошибок (Error Correction Model, ECM). ECM является классическим эконометрическим инструментом анализа нестационарных временных рядов, связанных коинтеграционным соотношением. Коинтеграция означает, что, несмотря на нестационарность отдельных рядов (например, экспорта и мировых цен), существует их долгосрочная линейная зависимость, то есть они не расходятся бесконечно во времени, а движутся вокруг общего равновесного уровня. ECM формализует это свойство: отклонение от долгосрочного равновесия в предыдущем периоде (так называемая ошибка коррекции) влияет на краткосрочную динамику зависимой переменной, постепенно возвращая её к равновесному уровню. В отличие от методов машинного обучения, ECM позволяет разделить влияние долгосрочных фундаментальных факторов (например, эластичность экспорта по мировым ценам) и краткосрочных отклонений от равновесия, вызванных временными шоками или сезонными колебаниями. Это делает ECM особенно полезной для анализа экономических процессов, где важна интерпретация не только прогнозных значений, но и механизмов восстановления равновесия.

В данной работе в качестве внешнего фактора, определяющего долгосрочную динамику экспорта белорусских какао-продуктов, использована мировая цена на какао-бобы. Выбор обусловлен высокой степенью интеграции белорусского экспорта в мировой рынок: рост мировых цен с лагом отражается на стоимости готовой продукции.

Определим этапы построения модели ЕСМ. Первый этап – оценка долгосрочного равновесного соотношения. Проанализирована связь между стоимостью экспорта какао-продуктов и мировой ценой на какао-бобы. Коэффициент долгосрочной эластичности составил 0,73 ($p < 0,01$), что указывает на значимую положительную зависимость: увеличение мировых цен на 1% в долгосрочном периоде сопровождается ростом экспортной выручки на 0,73%.

Второй этап – построение спецификации ЕСМ. Модель коррекции ошибок включает отклонение от долгосрочного равновесия в предыдущем периоде (ЕСТ), краткосрочные изменения экспорта и мировых цен, а также сезонные фиктивные переменные. Коэффициент коррекции ошибок (γ) оценивает скорость возвращения ряда к равновесию после шока.

Получили, что коэффициент коррекции ошибок составил -0,28 ($p < 0,01$). Это означает, что около 28% отклонения от долгосрочного равновесия устраняется в течение одного месяца. Отрицательный знак коэффициента подтверждает наличие механизма коррекции. Краткосрочная эластичность экспорта по цене составила 0,41. Для корректного моделирования в условиях изменения экономической ситуации в ЕСМ была добавлена фиктивная переменная, отражающая структурный сдвиг 2022 года.

Проанализируем результаты прогнозирования. Оценка качества моделей выполнена на тестовой выборке, охватывающей период с января 2021 по декабрь 2024 года (48 наблюдений). Используются три метрики: средняя абсолютная процентная ошибка (MAPE), средняя абсолютная ошибка (MAE) и корень из среднеквадратичной ошибки (RMSE). MAPE является основной метрикой, так как позволяет сравнивать точность прогнозов относительно масштаба фактических значений и интерпретировать ошибку в процентах. MAE и RMSE дополняют анализ, причём RMSE более чувствительна к крупным отклонениям, что важно при наличии выбросов. Результаты представлены в таблице 1.

Таблица 1 – Сравнение моделей прогнозирования

Модель	MAPE, %	MAE, тыс. долл. США	RMSE, тыс. долл. США
LASSO	8,42	443,65	722,23
ЕСМ	11,37	612,40	891,55

Модель LASSO продемонстрировала более высокую точность прогнозирования: значение MAPE составило 8,42%, что укладывается в целевой интервал 10–15%. Для ЕСМ ошибка оказалась выше и достигла 11,37%. На рисунке 2 представлено сравнение прогнозов моделей LASSO и ЕСМ с фактическими значениями экспорта за период 2021–2024 годов. Визуальный анализ подтверждает, что модель LASSO точно воспроизводит как сезонные колебания, так и общую тенденцию снижения экспорта после 2022 года. Модель ЕСМ также демонстрирует хорошее качество прогнозирования, однако её ошибка выше, чем у LASSO, что объясняется большей чувствительностью эконометрической модели к краткосрочным колебаниям и выбросам в данных.



Рисунок 2 – Динамика экспорта и прогнозы моделей LASSO и ЕСМ (2021–2024 гг.)

Преимущество LASSO перед ЕСМ в условиях структурного сдвига обусловлено тремя факторами. Во-первых, L1-регуляризация обеспечивает автоматический отбор признаков, обнуляя неинформативные переменные. Это позволяет модели выделять устойчивые закономерности, не адаптируясь к особенностям стабильного периода 2010–2020 годов. Во-вторых, LASSO устойчив к

мультиколлинеарности, эффективно выбирая релевантные признаки среди коррелированных лагов и скользящих средних. В-третьих, ECM как эконометрическая модель жёстко задаёт структуру зависимости (долгосрочное равновесие плюс краткосрочная динамика), что ограничивает её гибкость при резких изменениях режима. LASSO, напротив, самостоятельно определяет значимые факторы, что делает его более адаптивным к новым условиям.

В отличие от методов машинного обучения, ECM не имеет гиперпараметров в классическом понимании (как количество деревьев или глубина). Однако для повышения устойчивости ECM к структурным изменениям и улучшения точности прогнозирования в данном исследовании была проведена оптимизация состава признаков с помощью пошагового отбора (stepwise selection) на основе информационного критерия AIC.

Информационный критерий Акаике (AIC) – это мера относительного качества статистических моделей. Критерий основан на компромиссе между точностью подгонки модели под данные и её сложностью, которая оценивается по количеству используемых параметров. Чем меньше значение AIC, тем лучше модель балансирует между точностью и сложностью, что позволяет избежать переобучения. В отличие от классических статистических тестов (например, теста отношения правдоподобия), AIC позволяет сравнивать любые модели между собой, в том числе те, которые не являются вложенными друг в друга. Данный критерий применяется исключительно для выбора наилучшей модели из нескольких альтернатив и не предполагает содержательной интерпретации его абсолютных значений [4].

В результате оптимальная спецификация включила ECT с лагом 1, один лаг первой разности экспорта, текущее значение первой разности мировых цен, а также сезонные фиктивные переменные для декабря и марта. MAPE на тестовой выборке снизилась с 13,81% до 11,37%. Однако даже после оптимизации ECM уступает LASSO по точности прогнозирования, что объясняется принципиальной ограниченностью линейной параметрической структуры ECM по сравнению с гибким механизмом регуляризованного отбора признаков в LASSO.

В условиях структурного сдвига модель LASSO показала более высокую точность прогнозирования по сравнению с ECM. Ключевым фактором эффективности LASSO стал механизм автоматического отбора признаков, обеспечивающий устойчивость к изменению структуры данных. Анализ коэффициентов LASSO подтвердил, что наибольшее влияние на прогноз оказывает трёхмесячное скользящее среднее. Полученные выводы могут быть использованы при разработке инструментов поддержки управленческих решений в сфере планирования экспортной деятельности белорусских предприятий.

Список использованных источников:

1. РБК Компании [Электронный ресурс]. – Режим доступа : <https://companies.rbc.ru/news/wqIBjUn1fb/mirovoj-ryinok-kakao-sostoyanie-tsenyi-i-trendyi-2025-goda/> – Дата доступа : 13.03.2026.
2. belstat [Электронный ресурс]. – Режим доступа : <https://www.belstat.gov.by/> – Дата доступа : 13.03.2026.
3. Сурина, А. В. Анализ временных рядов. учеб. пособие / А.В. Сурина. – СПб. : Питер Ком, 2025. – 90 с.
4. Ежова, Е.О. Обобщение информационного критерия Акаике для выбора значений непрерывных параметров в моделях данных / Е.О. Ежова, В.В. Моттль, О.В. Красоткина // Таверический вестник информатики и математики. – 2009. – № 1. – С. 61–70.

UDC 004.8:[663.918.2+339.564]

COMPARATIVE ANALYSIS OF FORECASTING METHODS FOR COCOA PRODUCTS EXPORT

Tatarchuk V.A., student gr.372303

*Belarusian State University of Informatics and Radioelectronics,
Minsk, Republic of Belarus*

Ladkina M.I., Solodilova M.K. – Assistants of the Department of EI

Annotation. In the face of global cocoa price volatility, forecasting export revenue is critical for adapting business strategies. This paper presents a comparative analysis of two approaches for forecasting the value of cocoa products exports from the Republic of Belarus: a machine learning model with L1-regularization (LASSO) and an econometric error correction model (ECM). A structural break in the time series dynamics since 2022 is identified. It is shown that LASSO provides higher short-term forecasting accuracy, while ECM offers better interpretability and stability under non-stationary conditions.

Keywords. Machine learning, econometric modeling, error correction model, LASSO regression, export forecasting, structural break, cocoa products.