

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
(ФГБОУ ВО «АмГУ»)

Институт компьютерных и инженерных наук
Кафедра математического анализа и моделирования
Направление подготовки 01.03.02 – Прикладная математика и информатика
Направленность (профиль) образовательной программы «Прикладная математика и информатика»

ДОПУСТИТЬ К ЗАЩИТЕ
И.о. зав. кафедрой
_____ Н.Н. Максимова
« ____ » _____ 2025г.

БАКАЛАВРСКАЯ РАБОТА

на тему: Прогнозирование макроэкономических показателей с использованием моделей временных рядов

Исполнитель
студент группы 1101об

(подпись, дата)

С.О. Мельников

Руководитель
доцент, канд. физ.-мат. наук

(подпись, дата)

Е.М. Веселова

Нормоконтроль
ассистент

(подпись, дата)

В.О. Салмиянов

Благовещенск 2025

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
(ФГБОУ ВО «АмГУ»)

Институт компьютерных и инженерных наук
Кафедра математического анализа и моделирования

УТВЕРЖДАЮ
И.о. зав. кафедрой
_____ Н.Н. Максимова
« ____ » _____ 2025 г.

З А Д А Н И Е

К бакалаврской работе студента Мельникова Сергея Олеговича

1. Тема бакалаврской работы: Прогнозирование макроэкономических показателей с использованием моделей временных рядов (утверждена приказом от 03.04.2025 № 879-уч)
 2. Срок сдачи студентом законченной работы: 25.06.2025 г.
 3. Исходные данные к бакалаврской работе: отчет по преддипломной практике, научные статьи, учебно-методические работы, язык программирования Python, библиотеки: statsmodels, pmdarima, pandas
 4. Содержание бакалаврской работы (перечень подлежащих разработке вопросов): теоретические аспекты прогнозирования временных рядов, математические и инструментальные методы, применяемые для анализа данных; компьютерная визуализация результатов.
 5. Перечень материалов приложения: листинги программ
 6. Консультанты по бакалаврской работе: нормоконтроль – Салмиянов В.О., ассистент
 7. Дата выдачи задания: 14.04.2025 г.
- Руководитель бакалаврской работы: Веселова Елена Михайловна, доцент, канд. физ.-мат. наук, доцент
- Задание принял к исполнению (14.04.2025): _____ Мельников С.О.

РЕФЕРАТ

Бакалаврская работа содержит 79 с., 32 рисунка, 4 таблицы, 4 приложения, 20 источников.

ВРЕМЕННЫЕ РЯДЫ, МАКРОЭКОНОМИЧЕСКИЕ ПОКАЗАТЕЛИ, ПРОГНОЗИРОВАНИЕ, ВВП, ИНФЛЯЦИЯ, ЭКЗОГЕННЫЕ ФАКТОРЫ, PYTHON, SARIMAX, VAR, СТАЦИОНАРНОСТЬ, СЕЗОННОСТЬ

В выпускной квалификационной работе рассматривается задача прогнозирования ключевых макроэкономических показателей Российской Федерации – валового внутреннего продукта (ВВП) и инфляции на основе анализа временных рядов за период 2005–2021 гг. с проверкой результатов на данных 2022–2024 гг. В ходе работы реализованы и сравнены два подхода к моделированию: модель SARIMAX, учитывающая сезонные колебания и влияние экзогенных переменных, а также модель векторной авторегрессии (VAR), описывающая взаимосвязи между несколькими экономическими индикаторами.

Проведен комплексный анализ входных данных: тестирование на стационарность, декомпозиция рядов, выявление трендов и сезонности, а также учет влияния таких факторов, как цены на нефть, валютный курс, денежная масса и уровень заработной платы.

Цель работы заключается в возможности использования реализованных моделей для прогнозирования макроэкономической динамики России и формирования более обоснованных экономических сценариев в условиях внутренней и внешней нестабильности, а также в сравнении эффективности прогнозов выбранных моделей.

СОДЕРЖАНИЕ

Введение	5
1 Обзор макроэкономических показателей и подходов к их прогно- зированию	7
1.1 Ключевые макроэкономические показатели	7
1.2 Методы и подходы к прогнозированию экономических показате- телей	10
1.3 Временные ряды в макроэкономическом анализе	14
2 Модели временных рядов для прогнозирования	20
2.1 Классификация моделей временных рядов	20
2.2 Методы оценки параметров и проверки адекватности моделей	24
2.3 Краткая характеристика возможностей выбранного программ- ного средства	27
3 Практическое применение моделей временных рядов для прогно- зирования макроэкономических показателей	30
3.1 Описание и предварительный анализ исследуемых данных	30
3.2 Построение моделей и оценка их качества	37
3.2.1 Выбор моделей	37
3.2.2 Реализация и спецификация модели SARIMAX	38
3.2.3 Реализация и спецификация модели VAR	49
3.3 Анализ полученных результатов	61
Заключение	66
Библиографический список	68
Приложение А Прогноз ВВП и инфляции на основе модели SARI- MAX	71
Приложение Б Модель SARIMAX с варьированием переменной OIL	73
Приложение В Прогноз ВВП и инфляции на основе модели VAR	75
Приложение Г Модель VAR с варьированием переменной OIL	77

ВВЕДЕНИЕ

Современные экономические системы функционируют в условиях нестабильной внешней среды, где национальные рынки подвержены комплексному влиянию как внешних, так и внутренних кризисных факторов. Среди ключевых вызовов – резкие колебания сырьевых цен (особенно энергоресурсов), волатильность валютных пар, инфляционные скачки, геополитические противоречия, пандемии и санкционные режимы. Эти явления создают необходимость в оперативном реагировании со стороны государственных структур и бизнеса, основанном на точных и адаптивных прогнозах макроэкономических показателей.

Актуальность исследования обусловлена ростом потребности в методах прогнозирования, способных учитывать как внутренние закономерности экономической динамики, так и внешние шоки. В условиях глобальной неопределённости прогнозирование ВВП и инфляции становится критически важным для формирования антикризисных стратегий, бюджетного планирования и монетарной политики. Особую значимость приобретает разработка моделей, интегрирующих факторы, специфичные для российской экономики – такие как зависимость от нефтяного экспорта и валютных колебаний.

Цель работы – создание, программная реализация и сравнительный анализ двух подходов к прогнозированию макроэкономических индикаторов: модели SARIMAX (с учётом сезонности и экзогенных переменных) и VAR (для анализа межпеременных связей). Исследование направлено на выявление оптимальных методов прогнозирования ВВП и инфляции в условиях ограниченной выборки и высокой нестабильности внешней среды.

Задачи исследования:

–изучить структуру и динамику ключевых макроэкономических показателей (ВВП, инфляция, денежная масса, валютные курсы) на исторических данных за период 2005–2024 гг.;

- выполнить предварительную обработку временных рядов: проверку стационарности (тесты ADF/KPSS), удаление сезонных компонентов (STL-декомпозиция), коррекцию выбросов;

- реализовать прогнозные модели в Python с использованием библиотек statsmodels (SARIMAX/VAR), pmdarima (автоматический подбор параметров) и sklearn (метрики оценки);

- оценить влияние экзогенных факторов (цены на нефть, курс рубля, зарплаты) на точность прогнозов через сравнительный анализ метрик MAE и RMSE;

- провести сценарный анализ устойчивости моделей к изменениям ключевых переменных (например, $\pm 10\%$ к нефтяным ценам);

- сравнить эффективность SARIMAX и VAR на тестовой выборке (2022–2024 гг.) с акцентом на применимость в краткосрочном и среднесрочном горизонтах;

Практическая значимость работы заключается в создании инструментария для анализа макроэкономической устойчивости в условиях кризисов. Полученные модели могут быть адаптированы для государственных органов (например, Министерство экономического развития) и частных компаний, нуждающихся в прогнозировании налоговых поступлений, потребительского спроса или инвестиционной активности. Результаты исследования также полезны для совершенствования учебных курсов по эконометрике и машинному обучению в экономике.

1 ОБЗОР И АНАЛИЗ МАКРОЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ И ПОДХОДОВ К ПРОГНОЗИРОВАНИЮ

1.1 Ключевые макроэкономические показатели

Макроэкономика как самостоятельное направление экономической науки оформилась лишь в первой половине XX века. До этого основной акцент в экономических исследованиях делался на микроэкономических процессах – изучении поведения отдельных потребителей и производителей. Однако тяжелые экономические потрясения, в частности Великая депрессия 1929 года, продемонстрировали необходимость рассмотрения экономики на более высоком, агрегированном уровне.

Значительный вклад в формирование макроэкономической теории внёс британский экономист Джон Мейнард Кейнс. В своём труде «Общая теория занятости, процента и денег», опубликованном в 1936 году, он обосновал необходимость государственного регулирования экономики для сглаживания циклических колебаний. Согласно кейнсианской концепции, рыночный механизм не всегда способен самостоятельно обеспечить равновесие между спросом и предложением, вследствие чего требуется активная роль государства посредством фискальной и монетарной политики [19].

После окончания Второй мировой войны кейнсианские идеи легли в основу экономической политики многих развитых стран. Однако в 1970-х годах, когда мировая экономика столкнулась с феноменом стагфляции – сочетанием высокой инфляции и безработицы, стали набирать популярность альтернативные теории. В частности, монетаризм, представленный Милтоном Фридманом, акцентировал внимание на регулировании денежной массы как ключевом инструменте экономической стабилизации. Позднее появились и новые подходы, например, теория рациональных ожиданий, разработанная Робертом Лукасом и Томасом Сарджентом, подчёркивающая влияние ожиданий экономических субъектов на макроэкономическую динамику.

К концу XX века в экономической науке сформировались и иные школы мысли, такие как посткейнсианство, сосредоточенное на проблемах неопределённости и финансовой нестабильности, а также институциональная экономика, исследующая роль социальных институтов в экономическом развитии. Так, труды Элинора Остром по управлению общими ресурсами существенно обогатили понимание макроэкономических взаимодействий.

Современная макроэкономика также уделяет большое внимание процессам глобализации, таким как взаимосвязь национальных экономик, деятельность транснациональных корпораций и влияние цифровых платформ. Примером может служить мировой финансовый кризис 2008 года, когда нестабильность на ипотечном рынке США вызвала рецессию в глобальном масштабе.

Сегодня макроэкономика представляет собой многоплановую дисциплину, охватывающую широкий круг факторов – от мировых финансовых потоков до цифровой экономики и международной торговли. Особое значение в современных исследованиях приобретают методы математического моделирования и прогнозирования, которые позволяют строить обоснованные сценарии будущего экономического развития. Экономисты активно применяют инструменты, такие как теория игр, агентное моделирование и алгоритмы машинного обучения.

Глобальные вызовы:

- кризис 2008 года: крах ипотечного рынка США вызвал рецессию в странах Европейского союза и в Азии;

- цифровизация: стремительный рост криптовалют и развитие децентрализованных финансовых систем (DeFi) требуют переосмысления подходов к монетарной политике.

Оценка текущего состояния экономики базируется на ряде ключевых макроэкономических показателей. Одним из важнейших из них является валовой внутренний продукт (ВВП) [13], который отражает совокупную стоимость всех произведённых в стране товаров и услуг за определённый период. ВВП служит основным индикатором экономического роста и уровня благосостояния

граждан. Примером может служить устойчивый рост ВВП Китая в XXI веке, который в среднем составлял 6-8 % в год и оказывал значительное влияние на мировую экономику.

Среди других значимых макроэкономических показателей выделяют:

– Уровень инфляции – демонстрирует динамику цен на товары и услуги. Умеренная инфляция стимулирует экономический рост, тогда как её резкое повышение приводит к обесцениванию денежных средств и снижению покупательной способности населения. Примером разрушительного воздействия гиперинфляции является ситуация в Зимбабве в 2008 году, когда инфляция достигла колоссальных 79,6 млрд процентов.

– Уровень безработицы – характеризует долю экономически активного населения, не имеющего оплачиваемой работы. Высокие показатели безработицы ведут к снижению доходов населения и падению потребительского спроса, что, в свою очередь, замедляет экономический рост. Так, в странах ЕС в 2020 году уровень безработицы составил 7,5 % вследствие пандемии COVID-19, что негативно отразилось на производстве и уровне жизни.

– Платежный и торговый баланс – отражают движение международных финансовых потоков. Например, дефицит торгового баланса США в 2022 году (1,18 трлн долларов США) свидетельствует о высокой зависимости страны от импорта, делая экономику более чувствительной к изменениям глобальных рыночных условий. В то же время положительное сальдо торгового баланса способствует укреплению национальной валюты и финансовой устойчивости.

– Государственный долг – соотношение размера государственных заимствований к ВВП страны. Слишком высокий уровень государственного долга может вынудить правительство повысить налоги или сократить расходы. В Японии, например, в 2023 году этот показатель превысил 260 % ВВП, что создаёт долгосрочные риски для финансовой системы.

– Индекс потребительского доверия – отражает ожидания граждан относительно будущего состояния экономики. Рост данного индекса указывает на

высокий уровень уверенности потребителей, что способствует увеличению потребительского спроса и инвестиционной активности.

Современные экономические модели учитывают и новые индикаторы, среди которых:

– Индекс человеческого развития (ИЧР) – агрегированный показатель, объединяющий ВВП на душу населения, уровень образования и продолжительность жизни. Он позволяет оценивать не только экономическое развитие, но и качество жизни.

– Уровень цифровизации – демонстрирует степень проникновения современных технологий, таких как интернет-банкинг, электронная коммерция и использование искусственного интеллекта в экономике. В Китае, к примеру, более 40 % ВВП связано с цифровыми технологиями, что позволяет стране занимать лидирующие позиции в этой сфере.

– Экологические показатели – в их числе углеродный след, уровень загрязнения окружающей среды и эффективность использования природных ресурсов. В условиях глобальных климатических изменений значение этих показателей возрастает.

– Глобальные индексы конкурентоспособности и инноваций – позволяют оценить, насколько экономика способна адаптироваться к новым вызовам и технологическим изменениям.

Комплексный анализ макроэкономических показателей даёт возможность точнее прогнозировать экономические тенденции, разрабатывать эффективные стратегии развития и корректировать экономическую политику с учётом текущей ситуации. В условиях усиливающейся глобализации и ускоряющегося технологического прогресса сочетание традиционных и новых индикаторов становится важнейшим условием обеспечения устойчивого экономического роста.

1.2 Методы и подходы к прогнозированию экономических показателей

Формирование методов моделирования макроэкономических показателей насчитывает значительный исторический путь, в процессе которого было раз-

работано множество разнообразных подходов. В современной практике применяются эконометрические модели, модели общего равновесия, агентное моделирование, алгоритмы машинного обучения, а также гибридные решения, которые совмещают несколько методик [15, 16].

Одним из классических методов являются эконометрические модели, основанные на статистических взаимосвязях между переменными. Среди них модель ARIMA (Autoregressive Integrated Moving Average) широко используется для прогнозирования временных рядов макроэкономических данных, таких как инфляция, ВВП и безработица. Формальная запись модели выглядит следующим образом:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(1 - L)^d Y_t = c + (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t,$$

где Y_t – текущее значение временного ряда, ϕ и θ – коэффициенты модели, ε – ошибки случайных отклонений, d – порядок интегрирования, p – порядок авторегрессии, q – порядок скользящего среднего, а L – оператор лага, который сдвигает на 1 период назад.

Модели общего равновесия (CGE) широко используются для анализа последствий экономической политики и внешних шоков. С их помощью исследуется влияние налоговых реформ, изменений в монетарной политике, а также воздействие глобальных кризисов на экономику отдельной страны или региона.

Агентное моделирование представляет собой имитационный метод, позволяющий моделировать поведение большого числа агентов – домохозяйств, предприятий, финансовых институтов. Такой подход помогает учитывать сложные взаимодействия между экономическими субъектами и исследовать, каким образом их индивидуальные решения формируют макроэкономические эффекты [6]. Примером применения является проект EURACE, в рамках которого изучалось воздействие монетарной политики Европейского центрального банка (ЕЦБ) на уровень безработицы.

В числе современных инструментов анализа всё чаще встречаются методы, основанные на машинном обучении. Рекуррентные нейронные сети (LSTM), например, хорошо справляются с учётом временной структуры данных

и применяются для прогноза инфляции, динамики фондовых индексов и валютных курсов. Такие компании, как Google, успешно внедряют алгоритмы LSTM для макроэкономического прогнозирования, добиваясь при этом высокой точности результатов.

Гибридные модели сочетают в себе элементы классических эконометрических методов и искусственного интеллекта. Так, модель FRB/US, созданная Федеральной резервной системой США, объединяет традиционные макроэкономические уравнения с нейросетевыми компонентами, что позволяет более эффективно моделировать нелинейные взаимосвязи в экономике.

Кроме того, важное место в макроэкономическом прогнозировании занимают методы векторной авторегрессии (VAR), которые дают возможность учитывать взаимные связи между несколькими временными рядами. Также активно применяются байесовские методы моделирования, позволяющие включать в анализ априорные знания об экономических процессах и учитывать неопределённость прогноза.

Дополнительно в практике макроэкономического прогнозирования широко используются:

- модели динамического стохастического общего равновесия (DSGE);
- методы Монте-Карло для моделирования неопределенности;
- метод главных компонент (PCA) для снижения размерности данных;
- подходы ансамблевого обучения (Random Forest, XGBoost).

Выбор модели прогнозирования играет ключевую роль в экономических исследованиях. Именно от выбранной модели зависят точность прогноза и обоснованность выводов, которые могут стать основой для стратегического планирования и выработки экономической политики.

Основные факторы, влияющие на выбор методов прогнозирования:

- тип и доступность данных;
- сложность и объем обрабатываемых данных;
- требуемая точность и надежность прогнозов;
- объем доступных временных и вычислительных ресурсов;

- возможность интерпретации модели и прозрачность её выводов;
- гибкость модели и её способность адаптироваться к изменениям экономической среды.

В последние годы отмечается устойчивая тенденция к интеграции различных подходов, что позволяет повысить точность прогнозов и обеспечить адаптивность моделей к стремительно меняющимся экономическим условиям. В рамках бакалаврской работы акцент сделан на применении методов, основанных на анализе временных рядов.

Современные технологии создают новые перспективы для прогнозирования макроэкономических показателей. Рост вычислительных мощностей, доступность больших данных и прогресс в разработке алгоритмов машинного обучения открывают возможности создания более точных моделей, способных учитывать множество факторов, влияющих на экономику.

Одним из наиболее перспективных направлений является использование методов глубокого обучения. Рекуррентные нейронные сети (RNN) и их усовершенствованные варианты – Long Short-Term Memory (LSTM) и Gated Recurrent Unit (GRU) – применяются для анализа последовательных данных. Эти модели способны учитывать временные зависимости и выявлять скрытые закономерности в динамике макроэкономических показателей.

Основные преимущества глубокого обучения в экономическом прогнозировании заключаются в следующем:

- выявление нелинейных взаимосвязей между экономическими переменными;
- способность обрабатывать большие объёмы данных;
- высокая адаптивность к изменению экономических условий;
- возможность анализа многомерных данных с высокой степенью точности.

Важную роль также играют технологии Big Data. Современная экономика генерирует огромные объёмы информации, анализ которой существенно повышает точность прогнозов. Основными источниками таких данных служат соци-

альные сети, банковские транзакции, финансовая отчётность, показатели потребительского спроса и официальная статистика.

Работа с Big Data требует применения специализированных технологий, включая облачные вычисления, распределённые базы данных и алгоритмы машинного обучения.

Кроме того, блокчейн-технологии способствуют повышению прозрачности и достоверности экономических данных, минимизируя риски мошенничества и оптимизируя финансовые процессы.

Не менее актуальны и этические аспекты прогнозирования, в частности, проблема алгоритмических смещений, которые могут приводить к искажению результатов. Для противодействия таким эффектам применяются методы интерпретируемого машинного обучения, корректировка исходных данных и внедрение механизмов оценки справедливости моделей.

1.3 Временные ряды в макроэкономическом анализе

Анализ временных рядов является одним из важнейших инструментов макроэкономического исследования, позволяющим отслеживать динамику ключевых экономических показателей – таких как валовой внутренний продукт (ВВП), уровень инфляции, процентные ставки и государственный долг. Временной ряд представляет собой последовательность значений определённого показателя, упорядоченных по времени. Каждое значение связано с предыдущими и последующими наблюдениями, что придаёт этим данным особую структуру. Для их анализа требуются специальные методы, способные учитывать внутренние зависимости, изменение статистических характеристик во времени и влияние внешних факторов – от политических событий до технологических изменений или природных катастроф.

Временные ряды обладают рядом характерных особенностей, отличающих их от других типов данных. Прежде всего – это зависимость между наблюдениями. В отличие от данных в поперечных сечениях, где значения независимы, элементы временного ряда часто коррелируют с предыдущими значениями. К примеру, уровень инфляции в текущем месяце во многом обусловлен её ди-

намикой в прошлом. Такая зависимость называется автокорреляцией и играет ключевую роль при моделировании и прогнозировании временных рядов.

Вторая важная особенность – нестационарность. Она проявляется в изменении среднего значения, дисперсии или иных статистических свойств ряда с течением времени. Например, ВВП большинства стран демонстрирует устойчивый долгосрочный рост, что делает этот показатель нестационарным. Напротив, процентные ставки могут колебаться вокруг определённого среднего уровня. Работа с нестационарными рядами требует предварительной обработки – дифференцирования, логарифмирования или применения специальных фильтров для выделения тренда.

Ещё одна характерная черта временных рядов – чувствительность к внешним шокам. Внезапные события могут существенно изменить поведение ряда. Так, пандемия COVID-19 в 2020 году вызвала резкое сокращение мирового ВВП и рост безработицы, нарушив ранее устойчивые экономические тенденции. При анализе временных рядов необходимо учитывать не только регулярные компоненты (тренды, сезонность), но и влияние таких случайных факторов.

Существует несколько классификаций временных рядов, определяющих выбор методов их анализа:

– По временным интервалам выделяют дискретные и непрерывные ряды. Дискретные фиксируются через заданные промежутки времени (месяц, квартал, год), в то время как непрерывные отражают показатели, измеряемые постоянно (например, курсы валют в реальном времени). В макроэкономике чаще всего используются дискретные данные, поскольку официальная статистика публикуется с определённой периодичностью.

– По количеству переменных различают одномерные и многомерные ряды. Одномерные ряды описывают динамику одного показателя (например, уровень инфляции в России), тогда как многомерные отражают взаимосвязи между несколькими переменными (например, ВВП, инвестиции и потребление).

– По стационарности выделяют стационарные и нестационарные ряды. В стационарных рядах среднее значение и дисперсия остаются постоянными, а структура автокорреляции не меняется со временем. В нестационарных рядах наблюдаются тренды, сезонные колебания или структурные сдвиги, что требует использования специальных методов для корректной обработки данных.

– По типу данных выделяют количественные ряды (например, ВВП в долларах) и качественные, или категориальные (например, фазы экономического цикла – подъем, пик, спад, депрессия).

Источниками данных для построения временных рядов служат материалы официальной статистики, публикуемые национальными статистическими ведомствами, центральными банками и международными организациями (МВФ, Всемирный банк, ОЭСР).

При сборе и обработке временных рядов возникают определённые трудности. Пропуски в данных могут быть вызваны техническими сбоями или задержками публикаций. Кроме того, шум и выбросы способны исказить общую картину – например, резкие изменения цен на нефть, обусловленные геополитическими событиями, могут восприниматься как долгосрочный тренд. Для решения таких проблем применяются методы интерполяции, фильтрации и корректировки аномальных значений.

Отдельную сложность представляет агрегация данных. Макроэкономические показатели зачастую публикуются на уровне всей страны, однако для регионального анализа требуется их дезагрегация, что может снижать точность. Например, общий уровень безработицы не отражает различий между регионами и социальными группами.

Анализ временных рядов используется для решения широкого круга задач:

– Прогнозирование – ключевое направление, позволяющее правительствам и центральным банкам принимать обоснованные решения. Например, прогнозы роста ВВП помогают планировать бюджетные расходы, а модели инфляции используются для корректировки процентных ставок.

–Выявление структурных изменений – позволяет обнаруживать точки, где динамика ряда резко изменяется (например, переход от стагнации к росту после кризиса). Это особенно важно для оценки эффективности экономической политики.

–Анализ причинности – изучение взаимосвязей между переменными. Например, временные ряды помогают определить, влияет ли повышение процентной ставки на снижение инфляции с определенным лагом.

–Мониторинг экономических циклов – временные ряды позволяют идентифицировать фазы цикла (подъем, пик, спад, восстановление), что критично для антикризисного управления.

Применение временных рядов в макроэкономике имеет долгую историю. Первоначально использовались простые методы анализа трендов и сезонности, ещё в XIX веке применяемые для изучения циклов производства и занятости. В XX веке появились более формализованные модели, такие как ARIMA, значительно повысившие точность прогнозирования.

С развитием вычислительных технологий акцент сместился в сторону методов машинного обучения и анализа больших данных. Современные подходы, включая нейронные сети и гибридные модели, позволяют учитывать нелинейные зависимости и обрабатывать большие объёмы информации. Так, использование LSTM-сетей для прогнозирования ВВП стало возможно благодаря интеграции данных о потреблении, инвестициях и даже активности в социальных сетях.

При этом даже самые совершенные алгоритмы не заменяют глубокого понимания экономических процессов. Успешный анализ временных рядов требует сочетания математических методов, экономической интуиции и учёта контекста формирования данных.

Каждый временной ряд обычно рассматривается как сумма нескольких компонентов, отражающих различные аспекты динамики:

Тренд описывает долгосрочное направление изменения показателя, исключая краткосрочные колебания. Например, устойчивый рост ВВП Китая за

последние три десятилетия связан с процессами индустриализации, модернизацией инфраструктуры и глобализацией экономики. Для выявления тренда применяются методы полиномиальной регрессии, скользящего среднего или фильтра Ходрика-Прескотта, который разделяет ряд на трендовую и циклическую части.

Сезонная компонента отражает регулярные колебания, связанные с календарными периодами – месяцами, кварталами или днями недели. Например, увеличение розничных продаж в предпраздничные месяцы или сезонный спад сельскохозяйственного производства в зимний период. Сезонные эффекты могут быть аддитивными (когда амплитуда колебаний постоянна) или мультипликативными (когда амплитуда зависит от уровня ряда). Для их анализа используются методы сезонной декомпозиции, такие как STL (Seasonal and Trend decomposition using LOESS), позволяющие разделить временной ряд на тренд, сезонность и случайные остатки.

Циклическая компонента связана с экономическими циклами, такими как чередование фаз роста и рецессии. В отличие от сезонности, циклы не имеют фиксированной продолжительности и зависят от факторов, включая инвестиционную активность, потребительские настроения и политику центральных банков. Например, финансовый кризис 2008-2009 годов стал частью долгосрочного экономического цикла, включающего периоды подъема, пика, спада и восстановления.

Случайные колебания, или шум, представляют собой непредсказуемые отклонения, вызванные внешними шоками, такими как пандемии, войны или внезапные изменения цен на сырьевые товары. Эти компоненты моделируются с помощью стохастических процессов, включая «белый шум» или процессы с автокорреляцией, которые учитывают зависимость текущих значений от предыдущих.

Для обработки временных рядов применяются различные методы математического анализа [4, 5, 12]. Среди наиболее распространенных можно выделить сглаживание экспоненциальным средним (ETS), позволяющее устранять

шум и выявлять основные тенденции, и сезонную декомпозицию (STL), которая разделяет временной ряд на его компоненты: тренд, сезонность и остаток.

Одной из самых мощных моделей временных рядов является SARIMA (Seasonal ARIMA), которая учитывает сезонные компоненты и представляется уравнением:

$$(1 - \phi_1 L - \dots - \phi_p L^p)(1 - \Phi_1 L^s - \dots - \Phi_P L^{Ps})(1 - L)^d (1 - L^s)^D Y_t,$$

где запись эквивалентна данной:

$$c + (1 + \theta_1 L + \dots + \theta_q L^q)(1 + \Theta_1 L^s + \dots + \Theta_Q L^{Qs})\varepsilon_t,$$

где s – сезонный период, $\Phi_1, \Phi_2, \dots, \Phi_P$ – коэффициенты сезонной авторегрессии, $\Theta_1, \dots, \Theta_Q$ – коэффициенты сезонного скользящего среднего, P – порядок сезонной авторегрессии, D – порядок сезонного интегрирования, Q – порядок сезонного скользящего среднего, Y_t – текущее значение временного ряда, ϕ и θ – коэффициенты модели, ε – ошибки случайных отклонений, d – порядок интегрирования, p – порядок авторегрессии, q – порядок скользящего среднего, а L – оператор лага, который сдвигает на 1 период назад.

Особое место занимает модель SARIMA, активно применяемая в центральных банках и финансовых организациях для прогнозирования макроэкономических показателей. Европейский центральный банк, например, использует SARIMA при оценке инфляционных ожиданий и при принятии решений по монетарной политике. Применение данной модели позволило сократить ошибку прогноза инфляции на 15 %, что повысило эффективность управления процентными ставками.

Таким образом, временные ряды остаются важнейшим инструментом макроэкономического анализа, позволяющим выявлять ключевые закономерности, прогнозировать кризисы и разрабатывать стратегические решения. Современные вычислительные технологии и прогресс в алгоритмах прогнозирования делают эти методы всё более точными и востребованными в экономической практике.

2 МОДЕЛИ ВРЕМЕННЫХ РЯДОВ ДЛЯ ПРОГНОЗИРОВАНИЯ

2.1 Классификация моделей временных рядов

Прогнозирование макроэкономических показателей на основе временных рядов требует применения моделей, которые способны учитывать как долгосрочные тренды, так и краткосрочные колебания. В выпускной квалификационной работе основное внимание уделено использованию моделей SARIMA и VAR. При этом в рамках предварительного анализа применяется ряд классических методов статистического анализа временных рядов, включая принципы регрессионного подхода к исследованию зависимостей между экономическими показателями [1].

Перед построением моделей проводится предварительная обработка и анализ временных рядов. В первую очередь проверяется их стационарность с помощью теста Дики–Фуллера [2] и теста Филлипса–Перрона. При необходимости производится трансформация данных (например, дифференцирование) для достижения стационарности.

Модель SARIMA (Seasonal Autoregressive Integrated Moving Average) применяется для моделирования и прогнозирования одномерных временных рядов с учётом сезонных и трендовых компонентов. Этапы её построения включают определение оптимальной структуры модели (параметры ARIMA и сезонные компоненты), оценку параметров по методу максимального правдоподобия, проверку адекватности модели (анализ остатков, тест Дарбина–Уотсона) и построение прогноза. [7]

Модель VAR (Vector Autoregression) используется для моделирования взаимосвязанных временных рядов, позволяя учитывать влияние каждой переменной на другие переменные системы. Построение VAR-модели включает выбор оптимального числа лагов (по критериям AIC и BIC), оценку параметров с помощью метода наименьших квадратов, диагностику модели (устойчивость, автокорреляция остатков) и анализ импульсных откликов для изучения динамических взаимосвязей.

Применение моделей SARIMA и VAR позволяет учитывать особенности макроэкономических данных и обеспечивает возможность построения обоснованных прогнозов, что является важной задачей для разработки эффективной экономической политики.

Для учета динамики макроэкономических показателей используются различные подходы. Модели скользящего среднего (Moving Average, MA) применяются для сглаживания временных рядов и устранения случайных колебаний [11]:

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q},$$

где y_t – значение временного ряда в момент времени t , μ – среднее значение ряда, ε_t – белый шум (случайная ошибка), θ_i – параметры модели, q – порядок модели.

Авторегрессионные модели (AR) выражают текущее значение ряда через его предыдущие значения:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t,$$

где y_t – значение временного ряда в момент времени t , ϕ_0 – свободный член модели, ϕ_i – параметры авторегрессии, p – порядок авторегрессии, ε_t – случайная ошибка.

Для работы с нестационарными рядами используется модель ARIMA(p,d,q), где d – порядок дифференцирования для устранения тренда:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(1 - L)^d Y_t = c + (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t,$$

где Y_t – текущее значение временного ряда, ϕ и θ – коэффициенты модели, ε – ошибки случайных отклонений, d – порядок интегрирования, p – порядок авторегрессии, q – порядок скользящего среднего, а L – оператор лага, который сдвигает на 1 период назад.

В расширенной версии, учитывающей сезонные эффекты, используется модель SARIMA(p,d,q)(P,D,Q,s):

$$(1 - \phi_1 L - \dots - \phi_p L^p)(1 - \Phi_1 L^s - \dots - \Phi_P L^{Ps})(1 - L)^d (1 - L^s)^D Y_t,$$

где запись эквивалентна данной:

$$c + (1 + \theta_1 L + \dots + \theta_q L^q)(1 + \Theta_1 L^s + \dots + \Theta_Q L^{Qs})\varepsilon_t,$$

где s – сезонный период, $\Phi_1, \Phi_2, \dots, \Phi_P$ – коэффициенты сезонной авторегрессии, $\Theta_1, \dots, \Theta_Q$ – коэффициенты сезонного скользящего среднего, P – порядок сезонной авторегрессии, D – порядок сезонного интегрирования, Q – порядок сезонного скользящего среднего, Y_t – текущее значение временного ряда, ϕ и θ – коэффициенты модели, ε – ошибки случайных отклонений, d – порядок интегрирования, p – порядок авторегрессии, q – порядок скользящего среднего, а L – оператор лага, который сдвигает на 1 период назад.

Регрессия с распределенными лагами (DLM) позволяет оценить эффект запаздывающего воздействия факторов:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \beta_3 x_{t-2} + \dots + \beta_k x_{t-k} + \varepsilon_t,$$

где y_t – зависимая переменная, $x_t, x_{t-1}, \dots, x_{t-k}$ – объясняющие переменные с различными лагами, β_0 – свободный член модели, β_i – коэффициенты регрессии, k – максимальный лаг, ε_t – случайная ошибка.

Помимо традиционных моделей, важную роль в прогнозировании временных рядов играют гибридные подходы, комбинирующие регрессионные методы с машинным обучением. Например, регуляризация моделей (LASSO, Ridge) используется для борьбы с мультиколлинеарностью, а байесовские методы позволяют учитывать неопределенность параметров модели. Декомпозиция временных рядов (STL) позволяет выделить трендовую, сезонную и остаточную компоненты, что облегчает интерпретацию данных и выбор подходящей модели [8].

Модель VAR (Vector Autoregression) представляет собой многомерную статистическую модель временных рядов, предназначенную для описания и прогнозирования взаимосвязанных экономических переменных. Она особенно полезна, когда необходимо изучить динамическое взаимодействие между несколькими макроэкономическими показателями, такими как ВВП, инфляция, процентные ставки и другие [9].

Формально модель VAR(p), где p – это порядок модели (количество лагов), записывается следующим образом:

$$y_t = c + \sum_{i=1}^p A_i y_{t-i} + \varepsilon_t,$$

где y_t – вектор значений всех прогнозируемых переменных в момент времени t , c – вектор констант, A_i – матрицы коэффициентов при каждом лаге I , y_{t-I} – вектор значений переменных на шаге $t-I$, ε_t – вектор случайных ошибок.

Основные типы VAR-моделей:

- базовая VAR-модель: описывает взаимосвязь между переменными, не накладывая жестких ограничений на их взаимосвязь;
- структурная VAR-модель (SVAR): вводит определенные экономические ограничения для идентификации импульсных реакций;
- VAR с экзогенными переменными (VARX): расширяет базовую модель, включая внешние переменные, которые оказывают влияние на систему, но сами не зависят от текущих и прошлых значений эндогенных переменных.

Модель VAR позволяет не только прогнозировать поведение экономических переменных, но и проводить анализ импульсных реакций, которые показывают, как шок или изменение в одной переменной влияет на остальные переменные в системе с течением времени.

VAR-модели широко используются в макроэкономическом анализе для:

- оценки и прогнозирования экономических показателей;
- анализа и интерпретации динамических взаимодействий между показателями;
- изучения последствий экономической политики и шоков для экономики.

Регрессионный анализ временных рядов сталкивается с проблемами, такими как нестационарность, которая устраняется с помощью дифференцирования или сезонной декомпозиции; автокорреляция ошибок, учитываемая с помощью моделей ARMA, FGLS; мультиколлинеарность, снижаемая методами регуляризации. Различные тесты, такие как информационные критерии Акаике (AIC) и Шварца (BIC), помогают выбрать оптимальную модель. Таким образом, регрессионный анализ обеспечивает баланс между точностью прогнозов и ин-

терпретацией результатов, делая его важным инструментом макроэкономического прогнозирования.

2.2 Методы оценки параметров и адекватности моделей

В процессе анализа временных рядов особое внимание уделяется корректной настройке параметров моделей и проверке их соответствия реальным данным. Эти этапы играют решающую роль в обеспечении надежности прогнозов и интерпретации результатов. Оценка параметров позволяет определить модель под конкретный временной ряд, а проверка адекватности – убедиться, что модель не искажает закономерности, заложенные в данных. Ниже рассматриваются ключевые подходы к решению этих задач, их особенности и практическое применение в макроэкономике.

Параметры модели – это числовые коэффициенты, которые определяют ее поведение. Например, в модели SARIMA параметры авторегрессии (p) и скользящего среднего (q) задают степень зависимости текущего значения от прошлых наблюдений и ошибок соответственно. Их точное определение критично для успешного прогнозирования.

Метод максимального правдоподобия (Maximum Likelihood Estimation, MLE) заключается в поиске таких значений параметров, которые максимизируют вероятность наблюдения имеющихся данных при условии, что модель верна. Для временных рядов MLE часто используется в сочетании с предположением о нормальности распределения ошибок. Например, при оценке параметров модели ARIMA алгоритм итеративно подбирает коэффициенты, минимизируя логарифмическую функцию правдоподобия. Преимущество MLE – его статистическая эффективность, однако метод чувствителен к нарушению предположений (например, наличию выбросов).

Метод наименьших квадратов (Ordinary Least Squares, OLS) минимизирует сумму квадратов отклонений между фактическими значениями ряда и прогнозами модели. Он широко применяется в линейных моделях, таких как VAR (Vector Autoregression), где зависимость между переменными предполагается линейной. Например, для анализа взаимосвязи между ВВП, инфляцией и про-

центной ставкой OLS позволяет оценить коэффициенты, отражающие силу влияния каждой переменной. Недостаток метода – неспособность адекватно работать с нелинейными зависимостями и высокой мультиколлинеарностью.

Байесовский подход отличается от классических методов тем, что учитывает априорную информацию о параметрах. Например, если экономист уверен, что инфляция в стране редко превышает 5%, это знание может быть формализовано в виде априорного распределения. Современные алгоритмы, такие как MCMC (Markov Chain Monte Carlo), позволяют обновлять априорные представления на основе данных, получая апостериорное распределение параметров. Такой подход особенно полезен в условиях неопределенности, например, при прогнозировании показателей в развивающихся экономиках с ограниченной историей наблюдений.

Современные программные инструменты (например, библиотека `auto_arima` в Python) позволяют автоматически перебирать возможные комбинации параметров (p , d , q , P , D , Q) для модели SARIMA, выбирая оптимальную на основе критериев AIC или BIC. Этот подход снижает трудозатраты, но требует понимания, как интерпретировать результаты: автоматизация не исключает необходимости проверки адекватности модели.

Даже самая сложная модель бесполезна, если она не отражает реальных процессов. Для оценки адекватности используются два основных направления: анализ остатков и количественные метрики.

Остатки – это разница между фактическими значениями ряда и прогнозами модели. Идеальная модель должна оставлять «белый шум»: случайные колебания без автокорреляции и систематических отклонений.

Тест на автокорреляцию остатков (например, тест Льюнга-Бокса) проверяет, существует ли значимая зависимость между остатками на разных лагах. Если такая зависимость обнаружена, это свидетельствует о том, что модель не учла часть закономерностей (например, сезонность).

Тест на гетероскедастичность (ARCH-тест) выявляет изменения дисперсии остатков во времени. Это важно для финансовых временных рядов, где волатильность может быть нестационарной.

Графический анализ (например, коррелограмма остатков) помогает визуально оценить качество модели. Например, если остатки демонстрируют тренд, это указывает на недостаточное описание долгосрочных изменений в данных.

Критерии AIC (Akaike Information Criterion) и BIC (Bayesian Information Criterion) сравнивают модели, штрафую за избыточное количество параметров.

Математическая запись *AIC* выглядит следующим образом:

$$AIC = -2\ln(L) + 2k,$$

а *BIC*:

$$BIC = -2\ln(L) + k\ln(n),$$

где L – значение функции правдоподобия, k – количество параметров, n – размер выборки.

Модель с меньшим *AIC/BIC* считается более предпочтительной. Например, при сравнении двух моделей SARIMA для прогнозирования инфляции выбирают ту, у которой критерий ниже, даже если она сложнее.

Временные ряды нельзя разбивать на случайные подвыборки, как это делается в классической регрессии. Вместо этого используется последовательная кросс-валидация (time series split). Например, модель обучают на данных за 2000–2015 годы, тестируют на 2016 год, затем расширяют обучающую выборку до 2016 года и тестируют на 2017 году и т.д. Такой подход имитирует реальные условия прогнозирования и позволяет оценить устойчивость модели к изменениям в данных.

Для количественной оценки ошибок используются такие метрики как MAE (средняя абсолютная ошибка) и RMSE (среднеквадратичная ошибка).

MAE записывается в следующем виде:

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|,$$

где Y_t – фактическое значение временного ряда, $\hat{Y}_t$ – прогнозное значение, n – количество наблюдений.

RMSE записывается в следующем виде:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2},$$

где Y_t – фактическое значение временного ряда, $\hat{Y}_t$ – прогнозное значение, n – количество наблюдений [3].

2.3 Краткая характеристика возможностей выбранного продукта

Для решения задачи прогнозирования макроэкономических показателей требуется использование специализированных программных инструментов, обеспечивающих удобную работу с временными рядами, статистический анализ данных и построение прогнозных моделей. В рамках данного исследования применяются как универсальные среды разработки, так и специализированные библиотеки для анализа временных рядов.

–Python – один из наиболее популярных языков программирования в области анализа данных и машинного обучения. Он предоставляет широкий набор инструментов для работы с экономическими показателями, позволяет эффективно обрабатывать временные ряды, выполнять математические вычисления и строить прогнозные модели.

–Jupyter Notebook – интерактивная среда, в которой удобно разрабатывать и тестировать код, выполнять вычисления и визуализировать результаты. Она активно применяется в научных исследованиях благодаря удобству работы с данными и возможности их наглядного представления.

–Excel – программный инструмент, который, несмотря на свою простоту, используется для предварительного анализа данных, выполнения расчетов, построения диаграмм и выявления основных тенденций в экономических показателях [10].

–Google Colab – облачная среда разработки, которая позволяет работать с кодом на Python без необходимости установки дополнительных программ. Она

обеспечивает доступ к мощным вычислительным ресурсам, что особенно важно при обработке больших массивов данных.

– SQLite – легковесная система управления базами данных, используемая для хранения и обработки экономических данных. Она позволяет эффективно управлять структурированными данными и быстро извлекать необходимую информацию.

– Источники данных – для построения моделей временных рядов используются открытые базы данных и платформы, такие как Kaggle, World Bank Open Data, OECD, IMF Data, FRED Economic Data, а также национальные статистические агентства. Эти ресурсы предоставляют достоверные макроэкономические данные для анализа и прогнозирования.

– Библиотеки Python – в исследовании используются специализированные модули, такие как statsmodels, pmdarima, pandas, numpy, matplotlib и seaborn. Они обеспечивают комплексный функционал для обработки, анализа, моделирования и визуализации временных рядов с использованием SARIMA-моделей.

Выбранные программные средства обеспечивают выполнение ключевых этапов анализа и прогнозирования макроэкономических показателей. Используемые библиотеки для Python выполняют определенные функции, обеспечивая комплексный подход к решению поставленной задачи.

Pandas – одна из ключевых библиотек для работы с табличными данными. Она позволяет загружать информацию из различных источников (read_csv(), read_excel(), read_sql()), обрабатывать пропущенные значения (dropna(), fillna(), replace()) и выполнять сложные операции группировки (groupby(), pivot_table()). Кроме того, она поддерживает работу с временными рядами (rolling(), expanding(), resample()) и объединение данных (merge(), concat()) [18].

Numpy отвечает за числовые вычисления и работу с многомерными массивами. С его помощью можно создавать массивы (array(), linspace(), arange()), вычислять статистические показатели (mean(), median(), std()) и выполнять операции линейной алгебры (dot(), matmul(), linalg.solve()). Также библиотека ис-

пользуется для генерации случайных чисел (`random.normal()`, `random.uniform()`) и аппроксимации данных (`polyfit()`, `polyval()`) [14].

Для визуализации данных используются `matplotlib` и `seaborn`. Первая библиотека предоставляет базовые инструменты (`plot()`, `scatter()`, `bar()`) [17], а вторая – более удобные стили и расширенные возможности. Например, с помощью `heatmap()` можно построить корреляционную матрицу, а `set_style()` и `set_palette()` помогают настроить внешний вид графиков. Также можно создавать сложные композиции графиков через `figure()` и `subplots()`.

`Statsmodels` предлагает инструменты для анализа временных рядов и эконометрики. Ключевые функции:

- Проверка стационарности с помощью теста Дики-Фуллера (`adfuller()`).
- Построение моделей SARIMA через класс `SARIMAX()`.
- Сезонная декомпозиция временных рядов (`seasonal_decompose()`).
- Анализ автокорреляции и частичной автокорреляции (`plot_acf()`, `plot_pacf()`).

`Pmdarima` автоматизирует подбор параметров для SARIMA-моделей. Функция `auto_arima()` анализирует данные, определяет оптимальные значения $(p, d, q)(P, D, Q, s)$ и проверяет сезонные компоненты, что ускоряет процесс моделирования.

Комплексное использование этих библиотек позволяет:

- Строить SARIMA-модели с учетом сезонности и экзогенных факторов.
- Визуализировать тренды, прогнозы и остатки.
- Проводить диагностику модели и оценивать её точность.

Комплексное использование данных библиотек позволяет не только строить прогнозные модели, но и анализировать их результаты, выявлять ключевые тренды и принимать обоснованные решения на основе полученных данных.

– Рынок труда (Unemployment, WAGE, WAGEgr) – данные об уровне занятости и динамике зарплат, которые служат индикаторами социальной стабильности.

– Внешнеэкономическая конъюнктура (OIL, OILgr, USDRUB, EURRUB) – цены на нефть Brent и валютные курсы, оказывающие значительное влияние на экономику с экспортно-ориентированной моделью.

– Процентные ставки (Interest) – средняя 90-дневная межбанковская ставка, регулируемая ЦБ РФ для контроля над инфляцией и кредитованием.

Каждая переменная имеет чёткую экономическую интерпретацию, а их совокупность позволяет строить многомерные модели, учитывающие взаимосвязи между секторами экономики. Например, рост цен на нефть (OIL) напрямую влияет на валютные курсы (USDRUB, EURRUB) и, через них, на уровень инфляции (Inflation).

Обучение всех моделей осуществляется на базе этого основного датасета. Все преобразования, статистический анализ, отбор признаков и калибровка параметров моделей выполняются именно по этим данным.

Для объективной проверки прогностических возможностей моделей используется отдельный датасет, содержащий значения макроэкономических индикаторов за период с 2022 по 2024 год. Данный файл не используется при обучении, а служит исключительно для контроля качества прогнозирования: сравнение спрогнозированных значений ВВП и инфляции с их фактическими наблюдениями в этот период позволяет рассчитать метрики MAE, RMSE и построить графики для анализа ошибок моделей.

Структура второго (прогностического) датасета совпадает с первым по набору ключевых переменных, включая ВВП (GDP) и инфляцию (Inflation).

Период исследования охватывает значительные события, которые оставили отпечаток на данных:

Мировой финансовый кризис 2008–2009 гг.

– Падение ВВП: В IV квартале 2008 года темп роста ВВП составил – 9,4% (из таблицы данных), что стало самым резким спадом за весь период.

–Снижение цен на нефть: Brent упал с \$138,4 (III квартал 2008) до \$35,8 (IV квартал 2008), потеряв 61,7% стоимости.

–Девальвация рубля: USD/RUB вырос с 25,4 до 39,4 (+55%) за два квартала.

Геополитический кризис 2014-2015 гг.

–Усиление инфляции: В I квартале 2015 года годовая инфляция достигла 7,44%, что стало максимальным значением за весь период.

–Резкое укрепление доллара: USD/RUB вырос с 35,7 (IV квартал 2013) до 72,9 (I квартал 2015), увеличившись почти вдвое.

–Рост процентных ставок: Ставка рефинансирования ЦБ РФ поднялась с 7% до 17% в декабре 2014 года.

Пандемия COVID-19 (2020 г.)

–Резкий спад ВВП: Во II квартале 2020 года ВВП снизился на 9,3%, что сравнимо с кризисом 2008 года.

–Нефтяной коллапс: Цена Brent упала с \$67 (IV квартал 2019) до \$14,8 (I квартал 2020), потеряв 78% стоимости.

–Увеличение безработицы: Уровень безработицы вырос с 4,6% (IV квартал 2019) до 5,99% (II квартал 2020).

Эти события демонстрируют высокую чувствительность российской экономики к внешним шокам, особенно в секторах энергетики и валютного регулирования.

Для количественной оценки переменных, входящих в рассматриваемый набор данных, найдены статистические показатели (среднее, стандартное отклонение, минимум/максимум), представленные в таблице 1.

Таблица 1 – Статистические показатели

<i>Переменная</i>	<i>Среднее</i>	<i>Стандартное отклонение</i>	<i>Минимум</i>	<i>Максимум</i>
ВВП (GDP)	21,5 трлн Р	2,3 трлн Р	15,9 трлн Р (I кв. 2005)	23,7 трлн Р (IV кв. 2021)

Продолжение таблицы 1

НЕФТЬ (OIL)	65\$/баррель	20\$/баррель	35,8\$/баррель (II кв. 2020)	138,4\$/баррель (II кв. 2008)
Инфляция	2,0%	1,8%	–0,29% (III кв. 2017)	7,44% (I кв. 2015)
Безработица	6,5%	1,2%	4,4% (III кв. 2019)	9,1% (I кв. 2009)

Кроме того, выявлена сильная корреляция между переменными:

– GDP ↔ OIL: рост цен на нефть положительно влияет на ВВП.

– M0 ↔ BM: высокая мультиколлинеарность требует отбора предикторов при моделировании.

– WAGEgr ↔ Inflation: рост зарплат способствует ускорению инфляции.

Особенности и вызовы данных:

– Нестационарность: Большинство переменных (GDP, M0, BM, OIL) имеют выраженный тренд, что требует дифференцирования или логарифмирования для приведения к стационарному виду.

– Сезонность: ВВП демонстрирует рост в IV квартале (связано с увеличением потребительского спроса), а PPI — сезонные колебания, обусловленные цикличностью сельскохозяйственного производства.

– Аномалии: Экстремальные значения наблюдались в 2014–2015 гг. (резкое укрепление доллара) и 2020 г. (падение ВВП на 9,3%).

Для последующего прогнозирования были реализованы последовательные этапы подготовки данных, включающие работу как с основным обучающим датасетом, так и с тестовым набором, предназначенным для валидации результатов моделей.

Преобразование переменных. На этапе предварительной обработки данных, все преобразования – такие как логарифмирование отдельных переменных (например, OIL, WAGE), нормализация инфляционных индикаторов и масштабирование других экономических показателей – выполнялись исключительно

на базе обучающего датасета. Это обеспечивает корректное вычисление статистических параметров (среднее, стандартное отклонение, параметры шкалирования), которые впоследствии применяются к значениям соответствующих переменных в тестовом датасете. Такой подход предотвращает утечку информации между обучающей и тестовой выборками и гарантирует объективную оценку прогностической способности моделей на новых данных. Логарифмирование ряда цен на нефть и уровня заработной платы позволяет стабилизировать дисперсию и смягчить влияние аномальных выбросов, а нормализация инфляции обеспечивает сопоставимость временных рядов по масштабу.

Декомпозиция временного ряда. Для ключевых экономических показателей, таких как валовой внутренний продукт (GDP) и инфляция (Inflation), была проведена декомпозиция временного ряда на трендовую, сезонную и остаточную компоненты. Эта процедура осуществлялась на основе обучающего набора и позволила выявить скрытые закономерности, устойчивые тенденции и повторяющиеся сезонные эффекты. Модельные параметры сезонности и тренда, рассчитанные на обучающей части, были далее использованы при построении прогноза, чтобы обеспечить адекватную идентификацию структурных особенностей временного ряда на этапе тестирования (2022-2024 гг.).

Тестирование на стационарность. Для определения необходимости дифференцирования и дальнейшей корректной настройки моделей временных рядов были применены стандартные статистические тесты – тест Дики-Фуллера (ADF) и тест Квятковского-Филлипса-Шмидта-Шина (KPSS). Оценка стационарности проводилась исключительно на обучающем датасете. Для большинства исследуемых переменных выявлена нестационарность, что потребовало применения первого или второго порядка разностей для преобразования рядов к стационарному виду. Все параметры дифференцирования определялись на основе обучающих данных и затем применялись к тестовой части для получения сопоставимых прогнозов.

После выполнения указанных процедур подготовленные параметры нормализации, сезонности, тренда и дифференцирования, были применены к пере-

менным тестового датасета. Это обеспечивает воспроизводимость прогнозных сценариев и предотвращает попадание в модель информации о будущих значениях. Все прогнозы, построенные моделями, верифицировались исключительно по данным, что позволяет объективно оценить точность предсказаний вне обучающего периода.

На рисунках с 1 по 3 представлены графики ключевых переменных:

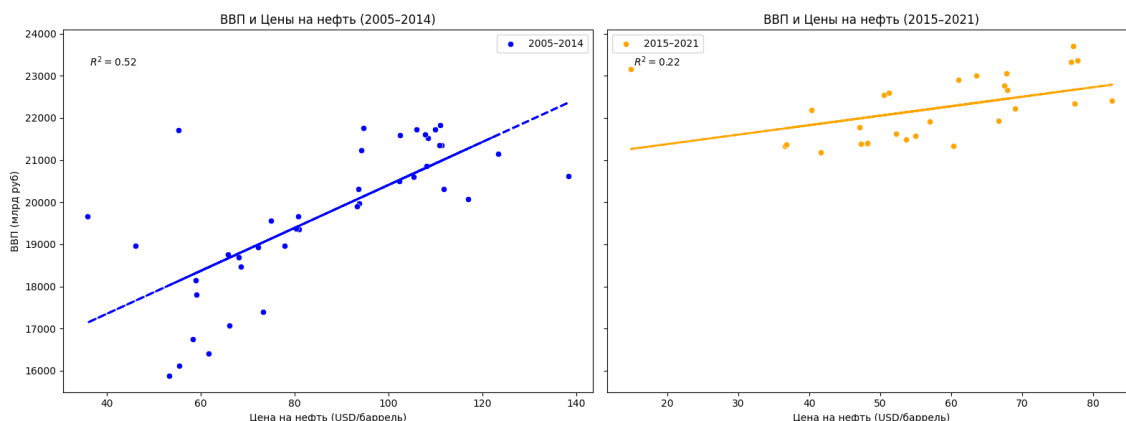


Рисунок 2 – ВВП и цены на нефть

ВВП и цены на нефть: В период 2005–2014 гг. наблюдается устойчивая корреляция ($R^2 \approx 0,7$), которая ослабевает после 2014 г. из-за санкций и девальвации рубля.

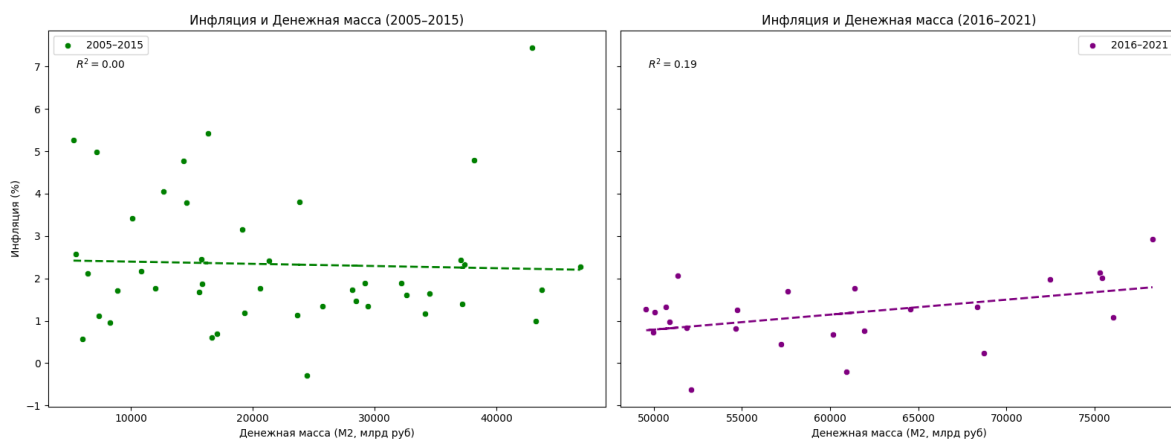


Рисунок 3 – Инфляция и денежная масса

Инфляция и денежная масса: Высокая связь ($R^2 \approx 0,6$) между ВМ и инфляцией в 2005–2015 гг., затем снижение из-за усиления контроля ЦБ РФ.

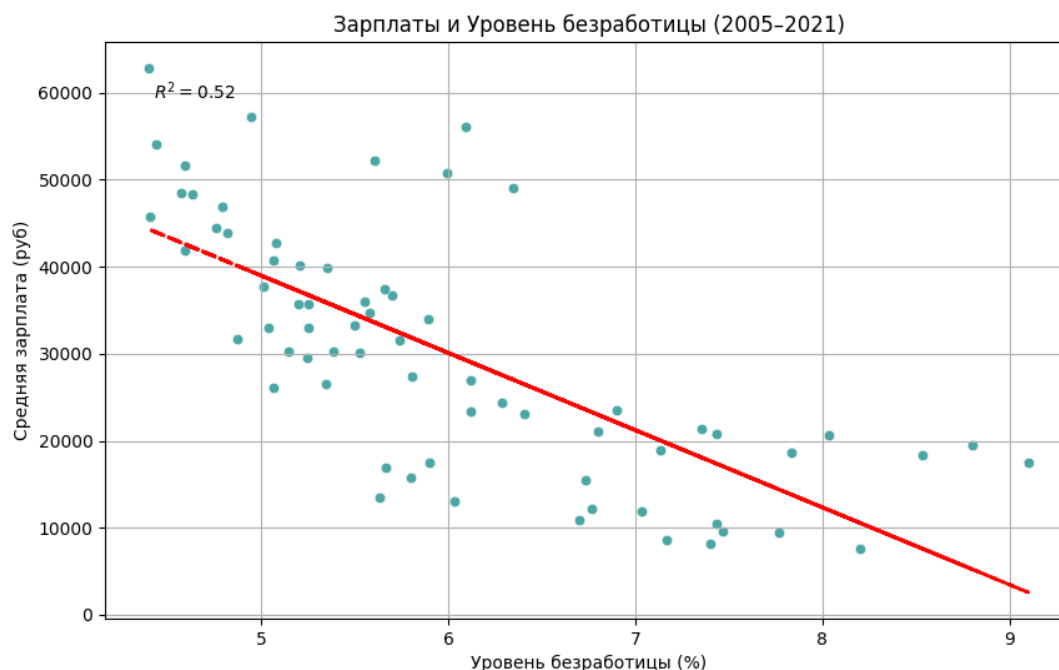


Рисунок 4 – Зарплаты и безработица

Зарплаты и безработица: Обратная связь (кривая Филлипса) выражена слабо, что может быть связано со структурными особенностями российского рынка труда.

Датасет предоставляет широкие возможности для анализа и прогнозирования макроэкономических показателей. Его ценность заключается в охвате различных экономических секторов и периодов, включающих как стабильный рост, так и кризисные явления. Однако успешное применение моделей требует учёта нестационарности, сезонности и мультиколлинеарности, а также интеграции внешних факторов, таких как цены на нефть и валютные курсы. Полученные в ходе анализа закономерности служат основой для выбора методов прогнозирования, включая ARIMA, SARIMA, машинное обучение и комбинированные подходы.

Кроме того, набор данных содержит «кластеры» высокой волатильности в периоды шоков, что требует применения гибридных моделей (например, GARCH или VAR с экзогенными регрессорами) для оценки динамики и рисков. Данные не содержат пропусков и готовы к детальной статистической обработ-

ке, включая тестирование на стационарность (ADF/KPSS) и анализ автокорреляции (ACF/PACF).

Для последующего построения и тестирования моделей:

- Данные до IV квартала 2021 года (data.csv) используются для обучения, подбора параметров и внутренней кросс-валидации;
- Данные за 2022-2024 годы (prognozresult.csv) служат для сравнения прогнозных значений с фактическими данными и для объективной оценки качества моделирования.

Такой подход обеспечивает корректное разделение обучающей и тестовой выборки, что необходимо для оценки реальной прогностической способности моделей.

3.2 Построение моделей и оценка их качества

3.2.1 Выбор моделей

Современный эконометрический анализ, направленный на прогнозирование макроэкономических индикаторов, невозможен без тщательного подбора моделей, способных отразить сложную структуру временных рядов. Выбор адекватной методологии – это не просто технический этап, а фундаментальный вопрос, от которого зависит точность прогнозов, возможность их интерпретации и применимость для реальных экономических решений.

На первом этапе моделирования анализировались различные классы временных моделей, включая как классические авторегрессионные и скользящие средние схемы (AR, MA, ARMA, ARIMA), так и их расширенные версии, способные учитывать сезонные и экзогенные эффекты (SARIMA, SARIMAX). Ключевым критерием для выбора инструментов являлось наличие сезонности в экономических показателях, их потенциальная нестационарность, а также зависимость от внешних и внутренних шоков. В процессе разработки прогнозных сценариев большое значение придавалось способности моделей не только повторять уже известные закономерности, но и эффективно справляться с новыми, нетривиальными колебаниями в экономике.

Все экспериментальные расчёты выполнялись с использованием современных языков программирования и пакетов анализа данных. Были организованы два чётко разделённых этапа: обучение моделей происходило исключительно на базе исторических данных, тогда как итоговая проверка осуществлялась на независимом тестовом наборе. Подобная схема позволяла объективно оценить не только внутреннюю, но и внешнюю валидность полученных результатов, исключить переобучение и повысить достоверность выводов.

Выбор моделей определялся результатами предварительного анализа данных, структурой исходных рядов, а также экономической интерпретируемостью факторов. В настоящей работе ключевое внимание было уделено моделям SARIMAX и VAR, как наиболее информативным и практическим для задач российской макроэкономики.

3.2.2 Реализация и спецификация модели SARIMAX

Одной из центральных моделей, применённых в настоящем исследовании для анализа и прогнозирования временных рядов макроэкономических показателей, стала модель SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eXogenous regressors). Применение этой модели обусловлено необходимостью не только выявлять внутренние закономерности динамики ВВП и инфляции, но и учитывать воздействие целого комплекса внешних и внутренних факторов, формирующих экономическую ситуацию в стране.

SARIMAX – это многоуровневая структура, которая позволяет гибко моделировать сезонные эффекты, долгосрочные тренды и влияние экзогенных переменных. Такая архитектура особенно востребована в макроэкономическом прогнозировании, где наблюдается выраженная сезонность, структурные сдвиги, а также регулярное влияние мировых цен на сырьевые товары, валютных курсов, объёмов денежной массы и прочих индикаторов.

В данном исследовании модели SARIMAX настраивались с учётом особенностей российских макроэкономических рядов: выбирались параметры для учета сезонных колебаний, осуществлялось тестирование на стационарность и, где это было необходимо, вводилось дифференцирование. Важной частью экс-

перимента стало включение различных экзогенных переменных, что дало возможность количественно измерять вклад внешних и внутренних факторов в динамику ВВП и инфляции.

Программная реализация всех экспериментов с SARIMAX осуществлялась на языке Python, преимущественно с использованием библиотеки statsmodels, а обработка и визуализация данных выполнялась при помощи пакетов pandas и matplotlib. Для калибровки и проверки моделей строго соблюдался принцип разделения обучающей и тестовой выборок.

Общая структура SARIMAX:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{k=1}^K \beta_k x_{k,t} + S_t + \varepsilon_t,$$

где y_t – значение прогнозируемого показателя (ВВП или инфляция) в момент времени t , $x_{k,t}$ – экзогенные переменные, φ_i , θ_j , β_k – коэффициенты модели, S_t – сезонная составляющая, ε_t – случайная ошибка.

Эксперимент 1 – SARIMAX без экзогенных переменных:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + S_t + \varepsilon_t,$$

где y_t – значение прогнозируемого показателя (ВВП или инфляция) в момент времени t , φ_i – коэффициенты авторегрессии, θ_j – коэффициенты скользящего среднего, S_t – сезонная составляющая, ε_t – случайная ошибка.

Данный вариант модели построен без учёта влияния каких-либо внешних факторов, что позволяет определить базовый уровень прогностической точности, достижимый только за счёт внутренних закономерностей макроэкономического ряда. Включение лишь собственной динамики и сезонности исследуемого показателя позволяет отделить эффект автокорреляции и сезонных волн, что важно для последующего анализа роли экзогенных факторов.

– MAE по ВВП: 2706

– RMSE по ВВП: 3194

– MAE по инфляции: 1.47

– RMSE по инфляции: 2.06

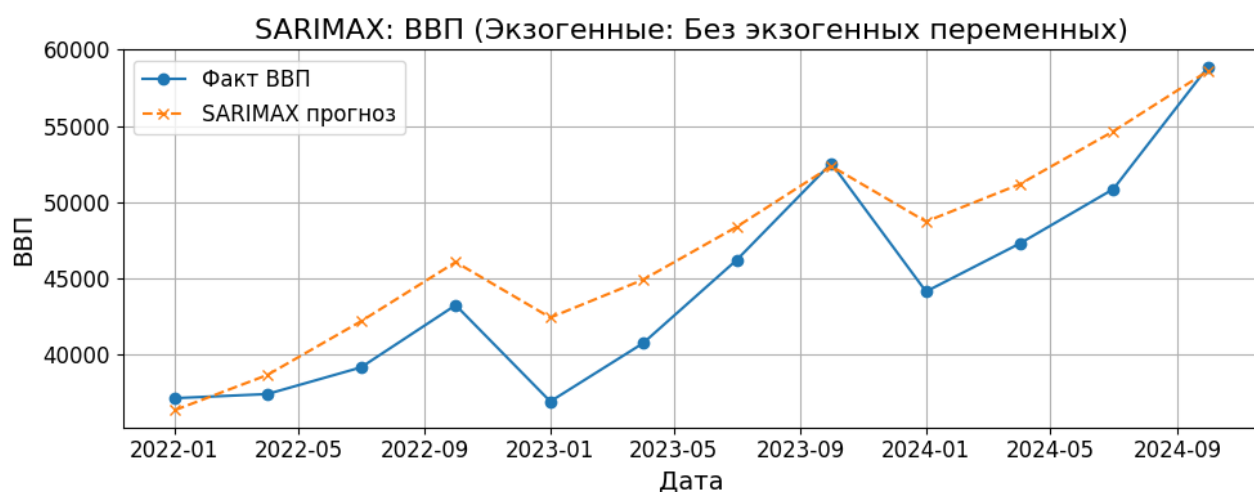


Рисунок 5 – Прогноз ВВП без экзогенных переменных

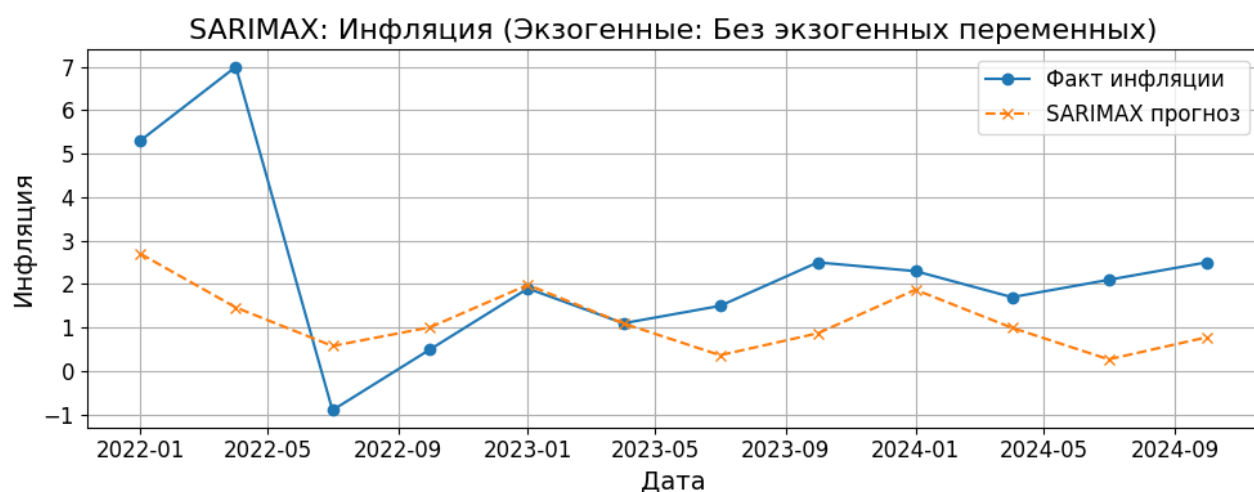


Рисунок 6 – Прогноз инфляции без экзогенных переменных

Эксперимент 2 – SARIMAX с экзогенной переменной OIL:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \beta_1 OIL_t + S_t + \varepsilon_t,$$

где y_t – значение прогнозируемого показателя (ВВП или инфляция) в момент времени t , φ_i – коэффициенты авторегрессии, θ_j – коэффициенты скользящего

среднего, S_t – сезонная составляющая, ε_t – случайная ошибка, OIL_t – цена на нефть Brent в момент времени t , β_1 – коэффициент при переменной OIL .

В этой модели в качестве дополнительного экзогенного фактора используется цена на нефть. Такой подход экономически обоснован для российской экономики, обладающей ярко выраженной сырьевой зависимостью. Включение OIL_t позволяет выявить прямую чувствительность ВВП и инфляции к глобальным изменениям в энергетическом секторе.

- MAE по ВВП: 1625
- RMSE по ВВП: 2142
- MAE по инфляции: 1.47
- RMSE по инфляции: 2.05

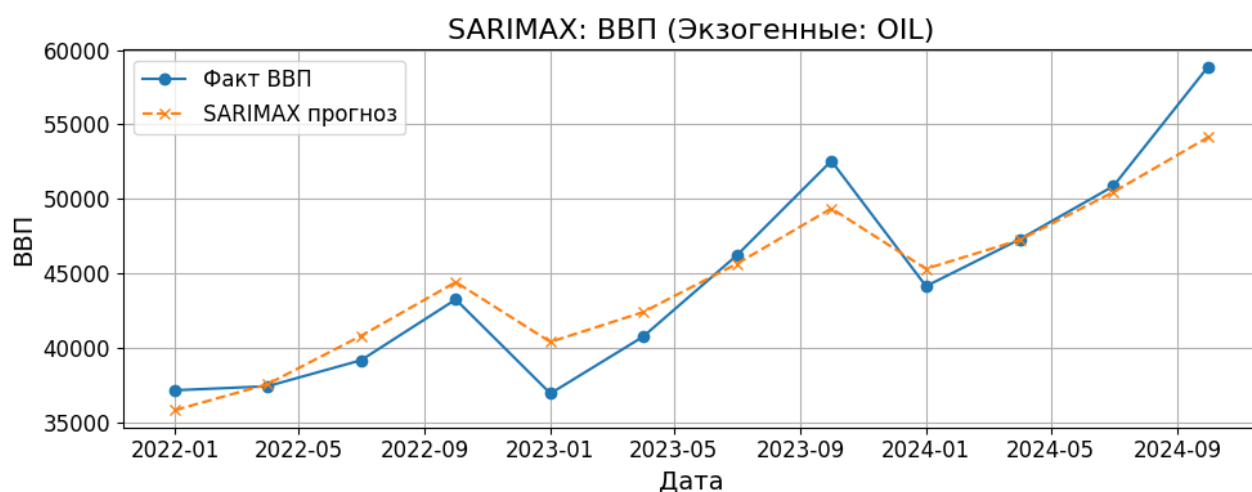


Рисунок 7 – Прогноз ВВП с экзогенной переменной OIL

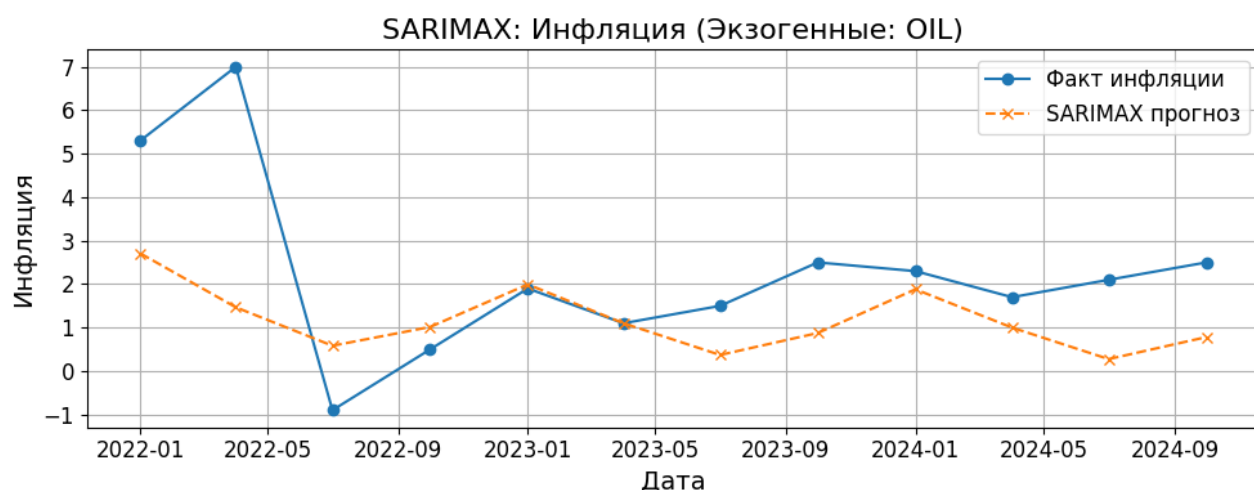


Рисунок 8 – Прогноз инфляции с экзогенной переменной OIL

Эксперимент 3 – SARIMAX с экзогенными переменными OIL и USDRUB:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \beta_1 OIL_t + \beta_2 USDRUB_t + S_t + \varepsilon_t,$$

где y_t – значение прогнозируемого показателя (ВВП или инфляция) в момент времени t , φ_i – коэффициенты авторегрессии, θ_j – коэффициенты скользящего среднего, S_t – сезонная составляющая, ε_t – случайная ошибка, OIL_t – цена на нефть Brent в момент времени t , $USDRUB_t$ – курс рубля к доллару, β_1 , β_2 – коэффициенты при экзогенных переменных.

Модель с двумя экзогенными переменными строит более комплексную картину влияния внешнеэкономических факторов, что особенно важно в периоды нестабильности валютного рынка и резких скачков сырьевых цен. Экономическая интерпретация результатов эксперимента позволяет судить о взаимном влиянии нефти и курса доллара на основные макроэкономические индикаторы.

– MAE по ВВП: 3249

– RMSE по ВВП: 3709

– MAE по инфляции: 1.62

–RMSE по инфляции: 2.15

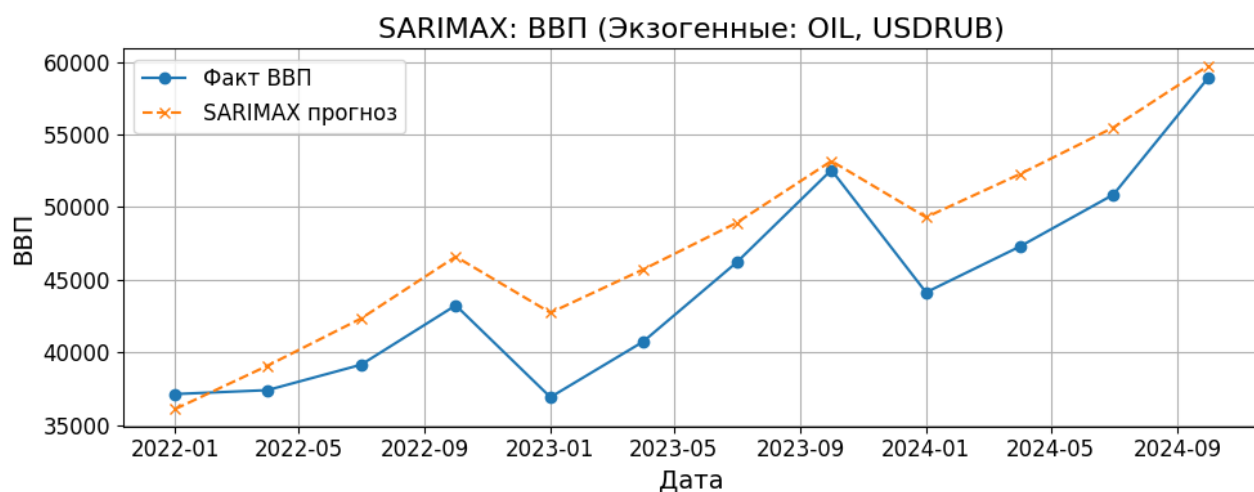


Рисунок 9 – Прогноз ВВП с двумя экзогенными переменными OIL и USDRUB

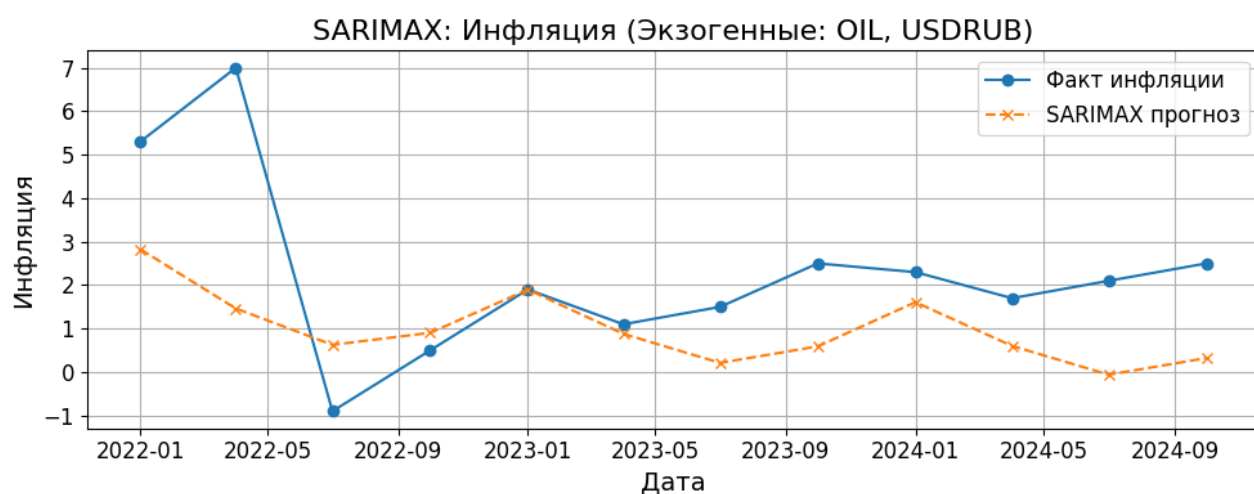


Рисунок 10 – Прогноз инфляции с двумя экзогенными переменными OIL и USDRUB

Эксперимент 4 – SARIMAX с экзогенными переменными BM и WAGE:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \beta_1 BM_t + \beta_2 WAGE_t + S_t + \varepsilon_t,$$

где y_t – значение прогнозируемого показателя (ВВП или инфляция) в момент времени t , φ_i – коэффициенты авторегрессии, θ_j – коэффициенты скользящего среднего, S_t – сезонная составляющая, ε_t – случайная ошибка, BM_t – денежная масса, $WAGE_t$ – уровень заработной платы, β_1, β_2 – коэффициенты при экзогенных переменных.

Включение внутренних макроэкономических факторов, таких как денежная масса и уровень заработной платы, позволяет учесть влияние структурных характеристик российской экономики на динамику исследуемых рядов. Этот подход важен для выявления и количественной оценки фундаментальных драйверов внутренней инфляции и экономического роста.

- MAE по ВВП: 3838
- RMSE по ВВП: 4236
- MAE по инфляции: 2.01
- RMSE по инфляции: 2.42

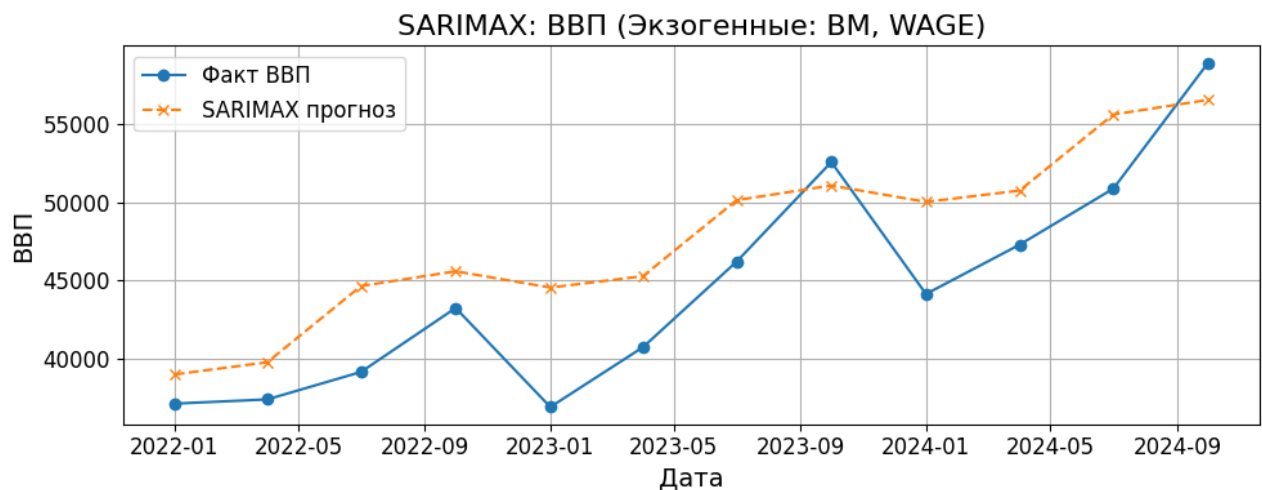


Рисунок 11 – Прогноз ВВП с двумя экзогенными переменными BM и WAGE

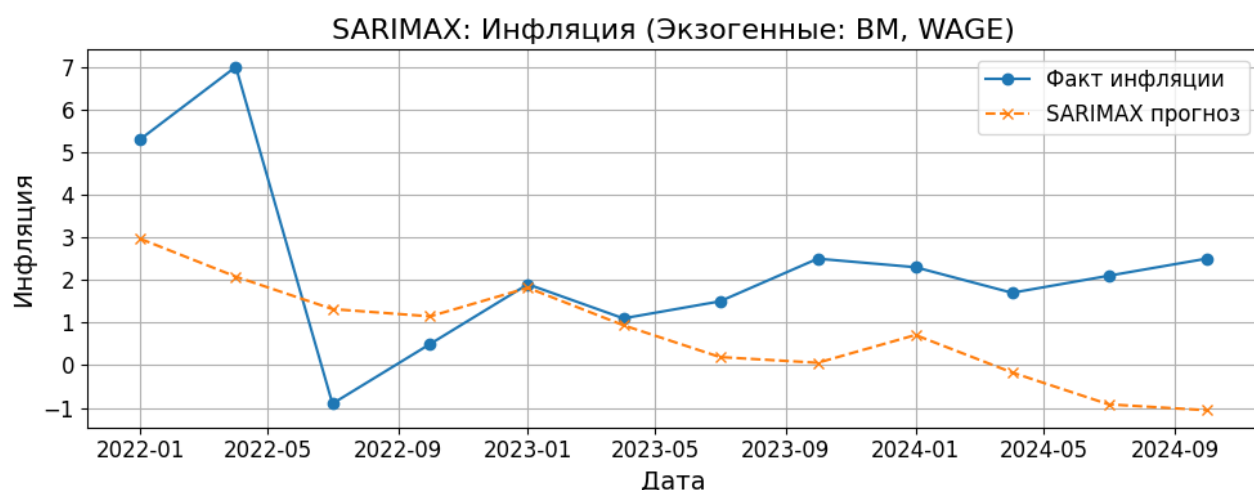


Рисунок 12 – Прогноз инфляции с двумя экзогенными переменными BM и WAGE

Эксперимент 5 – SARIMAX с экзогенными переменными OIL, USDRUB, BM, WAGE:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \beta_1 OIL_t + \beta_2 USDRUB_t + \beta_3 BM_t + \beta_4 WAGE_t + S_t + \varepsilon_t,$$

где y_t – значение прогнозируемого показателя (ВВП или инфляция) в момент времени t , φ_i – коэффициенты авторегрессии, θ_j – коэффициенты скользящего среднего, S_t – сезонная составляющая, ε_t – случайная ошибка, BM_t – денежная масса, $WAGE_t$ – уровень заработной платы, OIL_t – цена на нефть Brent в момент времени t , $USDRUB_t$ – курс рубля к доллару, $\beta_1, \beta_2, \beta_3, \beta_4$ – коэффициенты при экзогенных переменных.

Комплексная модель, в которую включены как внешние, так и внутренние экономические индикаторы, позволяет создать наиболее “гибкую” систему прогнозирования, учитывающую сложные взаимосвязи и перекрёстные эффекты между факторами. Это особенно актуально для построения сценариев макроэкономического развития в условиях многовекторных шоков.

–MAE по ВВП: 3252

–RMSE по ВВП: 3570

- MAE по инфляции: 1.6
- RMSE по инфляции: 2.00

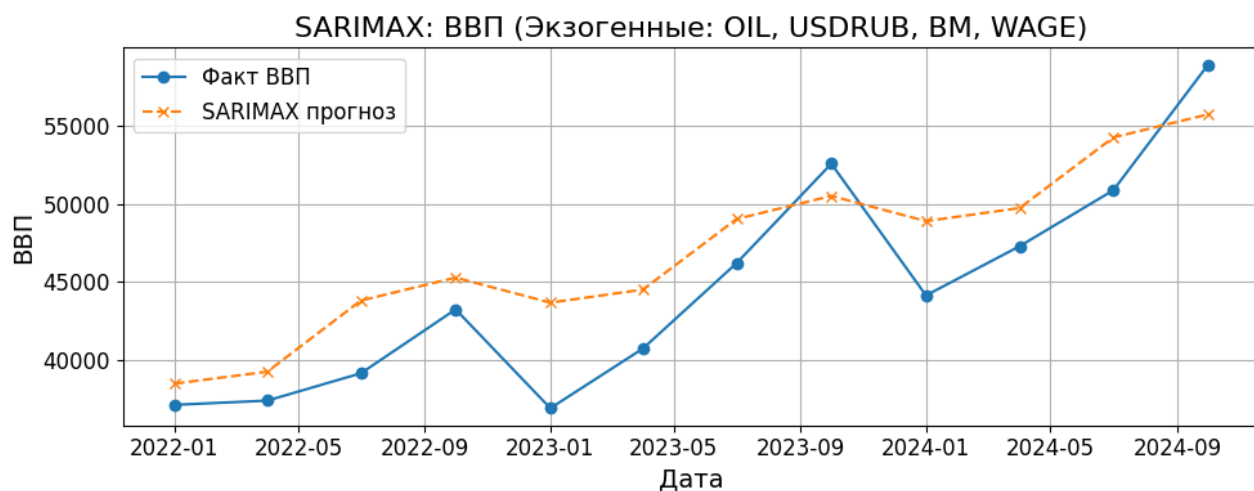


Рисунок 13 – Прогноз ВВП с экзогенными переменными OIL, USDRUB, BM, WAGE

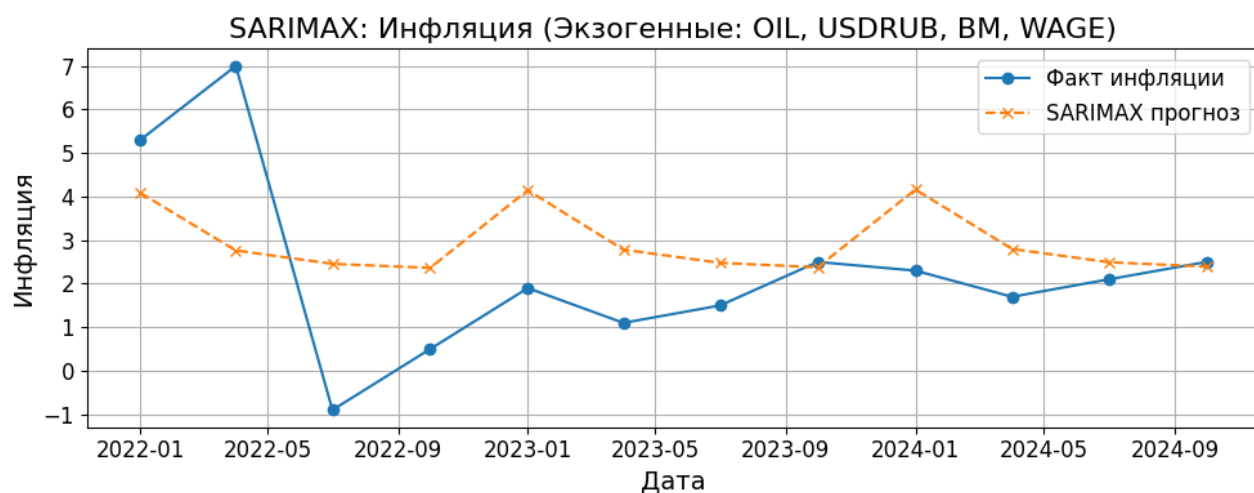


Рисунок 14 – Прогноз инфляции с экзогенными переменными OIL, USDRUB, BM, WAGE

Эксперимент 6 – Анализ чувствительности SARIMAX к изменению OIL:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \beta_1 OIL_t + S_t + \varepsilon_t,$$

где y_t – значение прогнозируемого показателя (ВВП или инфляция) в момент времени t , φ_i – коэффициенты авторегрессии, θ_j – коэффициенты скользящего среднего, S_t – сезонная составляющая, ε_t – случайная ошибка, OIL_t – искусственно изменённая цена на нефть (уменьшенная или увеличенная на 10% в обучающем периоде), β_1 – коэффициент при экзогенной переменной.

Данный эксперимент предназначен для анализа устойчивости прогноза к экстремальным изменениям цен на нефть. Полученные результаты позволяют оценить возможные сценарии развития экономики при внешних шоках и формировать политику макроэкономической устойчивости.

Таблица 2 – Метрики точности при изменении переменной OIL в модели SARIMAX

<i>Метрики</i>	<i>При уменьшении значений на 10%</i>	<i>При увеличении значений на 10%</i>
MAE по ВВП	2776	1612
Продолжение таблицы 2		
RMSE по ВВП	3303	2129
MAE по инфляции	1.47	1.47
RMSE по инфляции	2.05	2.06

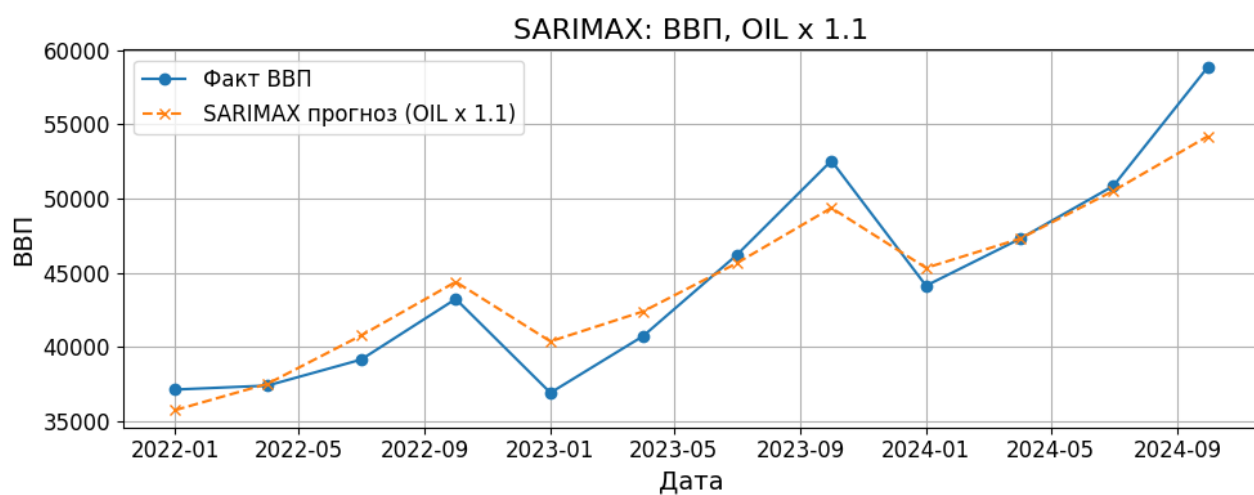


Рисунок 15 – Прогноз ВВП с увеличенной переменной OIL на 10%

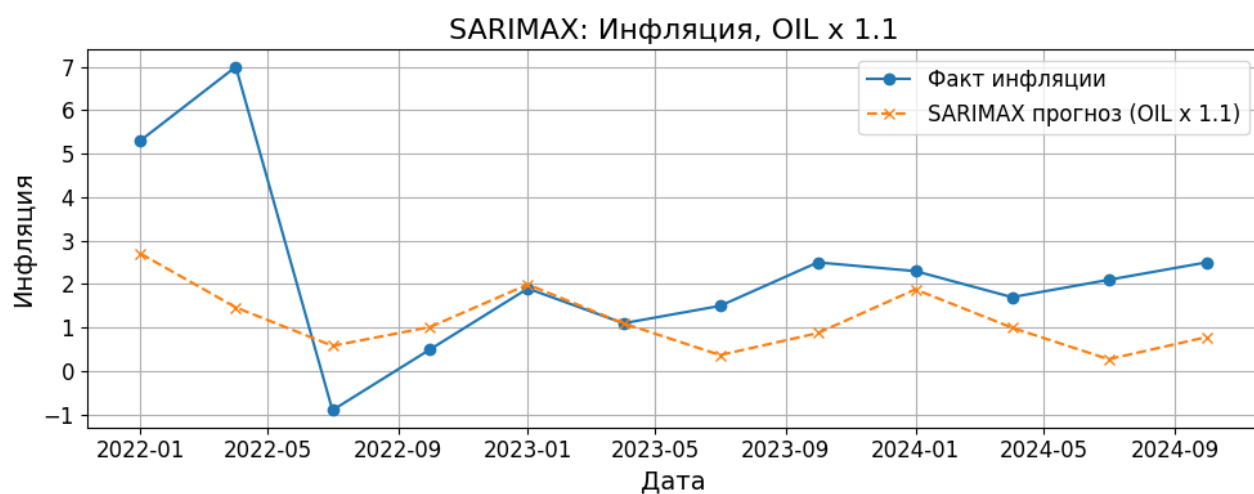


Рисунок 16 – Прогноз инфляции с увеличенной переменной OIL на 10%

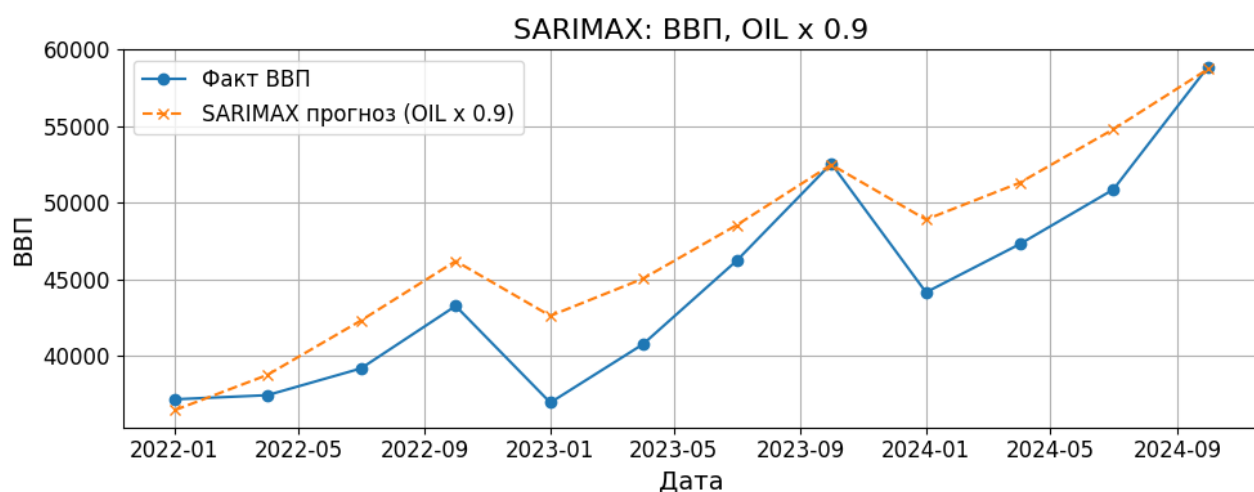


Рисунок 17 – Прогноз ВВП с уменьшенной переменной OIL на 10%

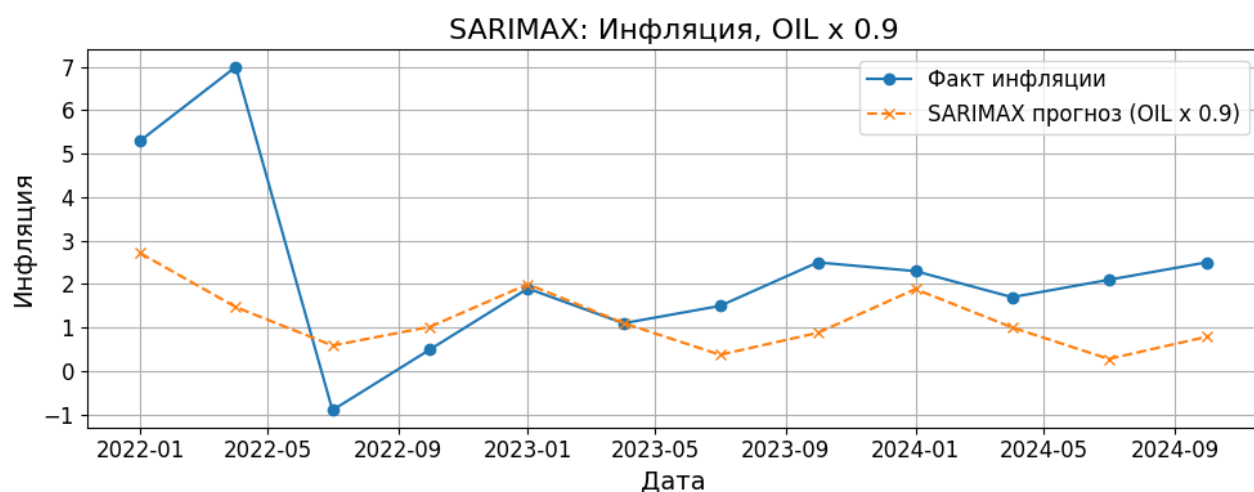


Рисунок 18 – Прогноз инфляции с уменьшенной переменной OIL на 10%

3.2.3 Реализация и спецификация модели VAR

Второй из моделей, реализованных в рамках данного исследования для прогнозирования макроэкономических индикаторов, стала модель векторной авторегрессии (VAR, Vector AutoRegression). Данная методология широко используется в мировой эконометрической практике для анализа сразу нескольких временных рядов, чьи значения на каждом временном интервале опреде-

ляются как их собственными лагами, так и лагами других переменных в системе.

Модель VAR даёт возможность не только учитывать автокорреляцию каждого показателя, но и моделировать сложные динамические связи между различными секторами экономики. Для российской экономики, где динамика ВВП, инфляции, нефтяных цен, валютных курсов и других макроэкономических индикаторов тесно взаимосвязаны, использование VAR позволяет выявить как прямое, так и косвенное влияние одних факторов на другие, а также отслеживать эффект обратных связей.

В рамках исследования VAR применялась для одновременного прогнозирования сразу нескольких показателей: валового внутреннего продукта (GDP), инфляции (Inflation), цены на нефть (OIL), курса рубля к доллару (USDRUB), денежной массы (BM) и средней заработной платы (WAGE). Это позволило реализовать комплексный подход и построить сценарии развития макроэкономики в условиях множественных внутренних и внешних шоков.

Математическая запись модели:

$$y_t = c + \sum_{i=1}^p A_i y_{t-i} + \varepsilon_t,$$

где y_t – вектор значений всех прогнозируемых переменных в момент времени t , c – вектор констант, A_i – матрицы коэффициентов при каждом лаге i , y_{t-i} – вектор значений переменных на шаге $t-i$, ε_t – вектор случайных ошибок.

В качестве параметра p (число лагов) выбирался оптимум по информационным критериям (AIC и др.) для каждой спецификации.

Эксперимент 1 – VAR для ВВП и инфляции:

В первом эксперименте VAR строится только по двум основным макроэкономическим индикаторам – валовому внутреннему продукту и инфляции. Каждая переменная в системе определяется как функция собственных прошлых значений, а также лагов второй переменной. Такой подход позволяет не только уловить автокорреляционные эффекты, но и понять, насколько динамика инфляции и экономического роста связана между собой исторически. Подобная

модель хорошо отражает внутренние процессы движения российской экономики без привлечения внешних факторов.

Формула:

$$\begin{pmatrix} GDP_t \\ Inflation_t \end{pmatrix} = c + \sum_{i=1}^p A_i \begin{pmatrix} GDP_{t-i} \\ Inflation_{t-i} \end{pmatrix} + \varepsilon_t,$$

где GDP_t – валовой внутренний продукт, $Inflation_t$ – уровень инфляции, c – вектор констант, A_i – матрицы коэффициентов при лаге i , p – порядок VAR-модели, ε_t – вектор ошибок, все значения – на момент времени t .

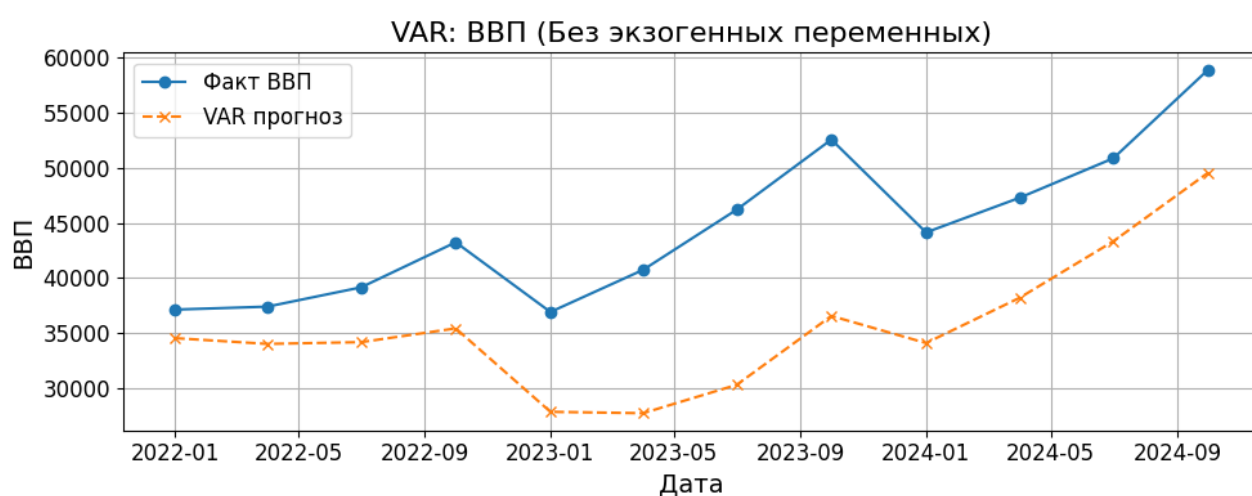


Рисунок 19 – Прогноз ВВП без экзогенных переменных

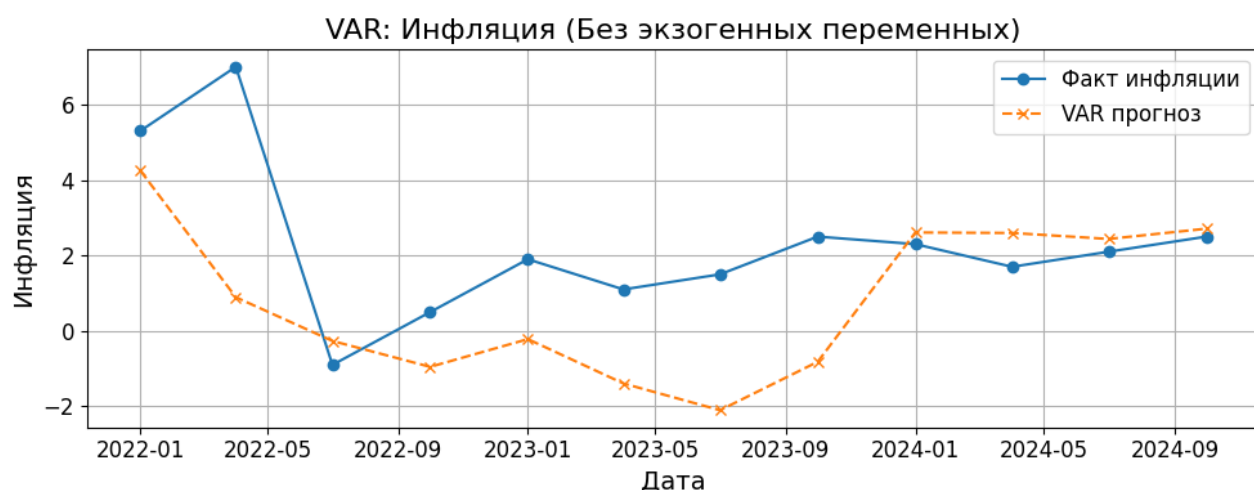


Рисунок 20 – Прогноз инфляции без экзогенных переменных

–MAE по ВВП: 9051

–RMSE по ВВП: 9954

–MAE по инфляции: 1.88

–RMSE по инфляции: 2.53

Эксперимент 2 – VAR для ВВП, инфляции и OIL:

В следующем эксперименте к системе добавляется цена на нефть, являющаяся важнейшим внешним драйвером для российской экономики. Модель позволяет учитывать, как глобальные сырьевые шоки передаются на внутренние макроэкономические процессы, и выявлять как прямое влияние изменений цен на нефть на динамику ВВП и инфляции, так и возможные обратные эффекты. Такая система хорошо подходит для оценки устойчивости экономики к внешним колебаниям.

Формула:

$$\begin{pmatrix} GDP_t \\ Inflation_t \\ OIL_t \end{pmatrix} = c + \sum_{i=1}^p A_i \begin{pmatrix} GDP_{t-i} \\ Inflation_{t-i} \\ OIL_{t-i} \end{pmatrix} + \varepsilon_t,$$

где GDP_t – валовой внутренний продукт, $Inflation_t$ – уровень инфляции, OIL_t – мировая цена на нефть, c – вектор констант, A_i – матрицы коэффициентов при

лаге i , p – порядок VAR-модели, ε_t – вектор ошибок, все значения – на момент времени t .

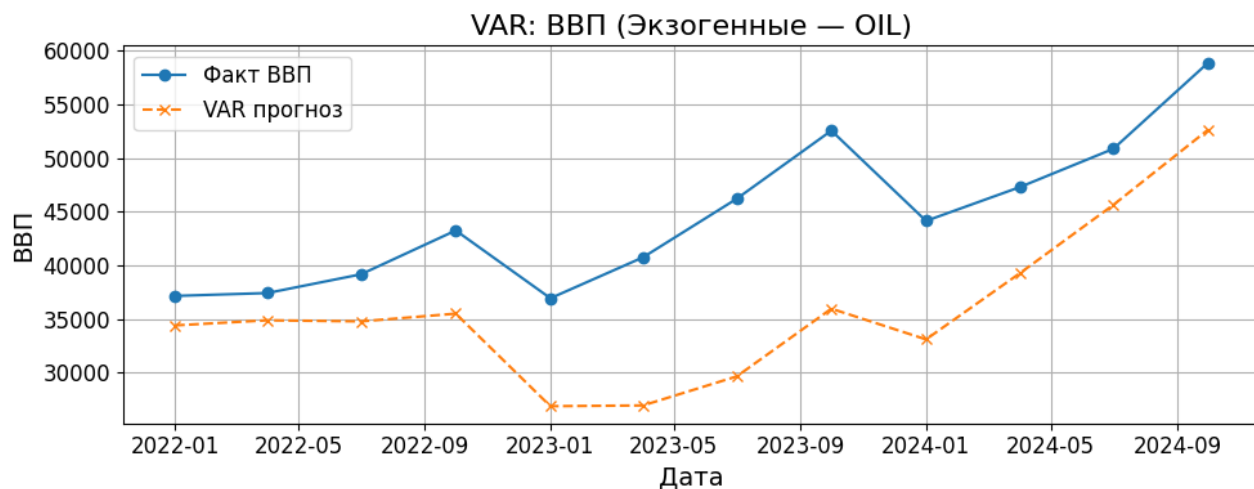


Рисунок 21 – Прогноз ВВП с OIL

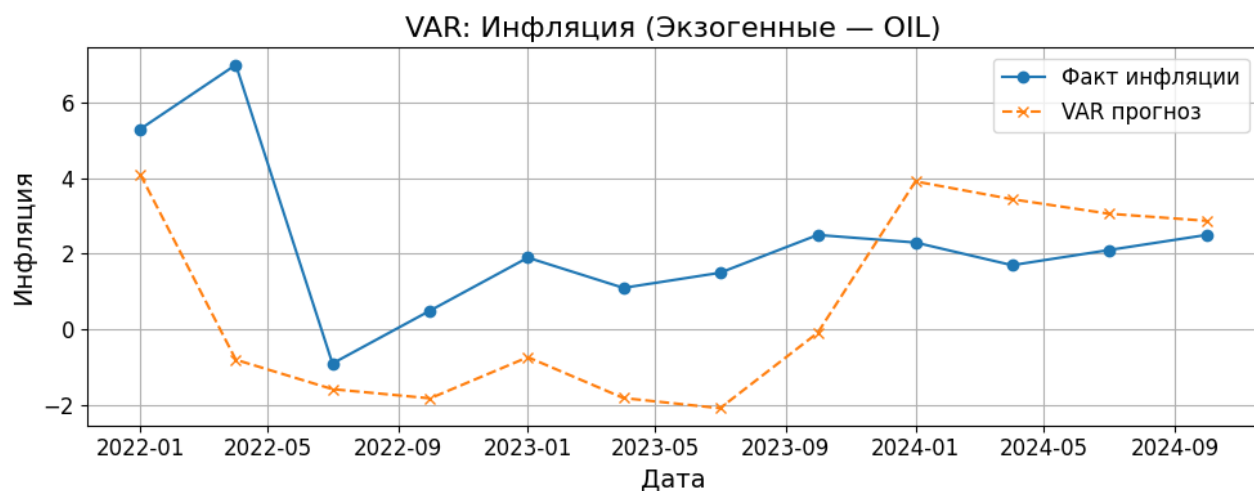


Рисунок 22 – Прогноз инфляции с OIL

- MAE по ВВП: 8755
- RMSE по ВВП: 9954
- MAE по инфляции: 2.37

–RMSE по инфляции: 3.02

Эксперимент 3 – VAR для ВВП, инфляции, OIL и USDRUB:

Здесь в систему VAR, помимо цен на нефть, включается курс рубля к доллару. Это позволяет одновременно оценивать воздействие валютных и сырьевых шоков на внутренние макроэкономические показатели. С помощью такой модели можно проследить, как внешние колебания передаются через несколько каналов, формируя сложные траектории развития ВВП и инфляции. Система уравнений учитывает задержку во влиянии всех четырёх показателей друг на друга.

Формула:

$$\begin{pmatrix} GDP_t \\ Inflation_t \\ OIL_t \\ USDRUB_t \end{pmatrix} = c + \sum_{i=1}^p A_i \begin{pmatrix} GDP_{t-i} \\ Inflation_{t-i} \\ OIL_{t-i} \\ USDRUB_{t-i} \end{pmatrix} + \varepsilon_t,$$

где GDP_t – валовой внутренний продукт, $Inflation_t$ – уровень инфляции, OIL_t – мировая цена на нефть, $USDRUB_t$ – курс рубля к доллару США, c – вектор констант, A_i – матрицы коэффициентов при лаге i , p – порядок VAR-модели, ε_t – вектор ошибок, все значения – на момент времени t .

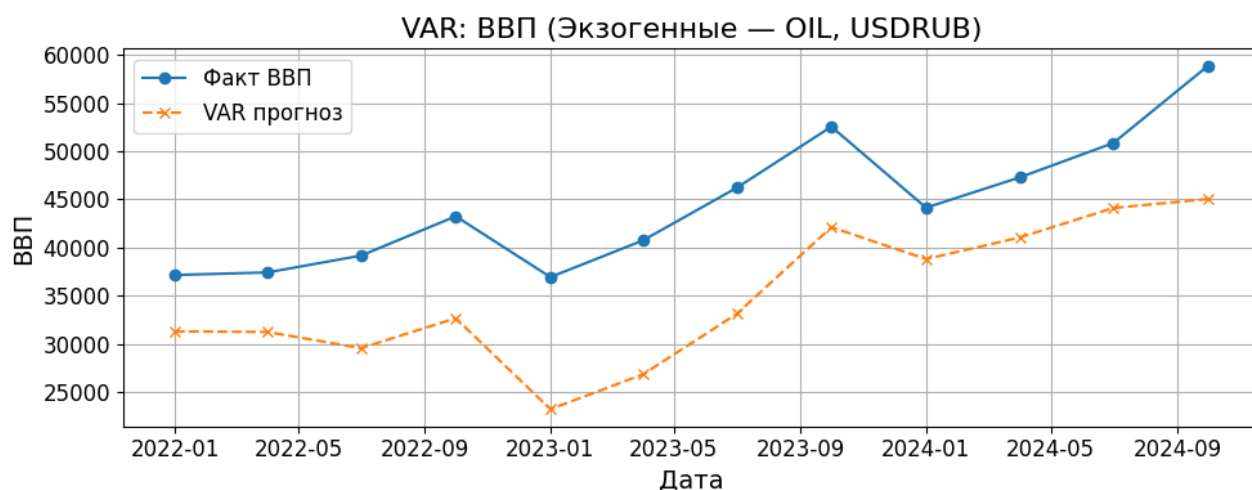


Рисунок 23 – Прогноз ВВП с OIL и USDRUB

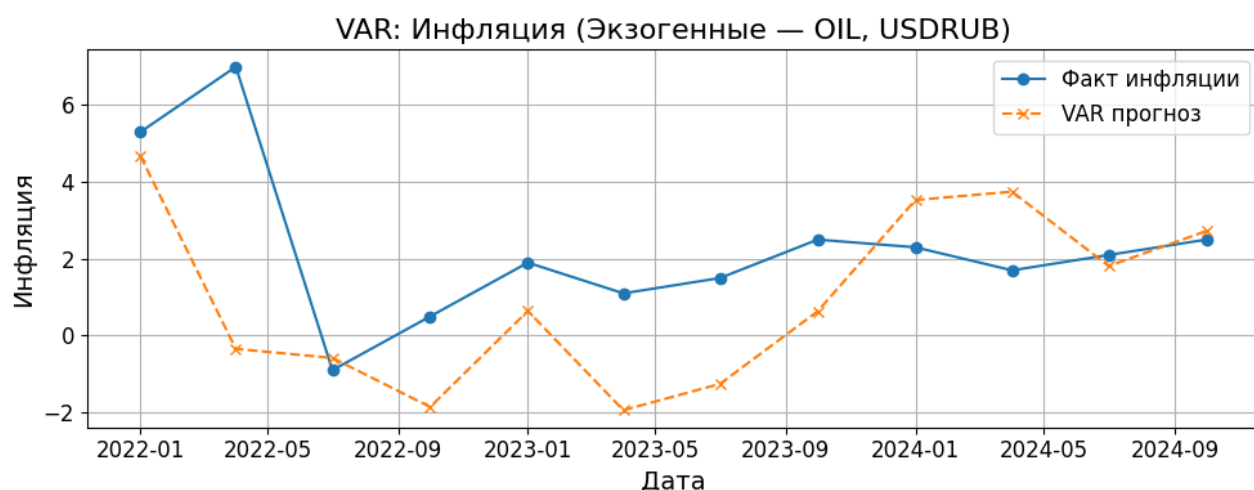


Рисунок 24 – Прогноз инфляции с OIL и USDRUB

–MAE по ВВП: 9636

–RMSE по ВВП: 10187

–MAE по инфляции: 1.95

–RMSE по инфляции: 2.70

Эксперимент 4 – VAR для ВВП, инфляции, денежной массы и зарплаты:

В этом варианте VAR учитываются только внутренние фундаментальные переменные – денежная масса и уровень средней заработной платы – вместе с ВВП и инфляцией. Такой эксперимент даёт возможность оценить роль монетарных и социальных факторов в формировании макроэкономических тенденций, а также проверить, насколько динамика экономики может быть объяснена без учёта внешних шоков.

Формула:

$$\begin{pmatrix} GDP_t \\ Inflation_t \\ BM_t \\ WAGE_t \end{pmatrix} = c + \sum_{i=1}^p A_i \begin{pmatrix} GDP_{t-i} \\ Inflation_{t-i} \\ BM_{t-i} \\ WAGE_{t-i} \end{pmatrix} + \varepsilon_t,$$

где GDP_t – валовой внутренний продукт, $Inflation_t$ – уровень инфляции, BM_t – денежная масса, $WAGE_t$ – средняя заработная плата, c – вектор констант, A_i –

матрицы коэффициентов при лаге i , p – порядок VAR-модели, ε_t – вектор ошибок, все значения – на момент времени t .

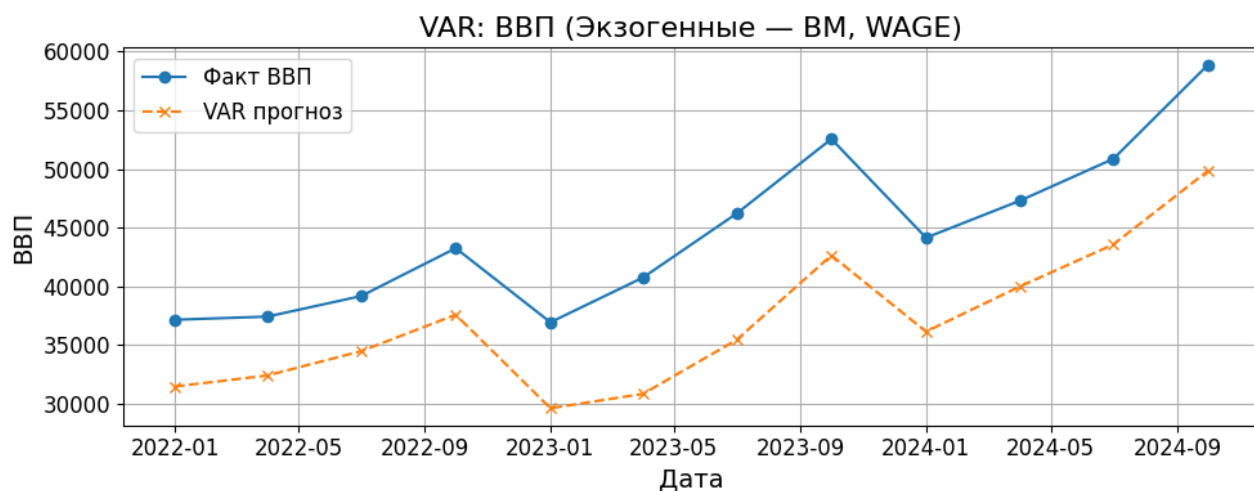


Рисунок 25 – Прогноз ВВП с ВМ и WAGE

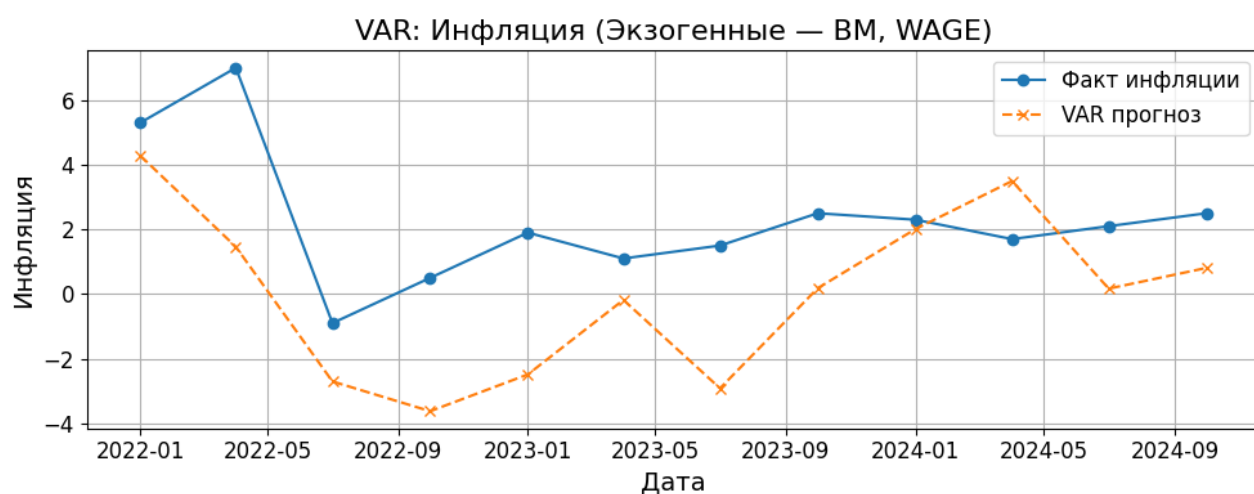


Рисунок 26 – Прогноз инфляции с ВМ и WAGE

- MAE по ВВП: 7558
- RMSE по ВВП: 7806
- MAE по инфляции: 2.55
- RMSE по инфляции: 3.00

Эксперимент 5 – VAR для ВВП, инфляции, нефти, курса доллара, денежной массы и зарплаты:

В самой расширенной конфигурации в VAR одновременно входят все ключевые внутренние и внешние макроэкономические переменные. Это позволяет анализировать наиболее полные сценарии развития российской экономики и выявлять сложные структуры взаимного влияния между всеми основными драйверами – от цен на нефть и валютных курсов до параметров денежного обращения и динамики заработной платы.

Формула:

$$\begin{pmatrix} GDP_t \\ Inflation_t \\ OIL_t \\ USDRUB_t \\ BM_t \\ WAGE_t \end{pmatrix} = c + \sum_{i=1}^p A_i \begin{pmatrix} GDP_{t-i} \\ Inflation_{t-i} \\ OIL_{t-i} \\ USDRUB_{t-i} \\ BM_{t-i} \\ WAGE_{t-i} \end{pmatrix} + \varepsilon_t,$$

где GDP_t – валовой внутренний продукт, $Inflation_t$ – уровень инфляции, BM_t – денежная масса, $WAGE_t$ – средняя заработная плата, OIL_t – мировая цена на нефть, $USDRUB_t$ – курс рубля к доллару США, c – вектор констант, A_i – матрицы коэффициентов при лаге i , p – порядок VAR-модели, ε_t – вектор ошибок, все значения – на момент времени t .

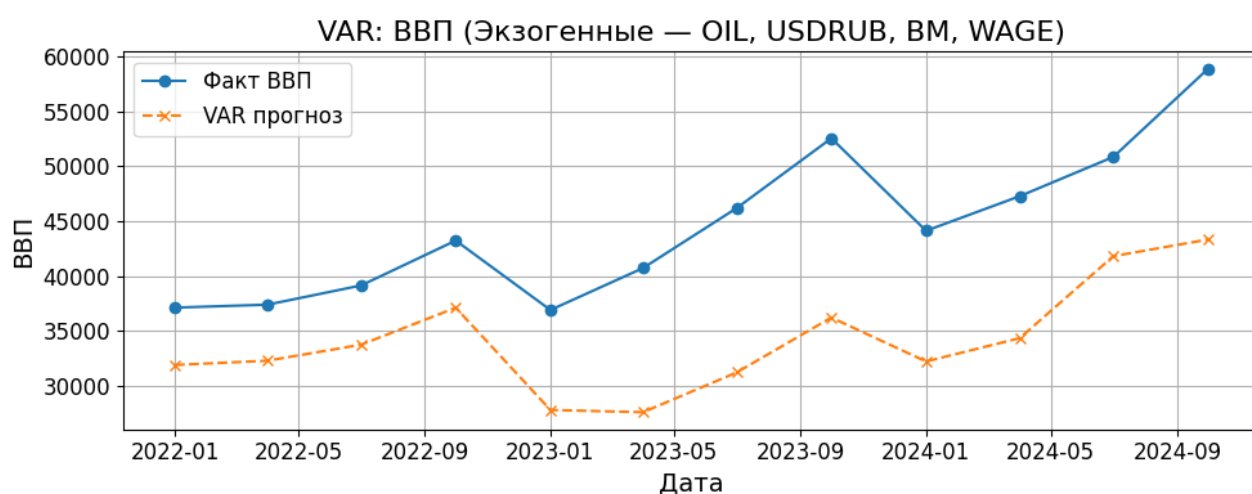


Рисунок 27 – Прогноз ВВП со всеми экзогенными переменными

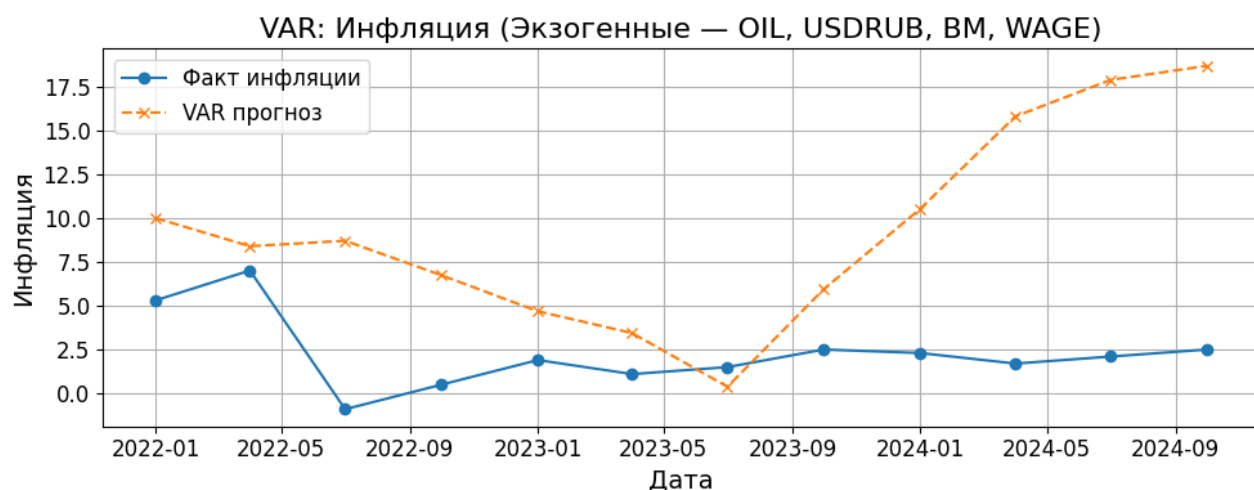


Рисунок 28 – Прогноз инфляции со всеми экзогенными переменными

–MAE по ВВП: 10395

–RMSE по ВВП: 11169

–MAE по инфляции: 7.15

–RMSE по инфляции: 8.93

Эксперимент 5 – Анализ чувствительности VAR к изменению OIL:

В дополнительном эксперименте проводится сценарное варьирование ключевой внешней переменной – цены на нефть (OIL). Для оценки устойчивости прогноза и степени влияния нефтяных шоков в обучающем наборе значения OIL искусственно увеличиваются или уменьшаются на фиксированный процент (например, $\pm 10\%$). Это позволяет количественно оценить, насколько сильно сырьевые сценарии могут изменить траекторию развития ВВП и инфляции в российской экономике, а также проследить, как эти изменения передаются внутри системы через внутренние взаимосвязи между переменными.

Формула:

$$\begin{pmatrix} GDP_t \\ Inflation_t \\ OIL_t \end{pmatrix} = c + \sum_{i=1}^p A_i \begin{pmatrix} GDP_{t-i} \\ Inflation_{t-i} \\ OIL_{t-i} \end{pmatrix} + \varepsilon_t,$$

где GDP_t – валовой внутренний продукт, $Inflation_t$ – уровень инфляции, OIL_t – мировая цена на нефть, OIL_t – искусственно изменённая цена на нефть (уменьшенная или увеличенная на 10% в обучающем периоде), c – вектор констант, A_i – матрицы коэффициентов при лаге i , p – порядок VAR-модели, ε_t – вектор ошибок, все значения – на момент времени t .

Таблица 3 – Метрики точности при изменении переменной OIL в модели VAR

<i>Метрики</i>	<i>При уменьшении значений на 10%</i>	<i>При увеличении значений на 10%</i>
MAE по ВВП	8755	8755
RMSE по ВВП	9954	9954
MAE по инфляции	2.37	2.37
RMSE по инфляции	3.02	3.02

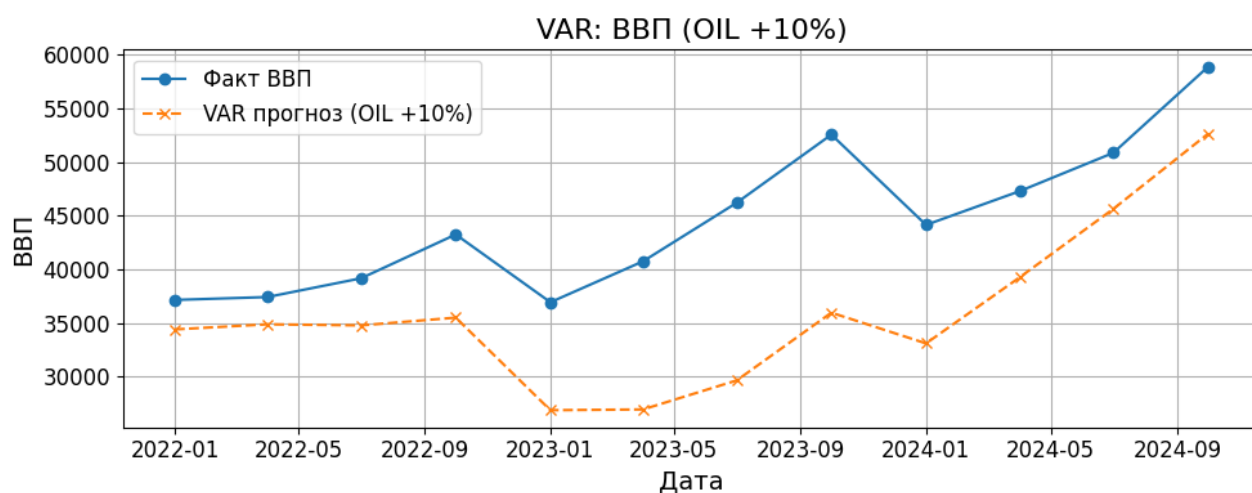


Рисунок 29 – Прогноз ВВП с увеличенной переменной OIL на 10%

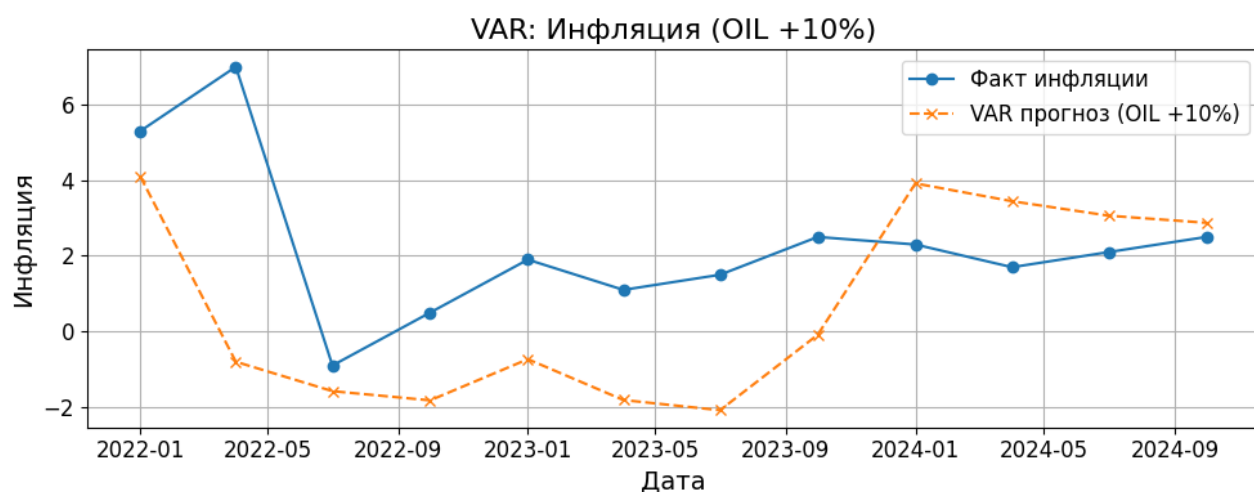


Рисунок 30 – Прогноз инфляции с увеличенной переменной OIL на 10%

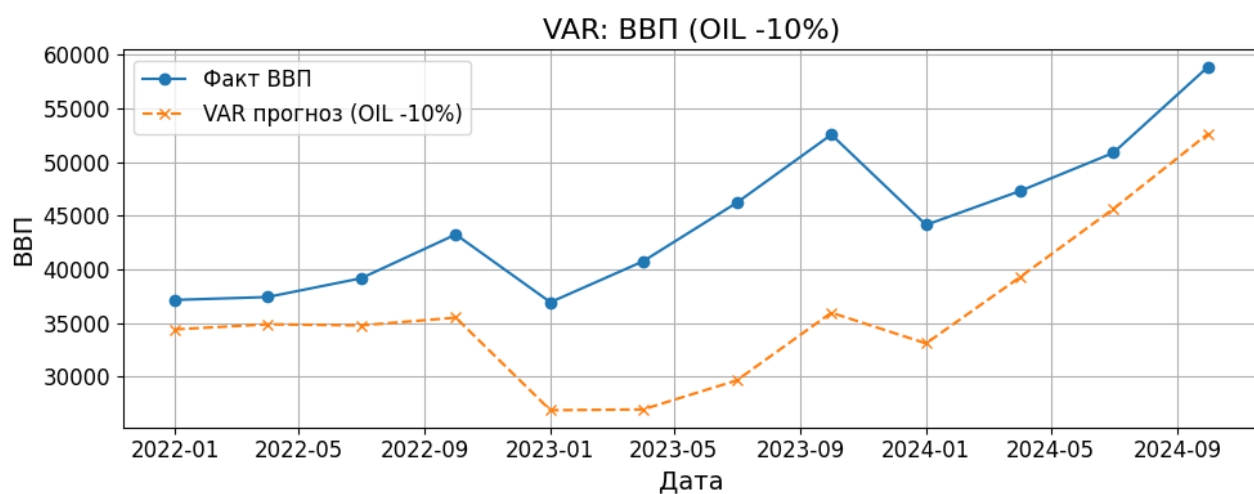


Рисунок 31 – Прогноз ВВП с уменьшенной переменной OIL на 10%

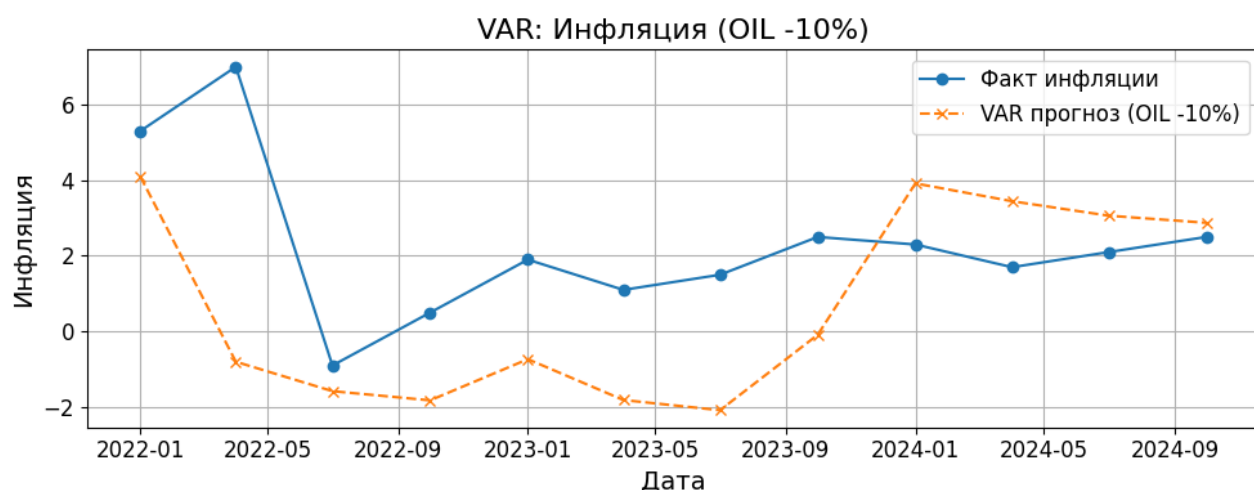


Рисунок 32 – Прогноз инфляции с уменьшенной переменной OIL на 10%

3.3 Анализ полученных результатов

В данном разделе представлен комплексный анализ результатов прогнозирования макроэкономических показателей, полученных с помощью моделей SARIMAX и VAR. Сравнение построено на основе итоговых значений метрик качества (MAE и RMSE) по основным прогнозируемым переменным – валовому внутреннему продукту (ВВП) и инфляции – на тестовом периоде 2022-2024 годов. Такой подход обеспечивает объективную оценку эффективности каждого из методов и позволяет сделать выводы о степени их пригодности для практического применения в условиях российской экономики.

Сводная таблица метрик прогнозирования:

Для обеспечения наглядности и прозрачности сравнения результатов ниже представлена таблица 3, отражающая значения средних абсолютных ошибок (MAE) и среднеквадратичных ошибок (RMSE) по ВВП и инфляции для всех рассмотренных конфигураций моделей.

Таблица 4 – Сравнение метрик точности моделей VAR и SARIMAX

Эксперимент	MAE ВВП	RMSE ВВП	MAE ин- фляция	RMSE ин- фляция
SARIMAX (без экзогенных переменных)	2706	3 194	1.47	2.06

SARIMAX (OIL)	1625	2142	1.47	2.05
SARIMAX (OIL, USDRUB)	3249	3709	1.62	2.15
SARIMAX (BM, WAGE)	3838	4236	2.01	2.42
SARIMAX (OIL, USDRUB, BM, WAGE)	3252	3570	1.60	2.00
SARIMAX (OIL +10%)	1612	2129	1.47	2.06
SARIMAX (OIL -10%)	2776	3303	1.47	2.05
VAR (GDP, Inflation)	9051	9954	1.88	2.53
VAR (GDP, Inflation, OIL)	8755	9954	2.37	3.02
VAR (GDP, Inflation, OIL, USDRUB)	9636	10187	1.95	2.70
VAR (GDP, Inflation, BM, WAGE)	7558	7806	2.55	3.00
VAR (GDP, Inflation, OIL, USDRUB, BM, WAGE)	10395	11169	7.15	8.93
VAR (OIL $\pm 10\%$)	8755	9954	2.37	3.02

Анализ результатов моделей SARIMAX

Применение моделей SARIMAX позволило выявить ряд ключевых закономерностей в динамике ВВП и инфляции. Основные особенности и преимущества данного подхода проявились в следующих аспектах.

Использование внешних факторов, прежде всего мировых цен на нефть, как экзогенных переменных, обеспечивает существенное повышение точности прогноза ВВП. Добавление переменной OIL приводит к наименьшим значениям ошибок ($MAE = 1625$; $RMSE = 2142$), что отражает глубокую структурную связь российской экономики с нефтяным сектором.

Базовая конфигурация SARIMAX (без экзогенных переменных) демонстрирует устойчивую точность, что свидетельствует о наличии выраженных внутренних трендов и сезонных закономерностей в динамике ключевых макроэкономических показателей.

Включение в модель дополнительных экзогенных переменных, таких как курс рубля к доллару США, денежная масса и уровень заработной платы, не способствует дальнейшему снижению ошибок прогнозирования. В ряде случаев наблюдается увеличение ошибок, что связано с усилением мультиколлинеарности, усложнением структуры модели и снижением обобщающей способности алгоритма.

Чувствительность прогноза ВВП к вариациям цены на нефть подтверждается экспериментами с искусственным изменением данной переменной. Повышение нефтяных цен на 10% приводит к дополнительному снижению ошибки ($MAE = 1612$), в то время как их уменьшение сопровождается увеличением ошибки до 2776. Это отражает сохраняющуюся зависимость российской экономики от конъюнктуры глобальных сырьевых рынков.

Для инфляции метрики ошибок остаются относительно стабильными во всех конфигурациях SARIMAX. Это свидетельствует о высокой инерционности ценовых процессов и значительном влиянии факторов, которые трудно формализовать в рамках выбранных моделей, например, административных мер или изменяющихся ожиданий экономических агентов.

Анализ результатов моделей VAR

Векторная авторегрессия применялась для моделирования взаимосвязей между несколькими макроэкономическими переменными и анализа их совместной динамики. Полученные результаты позволяют сделать следующие наблюдения.

Модели VAR во всех протестированных конфигурациях показали менее удовлетворительные результаты по сравнению с SARIMAX. Значения ошибок по ВВП значительно превышают соответствующие показатели SARIMAX, достигая MAE в диапазоне от 7558 до 10395.

Добавление дополнительных переменных, таких как цены на нефть, курс рубля, денежная масса и заработная плата, не только не улучшило, но в ряде случаев даже ухудшило точность прогнозирования. Это объясняется увеличением числа параметров модели и усилением взаимных корреляций между пре-

дикторами, что приводит к эффекту переобучения и потере способности модели к адекватному обобщению.

VAR-модели проявляют недостаточную гибкость в условиях резких структурных изменений, характерных для российской экономики. Прогнозы оказываются избыточно инерционными и слабо чувствительными к новым внешним шокам, что подтверждается экспериментами с варьированием нефтяных цен: ошибки остаются практически неизменными при изменении этого ключевого экзогенного фактора.

Наибольшие значения ошибок для инфляции наблюдаются именно в максимально насыщенных переменными конфигурациях VAR, что указывает на сложности корректного учета всех возможных взаимодействий внутри многомерной системы в условиях ограниченного объёма данных.

Сравнительный анализ и основные выводы

Сопоставление результатов обеих групп моделей позволяет сделать ряд выводов относительно их применимости для задач макроэкономического прогнозирования в условиях российской экономики.

Модели SARIMAX продемонстрировали высокую эффективность и достаточную точность при прогнозировании ВВП. Особенно успешной оказалась спецификация с одной экзогенной переменной – ценой на нефть, что соответствует структуре национальной экономики и её зависимости от внешних сырьевых факторов.

VAR-модели, несмотря на свой теоретический потенциал для улавливания сложных взаимосвязей между макроэкономическими показателями, в данной постановке задачи продемонстрировали менее удовлетворительные результаты. Ключевыми причинами этого явились высокая сложность модели, наличие мультиколлинеарности между переменными и относительно малый объём доступных исторических данных для устойчивой оценки всех параметров системы.

Для инфляции ни одна из моделей не показала существенного преимущества по качеству прогнозирования, что подчеркивает сложность данного пока-

зателя и необходимость дальнейшего поиска дополнительных факторов, способных повысить точность моделирования ценовой динамики.

В целом, результаты работы подтверждают, что выбор модели должен определяться не только исходя из её теоретической мощности, но и с учётом специфики исходных данных, экономической логики процесса и реального вклада каждой переменной в формирование прогнозируемого показателя.

Таким образом, проведённый анализ наглядно демонстрирует преимущества SARIMAX для краткосрочного прогнозирования ВВП в условиях российской экономики, а также показывает ограничения применения VAR в задачах с ограниченным числом наблюдений и высоким уровнем взаимных корреляций между переменными. Полученные результаты могут быть использованы для дальнейшей разработки методов эконометрического моделирования и уточнения подходов к отбору переменных и построению макроэкономических прогнозов.

ЗАКЛЮЧЕНИЕ

В бакалаврской работе представлен теоретический анализ и результаты тестирования моделей прогнозирования ключевых макроэкономических показателей Российской Федерации – валового внутреннего продукта и инфляции с использованием временных рядов. Исследование охватило все основные этапы: от подготовки исходных данных и их предварительной обработки до построения моделей, анализа полученных результатов и их интерпретации.

В ходе работы модель SARIMAX показала себя как надёжный инструмент для краткосрочного прогнозирования динамики ВВП. Наиболее высокая точность прогноза была достигнута при включении цены на нефть в качестве экзогенного фактора, что закономерно отражает специфику российской экономики. Вместе с тем добавление нескольких экзогенных переменных одновременно (OIL, USDRUB, BM, WAGE) не всегда приводило к улучшению прогноза. Это подчёркивает важность взвешенного выбора предикторов, чтобы избежать переобучения модели и проблем, связанных с мультиколлинеарностью.

Применение модели VAR позволило исследовать взаимосвязи между различными макроэкономическими показателями. Однако её прогностическая точность оказалась ниже. Это связано как с ограниченным объёмом доступных данных, так и со сложностью многомерных временных рядов, которые повышают чувствительность модели к шуму. Проведённые эксперименты показали, что для повышения эффективности VAR-моделей целесообразно рассмотреть возможность использования более сложных вариантов, включая структурные или гибридные модификации.

По итогам проведённого исследования можно выделить следующие ключевые выводы:

- модель SARIMAX демонстрирует высокую эффективность при краткосрочном прогнозировании ВВП;
- модель VAR даёт ценные результаты при анализе взаимосвязей, но предъявляет высокие требования к объёму и качеству данных;

- включение экзогенных факторов, прежде всего цен на нефть, способствует повышению точности прогнозов;

- стабильность прогностических моделей возрастает при использовании сценарного анализа и варьировании ключевых переменных.

В целом, результаты работы подтвердили значимость и перспективность применения методов анализа временных рядов для прогнозирования макроэкономических показателей. Полученный опыт и практические наработки могут быть использованы как в дальнейшем научном исследовании, так и в прикладной практике государственного и корпоративного планирования. В перспективе дальнейшее развитие данного направления, по моему мнению, будет связано с внедрением нейросетевых подходов, гибридных моделей и байесовских методов, что позволит повысить гибкость и адаптивность прогностических систем к изменяющимся условиям внешней среды.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Аксянова, А.В. Математическое моделирование экономических процессов: учебное пособие / А.В. Аксянова, А.Н. Валеева, Д.Н. Валеева, А.М. Гумеров. – Казань: КНИТУ, 2016. – 92 с. – URL: <https://e.lanbook.com/book/102125> (дата обращения: 21.02.2025).
2. Андриевская, Н.В. Статистические и интеллектуальные методы прогнозирования: учебное пособие / Н.В. Андриевская. – Пермь: ПНИПУ, 2022. – 222 с. – URL: <https://e.lanbook.com/book/328751> (дата обращения: 03.04.2025).
3. Бурнашев, Р.А. Анализ и прогнозирование временных рядов: путь от классических моделей к современным решениям. Часть 1. Базовые инструменты и методы: учебное пособие для вузов / Р.А. Бурнашев, М.К. Арабов, М.Д. Мисаров. – Санкт-Петербург: Лань, 2025. – 216 с – URL: <https://e.lanbook.com/book/488951> (дата обращения: 09.06.2025).
4. Вержбицкий, В.М. Численные методы (линейная алгебра и нелинейные уравнения): учеб. пособие для вузов: рек. Мин. обр. РФ / В.М. Вержбицкий. – М.: Высш. шк., 2000. – 268 с.
5. Вержбицкий, В.М. Численные методы (математический анализ и обыкновенные дифференциальные уравнения): учеб. пособие для вузов / В.М. Вержбицкий. – М.: Высш. шк., 2001. – 382с
6. Горожанина, Е.И. Имитационное моделирование: учебник / Е. И. Горожанина, Е.А. Богданова. – 2-е изд. – Самара: ПГУТИ, 2023. – 300 с. – URL: <https://e.lanbook.com/book/411686> (дата обращения: 03.04.2025).
7. Золкин, А.Л. Программирование эконометрических моделей: учебное пособие для вузов / А.Л. Золкин. – Санкт-Петербург: Лань, 2025. – 140 с. – URL: <https://e.lanbook.com/book/482948> (дата обращения: 09.06.2025).
8. Кабанов, А.Н. Технологии обработки информации: учебное пособие / А.Н. Кабанов, Д.Н. Фоломкин. – Рязань: РГРТУ, 2017. — 48 с – URL: <https://e.lanbook.com/book/168325> (дата обращения: 09.06.2025).

9. Катаргин, Н.В. Эконометрическое моделирование / Н.В. Катаргин. – 2-е изд., стер. – Санкт-Петербург: Лань, 2023. – 124 с. – URL: <https://e.lanbook.com/book/306797> (дата обращения: 09.06.2025).
10. Колмогорова, С.М. Табличный процессор «Excel»: учебно-методическое пособие / С.М. Колмогорова. – Екатеринбург: УГГУ, 2023. – 172 с. – URL: <https://e.lanbook.com/book/453554> (дата обращения: 03.04.2025).
11. Кизбикенов, К.О. Прогнозирование и временные ряды: учебное пособие / К.О. Кизбикенов. – Барнаул: АлтГПУ, 2017. – 115 с. – <https://e.lanbook.com/book/112174> (дата обращения: 03.04.2025).
12. Ландовский, В.В. Численные методы: учебное пособие / В.В. Ландовский. – Новосибирск: НГТУ, 2023. – 72 с. – URL: <https://e.lanbook.com/book/404582>. (дата обращения: 27.11.2024).
13. Лихачев, М.О. Макроэкономика: учебно-методическое пособие / М.О. Лихачев. – Москва: МПГУ, 2017. – 116 с. – URL: <https://e.lanbook.com/book/107350> (дата обращения: 09.06.2025).
14. Митяков, Е.С. Искусственный интеллект и машинное обучение: учебное пособие для вузов / Е.С. Митяков, А.Г. Шмелева, А.И. Ладынин. – Санкт-Петербург: Лань, 2025. – 252 с. – URL: <https://e.lanbook.com/book/450827> (дата обращения: 03.04.2025).
15. Салмина, Н.Ю. Моделирование социально-экономических систем и процессов: учебное пособие / Н.Ю. Салмина. – Томск: Томский государственный университет систем управления и радиоэлектроники, 2016. – 198 с. – URL: <https://www.iprbookshop.ru/72139.html>. (дата обращения: 11.12.2024).
16. Салмина, Н.Ю. Экономическое моделирование: учебное пособие / Н. Ю. Салмина. – Томск: Томский государственный университет систем управления и радиоэлектроники, Эль Контент, 2011. – 108 с. – URL: <https://www.iprbookshop.ru/13916.html>. (дата обращения: 11.12.2024).
17. Титов, А.Н. Визуализация данных в Python. Работа с библиотекой Matplotlib: учебно-методическое пособие / А.Н. Титов, Р.Ф. Тазиева. – Ка-

заны: КНИТУ, 2022. – 92 с. – URL: <https://e.lanbook.com/book/331025> (дата обращения: 03.04.2025).

18. Титов, А.Н. Обработка данных в Python. Основы работы с библиотекой Pandas: учебно-методическое пособие / А.Н. Титов, Р.Ф. Тазиева. – Казань: КНИТУ, 2022. – 116 с. – URL: <https://e.lanbook.com/book/331013> (дата обращения: 03.04.2025).

19. Шашкова, С.Н. Функционирование рынка труда в современном обществе: учебник / С.Н. Шашкова. – Тула: ТулГУ, 2022. – 364 с. – URL: <https://e.lanbook.com/book/331112> (дата обращения: 03.04.2025).

20. Kaggle: Платформа по исследованию данных [Электронный ресурс]: – 01.02.2025. – Режим доступа: <https://www.kaggle.com/datasets/adilshamim8/economic-indicators-and-inflation/data>. (дата обращения: 03.04.2025).

ПРИЛОЖЕНИЕ А

Прогноз ВВП и инфляции на основе модели SARIMAX

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.statespace.sarimax import SARIMAX
from sklearn.metrics import mean_absolute_error, mean_squared_error
import warnings

warnings.filterwarnings('ignore')

data = pd.read_csv("data.csv", parse_dates=["Date"], dayfirst=True)
future = pd.read_csv("prognozresult.csv", parse_dates=["Date"], dayfirst=True)
data.set_index("Date", inplace=True)
future.set_index("Date", inplace=True)

full = pd.concat([data, future])
full = full.replace([np.inf, -np.inf], np.nan).ffill().dropna()

experiments = [
    [], # Без экзогенных переменных
    ['OIL'],
    ['OIL', 'USDRUB'],
    ['BM', 'WAGE'],
    ['OIL', 'USDRUB', 'BM', 'WAGE']
]

def run_sarimax(target, exog_vars):
    train = full[target][full.index < "2022-01-01"]
    test = full[target][full.index >= "2022-01-01"]
    if exog_vars:
        exog_train = full[exog_vars][full.index < "2022-01-01"]
        exog_test = full[exog_vars][full.index >= "2022-01-01"]
    else:
        exog_train = None
        exog_test = None
    model = SARIMAX(train, order=(1,1,1), seasonal_order=(1,1,1,4), exog=exog_train)
    result = model.fit(dispatch=False)
    forecast = result.predict(start=test.index[0], end=test.index[-1], exog=exog_test)
    mae = mean_absolute_error(test, forecast)
    rmse = np.sqrt(mean_squared_error(test, forecast))
    return test, forecast, mae, rmse

print("Начинаю эксперименты...\n")

results = []

for i, exog_vars in enumerate(experiments, 1):
```

```

var_str = ', '.join(exog_vars) if exog_vars else "Без экзогенных переменных"
print(f"\n== Эксперимент {i}: {var_str} ==")

gdp_test, gdp_forecast, gdp_mae, gdp_rmse = run_sarimax('GDP', exog_vars)
inf_test, inf_forecast, inf_mae, inf_rmse = run_sarimax('Inflation', exog_vars)

print(f"[ВВП]      MAE: {gdp_mae:.2f}, RMSE: {gdp_rmse:.2f}")
print(f"[Инфляция] MAE: {inf_mae:.2f}, RMSE: {inf_rmse:.2f}")

results.append({
    "Эксперимент": i,
    "Экзогенные": ', '.join(exog_vars) if exog_vars else "Без экзогенных",
    "MAE ВВП": gdp_mae,
    "RMSE ВВП": gdp_rmse,
    "MAE инфляция": inf_mae,
    "RMSE инфляция": inf_rmse
})

# График ВВП
plt.figure(figsize=(10, 4))
plt.plot(gdp_test.index, gdp_test, label="Факт ВВП", marker='o')
plt.plot(gdp_test.index, gdp_forecast, label="SARIMAX прогноз", linestyle='--',
marker='x')
plt.title(f"SARIMAX: ВВП (Экзогенные: {var_str})")
plt.xlabel("Дата")
plt.ylabel("ВВП")
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

# График инфляции
plt.figure(figsize=(10, 4))
plt.plot(inf_test.index, inf_test, label="Факт инфляции", marker='o')
plt.plot(inf_test.index, inf_forecast, label="SARIMAX прогноз", linestyle='--',
marker='x')
plt.title(f"SARIMAX: Инфляция (Экзогенные: {var_str})")
plt.xlabel("Дата")
plt.ylabel("Инфляция")
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

print("\nИтоговая таблица по всем экспериментам:\n")
df_results = pd.DataFrame(results)
print(df_results.to_string(index=False))

```

ПРИЛОЖЕНИЕ Б

Модель SARIMAX с варьированием переменной OIL

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.statespace.sarimax import SARIMAX
from sklearn.metrics import mean_absolute_error, mean_squared_error
import warnings

warnings.filterwarnings('ignore')

# 1. Загрузка данных

data = pd.read_csv("data.csv", parse_dates=["Date"], dayfirst=True)
future = pd.read_csv("prognozresult.csv", parse_dates=["Date"], dayfirst=True)
data.set_index("Date", inplace=True)
future.set_index("Date", inplace=True)

# 2. Объединение

full = pd.concat([data, future])
full = full.replace([np.inf, -np.inf], np.nan).ffill().dropna()

# 3. Подготовка экзогенных (OIL только до 2021, далее фиксируем последним)
key_var = 'OIL'
exog_vars = ['OIL'] # Тестируем только нефть как фактор

# Получим последнее значение нефти до 2022 года (последнее реальное наблюдение)
last_oil_value = full.loc[full.index < "2022-01-01", 'OIL'].iloc[-1]

# Сценарии анализа
change_values = [0.9, 1.1] # -10% и +10%

for factor in change_values:
    print(f"\n=== Анализ: {key_var} x {factor:.1f} в 2021 году ===")
    full_mod = full.copy()

    # Модифицируем только OIL до 2022 года (в тесте она всё равно зафиксирована)
    full_mod.loc[full_mod.index < "2022-01-01", key_var] *= factor

    # Обеспечиваем экзогенные для теста (фиксируем OIL на последнем значении)
    exog_train = full_mod[[key_var]].loc[full_mod.index < "2022-01-01"]
    exog_test = full_mod[[key_var]].loc[full_mod.index >= "2022-01-01"].copy()

    # Если в тесте есть пропуски по нефти – заполняем последним известным значением
    exog_test[key_var] = exog_test[key_var].fillna(last_oil_value)

    # Целевые
    gdp_train = full_mod['GDP'].loc[full_mod.index < "2022-01-01"]
    gdp_test = full_mod['GDP'].loc[full_mod.index >= "2022-01-01"]
    inf_train = full_mod['Inflation'].loc[full_mod.index < "2022-01-01"]
```

```

inf_test = full_mod['Inflation'].loc[full_mod.index >= "2022-01-01"]

# ВВП
model_gdp = SARIMAX(gdp_train, order=(1,1,1), seasonal_order=(1,1,1,4), ex-
og=exog_train)
result_gdp = model_gdp.fit(dispatch=False)
gdp_forecast = result_gdp.predict(start=gdp_test.index[0], end=gdp_test.index[-
1], exog=exog_test)
gdp_mae = mean_absolute_error(gdp_test, gdp_forecast)
gdp_rmse = np.sqrt(mean_squared_error(gdp_test, gdp_forecast))

# Инфляция
model_inf = SARIMAX(inf_train, order=(1,1,1), seasonal_order=(1,1,1,4), ex-
og=exog_train)
result_inf = model_inf.fit(dispatch=False)
inf_forecast = result_inf.predict(start=inf_test.index[0], end=inf_test.index[-
1], exog=exog_test)
inf_mae = mean_absolute_error(inf_test, inf_forecast)
inf_rmse = np.sqrt(mean_squared_error(inf_test, inf_forecast))

print(f"[ВВП]      MAE: {gdp_mae:.2f}, RMSE: {gdp_rmse:.2f}")
print(f"[Инфляция] MAE: {inf_mae:.2f}, RMSE: {inf_rmse:.2f}")

# График ВВП
plt.figure(figsize=(10, 4))
plt.plot(gdp_test.index, gdp_test, label="Факт ВВП", marker='o')
plt.plot(gdp_test.index, gdp_forecast, label=f"SARIMAX прогноз (OIL x {fac-
tor:.1f})", linestyle='--', marker='x')
plt.title(f"SARIMAX: ВВП, OIL x {factor:.1f}")
plt.xlabel("Дата")
plt.ylabel("ВВП")
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

# График инфляции
plt.figure(figsize=(10, 4))
plt.plot(inf_test.index, inf_test, label="Факт инфляции", marker='o')
plt.plot(inf_test.index, inf_forecast, label=f"SARIMAX прогноз (OIL x {fac-
tor:.1f})", linestyle='--', marker='x')
plt.title(f"SARIMAX: Инфляция, OIL x {factor:.1f}")
plt.xlabel("Дата")
plt.ylabel("Инфляция")
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

print(f"Чувствительность SARIMAX к изменению {key_var} x {factor:.1f}
построена.\n" + "-"*65)

```

ПРИЛОЖЕНИЕ В

Прогноз ВВП и инфляции на основе модели VAR

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.api import VAR
from sklearn.metrics import mean_absolute_error, mean_squared_error
import warnings

warnings.filterwarnings('ignore')

# 1. Загрузка данных

data = pd.read_csv("data.csv", parse_dates=["Date"], dayfirst=True)
future = pd.read_csv("prognozresult.csv", parse_dates=["Date"], dayfirst=True)
data.set_index("Date", inplace=True)
future.set_index("Date", inplace=True)

# 2. Объединение

full = pd.concat([data, future])
full = full.replace([np.inf, -np.inf], np.nan).ffill().dropna()

# 3. Эксперименты: только экзогенные переменные (GDP и Inflation всегда есть)
exog_variants = [
    [], # Без экзогенных
    ['OIL'],
    ['OIL', 'USDRUB'],
    ['BM', 'WAGE'],
    ['OIL', 'USDRUB', 'BM', 'WAGE']
]

print("Начинаю эксперименты...\n")
results = []

for i, exog in enumerate(exog_variants, 1):
    all_vars = ['GDP', 'Inflation'] + exog
    exog_str = "Без экзогенных переменных" if not exog else "Экзогенные – " + ", ".join(exog)
    print(f"\n=== Эксперимент {i}: {exog_str} ===")
    full_data = full[all_vars].replace([np.inf, -np.inf], np.nan).ffill().dropna()

    train = full_data[full_data.index < "2022-01-01"]
    test = full_data[full_data.index >= "2022-01-01"]
    n_forecast = len(test)

    try:
        model = VAR(train)
        lag_order = model.select_order(maxlags=8).aic
        var_result = model.fit(lag_order)
        forecast = var_result.forecast(train.values[-lag_order:], steps=n_forecast)
        forecast_df = pd.DataFrame(forecast, index=test.index, columns=train.columns)

        # Метрики
        gdp_mae = mean_absolute_error(test['GDP'], forecast_df['GDP'])
        gdp_rmse = np.sqrt(mean_squared_error(test['GDP'], forecast_df['GDP']))
        inf_mae = mean_absolute_error(test['Inflation'], forecast_df['Inflation'])
        inf_rmse = np.sqrt(mean_squared_error(test['Inflation'], forecast_df['Inflation']))
```

```

print(f"MAE по ВВП: {gdp_mae:.2f}, RMSE: {gdp_rmse:.2f}")
print(f"MAE по инфляции: {inf_mae:.2f}, RMSE: {inf_rmse:.2f}")

results.append({
    "Эксперимент": i,
    "Экзогенные": ', '.join(exog) if exog else "Без экзогенных",
    "MAE ВВП": gdp_mae,
    "RMSE ВВП": gdp_rmse,
    "MAE инфляция": inf_mae,
    "RMSE инфляция": inf_rmse
})

# График ВВП
plt.figure(figsize=(10, 4))
plt.plot(test.index, test['GDP'], label='Факт ВВП', marker='o')
plt.plot(forecast_df.index, forecast_df['GDP'], label='VAR прогноз', lin-
estyle='--', marker='x')
plt.title(f"VAR: ВВП ({exog_str})")
plt.xlabel("Дата")
plt.ylabel("ВВП")
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

# График инфляции
plt.figure(figsize=(10, 4))
plt.plot(test.index, test['Inflation'], label='Факт инфляции', marker='o')
plt.plot(forecast_df.index, forecast_df['Inflation'], label='VAR прогноз',
linestyle='--', marker='x')
plt.title(f"VAR: Инфляция ({exog_str})")
plt.xlabel("Дата")
plt.ylabel("Инфляция")
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

print("-" * 65)

except Exception as e:
    print(f"Ошибка для набора {all_vars}: {e}")
    print("-" * 65)

# Итоговая таблица
print("\nИтоговая таблица по всем экспериментам:\n")
df_results = pd.DataFrame(results)
print(df_results.to_string(index=False))

```

ПРИЛОЖЕНИЕ Г

Модель VAR с варьированием переменной OIL

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.api import VAR
from sklearn.metrics import mean_absolute_error, mean_squared_error
import warnings

warnings.filterwarnings('ignore')

# Загрузка данных
data = pd.read_csv("data.csv", parse_dates=["Date"], dayfirst=True)
future = pd.read_csv("prognozresult.csv", parse_dates=["Date"], dayfirst=True)
data.set_index("Date", inplace=True)
future.set_index("Date", inplace=True)

# Объединение данных
full = pd.concat([data, future])
full = full.replace([np.inf, -np.inf], np.nan).ffill().dropna()

# Эксперимент: анализ чувствительности к варьированию OIL (+10%, -10%)
oil_var_scenarios = {
    "OIL +10%": 1.10,
    "OIL -10%": 0.90
}

exog_vars = ['GDP', 'Inflation', 'OIL'] # Можешь добавить сюда другие, если делаешь
более сложную систему

for scenario_name, oil_factor in oil_var_scenarios.items():
    print(f"\n=== Анализ чувствительности VAR: {scenario_name} ===")

    # Копируем данные и изменяем только обучающий период
    df_mod = full.copy()
    mask_train = df_mod.index < "2022-01-01"
    df_mod.loc[mask_train, 'OIL'] = df_mod.loc[mask_train, 'OIL'] * oil_factor

    full_data = df_mod[exog_vars].replace([np.inf, -np.inf], np.nan).ffill().dropna()
    train = full_data[full_data.index < "2022-01-01"]
    test = full_data[full_data.index >= "2022-01-01"]
    n_forecast = len(test)

    try:
        model = VAR(train)
        lag_order = model.select_order(maxlags=8).aic
        var_result = model.fit(lag_order)
        forecast = var_result.forecast(train.values[-lag_order:], steps=n_forecast)
        forecast_df = pd.DataFrame(forecast, index=test.index, columns=train.columns)

        # Метрики
        gdp_mae = mean_absolute_error(test['GDP'], forecast_df['GDP'])
        gdp_rmse = np.sqrt(mean_squared_error(test['GDP'], forecast_df['GDP']))
        inf_mae = mean_absolute_error(test['Inflation'], forecast_df['Inflation'])
        inf_rmse = np.sqrt(mean_squared_error(test['Inflation'], forecast_df['Inflation']))

        print(f"MAE по ВВП: {gdp_mae:.2f}, RMSE: {gdp_rmse:.2f}")
        print(f"MAE по инфляции: {inf_mae:.2f}, RMSE: {inf_rmse:.2f}")
```

```

# График ВВП
plt.figure(figsize=(10, 4))
plt.plot(test.index, test['GDP'], label='Факт ВВП', marker='o')
plt.plot(forecast_df.index, forecast_df['GDP'], label=f'VAR прогноз ({scenario_
io_name})', linestyle='--', marker='x')
plt.title(f"VAR: ВВП ({scenario_name})")
plt.xlabel("Дата")
plt.ylabel("ВВП")
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

# График инфляции
plt.figure(figsize=(10, 4))
plt.plot(test.index, test['Inflation'], label='Факт инфляции', marker='o')
plt.plot(forecast_df.index, forecast_df['Inflation'], label=f'VAR прогноз
({scenario_name})', linestyle='--', marker='x')
plt.title(f"VAR: Инфляция ({scenario_name})")
plt.xlabel("Дата")
plt.ylabel("Инфляция")
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

print("Графики построены. Сравните визуально и по метрикам.")
print("-" * 65)
except Exception as e:
    print(f"Ошибка для сценария {scenario_name}: {e}")
    print("-" * 65)

print("\nВсе сценарии чувствительности VAR по OIL завершены!")

```