МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра математического моделирования и анализа данных

БАЖАНОВА

Надежда Дмитриевна

КРАТКОСРОЧНОЕ ПРОГНОЗИРОВАНИЕ И НАУКАСТИНГ МАКРОЭКОНОМИЧЕСКИХ ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ РЕГРЕССИОННЫХ МОДЕЛЕЙ MIDAS

Дипломная работа

Научный руководитель:
зав. кафедрой ММАД,
доктор экономических наук,
профессор В. И. Малюгин

Допущен к защите
«___» _____ 2025 г.
Зав. кафедрой математического моделирования и анализа данных
доктор экономических наук, профессор В. И. Малюгин

Минск, 2025

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет прикладной математики и информатики
Кафедра математического моделирования и анализа данных

УТВЕРЖДАЮ
Заведующий кафедрой
Малюгин В. И.
(подпись) (фамилия, инициалы)

14.11.2024 г.

ЗАДАНИЕ
на дипломную работу

Обучающемуся Бажанова Н. Д.
Курс 4 Учебная группа 7
Специальность Прикладная математика.

Тема дипломной работы: Краткосрочное прогнозирование и наукастинг макроэкономических временных рядов на основе регрессионных моделей MIDAS
Утверждена приказом ректора БГУ от 12.11.2024 № 1295-ПС

Исходные данные к дипломной работе:

1. Малюгин В. И. Краткосрочное прогнозирование и наукастинг темпов роста инфляции на основе моделей по смешанным данным — Журнал «Банковский вестник» №1/726
2. Макеева, Н.М. Наукастинг элементов использования ВВП России / Н.М. Макеева, И.П. Станкевич – Статья 2022/10, Экономический журнал ВШЭ.
3. Макеева Н.М., Станкевич И. П. Наукастинг макроэкономических показателей экономики России на основе анализа новостного фона и регулярных данных Росстата

Перечень подлежащих разработке вопросов или краткое содержание расчетно-пояснительной записки:

1. Подготовить аналитический обзор эконометрических моделей и методов машинного обучения на основе данных смешанной частоты, а также алгоритмов комбинирования и оценки точности модельных прогнозов.
2. Сформировать временные ряды экономических показателей и провести их предварительный статистический анализ с целю использования в рассматриваемых эконометрических моделях.
3. Разработать альтернативные варианты моделей для краткосрочного прогнозирования и наукастинга квартальных темпов роста реального ВВП Республики Беларусь по экономическим переменным с месячной частотой наблюдения на основе регрессионной модели MIDAS по данным смешанной частоты,

различающиеся составом используемых экзогенные переменных и методом оценивания.
4. Разработать программный инструментарий, реализующий алгоритмы построения и оценки точности вневыборочных прогнозов на основе одиночных моделей и различных способов комбинирования моделей.
5. Провести сравнительный анализ точности прогнозов на основе построенных моделей по смешанным и агрегированным данным.

Примерный календарный график выполнения дипломной работы:

– **февраль (1-ая неделя)** – получение задания, ознакомление с условиями работы, изучение основных теоретических вопросов;

– **февраль (2-3-я неделя)** – описание временных рядов для моделирования годовых темпов прироста ВВП и проведение статистического анализа;

– **март (4-5-ая неделя)** – обоснование выбора спецификации моделей MIDAS и их построение;

– **март (6-7-ая неделя)** – написание программы с реализации метода расширяющегося окна для тестирования прогностических способностей моделей;

– **апрель (8-9-ая неделя)** – проведение сравнительного анализа прогностических способностей построенных моделей;

– **апрель (10-11-ая неделя)** – описание, оформление теоретической части работы;

– **май (12-15-ая неделя)** – подготовка презентации, прохождение системы «Антиплагиат», получение допуска, рецензирование дипломной работы.

Дата выдачи задания 14.11.2024 г.
Срок сдачи законченной дипломной работы 27.05.2025 г.

Руководитель дипломной работы

_____ В. И. Малюгин
(подпись) (инициалы, фамилия)

Подпись обучающегося _____

Дата 14.11.2024 г.

Проинформирован о недопустимости привлечения третьих лиц к выполнению дипломной работы, плагиата, фальсификации или подлога материалов.

_____ Н. Д. Бажанова
(подпись) (инициалы, фамилия обучающегося)

ОГЛАВЛЕНИЕ

РЕФЕРАТ **4**

ВВЕДЕНИЕ **7**

1 МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ РАССМАТРИВАЕМЫХ МОДЕЛЕЙ И АЛГОРИТМОВ **9**

1.1 Временные ряды и лаговый оператор 9

1.2 Модели с распределенным лагом и авторегрессии с распределенным лагом 10

1.2.1 Модели с распределенным запаздыванием 10

1.2.2 Модели авторегрессии с распределенным запаздыванием 11

1.3 Описание MIDAS моделей 12

1.3.1 Базовые MIDAS модели 12

1.3.2 Линейные модели с регуляризацией 16

1.3.3 U-MIDAS модели 17

1.3.4 MIDAS-GETS и MIDAS-GETSIS модели 18

1.4 Комбинирование прогнозов 18

1.4.1 Простое среднее 19

1.4.2 Взвешивание с помощью весов, вычисленных методом наименьших квадратов 19

1.4.3 Взвешивание с помощью средней квадратичной ошибки и рангов MSE 20

2 ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА И АНАЛИЗ ВРЕМЕННЫХ РЯДОВ **21**

2.1 Выбор переменных для прогноза ВВП 21

2.2 Выбор периода исследования 26

2.3 Сезонная корректировка данных 28

2.4 Тестирование стационарности 29

2.5 Корреляционный анализ 31

3 МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ТЕМПОВ ПРИРОСТА ВВП БЕЛАРУСИ **33**

3.1 Построение MIDAS моделей 34

3.1.1 Построение моделей с переменной $CESI_SA$ 34

3.1.2 Построение моделей с переменной PP_G 37

3.1.3 Построение моделей с переменной INV_G 39

3.1.4 Построение моделей с переменной RET_G 40

3.1.5 Построение моделей с переменной $AGRO_G$ 42

3.1.6 Построение моделей с переменной BLD_SA 44

3.2 Комбинирование прогнозов моделей MIDAS 47
3.3 Построение ARDL моделей 50
3.4 Сравнительный анализ точности полученных прогнозов и наукастинг 53

ЗАКЛЮЧЕНИЕ **55**

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ **56**

ПРИЛОЖЕНИЕ А **58**

ПРИЛОЖЕНИЕ Б **59**

РЕФЕРАТ

В дипломной работе 60 страниц, 28 иллюстраций, 20 таблиц, 25 источников, 2 приложения.

Ключевые слова: ВЫСОКОЧАСТОТНЫЕ ДАННЫЕ, РЕГРЕССИОННАЯ МОДЕЛЬ MIDAS, КРАТКОСРОЧНОЕ ПРОГНОЗИРОВАНИЕ И НАУКАСТИНГ РЕАЛЬНОГО ВВП, КОМБИНИРОВАННЫЕ ПРОГНОЗЫ, ЭКОНОМИКА БЕЛАРУСИ.

Объект исследования — валовый внутренний продукт Республики Беларусь по источникам использования доходов, начиная с конца 2016 года по настоящее время.

Цель работы — разработка регрессионных моделей MIDAS для краткосрочного прогнозирования и наукастинга темпов прироста реального ВВП Республики Беларусь на основе экономических показателей, доступных с месячной частотой наблюдения. Проводится сравнительный анализ точности прогнозов на основе построенных моделей по смешанным и агрегированным данным, а также на основе комбинированных прогнозов.

Методы исследования — модели по смешанным и агрегированным данным MIDAS и ARDL соответственно, методы статистического анализа, тестирования стационарности временных рядов.

Результаты работы: разработанные различные модификации модели по смешанным данным MIDAS, агрегированным данным ARDL и комбинированные прогнозы; проведён сравнительный анализ точности самостоятельных и комбинированных прогнозов исследуемого показателя.

Область применения — в настоящий момент времени полученные результаты могут быть применены в экономической политике для оценки и прогнозирования экономических тенденций. Прогнозы могут служить основой для научных исследований в области эконометрики.

Дипломная работа выполнена автором самостоятельно.

РЭФЕРАТ

У дыпломнай працы 60 старонак, 28 малюнкаў, 20 табліц, 25 крыніц, 2 прыкладання.

Ключавыя словы:ВЫСОКАЧАСТОТНЫЯ ДАДЗЕНЫЯ, РЭГРЭСІЙНАЯ МАДЭЛЬ MIDAS, КАРОТКАТЭРМІНОВАЕ ПРАГНАЗАВАННЕ І НАВУКАСТЫНГ РЭАЛЬНАГА ВУП, КАМБІНАВАНЫЯ ПРАГНОЗЫ, ЭКАНОМІКА БЕЛАРУСІ.

Аб'ект даследавання — валавы ўнутраны прадукт Рэспублікі Беларусі ў адпаведнасці з крыніцамі выкарыстання даходаў, пачынаючы з канца 2016 года да сённяшняга дня.

Мэта працы — распрацоўка рэгрэсійных мадэляў MIDAS для кароткатэрміновага прагназавання і навукастынгу тэмпаў прыросту рэальнага ВУП Рэспублікі Беларусь на аснове эканамічных паказчыкаў, даступных з месячнай частатой назірання. Праводзіцца параўнальны аналіз дакладнасці прагнозаў на аснове пабудаваных мадэляў па змяшаных і агрэгаваных дадзеных, а таксама на аснове камбінаваных прагнозаў.

Метады даследавання — мадэлі па змешаных і агрэгаваных дадзеных MIDAS і ARDL адпаведна, метады статыстычнага аналізу, тэставанні стацыянарнасці часавых шэрагаў.

Вынікі працы: распрацаваныя розныя мадыфікацыі мадэлі па змяшаных дадзеных MIDAS, агрэгаваных дадзеных ARDL і камбінаваныя прагнозы; праведзены параўнальны аналіз дакладнасці самастойных і камбінаваных прагнозаў доследнага паказчыка.

Вобласть прымяненння — у сапраўдны момант часу атрыманыя вынікі могуць быць ужытыя ў эканамічнай палітыцы для ацэнкі і прагназавання эканамічных тэндэнцый. Прагнозы могуць служыць асновай для навуковых даследаванняў у галіне эканометрыкі.

Дыпломная праца выканана аўтарам самастойна.

ABSTRACT

The thesis project is presented in the form of an explanatory note of 60 pages, 28 figures, 20 tables, 25 references, 2 appendices.

Keywords: HIGH-FREQUENCY DATA, REGRESSION MODEL MIDAS, REAL GDP SHORT-TERM FORECASTING AND NOWCASTING, COMBINED FORECASTS, BELARUSIAN ECONOMY.

The research object is the gross domestic product of the Republic of Belarus by sources of income utilization, starting from the end of 2016 to the present.

The aim of the work is to develop MIDAS regression models for short-term forecasting and nowcasting the growth rates of real GDP of the Republic of Belarus based on economic indicators available with monthly observation frequency. A comparative analysis of the accuracy of forecasts based on the constructed models using mixed and aggregated data, as well as combined forecasts, is conducted.

The research methods include MIDAS models for mixed and aggregated data and ARDL models, as well as statistical analysis techniques and tests for the stationarity of time series.

The results include the development of various modifications of MIDAS models for mixed data, ARDL models for aggregated data, and combined forecasts. A comparative analysis of the accuracy of independent and combined forecasts of the studied indicator was conducted.

The obtained results can be applied in economic policy for assessing and forecasting economic trends. The forecasts may also serve as a foundation for scientific research in the field of econometrics.

The thesis project was done solely by the author.

ВВЕДЕНИЕ

В данной работе рассматривается применение моделей по смешанным данным для анализа прогнозирования валового внутреннего продукта Республики Беларусь.

Многие фундаментальные макроэкономические показатели формируются Национальным статистическим комитетом Республики Беларусь на квартальной и годовой частоте. В то же время доступны более оперативные показатели, которые публикуются с месячной и дневной периодичностью. В частности, показатель валового внутреннего продукта (ВВП) по методу использования доходов формируется на квартальной частоте, а отраслевые показатели и статистика цен — на месячной частоте, обменные курсы валют и денежно-кредитные показатели — на дневной частоте. Причём в соответствии с календарём публикации данных квартальная оценка ВВП публикуется на 90 день после окончания отчётного квартала. В связи с этим становится актуальным вопрос об использовании более оперативной информации при прогнозирования показателя ВВП.

Все классические регрессионные модели предназначены для работы с данными одинаковой периодичности. Существуют следующие подходы для решения данной задачи:

- агрегирование (усреднение, среднее взвешенное, нормирование по долям) высокочастотных переменных;
- интерполяция низкочастотных переменных.

Однако второй способ используется крайне редко, в случае агрегации возникает потеря информации о динамике объясняющих переменных, что может привести к снижению точности моделей.

В связи с этим возникает вопрос: как можно без преобразования данных и потери какой-либо информации строить регрессионную модель для предсказания исследуемых показателей. Одним из методов работы с данными смешанной частоты, рассматриваемый в данной работе, являются MIDAS модели (mixed data sampling), впервые предложенные Ghysels, Santa-Clara и Valkanov [20]. Этот метод, изначально применявшийся в финансовых исследованиях, сейчас активно используется для прогнозирования макроэкономических временных рядов.

Цель работы — разработка регрессионных моделей MIDAS для краткосрочного прогнозирования и наукастинга темпов прироста реального ВВП Республики Беларусь на основе экономических показателей, доступных с месячной частотой наблюдения.

В ходе выполнения работы для достижения цели были поставлены следующие *задачи*:

1. подготовить аналитический обзор эконометрических моделей и методов машинного обучения на основе данных смешанной частоты, а также алгоритмов комбинирования и оценки точности модельных прогнозов;

2. сформировать временные ряды экономических показателей и провести их предварительный статистический анализ с целю использования в рассматриваемых эконометрических моделях;
3. разработать альтернативные варианты моделей для краткосрочного прогнозирования и наукастинга квартальных темпов роста реального ВВП Республики Беларусь по экономическим переменным с месячной частотой наблюдения на основе регрессионной модели MIDAS по данным смешанной частоты, различающиеся составом используемых экзогенные переменных и методом оценивания;
4. разработать программный инструментарий, реализующий алгоритмы построения и оценки точности вневыборочных прогнозов на основе одиночных моделей и различных способов комбинирования моделей;
5. провести сравнительный анализ точности прогнозов на основе построенных моделей по смешанным и агрегированным данным.

ГЛАВА 1
МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ РАССМАТРИВАЕМЫХ МОДЕЛЕЙ И АЛГОРИТМОВ

1.1 Временные ряды и лаговый оператор

• *Случайная функция* — это параметрическое семейство случайных векторов $x(t) = x(\omega, t) \in \mathbb{R}^n$, определённых на одном и том же вероятностном пространстве (Ω, F, P), $\omega \in \Omega$, где $t \in T \subset \mathbb{R}^m$ — параметр, изменяющийся на множестве T.

• *Временной ряд* — это упорядоченные во времени статистические наблюдения над одним и тем же объектом в динамике, то есть это случайная функция $x_t = x(t) = x(\omega, t) \in \mathbb{R}$, где $t \in T \subset \mathbb{R}$ — это время.

• *Отсчетом* временного ряда назовем случайное значение $x(t)$ этого временного ряда в момент времени $t \in T$. Расстояние между отсчетами по времени будем называть *лагом*. Совокупность всех отсчетов при фиксированном случайном эксперименте $\omega \in \Omega$

$$X = \{x(t) = x(\omega, t),\ t \in T\}_{\omega \in \Omega} \tag{1.1.1}$$

будем называть *реализацией* временного ряда.

Теперь введем определение, на основании которого и будут строиться все рассматриваемые нами модели.

• *Лаговый оператор* — это оператор сдвига, позволяющий получить значения элементов временного ряда на основании ряда предыдущих значений (Обозначение: L).

То есть для реализации временного ряда $X = \{x_1, x_2, \ldots, x_t\}$ лаговый оператор будет действовать следующим образом

$$L^k x_t = x_{t-k}. \tag{1.1.2}$$

Этот оператор обладает следующими свойствами:

1. $L^0 = 1$;
2. $Lc = c$, $c \in \mathbb{R}$;
3. $L^{-1} x_t = x_{t+1}$.

Наряду с лаговым оператором определяются лаговые многочлены. Например, лаговый многочлен p-ой степени может быть записан как

$$b(L) = b_0 + b_1 L + \ldots + b_p L^p. \tag{1.1.3}$$

Тогда

$$b(L)x_t = b_0x_t + b_1x_{t-1} + \ldots + b_px_{t-p}. \tag{1.1.4}$$

1.2 Модели с распределенным лагом и авторегрессии с распределенным лагом

Для начала рассмотрим случай, когда необходимо агрегировать данные. В эконометрическом анализе часто встречается ситуация, когда исследуемая переменная зависит не только от текущего значения, но и от значений в прошлом. Это может быть связано с тем, что экономические процессы имеют инерционный характер, и изменения в одной переменной могут оказывать влияние на её будущие значения. В таких случаях модели с распределенными лагами и авторегрессии с распределенными лагами становятся важными инструментами для анализа временных рядов. В данном разделе мы более подробно рассмотрим оба типа моделей, их теоретические основы.

1.2.1 Модели с распределенным запаздыванием

Рассмотрение моделей начнём с простейшей модели с распределенным запаздыванием (distributed lag, DL), поскольку модели MIDAS регрессии имеют общие черты с этими моделями. Однако модели с распределенным запаздыванием проще по своей структуре.

Введём следующие обозначения для того, чтобы сформулировать общий вид моделей с распределенным запаздыванием:

- y_t — эндогенная переменная, соответствующая прогнозируемому показателю;
- x_t — экзогенная переменная, по которой будет прогнозироваться зависимое значение;
- ε_t — белый шум, случайная величина, имеющая стандартное нормальное распределение;
- β_0 — свободный член.

Также введём полиномиальный лаговый оператор

$$b(L) = \sum_{i=0}^{p} b_iL^i \tag{1.2.1}$$

p-ой степени по формуле (1.1.3). Тогда, если мы имеем реализации временных рядов $X = \{x_1, \ldots, x_t, \ldots\}$ и $Y = \{y_1, \ldots, y_t, \ldots\}$, $t > p$, то модель с распределенным запаздыванием, или DL-модель, может быть записана в следующем виде:

$$y_t = \beta_0 + b(L)x_t + \varepsilon_t. \tag{1.2.2}$$

Эту модель можно также, используя формулы (1.2.1) и (1.1.2), переписать

в виде

$$y_t = \beta_0 + \sum_{i=0}^{p} b_i x_{t-i} + \varepsilon_t, \tag{1.2.3}$$

или, используя формулу (1.1.4), можно представить DL-модель в виде

$$y_t = \beta_0 + b_0 x_t + b_1 x_{t-1} + \ldots + b_p x_{t-p} + \varepsilon_t. \tag{1.2.4}$$

Однако построенная модель может работать лишь с данными одной частоты, поэтому для использования этой модели нам надо агрегировать объясняющие показатели x_t, чтобы они имели одну частоту с прогнозируемым показателем y_t.

Вообще говоря, DL-модель является частным случаем модели авторегрессии с распределенным запаздыванием (autoregressive distributed lag, ARDL), которую мы также рассмотрим позже.

В качестве модификации можно рассмотреть DL-модель с лагами Алмон. Оценивая модель (1.2.3), можно предположить, что коэффициент при лаговой переменной меняется в некотором смысле плавно, и приблизить его с помощью многочлена

$$b_i = \sum_{j=0}^{q} c_j i^j. \tag{1.2.5}$$

Тогда DL-модель с лагами Алмон имеет вид

$$y_t = \beta_0 + \sum_{i=0}^{p} \sum_{j=0}^{q} c_j i^j x_{t-i} + \varepsilon_t, \tag{1.2.6}$$

или

$$y_t = \beta_0 + \sum_{j=0}^{q} c_j z_j + \varepsilon_t, \quad z_j = \sum_{i=0}^{p} i^j x_{t-i}. \tag{1.2.7}$$

Также обратим внимание на тот факт, что параметры p, q подбираются вручную. Налагая различные ограничения, такие как максимальная степень максимальная степень q, начальные и конченые условия, на многочлены, можно сконструировать наиболее удовлетворительную модель. Однако такой подход оставляет место для ошибок спецификации и субъективной подгонки моделей, так как статистического способа определить необходимую форму многочлена не существует. Это может привести к выбору неподходящей модели, что повлияет на результаты анализа и сделает выводы менее надёжными.

1.2.2 Модели авторегрессии с распределенным запаздыванием

Дополним модель AR(p) заданной формулой (1.2.8) некоторыми экзогенными переменными и их лагами, например, до порядка q.

$$\sum_{i=0}^{p} \beta_i y_{t-i} = \varepsilon_t, \tag{1.2.8}$$

где $\beta_0 = 1$, а $\beta_1, \ldots, \beta_p$ — коэффициенты авторегрессии, причем $\beta_p \neq 0$. Такую модель называют открытой или же моделью авторегрессии с распределенным запаздыванием (autoregressive distributed lags, ARDL(p, q))

$$\sum_{i=0}^{p} \beta_i y_{t-i} = \sum_{j=0}^{q} \alpha_j x_{t-j} + \varepsilon_t. \tag{1.2.9}$$

Рассмотрим случай модели VAR(p) заданной формулой (1.2.10).

$$y_t = \beta_0 + \sum_{i=1}^{p} B_i y_{t-i} + \varepsilon_t, \tag{1.2.10}$$

где $B_i = (\beta_{ij})$ — матрица. Если мы добавим экзогенные переменные к модели VAR(p), то получим модель векторной авторегрессии с распределенным запаздыванием (vector autoregressive distributed lags, VARDL(p, q))

$$y_t = \beta_0 + \sum_{i=1}^{p} B_i y_{t-i} + \sum_{i=0}^{q} A_i x_{t-i} + \varepsilon_t, \tag{1.2.11}$$

где $A_i = (\alpha_{ij})$ — матрица. Эти модели, как и DL-модели, можно модифицировать, изменяя лаговые коэффициенты. Но как и в случае DL-моделей, эти модели не могут работать с данными смешанной частоты.

Таким образом, мы можем сделать вывод, что модели авторегресии позволяют нам определять значение эндогенной переменной с помощью лагов самой эндогенной переменной и лагов экзогенной переменной.

1.3 Описание MIDAS моделей

1.3.1 Базовые MIDAS модели

С целью введения базовой модели регрессии по данным смешанной частоты (mixed data sampling, MIDAS) введём обозначения для переменных. Пусть эндогенная переменная y_t имеет фиксированную частоту. Она может быть годовая, квартальная, месячная и так далее. Для конкретики возьмём квартальную частоту. Кроме того, пусть независимая переменная замерена в m раз чаще. Например, если у эндогенной переменной квартальная частота, то для экзогенной переменной с месячной частотой возьмём $m = 3$.

Таким образом, у нас имеются следующие обозначения:

- $t = 1, \ldots, T$ — единицы времени;
- y_t — эндогенная переменная, измеренная ежеквартально;

- $x_t^{(m)}$ — экзогенная переменная, измеренная ежемесячно;
- $\varepsilon_t^{(m)}$ — белый шум;
- β_0 — свободный член;
- β_1 — действительное число.

Введем полиномиальный лаговый оператор следующего вида

$$b(L^{1/m}, \Theta) = \sum_{i=0}^{p} b(i, \Theta) L^{i/m}, \tag{1.3.1}$$

где

$$L^{i/m} x_t^{(m)} = x_{(t-i)/m}^{(m)}. \tag{1.3.2}$$

Ключевую роль в результатах прогнозирования моделью MIDAS играет функция лаговых коэффициентов

$$b(i, \Theta), \quad i = \overline{0, p}. \tag{1.3.3}$$

Данную функцию можно задавать по-разному. Тривиально можно предполагать, что все лаги экзогенной переменной участвуют в модели с одинаковым весом, то есть

$$b(\Theta) = \frac{1}{p}. \tag{1.3.4}$$

Однако этот способ не является эффективным и вызывает эффект мультиколлинеарности факторов в силу большого числа лагов в модели.

Наиболее распространёнными являются следующие виды функции лаговых коэффициентов:

- экспоненциальные лаги Алмон

$$b(i, \Theta) = \frac{e^{\Theta_1 i + \cdots + \Theta_q i^q}}{\sum\limits_{j=0}^{p} e^{\Theta_1 j + \cdots + \Theta_q j^q}}, \tag{1.3.5}$$

  где значение q либо задано априорно в самой программе, либо задаётся вручную;
- бета лаги (они требуют уже два параметра Θ)

$$b(i, \Theta_1, \Theta_2) = \frac{f(\frac{i}{p}, \Theta_1, \Theta_2)}{\sum\limits_{j=0}^{p} f(\frac{j}{p}, \Theta_1, \Theta_2)}, \tag{1.3.6}$$

  где

$$f(x, \Theta_1, \Theta_2) = \frac{x^{a-1}(1-x)^{b-1}\Gamma(\Theta_1 + \Theta_2)}{\Gamma(\Theta_1)\Gamma(\Theta_2)}; \tag{1.3.7}$$

Приведём ещё примеры трёх менее распространённых способов задания этих функций:

- неэкспоненциальные лаги Алмона

$$b(i,\Theta) = \sum_{j=0}^{p} \Theta_j i^j \tag{1.3.8}$$

- гиперболический способ

$$b(i,\Theta) = \frac{g(\frac{i}{p},\Theta)}{\sum_{j=0}^{p} g(\frac{j}{K}p,\Theta)}, \quad g(x,\Theta) = \frac{\Gamma(x+\Theta)}{\Gamma(x+1)\Gamma(\Theta)}; \tag{1.3.9}$$

- геометрический способ

$$b(i,\Theta) = \frac{\Theta^i}{\sum_{j=0}^{\infty} \Theta^j}, \quad |\Theta| \leq 1. \tag{1.3.10}$$

Очевидно, что, таким образом, функцию лаговых коэффициентов можно считать гиперпараметром. Различные варианты задания этих функций будут по-разному справляться с решением задач. Фактически задание такой функции определяет способ агрегации данных высокой частоты в ряд более низкой частоты (например, данные месячной частоты в данные квартальной частоты).

Отметим, что различные ограничения для MIDAS-модели используются с целью снижения размерности данных вследствие наложения ограничения на некоторые параметры рассматриваемой модели. Снижение размерности позволяет предотвратить переобучение модели и не потерять имеющуюся вариативность для правильной подстройки под имеющийся набор данных.

В силу всех введённых обозначений, можем записать базовую модель MIDAS в следующем виде

$$y_t = \beta_0 + \beta_1 \cdot b(L^{1/m},\Theta) x_t^{(m)} + \varepsilon_t^{(m)}. \tag{1.3.11}$$

Также в силу формулы (1.3.1) можем записать это уравнение в виде

$$y_t = \beta_0 + \beta_1 \sum_{i=0}^{p} b(i,\Theta) L^{i/m} x_t^{(m)} + \varepsilon_t^{(m)}, \tag{1.3.12}$$

а в силу формулы (1.3.2)

$$y_t = \beta_0 + \beta_1 \sum_{i=0}^{p} b(i,\Theta) x_{(t-i)/m}^{(m)} + \varepsilon_t^{(m)}. \tag{1.3.13}$$

Случай многих экзогенных переменных

Мы можем усложнить структуру модели, включив в неё и два, и три, и любое другое число объясняющих факторов. Аналогично мы усложним структуру, включая в модель новые лаги. Таким образом, более общий вид модели MIDAS регрессии может быть записан как

$$y_t = \beta_0 + \beta_1 \sum_{j=0}^{q} \sum_{i=0}^{p} b_{ij}(i, \Theta) L^{i/m_j} x_t^{(m_j)} + \varepsilon_t^{(m)}. \tag{1.3.14}$$

Внимательно оценив эту структуру, можно сделать заключить, что, вообще говоря, возможно включать объясняющие факторы разных частот, поскольку каждому фактору соответствует своя собственная полиномиальная параметризация. Это позволяет решать такие задачи, как, например, задача объяснения квартальной переменной одновременно по ежемесячному объясняющему фактору и ежедневному.

Наиболее распространенным является частный случай базовой MIDAS модели с двумя экзогенными переменными

$$y_t = \beta_0 + \beta_1 \sum_{i=0}^{p} b_{i1}(i, \Theta) L^{i/m_1} x_t^{(m_1)} + \beta_1 \sum_{i=0}^{p} b_{i2}(i, \Theta) L^{i/m_2} x_t^{(m_2)} + \varepsilon_t^{(m)}. \tag{1.3.15}$$

Нелинейный случай

В работах Ghysels, Sinko и Valkanov также представлен и нелинейный случай MIDAS модели многих экзогенных переменных в виде

$$y_t = \beta_0 + f\Big(\sum_{j=1}^{q} \sum_{i=1}^{p} b_{ij}(L^{i/m_j}, \Theta) g\big(x_t^{(m_j)}\big) \Big) + \varepsilon_t^{(m)}, \tag{1.3.16}$$

где функции f и g могут быть целиком известны или зависеть от параметров. Такая модель может быть полезна, особенно в приложениях с волатильностью и исследованиях соотношения риска и доходности [19].

Многомерный случай

Мы можем продолжать обобщать модель многих экзогенных переменных (1.3.14) и ещё больше усложнить ее структуру. Пусть у нас теперь $\mathcal{B}_{ij}$ — это матрицы полиномов размерности $n \times n$, $\mathcal{B}_0$ — это n-мерный вектор, а Y_t, ε_t и X_t — это n-мерные векторные процессы. Тогда многомерную модель MIDAS регрессии с многими объясняющими факторами мы можем записать в виде

$$Y_t = \mathcal{B}_0 + \sum_{i=1}^{p} \sum_{j=1}^{q} \mathcal{B}_{ij}(i, \Theta) L^{i/m_j} X_t^{(m_j)} + \varepsilon_t^{(m)}. \tag{1.3.17}$$

Основная проблема заключается в том, как справиться с распространением параметров в многомерном контексте. Один из подходов заключается в рассмотрении всех недиагональных элементов, соответствующих одному многочлену, в то время как диагональные элементы — второму. Ограничения могут быть недействительными и будут выбраны в зависимости от приложения. Рассмотрение многомерных регрессий MIDAS позволяет решить проблемы причинно-следственной связи Грейнджера, избегая ошибок временной агрегации, которые могут маскировать или создавать ложные долговые расписки.

Однако при рассмотрении всех последующих модификаций MIDAS модели мы все же остановимся на одномерном случае с одним объясняющим фактором и с линейной структурой.

1.3.2 Линейные модели с регуляризацией

В качестве подхода, сохраняющего идею исключения «лишних» переменных, но более свойственного машинному обучению, рассматриваются линейные модели с регуляризацией. Мы рассмотрим модели с L1 регуляризацией (LASSO регрессии), построенные на базе MIDAS модели. На практике при применении MIDAS моделей, как правило, используется только одна объясняющая переменная, что обуславливается небольшим объёмом данных и необходимостью целого набора дополнительных регрессоров при введении в модель одной новой объясняющей переменной. Так, при использовании месячных данных для объяснения квартальных — это три дополнительных регрессора, при использовании данных более высокой частоты число регрессоров возрастает. Регуляризация позволяет снизить остроту этой проблемы и использовать наборы из нескольких объясняющих переменных в рамках одной модели.

LASSO-MIDAS модель может быть записана как

$$y_t = \beta_0 + \sum_{i=0}^{p} b_i x^{(m)}_{(t-i)/m} + \varepsilon^{(m)}_t. \tag{1.3.18}$$

При этом функция потерь (целевая функция) записывается как

$$\text{Loss} = \sum_{t=1}^{T} (y_t - \widehat{y}_t)^2 + \lambda \sum_{i=0}^{p} |b_i| \to \min, \tag{1.3.19}$$

то есть решается задача минимизации этого функционала. Параметр регуляризации λ подбирается с помощью кросс-валидации.

Одним из самых серьёзных ограничений стандартной LASSO регрессии является отсутствие у неё oracle property (способность модели корректно отбирать и состоятельно оценивать ненулевые коэффициенты). В качестве решения этой проблемы предлагается использовать адаптивное LASSO, которое обладает oracle property. Адаптивное LASSO подразумевает использование весов в функции потерь

$$\text{Loss} = \sum_{t=1}^{T}(y_t - \widehat{y}_t)^2 + \lambda \sum_{i=0}^{p} \omega_i |b_i| \to \min, \tag{1.3.20}$$

где ω_i — веса, полученные при помощи состоятельных оценок коэффициентов (например, МНК-оценки).

1.3.3 U-MIDAS модели

Менее распространённым вариантом моделей MIDAS регрессии является неограниченный MIDAS, или U-MIDAS (unrestricted mixed data sampling). В отличие от базовой MIDAS модели, в U-MIDAS модели не накладываются ограничения на полиномиальный лаговый оператор, то есть такой подход не прибегает к функциональным многочленам с распределённым запаздыванием. В своей работе Foroni, Marcellino и Schumacher описали процесс построения данной модели, которая принимает следующий вид [18]:

$$c(L^k)\omega(L)y_t = \sum_{j=1}^{K} \delta_j(L) x_{j,t}^{(m)} + \varepsilon_t^{(m)}, \tag{1.3.21}$$

где заданы следующие лаговые операторы $c(L^k) = 1 - c_1 L^k - \ldots - c_s L^{ks}$, $\omega(L) = \omega_0 + \omega_1 L + \ldots + \omega_{k-1} L^{k-1}$, $\delta_j(L) = \delta_{j,0} + \delta_{j,1} L + \ldots + \delta_{j,v} L^v$.

Обратим внимание, что если мы предположим, что порядки запаздывания s и v достаточно велики, чтобы сделать член ошибки $\varepsilon_t^{(m)}$ некоррелированным, то все параметры в модели U-MIDAS могут быть оценены простым методом наименьших квадратов.

Если мы упростим модель, задав $c(L^k) = 1$, $N = 1$, $\omega(L) = 1$ в формуле (1.3.21), то она будет задаваться формулой

$$y_t = \delta_1(L) x_t^{(m)} + \varepsilon_t, \tag{1.3.22}$$

что как раз и представляет собой модель MIDAS регрессии без ограничений на лаговый многочлен.

Следовательно, можем заключить, что базовая модель MIDAS — это частный случай U-MIDAS модели, поскольку она получена путём наложения определённых ограничений. Ключевым преимуществом базовой MIDAS модели является то, что она допускает длинные лаги при ограниченном количестве параметров, что может быть особенно полезно в финансовых приложениях с большим несоответствием между частотами выборки y и x. Например, когда y является ежемесячной переменной, а x — ежедневной. Однако для макроэкономических приложений с небольшими различиями в частотах выборки, например, для ежемесячных и квартальных данных, базовая модель может иметь определённые недостатки. Например, базовая модель сильно нелинейна по параметрам, так что она не может быть оценена с помощью OLS. В целом предполагается, что U-MIDAS модель должна работать лучше, чем базовая

MIDAS модель, пока частота агрегации мала и U-MIDAS модель не слишком сильно параметризована.

1.3.4 MIDAS-GETS и MIDAS-GETSIS модели

В данных моделях реализуется подход «от общего к частному» [14]. Повторяющийся алгоритм поочерёдно из модели убирает по одной переменной и проверяет на значимость с помощью диагностических тестов (нормальность, автокорреляцию остатков и т. д.). В случае незначимости переменные удаляются и формируются итоговые комбинации модели. После расчёта всех различных сочетаний, итоговые моде-ли сравниваются с помощью выбора информационного критерия. Затем выбирается наилучшая модель. В случае модели MIDAS-GETSIS, если в анализируемых данных имеются выборы и структурные разрывы моделирование вида «от общего к частному» с использованием индикаторов насыщения позволяет оценивать показатели, разделяя их на определённые выборки и добавляя фиктивные переменные.

1.4 Комбинирование прогнозов

В практике экономического прогнозирования аналитики часто располагают множеством моделей и соответствующих прогнозов одной и той же переменной. Эти прогнозы могут основываться на различных предпосылках и использовать разную информацию. Традиционно выбор прогнозов сводился к определению «лучшего» из имеющегося.

Однако стоит заметить, что многочисленные исследования, включая работы Тиммерманна (2006), Клемена (1989) и Макридакиса с Хибоном (2000), продемонстрировали, что объединение нескольких прогнозов может обеспечить большую точность, чем выбор единственного прогноза. Комбинирование прогнозов представляет собой метод, направленный на создание единого прогноза на основе нескольких исходных оценок. Вместо того чтобы просто оценивать качество каждого прогноза и отбирать лучший, комбинирование использует методы оценки для вычисления весов, которые применяются к прогнозам для получения средневзвешенного значения.

Методы комбинирования прогнозов обычно основываются на расчёте средневзвешенной величины: для каждой выборки прогнозов определяется среднее значение кросс-прогноза, где разным прогнозам присваиваются различные веса. Также является возможным применение более сложных схем взвешивания, которые позволяют весам изменятся в процессе прогнозирования. Однако мы рассмотрим только постоянные веса, включая методы простого среднего, наименьшего квадратов, средней квадратичной ошибки, рангов среднеквадратичной ошибки. В данных случаях объединённый прогноз

F_{t+h} определяется как взвешенная сумма отдельных прогнозов

$$F_{t+h} = \sum_{i=1}^{n} w_i \widetilde{y}_{t+h,i}, \tag{1.4.1}$$

где w_i — вес i-го прогноза, n — число прогнозов в массиве, $\widetilde{y}_{t+h,i}$ — прогноз от i-ой модели.

1.4.1 Простое среднее

Метод простого среднего представляет собой один из наиболее распространённых способов комбинирования прогнозов, который основывается на вычислении арифметического среднего значений прогнозов для каждого наблюдения. В этом подходе все прогнозы получают одинаковые веса, что делает его простым в применении и понимании. Комбинированный прогноз F_{t+h}, созданный в момент времени t для h шагов вперёд, можно записать следующим видом:

$$F_{t+h} = \frac{1}{n} \sum_{i=1}^{n} \widetilde{y}_{t+h,i}, \tag{1.4.2}$$

где $\widetilde{y}_{t+h,i}$ обозначает прогноз от i-ой модели на момент $t + h$.

Данный метод имеет несколько значительных преимуществ. Он отличается простотой реализации и требует минимальных вычислительных ресурсов. Использование равных весов позволяет избежать чрезмерного влияния отдельных прогнозов, что особенно важно в условиях неопределённости. Исследования показывают, что в ряде случаев простое среднее превосходит более сложные методы, такие как метод наименьших квадратов и другие статистические методы [22]. Это может быть объяснено тем, что простое усреднение помогает снизить дисперсию ошибок прогноза, поскольку веса не зависят от вторых моментов, что является основной многих более сложных методов [15].

Однако метод простого среднего не лишён недостатков. Он не учитывает историческую динамику прогнозов, что может привести к потере важной информации, и полагается на предположение о равнозначности всех прогнозов, что не всегда соответствует действительности. Кроме того, метод эффективен только при условии нормальности распределения ошибок. Если это предположение нарушается, линейные методы комбинирования могут давать неудовлетворительные результаты.

1.4.2 Взвешивание с помощью весов, вычисленных методом наименьших квадратов

Метод наименьших квадратов (МНК) представляет собой подход к вычислению весов для комбинирования прогнозов, который основывается на регрессии прогнозов против фактических значений. Для его применения необходимо знать истинные значения прогнозируемой для некоторых периодов прогноза. роцедура вычисления весовых коэффициентов подразумевает со-

здание линейной регрессионной модели, где целевой переменной выступают реальные значения. Полученные коэффициенты этой регрессии и используются как веса. Формула линейной регрессии выглядит следующим образом:

$$y_t = \alpha + \sum_{i=1}^{n} w_i \widetilde{y}_{t,i} + \varepsilon_t, \tag{1.4.3}$$

где y_t — фактическое значение, $\widetilde{y}_{t,i}$ — прогноз от i-ой модели, α — свободный член, а ε_t — белый шум. В результате регрессионного анализа, найденные значения w_i становятся весами, определяющими значимость каждого отдельного прогноза в финальном скомбинированном результате.

1.4.3 Взвешивание с помощью средней квадратичной ошибки и рангов MSE

Подход, разработанный Стокком и Уотсоном (2001) и использующий метод весов средней квадратичной ошибки (MSE), предполагает сопоставление индивидуальных прогнозов с наблюдаемыми данными в течение заданного временного интервала. На основе этого сравнения для каждого прогноза вычисляется средняя квадратичная ошибка, которая служит основой для определения весов данных прогнозов:

$$w_i = \frac{\dfrac{1}{\mathrm{MSE}_i^k}}{\sum\limits_{j=1}^{n} \dfrac{1}{\mathrm{MSE}_j^k}}, \tag{1.4.4}$$

где k используется для возведения MSE в различные степени. Значение k может быть различным, но наиболее распространенным значением является $k = 1$.

Метод рангов MSE Айолфи и Тиммерманна (2006) похож на метод весов MSE, описанный выше, но вместо вычисления отношения значений MSE этот метод вычисляет MSE каждого прогноза, ранжирует их, а затем вычисляет отношение обратных рангов.

Вес w_i для i-го прогноза в этом методе определяется как отношение значения ранга к сумме всех рангов:

$$w_i = \frac{r_i}{\sum\limits_{j=1}^{n} r_j}, \tag{1.4.5}$$

где r_i — ранг i-го прогноза. Этот метод даёт возможность оценить достоверность предсказаний, поскольку прогнозы с большей точностью получают более высокий ранг и, как следствие, больший вес.

ГЛАВА 2
ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА И АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

2.1 Выбор переменных для прогноза ВВП

Валовый внутренний продукт (ВВП) является ключевым индикатором экономической активности, отражающим общую стоимость товаров и услуг, произведённых в стране. Понимание его структуры и факторов, влияющих на его формирование, играет одну из ключевых значений для создания эффективных моделей для прогнозирования. В данном разделе мы проанализируем вклад различных отраслей экономики в ВВП с целью дальнейшего выявления объясняющих факторов, которые будут применятся в процессе построения моделей в качестве экзогенных переменных высокой частоты.

На практике применяются три основных метода расчёта ВВП [6]:

1. ***производственный метод.*** При расчёте производственным методом ВВП формируется как сумма валовой добавленной стоимости по видам экономической деятельности и чистых (за вычетом субсидий) налогов на продукты;
2. ***по использованию доходов.*** ВВП, рассчитанный данным методом представляет собой сумму расходов на конечное потребление домашних хозяйств, государственных организаций, некоммерческих организаций, обслуживающих домашние хозяйства, валового накопления и чистого экспорта товаров и услуг;
3. ***по источникам доходов.*** Метод формирования ВВП по источникам доходов складывается из оплаты труда наёмных работников, валовой прибыли всех отраслей и чистых налогов на производство и импорт.

Анализ вкладов компонентов, входящих в состав валового внутреннего продукта, позволяет выявить ключевые сектора экономики, которые оказывают значительное влияние на его формирование и развитие. Рассмотрим подробнее доли компонентов, используя два метода: метод использования доходов и производственный метод. При этом акцентируем внимание на тех секторах, которые составляют 5% от общего объёма ВВП, так как они могут быть особенно значимыми для анализа экономической активности.

На рисунках 2.1 и 2.2 представлены круговые диаграммы, которые наглядно отображают доли различных отраслей экономики. Эти диаграммы иллюстрируют, как распределяются доли отраслей, используемых в производственном методе, направлений расходов, применяемых в методе по использованию доходов, в структуре валового внутреннего продукта. Эти диаграммы основаны на данных, собранных за период с 2020 по 2023 годы и представляют собой средние значения, что позволяет лучше понять структуру экономики в

этот временной промежуток.

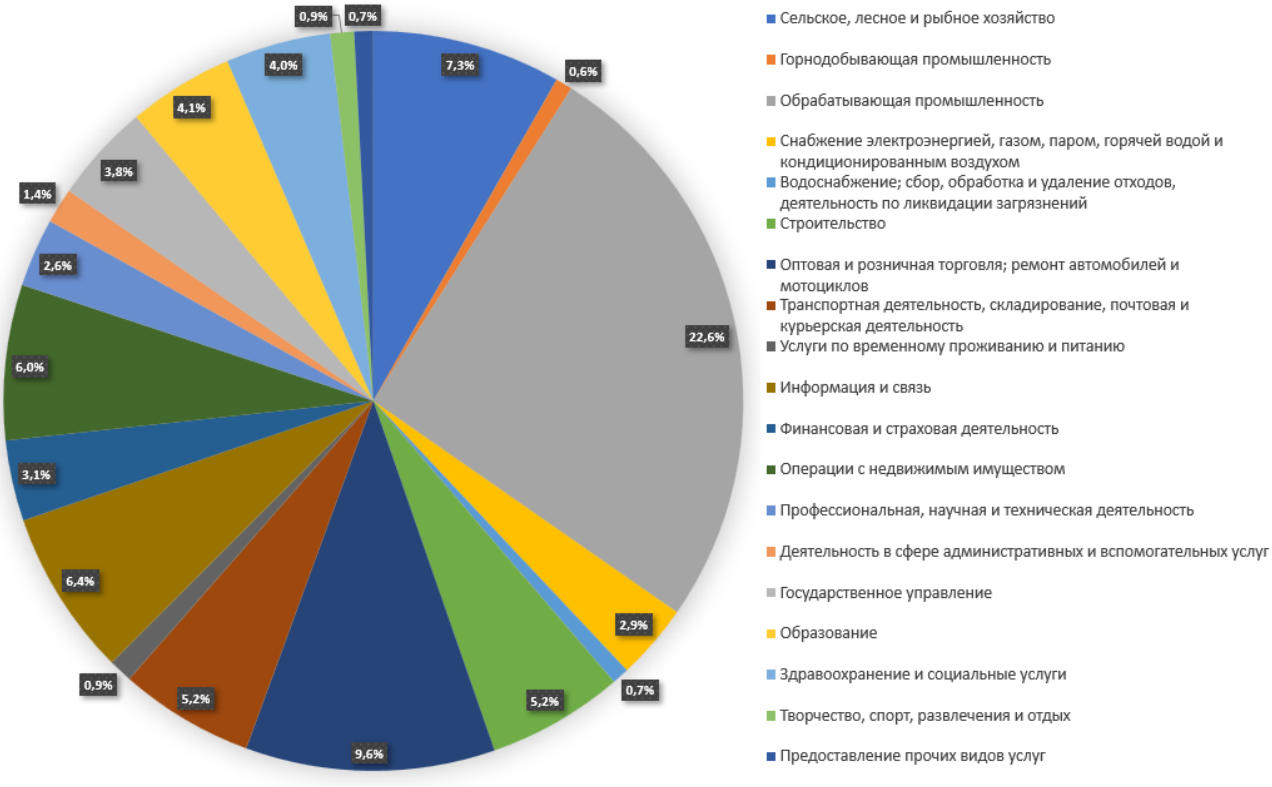


Рисунок 2.1 — Доля отраслей экономики в ВВП (производственный метод)

Рисунок 2.2 — Доля направлений использования доходов в ВВП (метод по использованию доходов)

Из графика можно заметить, что при использовании производственного метода основными компонентами, формирующими ВВП, являются:

- промышленность $> 50\%$;
- оптовая и розничная торговля 9,6%;
- сельское, лесное и рыбное хозяйство 7,3%;
- информация и связь 6,4%;
- операции с недвижимостью 6,0%;
- строительство 5,2%;
- транспорт 5,2%.

Аналогично, в методе по использованию доходов основными компонентами, формирующими ВВП, являются:

- расходы на конечное потребление 69,8%;
- валовое накопление основного капитала 22,5%.

Помимо вклада в ВВП, важно также обращать внимание на своевременность и регулярность выпуска данных, основываясь на информации из календаря публикаций данных Белстата [1, 7]. В исследовании используются только данные на месячной частоте, так как применяется MIDAS модель для прогнозирования квартального ВВП на основе месячных данных. Использование данных с поздним выпуском или тех, которые отсутствуют в открытом доступе, нецелесообразно для анализа. Данные, публикуемые с квартальной или годовой частотой, не соответствуют поставленных задачам и не будут учтены в данной работе. В таблице 2.1 представлен календарь выпуска месячных данных Белстата.

Таблица 2.1 — Календарь выпуска данных

Название показателя	Выпуск данных[1]			
	Март 2025	Апрель 2025	Май 2025	Июнь 2025
Объём промышленного производства	17-ое (02.25)	16-ое (03.25)	16-ое (04.25)	16-ое (05.25)
Объём продукции сельского хозяйства	17-ое (02.25)	15-ое (03.25)	15-ое (04.25)	17-ое (05.25)
Объём розничной и оптовой торговли	18-ое (02.25)	18-ое (03.25)	19-ое (04.25)	18-ое (05.25)
Объём инвестиций в основной капитал	24-ое (02.25)	23-ое (03.25)	22-ое (04.25)	23-ое (05.25)
Индекс объёма строительно-монтажных работ	21-ое (02.25)	21-ое (03.25)	21-ое (04.25)	23-ое (05.25)
Индекс объёма перевозок пассажиров	17-ое (02.25)	17-ое (03.25)	19-ое (04.25)	17-ое (05.25)

[1]Примечание: указана конкретная дата выпуска, а также отчётный период (в круглых скобках).

Необходимо заметить, что валовое накопление основного капитала включает следующие типы активов [8, с. 15]:

1. инвестиции в основной капитал;
2. разведка и оценка запасов полезных ископаемых;
3. расходы государства на приобретение основных средств;
4. затраты на улучшение непроизведенных материальных активов;
5. расходы на создание и приобретение отдельных объектов нематериальных активов;
6. расходы на создание и приобретение объектов интеллектуальной собственности.

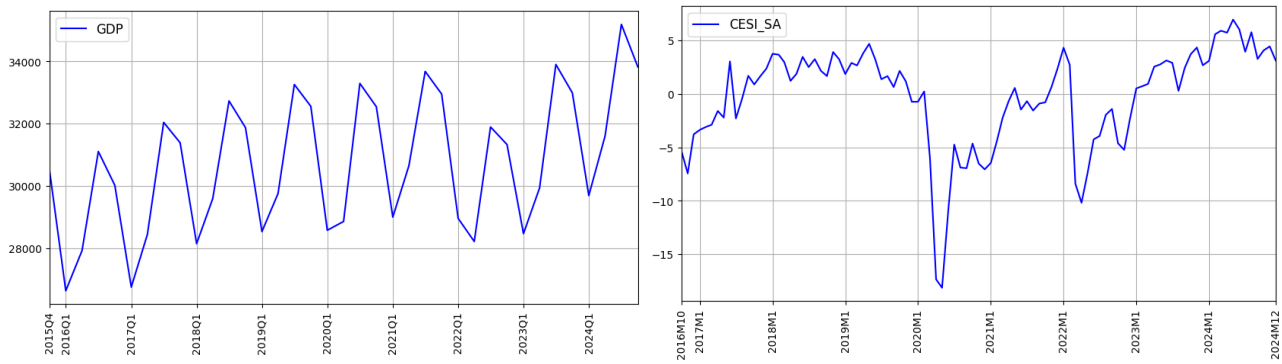
Из перечисленных активов Белстат официально выкладывает месячные данные только по объёму инвестиций в основной капитал. В дальнейшей работе не будут использоваться данные по оптовой торговле, а только по розничной, поскольку он является прямым отражением потребительского спроса и менее подвержен колебаниям внешних рынков, что делает его более стабильным индикатором внутренней экономической активности.

Вместо индекса объёма перевозок пассажиров в качестве независимой переменной для прогнозирования ВВП был добавлен сезонно-скорректированный сводный индекс экономических настроений (СИЭН), что дополнительно повысит точность анализа. Значение СИЭН представляет собой среднее геометрическое взвешенное значение индексов экономических настроений промышленности, строительства, торговли и транспорта. При этом весами для взвешивания служат доли указанных секторов в структуре ВВП, нормируемые до 100 процентов [5, с. 14].

Таким образом, для проведения исследований используются следующие экономические временные ряды:

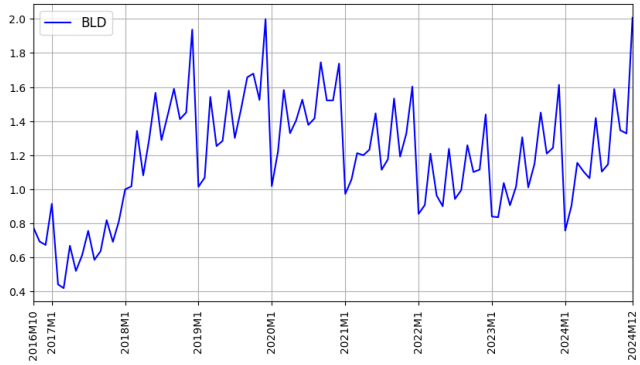
1. на квартальной частоте:
   - *GDP* — ВВП Беларуси по источникам использования доходов в среднегодовых ценах 2018 г.;
2. на месячной частоте:
   - *CESI_SA* — сезонно скорректированный сводный индекс экономических настроений (СИЭН);
   - *BLD* — базисный индекс объёма строительно монтажных работ;
   - *PP* — объём промышленного производства в среднегодовых ценах 2018 г.;
   - *INV* — объём инвестиций в основной капитал в среднегодовых ценах 2018 г.;
   - *AGRO* — объём продукции сельского хозяйства в среднегодовых ценах 2018 г.;
   - *RET* — объём розничного товарооборота в среднегодовых ценах 1995 г.

Графики временных рядов, указанных выше, приведены на рисунке 2.3.

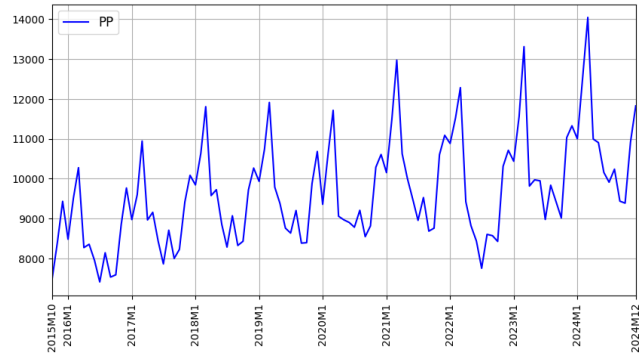


а)

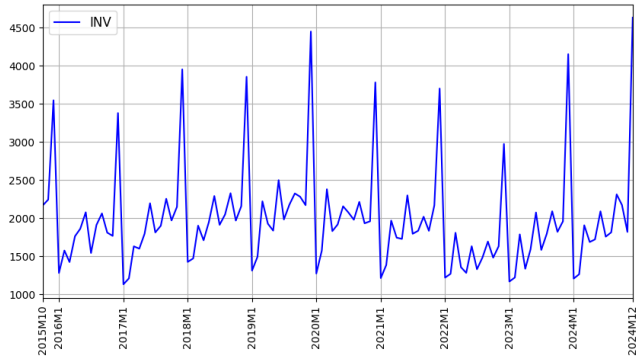
б)



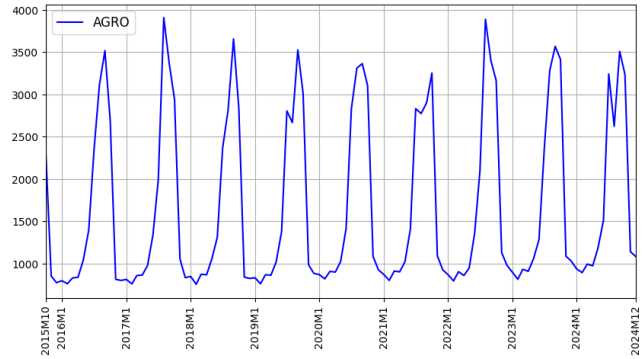
в)



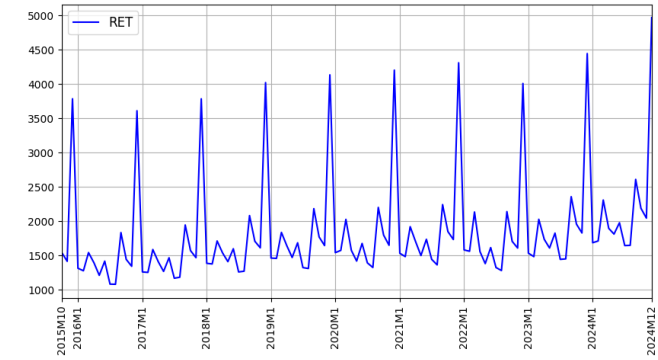
г)



д)



е)



ж)

Рисунок 2.3 — Исходные временные ряды:
а) *GDP*; б) *CESI_SA*; в) *BLD*; г) *PP*; д) *INV*; е) *AGRO*; ж) *RET*

2.2 Выбор периода исследования

Для дальнейшего анализа был построен график годовых темпов прироста ВВП на квартальной основе (рис. 2.4).

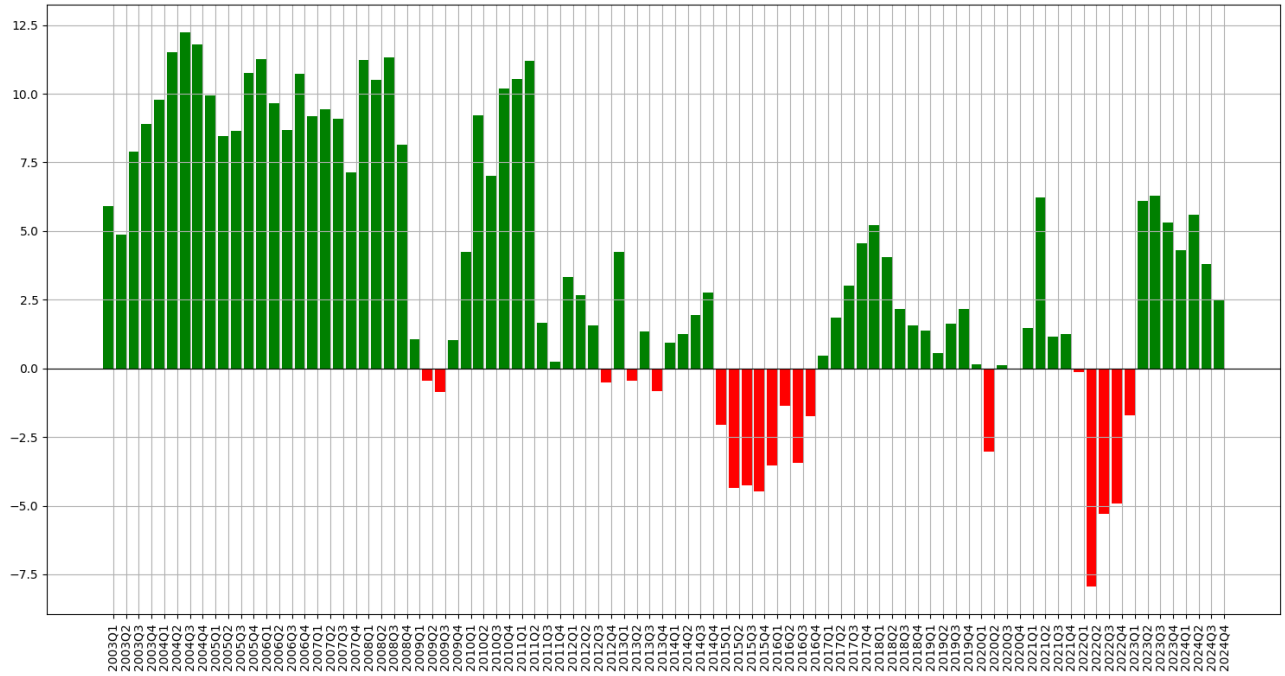


Рисунок 2.4 — Годовые темпы прироста ВВП

В 2001 году в Республике Беларусь был выбран режим таргетирования обменного курса. Валютные кризисы 2009 и 2011 годов продемонстрировали неэффективность этого монетарного режима на фоне политики, направленной на рост ВВП. На рисунке 2.4 видно, что в период с 2003 по 2011 год, за исключением 2009 и 2011 года, наблюдался высокий экономический рост.

С 2012 по 2014 год Национальный банк Республики Беларусь (НБ РБ) отказался от фиксирования валютного курса. В это время реальные процентные ставки по вкладам и кредитам превышали нейтральные уровни, что сдерживало экономическую активность. Однако недооцененность белорусского рубля смягчала негативное влияние ставок на ВВП. Падение цен на нефть и санкции против России в 2014 году привели к переоцененности рубля, что стало одной из причин экономического кризиса 2015 года.

В начале 2015 года НБ РБ перешел к монетарному таргетированию, сначала управляя денежным предложением, а затем, с середины 2016 года, и процентными ставками. Период 2017-2023 годов оказался благоприятным для экономической активности. Валютный курс поддерживал ВВП, а недооцененность рубля способствовала росту экспорта. Однако в 2020 году, из-за пандемии, и в 2022 году, из-за санкций, наблюдалось падение ВВП. В 2017-2020 годах влияние процентных ставок на ВВП было нейтральным, но затем, с восстановлением экономики, стало стимулирующим.

На основе исторического обзора можно выделить следующие временные периоды:

1. 2003q1-2011q4;
2. 2012q1-2014q4;
3. 2015q1-2016q3;
4. 2016q4-2024q4.

Для каждого периода были вычислены описательные статистики: среднее, максимальное и минимальное значения, стандартное отклонение. Годы 2009, 2011, 2020 и 2022 не учитывались из-за экономических кризисов. Расчитанные статистики представлены в таблице 2.2.

Таблица 2.2 — Описательные статистики годовых темпов прироста ВВП в разные периоды

Временной период	Среднее	Минимальное значение	Максимальное значение	Стандартное отклонение
2003q1-2011q4	9,203717	4,251913	12,235506	2,020073
2012q1-2014q4	1,521488	-0,823916	4,249368	1,582639
2015q1-2016q3	-3,353831	-4,480990	-1,356470	1,211139
2016q4-2024q4	2,764198	-1,749545	6,299995	2,306796

Анализ временного ряда годовых темпов прироста ВВП показывает, что его поведение значительно изменяется в выделенные периоды, что имеет важное значение для выбора временного интервала для построения прогнозных моделей.

1. **Период 2003q1-2011q4.** В этом периоде наблюдается стабильный и высокий рост годовых темпов прироста ВВП, средний темп — 9,2%.
2. **Период 2012q1-2014q4.** Здесь фиксируется значительное замедление роста до 1,52%, что указывает на экономические трудности и неопределённость. Наличие отрицательных значений в этом временном интервале демонстрирует высокую волатильность и нестабильность.
3. **Период 2015q1-2016q3.** Данный период характеризуется глубокой рецессией, со средним темпом прироста -3,35%. Отрицательные значения и высокая волатильность делают данные крайне неблагоприятными для моделирования.
4. **Период 2016q4-2024q4.** Начиная с этого момента, наблюдается восстановление экономики, с положительным средним темпом 2,76%. Несмотря на остающуюся волатильность, данный период демонстрирует более стабильное и позитивные тенденции, что делает его более подходящим для построения моделей.

Таким образом, в качестве временного интервала для моделирования был выбран период, начинающийся с 2016q4, который демонстрирует положительные и устойчивые тренды, отличающиеся от предыдущих периодов с их высокой волатильностью и негативными значениями.

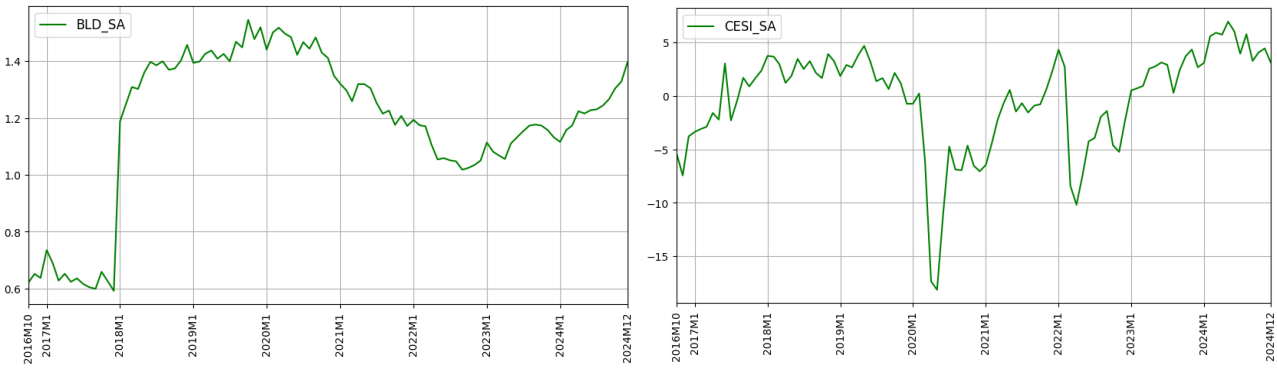
2.3 Сезонная корректировка данных

Сезонная корректировка является важным этапом в анализе временных рядов, особенно при прогнозировании экономических показателей, таких как валовый внутренний продукт. Для устранения сезонного фактора временные ряды, отражающие объём в среднегодовых ценах, были приведены к годовым темпам прироста, которые вычисляются по формуле (2.3.1). Они позволяют легче воспринимать изменения в экономике, так как представляют относительно увеличение или уменьшение показателя по сравнению с соответствующим периодом прошлого года. Базисные индексы были сезонно скорректированы с помощью метода TRAMO/SEATS по причине того, что данные временные ряды отражают уровень показателя относительно базового периода.

$$\text{Годовой прирост} = \begin{cases} \left(\dfrac{x_t}{x_{t-4}} - 1\right) \times 100\%, & \text{данные квартальные} \\ \left(\dfrac{x_t}{x_{t-12}} - 1\right) \times 100\%, & \text{данные месячные} \end{cases} \tag{2.3.1}$$

где x_t — значение в текущем квартале (месяце), $x_{t-4(t-12)}$ — значение в том же квартале (месяце) предыдущего года.

Сезонно скорректированные временные ряды были названы по следующему принципу: к исходному названию ряда было добавлено $_G$, если данные были приведены к годовым приростам, и $_SA$, в случае использования метода TRAMO/SEATS. Исключением является ВВП, он был приведён к приростам и назван *GGDP*. Также необходимо заметить, что временной ряд *CESI_SA* уже был сезонно скорректированным, поэтому работа с ним не проводилась. Полученные временные ряды приведены на рисунках 2.5 и 2.6.



а) б)

Рисунок 2.5 — Сезонно скорректированные базисные индексы: а) *BLD_SA*; б) *CESI_SA*

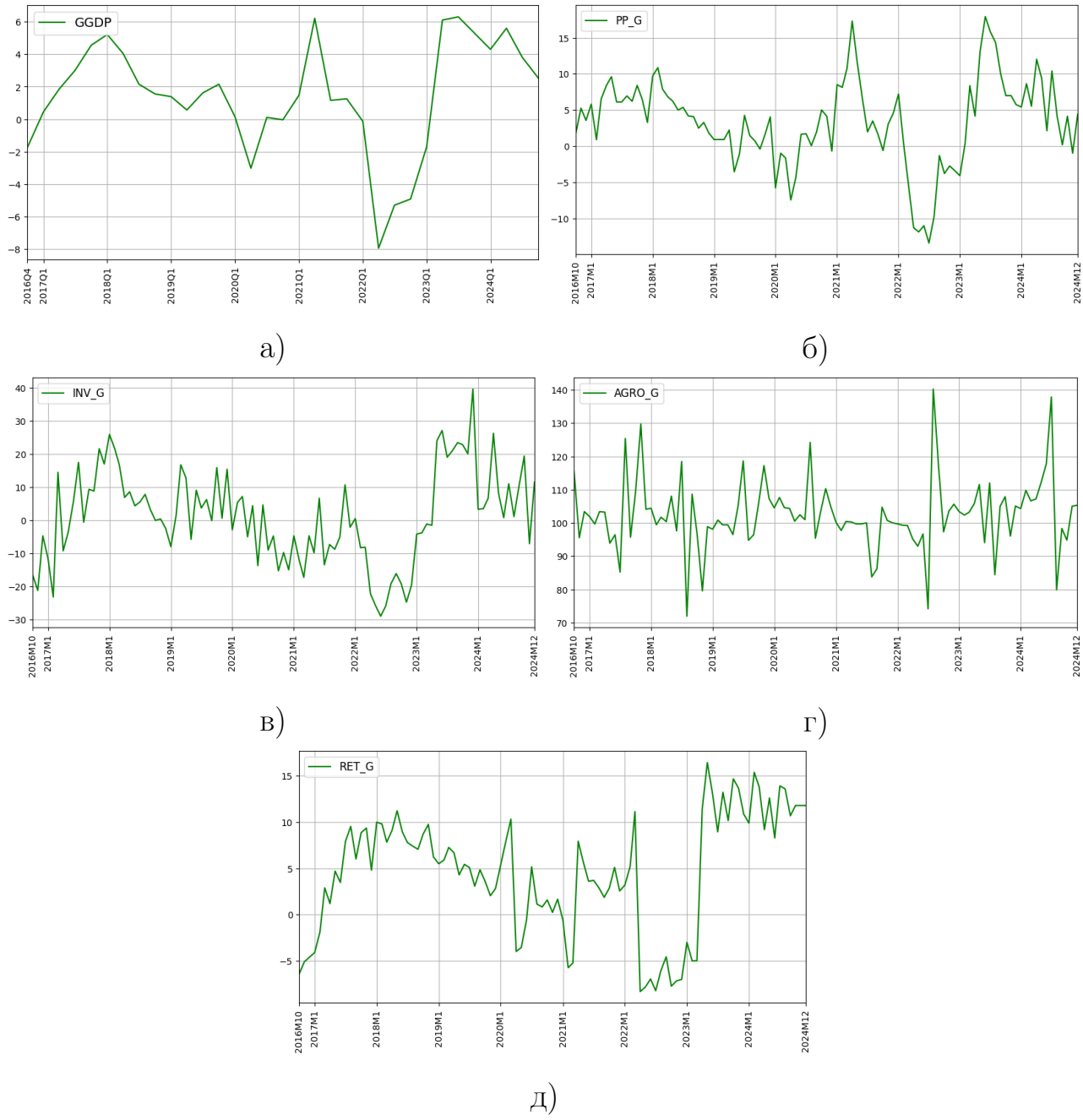


Рисунок 2.6 — Годовые темпы прироста временных рядов, отражающие объём в среднегодовых ценах:
а) $GGDP$; б) PP_G; в) INV_G; г) $AGRO_G$; д) RET_G

2.4 Тестирование стационарности

Для тестирования стационарности и определения типа стационарности обычно используют «тесты единичного корня», например расширенный тест Дики-Фуллера (ADF-тест). Однако следует отметить, что он может быть склонен к ошибочным выводам в случае наличия структурных изменений. Альтернативой является модифицированный тест Дики-Фуллера, известный как BPUR-тест (Break Point Unit Root). Достоинством данного теста является

возможность выявления одного наиболее выраженного структурного изменения во временном ряде, описываемом DS-моделью. BPUR-тест помогает избежать ошибок и обеспечивает более надёжные выводы о стационарности ряда. Результаты тестирования данных с помощью указанных тестов приведены в таблицах 2.3 и 2.4 соответственно.

Таблица 2.3 — Результаты тестирования временных рядов с помощью ADF-теста

Временной ряд	В модель включены	Значение ADF-статистики	Пороговые значения, уровень значимости $\varepsilon = 0{,}05$	Тип модели
$GGDP_t$	—	-2,184	-1,951	TS
PP_G_t	c	-3,502	-2,891	TS
INV_G_t	c	-2,901	-2,891	TS
$AGRO_G_t$	c	-14,192	-2,891	TS
RET_G_t	c	-3,348	-2,891	TS
$CESI_SA_t$	—	-2,615	-1,944	TS
BLD_SA_t	c	-1,780	-2,891	DS

Таблица 2.4 — Результаты тестирования временных рядов с помощью BPUR-теста

Временной ряд	В модель включены	Значение ADF-статистики	Пороговые значения, уровень значимости $\varepsilon = 0.05$	Моменты структурных изменений	Тип модели
$GGDP_t$	c	-6,134	-4,444	2022q3 (AO)	TS
PP_G_t	c, t	-4,742	-4,860	2023m1 (AO)	DS
INV_G_t	c	-4,918	-4,444	2017m2 (AO)	TS
$AGRO_G_t$	c	-11,936	-4,444	2024m7 (AO)	TS
RET_G_t	c	-5,071	-4,444	2023m3 (AO)	TS
$CESI_SA_t$	c, t	-5,843	-4,860	2019m10 (AO)	TS
BLD_SA_t	c	-4,885	-4,444	2017m12 (IO)	TS

Результаты тестирования (табл. 2.4) указывают, что временные ряды на временном интервале 2016q4-2024q4 стационарны, за исключением данных по промышленному производству. Следует учитывать, что период тестирования короткий, что увеличивает вероятность ложных выводов. Приведение

к годовым приростам является методом, который позволяет привести временные ряды к стационарному виду. Следовательно, временной ряд PP_G будем считать стационарным, и тест выдал ошибку второго рода. Это можно подтвердить увеличив период тестирования.

Таким образом, все временные ряды являются стационарными. Это означает, что их статистические свойства, такие как среднее и дисперсия, остаются постоянными во времени.

2.5 Корреляционный анализ

Корреляционный анализ является статистическим методом, который позволяет определить наличие зависимости между переменными и оценить её силу с помощью коэффициентов корреляции [10, с. 349]. В данной работе используются временные ряды с различной частотой, что требует преобразования данных высокой частоты (месячных) в данные низкой частоты (квартальную). Исходные временные ряды, отражающие объем в среднегодовых ценах, были агрегированы путём суммирования и преобразованы в годовые темпы прироста. Сезонно скорректированные временные ряды, представляющие базисные индексы, были агрегированы с использованием усреднения. В результате все временные ряды были приведены к квартальной частоте, что позволило провести корреляционный анализ. В итоге была построена тепловая карта для визуализации корреляционной матрицы временных рядов за период 2016q4-2024q4, которая представлена на рисунке 2.7.

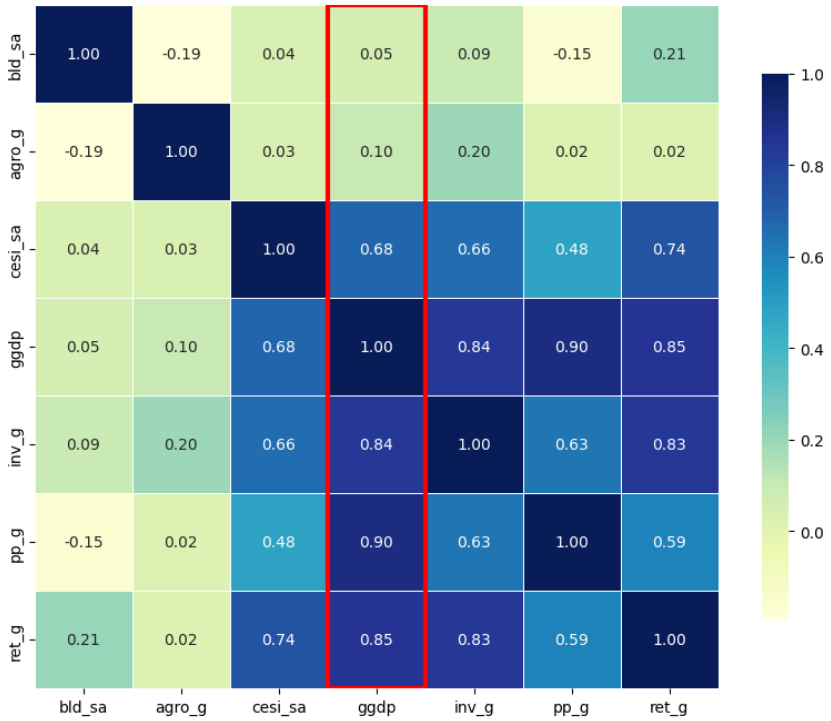


Рисунок 2.7 — Тепловая карта

Анализ, проведённый на основе тепловой карты корреляций, позволяет сделать несколько ключевых выводов о взаимосвязях между различными переменными:

1. **Сильная корреляция**
   Временной ряд $GGDP$ демонстрирует высокую положительную корреляцию с переменными INV_G (0,84), PP_G (0,9) и RET_G (0,85). Это может свидетельствовать о том, что рост значений $GGDP$ связано с ростом значений указанных переменных.
2. **Умеренная корреляция**
   Наблюдается умеренная корреляция $GGDP$ с переменной $CESI_SA$ (0,68). Эти результат указывает на наличие взаимосвязи, хотя она менее выраженная по сравнению с указанными выше.

Сильная и умеренная корреляция между переменными указывает на значительное и предсказуемое влияние одной переменной на другую, что позволяет делать уверенные прогнозы.

Мультиколлинеарность возникает, когда две или более независимых переменных в регрессионной модели сильно коррелируют друг с другом. Это может привести к проблемам при оценке коэффициентов модели, увеличивая их стандартные ошибки и делая их оценку ненадёжной. На основе представленной тепловой карты можно сделать следующие выводы о мультиколлинеарности:

1. **Сильная корреляция между переменными**
   Переменные INV_G и RET_G имеют корреляцию 0,83, что указывает на сильную взаимосвязь. Включение обеих переменных в модель может привести к проблемам с мультиколлинеарностью, так как они могут в значительной степени объяснять одно и то же изменение в зависимой переменной [4].
2. **Умеренные корреляции**
   Переменные INV_G и RET_G имеют корреляцию 0,63 и 0,59 с PP_G и 0,66 и 0,74 с $CESI_SA$. Хотя эта корреляция не такая высокая, она всё равно может привести к умеренной мультиколлинеарности, если обе переменные будут включены в модель.

ГЛАВА 3
МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ТЕМПОВ ПРИРОСТА ВВП БЕЛАРУСИ

Моделирование и прогнозирование темпов прироста ВВП имеют важное значение в образовании эффективной экономической политики в современных экономических исследованиях. Для сравнения эффективности различных методов прогнозирования в качестве базовой модели была выбрана модель авторегрессии с распределенными лагами ARDL (Autoregressive Distributed Lag). Она служит надёжным ориентиром для оценки и сопоставления альтернативных подходов.

В ходе работы были рассмотрены различные варианты модели MIDAS, в частности, неограниченная версия U-MIDAS, а также модификации MIDAS-GETS и MIDAS-GETSIS, которые применяют техники насыщения индикаторов. Эти варианты позволяют учитывать аномалии и изменения в структуре данных, что положительно сказывается на качестве прогнозов.

Чтобы обеспечить достоверность и обоснованность прогнозов роста ВВП, был применён метод расширяющегося окна. Этот подход позволяет постоянно обновлять обучающую выборку и проводить более точную оценку моделей. Метод расширяющегося окна состоит из следующих этапов [11, с. 19]:

1. Вся выборка данных первоначально делится на две части — обучающую и тестируемую. В результате в обучающую выборку первоначально вошли наблюдения за период с 2016q4 по 2023q2, а в тестируемую — с 2023q3 по 2024q4.
2. Основываясь на фактических временных рядах обучающей выборки осуществляется оценка модели.
3. С помощью оцененной модели производится прогноз на один квартал вперед.
4. Увеличивается длина обучающей выборки на один квартал, и заново оцениваются параметры модели.
5. Используя переоценную модель с учетом увеличенной выборки, производится прогноз на один месяц вперед.
6. Аналогично шагам 4 и 5 происходит увеличение выборки, переоценка модели и построение прогноза на один квартал вперед до тех пор, пока обучающая выборка не достигнет периода с 2016q4 по 2024q3.
7. В результате полученные 6 точечных прогнозов темпов прироста ВВП сравниваются с фактическими данными, и рассчитываются показатели точности прогноза: MAE, RMSE.

На рисунке 3.1 представлена общая схема алгоритма, иллюстрирующая процесс деления данных на обучающую и тестируемую выборки, а также последовательность оценивания и прогноза. Это наглядное представление помогает лучше понять структуру и динамику расширяющегося окна.

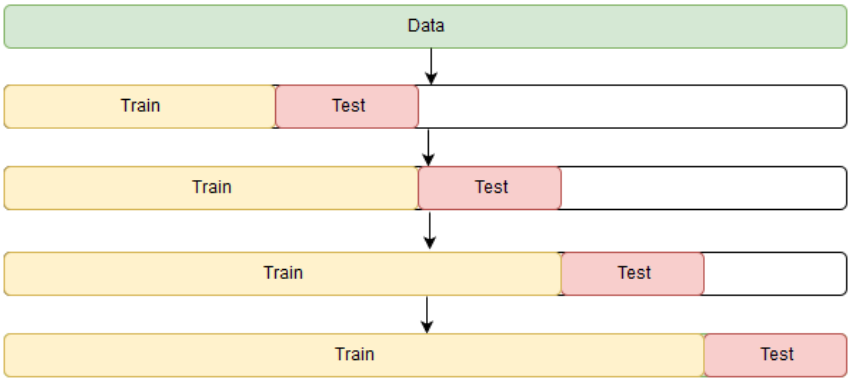


Рисунок 3.1 — Схема реализации метода расширяющегося окна

3.1 Построение MIDAS моделей

В рамках настоящего исследования модели MIDAS были построены для одной экзогенной переменной. Такой подход позволяет сосредоточится на детальном анализе влияния данной переменной на годовые темпы прироста ВВП. Каждая модель адаптировано к специфическим характеристикам данных, что даёт возможность учитывать различные аспекты и динамику временных рядов.

Фокусировка на одной экзогенной переменной позволяет глубже понять ее влияние на экономические процессы и оценить её вклад в формирование эффективной экономической политики. В следующих подразделах представлены результаты построения моделей для выбранной переменной, а также проведён анализ их прогностических способностей.

Кроме того, в дальнейшем была осуществлена комбинация полученных прогнозов, что позволило выявить синергетический эффект и повысить точность прогнозирования.

3.1.1 Построение моделей с переменной *CESI_SA*

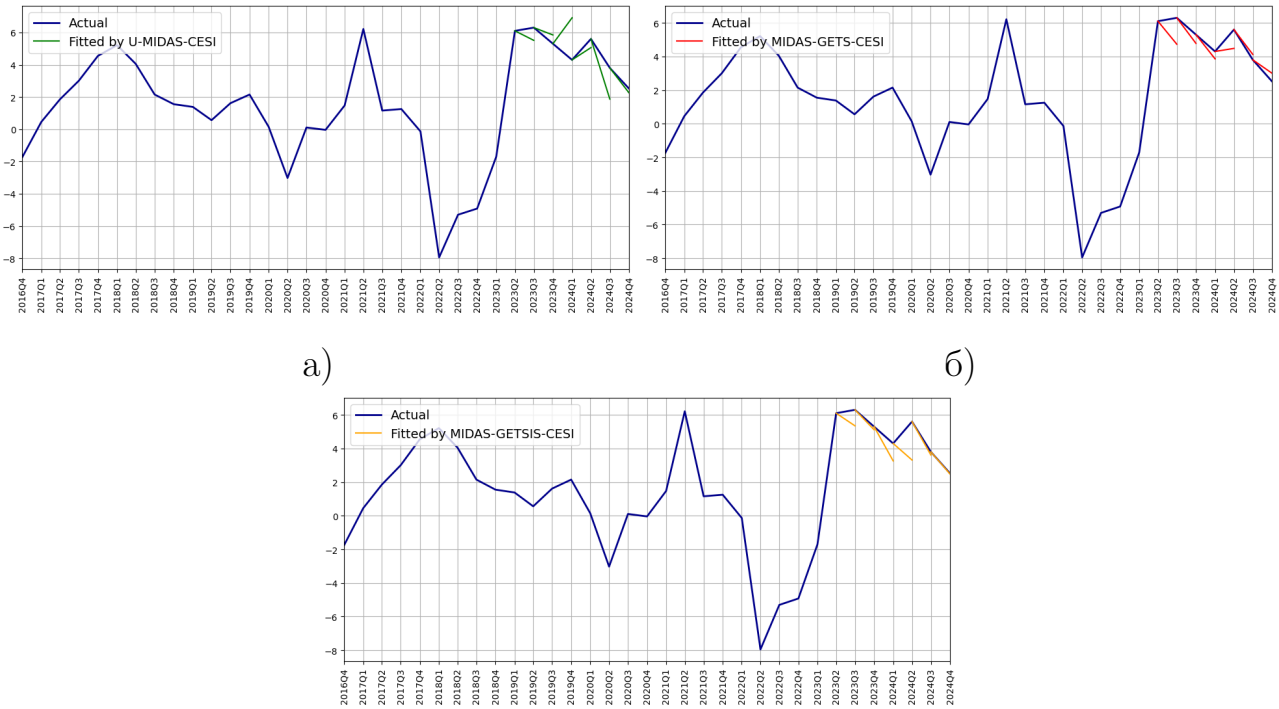
В процессе построения моделей U-MIDAS, MIDAS-GETS, MIDAS-GETSIS в качестве эндогенной переменой выступал временной ряд годовых темпов прироста ВВП, что является основной целью исследования.

Модели U-MIDAS и MIDAS-GETS включали набор квартальных регрессоров, среди которых присутствовали константа, лаг переменной *GGDP* и дамми-переменная *DUM2022Q2*. Несмотря на то что тест BPUR указал на *2022Q3*, дамми-переменная *DUM2022Q2* оказалась статистически значимой, учитывая возможные задержки в выявлении структурных изменений. В качестве регрессора более высокой частоты использовалась переменная *CESI_SA* без лага, что позволяет учитывать влияние месячных данных на динамику ВВП. Установлено значение лагов на уровне двенадцати месяцев, что помогает учесть длительные зависимости в данных.

Модель MIDAS-GETSIS повторяет спецификацию модели MIDAS-GETS,

за исключением отсутствия дамми-переменной. Этот подход MIDAS автоматически учитывает выбросы и структурные изменения (Приложение А).

На рисунке 3.2 представлены «волосатые графики», демонстрирующие результаты прогнозирования на один квартал вперёд с помощью метода расширяющегося окна (Приложение Б). В таблице 3.1 приводятся показатели точности прогнозов RMSE, MAE.



а) б)

в)

Рисунок 3.2 — Наблюдаемые и прогнозные значения *GGDP* на основе моделей:
а) U-MIDAS-CESI; б) MIDAS-GETS-CESI; в) MIDAS-GETSIS-CESI

Таблица 3.1 — Результаты оценки точности краткосрочного прогноза годовых темпов прироста ВВП с помощью моделей MIDAS

Модель	RMSE	MAE
U-MIDAS-CESI	1,403011	1,112751
MIDAS-GETS-CESI	0,867813	0,742120
MIDAS-GETSIS-CESI	1,100931	0,780526

Выбирая модель для дальнейшего использования, важно обращать внимание на то, какая модель обеспечивает минимальные значения показателей точности. Однако со временем другой метод может стать более точным, поэтому использование комбинированного прогноза может стать оптимальным решением. Это особенно в условиях изменяющихся данных и внешних факторов, влияющих на точность прогнозирования.

Таким образом, были построены комбинированные прогнозы, которые осуществляются путём взвешивания на основе различных подходов. В частности, использованы следующие методы [11, с. 21]:

- равных весов (MIDAS-AVER);
- весов полученных с помощью МНК (MIDAS-OLS);
- обратных значений функции потерь (MIDAS-MSE);
- обратных значений рангов функции потерь (MIDAS-RANKS).

Для построения комбинированных прогнозов шесть точечных прогнозов были помещены в один временной ряд и далее была проведено комбинирование, что позволяет повысить точности прогнозов и помогает компенсировать недостатки каждой из модели MIDAS. На рисунке 3.3 представлены полученные комбинированные прогнозы по переменной *CESI_SA*, позволяя визуально оценить их эффективность. Кроме того, в таблице 3.2 показаны показатели точности прогнозов RMSE, MAE, что позволяет сравнить результаты разных методов комбинирования и отдельных моделей.

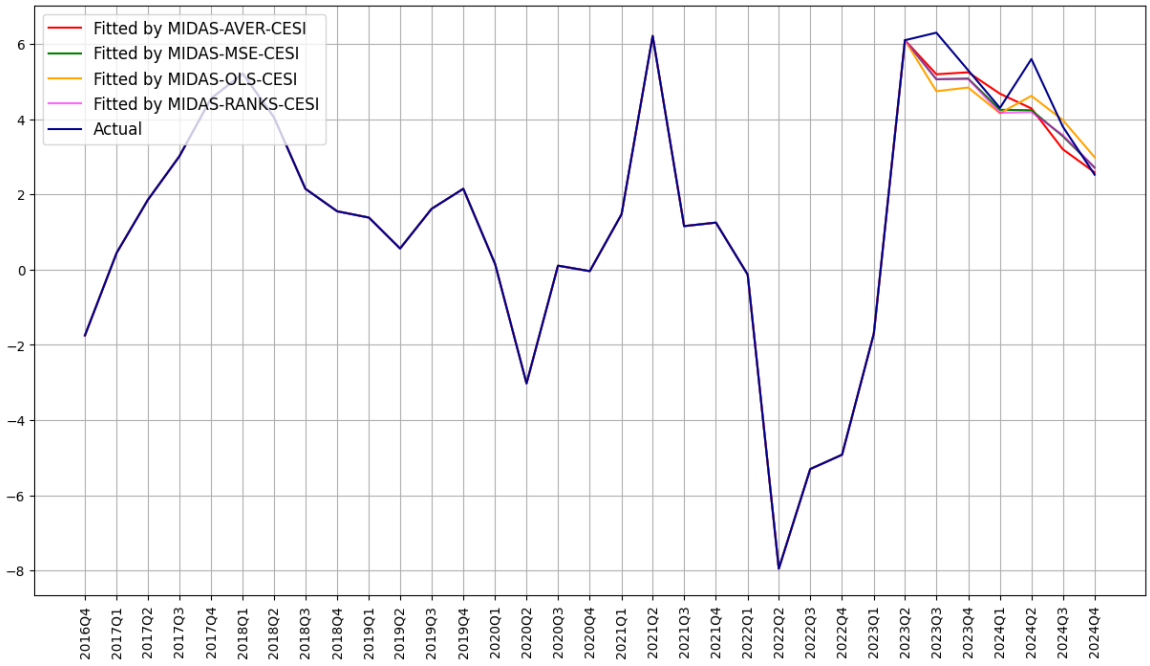


Рисунок 3.3 — Комбинированные прогнозы темпов прироста ВВП

Таблица 3.2 — Результаты оценки точности комбинированных прогнозов моделей MIDAS

Модель	RMSE	MAE
MIDAS-AVER-CESI	0,759032	0,585382
MIDAS-OLS-CESI	0,801509	0,628978
MIDAS-MSE-CESI	0,768115	0,552095
MIDAS-RANKS-CESI	0,782078	0,569598

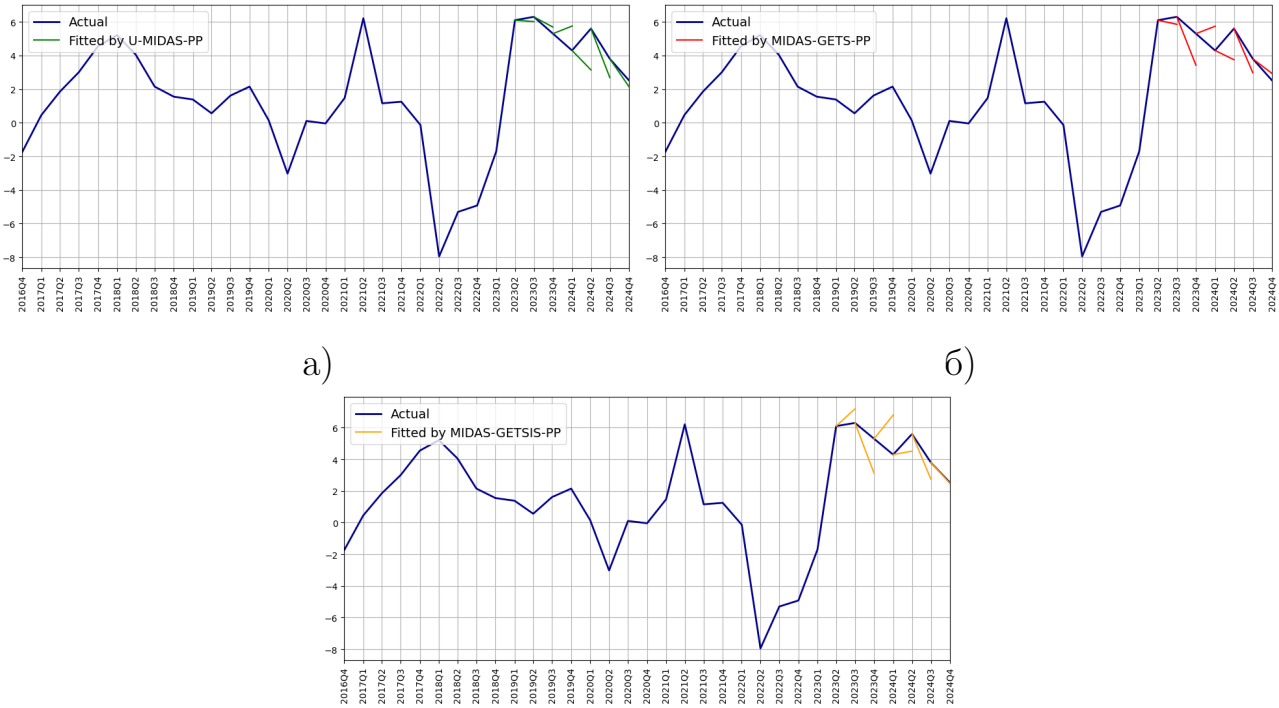
3.1.2 Построение моделей с переменной *PP_G*

В соответствии с ранее описанным подходом построения моделей MIDAS. В моделях U-MIDAS, MIDAS-GETS, MIDAS-GETSIS в качестве зависимой переменной использовался временной ряд *GGDP*.

В моделях U-MIDAS и MIDAS-GETS была сформирована совокупность низкочастотных переменных, включая лаг переменной *GGDP* и дамми-переменную *DUM2022Q2*. Хотя тест BPUR указал на *2023Q3* как на период структурных изменений, дамми-переменная *DUM2022Q2* оказалась статистически значимой, что может быть связано с задержками в обнаружении этих изменений тестом. В качестве высокочастотной переменной использовалась переменная *PP_G* без лага, что позволяет учесть влияние месячных показателей на динамику ВВП. Значение лагов было установлено на уровне двенадцати месяцев, что помогает учитывать долгосрочные зависимости в данных.

Модель MIDAS-GETSIS следует аналогичной спецификации, как и в MIDAS-GETS, но без включения дамми-переменной. Этот подход автоматически обрабатывает выбросы и структурные изменения в данных.

На рисунке 3.4 представлены «волосатые графики», демонстрирующие 6 точечных прогнозов на один квартал вперед, полученные с использованием метода расширяющегося окна. В таблице 3.3 содержатся показатели точности прогнозов.



а) б)

в)

Рисунок 3.4 — Наблюдаемые и прогнозные значения *GGDP* на основе моделей:
а) U-MIDAS-PP; б) MIDAS-GETS-PP; в) MIDAS-GETSIS-PP

Таблица 3.3 — Результаты оценки точности краткосрочного прогноза годовых темпов прироста ВВП с помощью моделей MIDAS

Модель	RMSE	MAE
U-MIDAS-PP	1,274650	1,011320
MIDAS-GETS-PP	1,295523	1,142075
MIDAS-GETSIS-PP	1,531529	1,292890

Используя шесть точечных прогнозов каждой модели MIDAS, мы создаем более точные комбинированные результаты, что позволяет улучшить качество предсказаний. Такой подход позволяет учесть сильные стороны каждой модели и минимизировать их недостатки. На рисунке 3.5 иллюстрируются эти комбинированные прогнозы, демонстрируя их динамику и точность. В таблице 3.4 представлены показатели точности прогнозов, такие как RMSE, MAE, которые подтверждают преимущество комбинированного подхода.

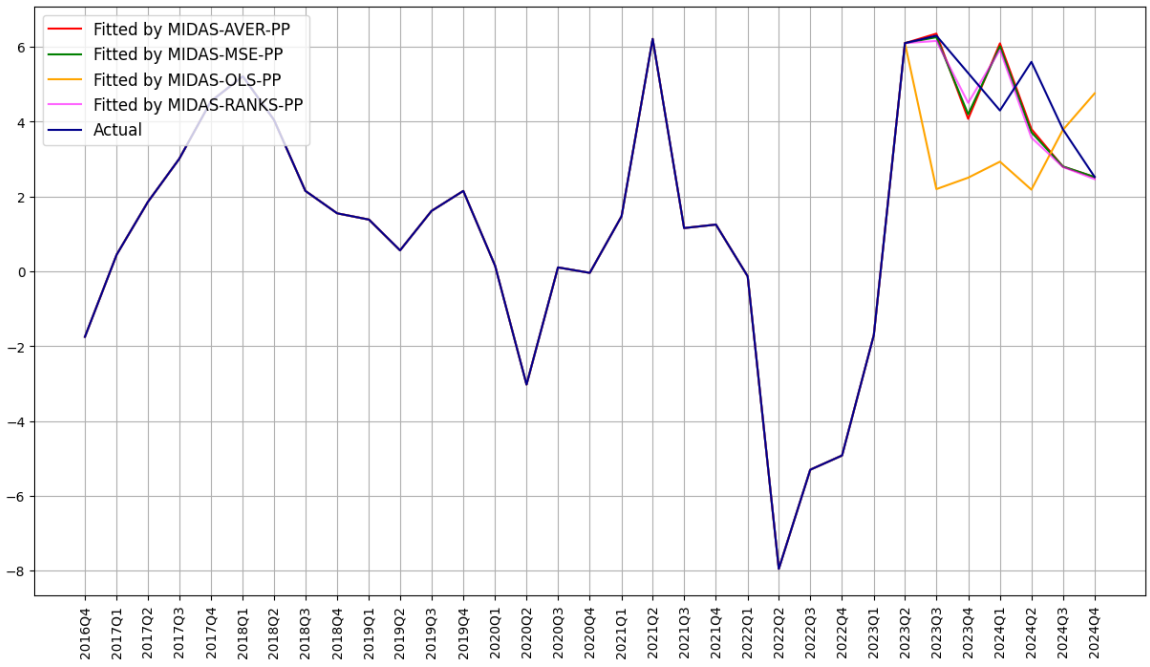


Рисунок 3.5 — Комбинированные прогнозы темпов прироста ВВП

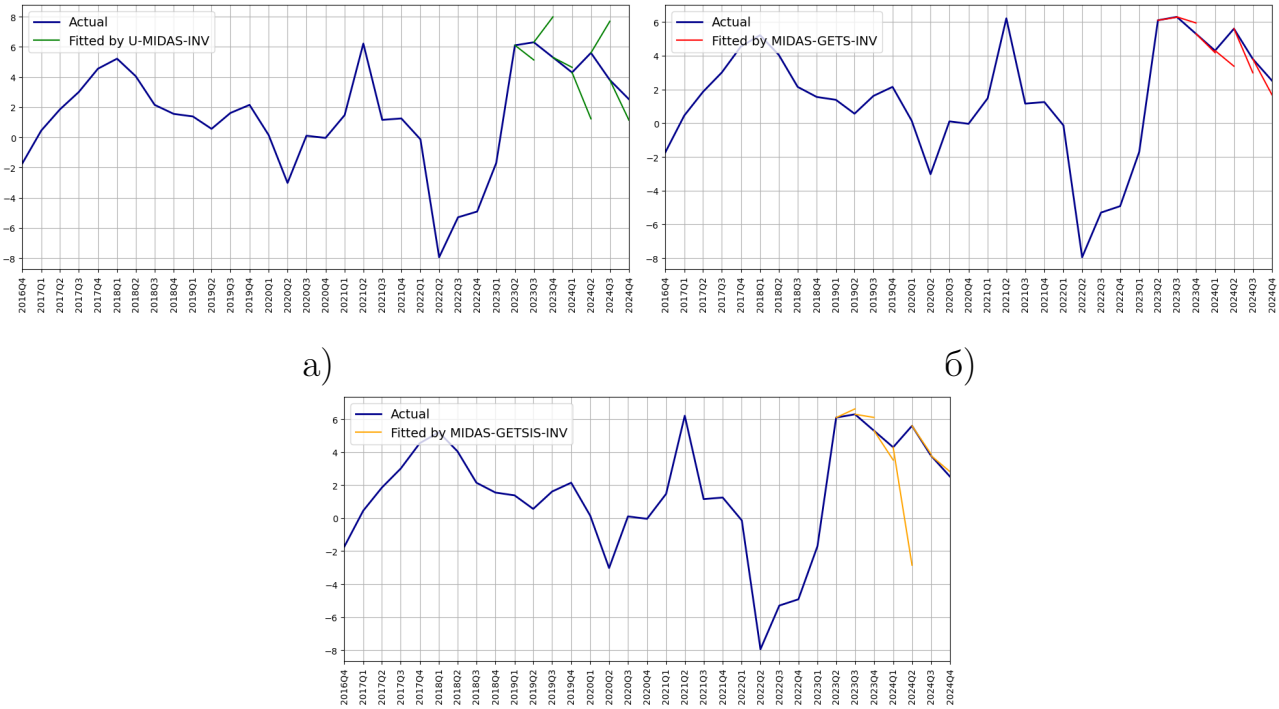
Таблица 3.4 — Результаты оценки точности комбинированных прогнозов моделей MIDAS

Модель	RMSE	MAE
MIDAS-AVER-PP	1,220950	0,978648
MIDAS-OLS-PP	2,682217	2,322494
MIDAS-MSE-PP	1,202868	0,955878
MIDAS-RANKS-PP	1,183157	0,941489

3.1.3 Построение моделей с переменной *INV_ G*

Спецификации моделей U-MIDAS, MIDAS-GETS и MIDAS-GETSIS аналогичны ранее описанным, но в качестве высокочастотной переменной используется лаг третьего порядка *INV_ G*(-3). В качестве зависимой переменной выступал временной ряд *GGDP*.

На рисунке 3.6 представлены «волосатые графики», показывающие прогнозы полученные с помощью моделей MIDAS. В таблице 3.5 указаны показатели точности RMSE, MAE.



а) б)

в)

Рисунок 3.6 — Наблюдаемые и прогнозные значения *GGDP* на основе моделей:
а) U-MIDAS-INV; б) MIDAS-GETS-NV; в) MIDAS-GETSIS-INV

Таблица 3.5 — Результаты оценки точности краткосрочного прогноза годовых темпов прироста ВВП с помощью моделей MIDAS

Модель	RMSE	MAE
U-MIDAS-INV	2,734497	2,307378
MIDAS-GETS-INV	1,059395	0,776326
MIDAS-GETSIS-INV	3,484022	1,792956

В процессе работы были использованы полученные шесть точечных прогнозов каждой модели MIDAS для создания более надёжных комбинированных результатов, что позволяет улучшить общую достоверности предсказаний.

Применение комбинированных методов особенно актуально в условиях высокой изменчивости данных. Рисунок 3.7 иллюстрирует полученные комбинированные прогнозы, позволяя наглядно оценить их эффективность. Были рассчитаны показатели точности прогнозов RMSE, MAE (табл. 3.6), демонстрирующие, что комбинированные прогнозы обеспечивают более высокую точность по сравнению с отдельными моделями. Таким образом, полученные результаты подчёркивают важность использования комбинированных подходов в прогнозировании.

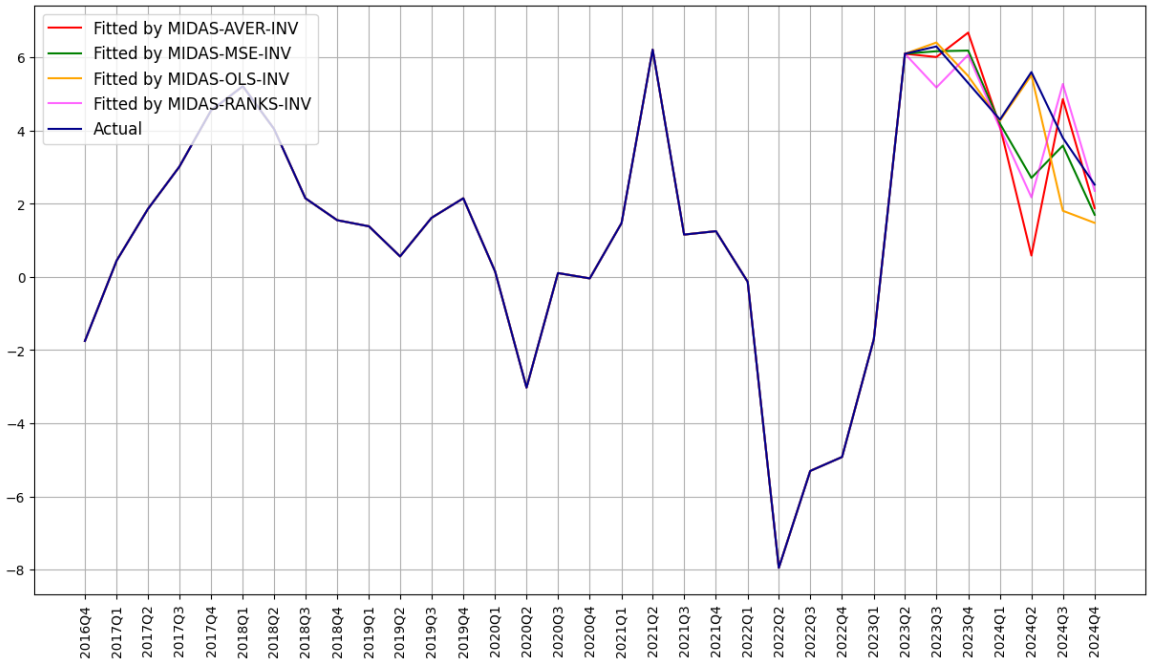


Рисунок 3.7 — Комбинированные прогнозы темпов прироста ВВП

Таблица 3.6 — Результаты оценки точности комбинированных прогнозов моделей MIDAS

Модель	RMSE	MAE
MIDAS-AVER-INV	2,187114	1,427827
MIDAS-OLS-INV	0,923301	0,574809
MIDAS-MSE-INV	1,284776	0,842558
MIDAS-RANKS-INV	1,6246222	1,199027

3.1.4 Построение моделей с переменной *RET_G*

В ходе разработки моделей U-MIDAS, MIDAS-GETS и MIDAS-GETSIS в качестве экзогенной переменной выступал временной ряд *GGDP*, отражающий годовые темпы прироста ВВП. Модели U-MIDAS и MIDAS-GETS включали набор низкочастотных переменных, таких как лаг переменной *GGDP* и дамми-переменная *DUM2022Q2*. В качестве высокочастотной переменной

использовалась RET_G без лага. Лаги установлены на уровне двенадцати месяцев. Модель MIDAS-GETSIS имеет аналогичную спецификацию, но без дамми-переменной.

«Волосатые графики», отображающие одноточечные прогнозы, полученные с использованием метода расширяющего окна, проиллюстрированы на рисунке 3.8. Показатели точности прогнозов RMSE, MAE указаны в таблице 3.7.

а) б)

в)

Рисунок 3.8 — Наблюдаемые и прогнозные значения *GGDP* на основе моделей:
а) U-MIDAS-RET; б) MIDAS-GETS-RET; в) MIDAS-GETSIS-RET

Таблица 3.7 — Результаты оценки точности краткосрочного прогноза годовых темпов прироста ВВП с помощью моделей MIDAS

Модель	RMSE	MAE
U-MIDAS-RET	2,561823	2,211921
MIDAS-GETS-RET	2,000619	1,861401
MIDAS-GETSIS-RET	1,861094	1,568906

В процессе работы были построены комбинированные прогнозы, полученные на основе различных моделей MIDAS, а именно U-MIDAS, MIDAS-GETS, MIDAS-GETSIS. На их основе, используя шесть точечных прогнозов, удалось получить более точные результаты, существенно улучшив качество предсказа-

ний. В данном случае объединение прогнозов позволило уравновесить разные подходы и использовать их лучшие характеристики.

Динамика комбинированных прогнозов, отображённая на рисунке 3.9, даёт возможность визуально оценить их точность в сопоставлении с отдельными прогнозами. В таблице 3.8 представлены основные показатели точности, включая RMSE и MAE, служащие доказательством эффективности комбинированного подхода, а также подчёркивающие его преимущества перед использованием изолированных моделей.

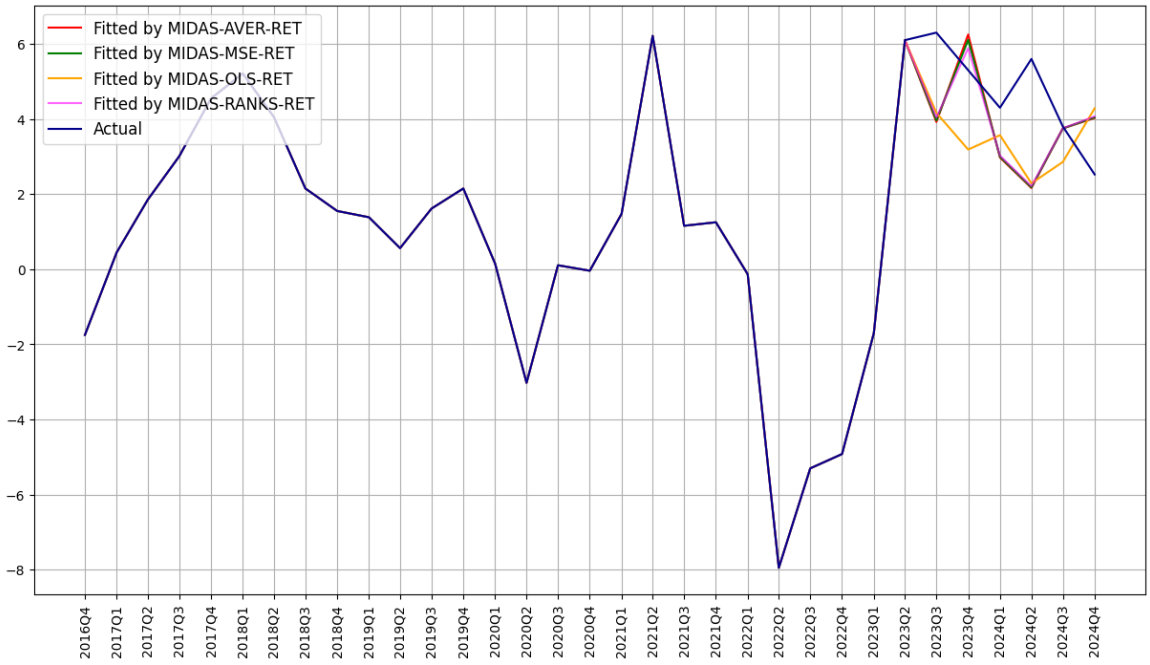


Рисунок 3.9 — Комбинированные прогнозы темпов прироста ВВП

Таблица 3.8 — Результаты оценки точности комбинированных прогнозов моделей MIDAS

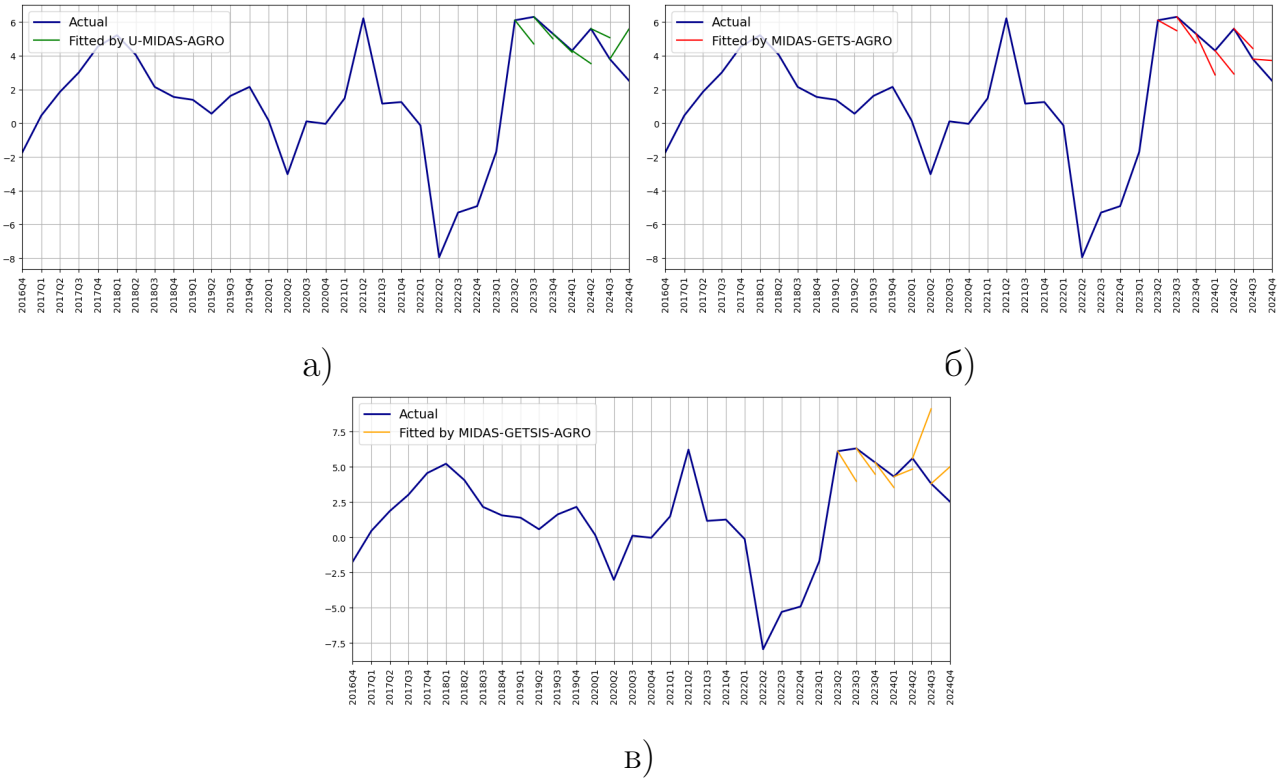
Модель	RMSE	MAE
MIDAS-AVER-RET	1,931659	1,607066
MIDAS-OLS-RET	2,021438	1,832350
MIDAS-MSE-RET	1,906980	1,573402
MIDAS-RANKS-RET	1,869118	1,517074

3.1.5 Построение моделей с переменной *AGRO_G*

При математическом моделировании ВВП были выбраны следующие спецификации моделей U-MIDAS, MIDAS-GETS, MIDAS-GETSIS: в качестве зависимой переменной выступал временной ряд годовых темпов прироста ВВП *GGDP*. U-MIDAS и MIDAS-GETS включали набор низкочастотных эндогенных переменных, таких так лаг переменной *GGDP* и дамми-переменная

DUM2022Q2, и высокочастотной независимой переменной — временной ряд *AGRO_G* без лага, чтобы учитывать влияние месячного показателя на динамику ВВП. Модель MIDAS-GETSIS имеет такую же спецификацию как и в MIDAS-GETS, но без включения дамми-переменной.

На рисунке 3.10 изображены так называемые «волосатые графики», которые отображают точечные прогнозы, полученные методом расширяющегося окна. В таблице 3.9 представлены показатели точности шести точечных прогнозов, включая такие метрики, как RMSE, MAE, позволяющие оценить качество прогнозов и сопоставить эффективность разных моделей. Высокая точность прогнозов является важнейшим аспектом для принятия обоснованных решений в экономике, где даже незначительные перемены способны значительно повлиять на результаты.



а) б)

в)

Рисунок 3.10 — Наблюдаемые и прогнозные значения *GGDP* на основе моделей:
а) U-MIDAS-AGRO; б) MIDAS-GETS-AGRO; в) MIDAS-GETSIS-AGRO

Таблица 3.9 — Результаты оценки точности краткосрочного прогноза годовых темпов прироста ВВП с помощью моделей MIDAS

Модель	RMSE	MAE
U-MIDAS-AGRO	1,726421	1,397109
MIDAS-GETS-AGRO	1,417249	1,217145
MIDAS-GETSIS-AGRO	2,631435	2,079608

Далее комбинированные прогнозы были сформированы на основе шести точечных прогнозов, что позволяет значительно повысить точность предсказаний. Такой подход позволяет учитывать разнообразие моделей и улучшать общую надёжность результатов. На рисунке 3.11 отображены полученные комбинированные прогнозы, что помогает визуально оценить их динамику.

Однако стоит отметить, что в данном случае комбинирование не дало ожидаемого улучшения в качестве прогнозов и даже отработало хуже, чем отдельные модели. Это может быть связано с тем, что разные модели давали противоречивые результаты, что ухудшало общую точность. Также возможным факторов является неправильное распределение весов при комбинировании. В таблице 3.10 представлены показатели точности, такие как RMSE и MAE, которые подтверждают эти выводы.

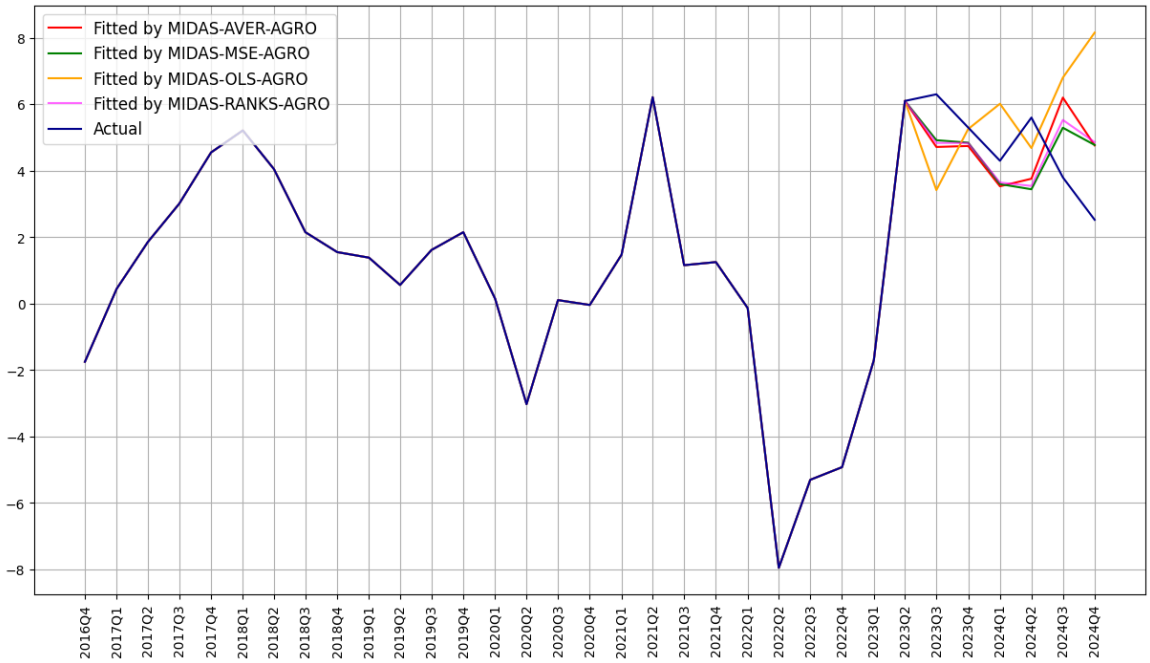


Рисунок 3.11 — Комбинированные прогнозы темпов прироста ВВП

Таблица 3.10 — Результаты оценки точности комбинированных прогнозов моделей MIDAS

Модель	RMSE	MAE
MIDAS-AVER-AGRO	1,711210	1,564621
MIDAS-OLS-AGRO	2,967154	2,365587
MIDAS-MSE-AGRO	1,558026	1,406887
MIDAS-RANKS-AGRO	1,605228	1,450871

3.1.6 Построение моделей с переменной *BLD_SA*

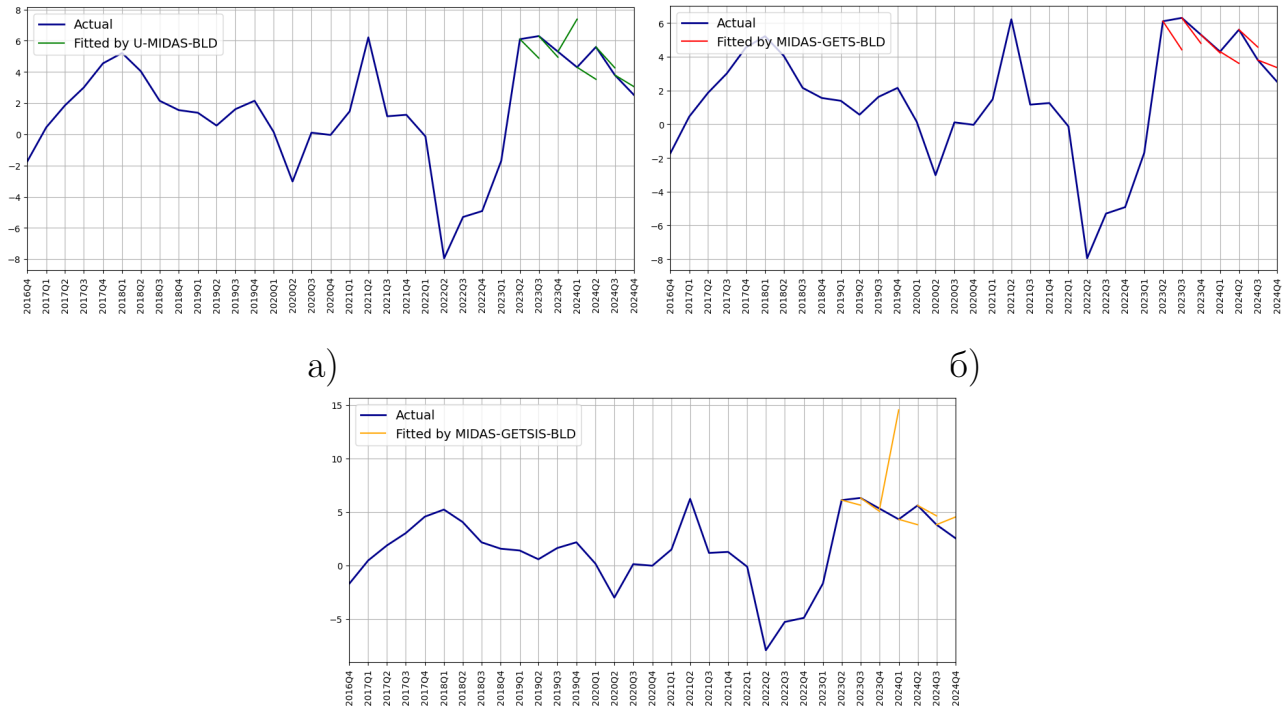
Спецификации моделей U-MIDAS, MIDAS-GETS, MIDAS-GETSIS совпадают с вышеописанными случаями, но в качестве высокочастотной переменной

выступает лаг первого порядка временного ряда *BLD_SA*.

Тем не менее нужно отметить, что при проведении анализа с применением модели MIDAS-GETS-BLD из уравнения была удалена переменная высокой частоты *BLD_SA*. Это решение основывается на итогах диагностических тестов, продемонстрировавших, что указанная переменная не оказывает существенного воздействия на зависимую переменную.

Обоснованием исключения является то, что временной ряд *BLD_SA* не продемонстрировал статистической значимости по отношению к модели. Этот факт может предполагать, что информация по строительству не способна принимать важный вклад в объяснение годовых вариаций темпов прироста ВВП, как это демонстрируется в разделе 2.1. Это также подчёркивает значение выполнения детального анализа переменных до включения в модель. Не все данные являются полезными или актуальными, и добавление лишних переменных может привести к ухудшению результатов. Кроме того, удаление нерелевантных переменных помогает определить приоритеты более важных факторах, которые могут способствовать повышению прогностической способности модели.

На рисунке 3.12 представлены «волосатые графики» построенных прогнозов с помощью моделей MIDAS, что позволяет наглядно оценить динамику и точность прогнозов. В таблице 3.11 собраны показатели точности прогнозов.



а) б)

в)

Рисунок 3.12 — Наблюдаемые и прогнозные значения *GGDP* на основе моделей:
а) U-MIDAS-BLD; б) MIDAS-GETS-BLD; в) MIDAS-GETSIS-BLD

Таблица 3.11 — Результаты оценки точности краткосрочного прогноза годовых темпов прироста ВВП с помощью моделей MIDAS

Модель	RMSE	MAE
U-MIDAS-BLD	1,653827	1,320271
MIDAS-GETS-BLD	1,230865	1,010831
MIDAS-GETSIS-BLD	4,346243	2,633591

В процессе исследования были формированы комбинированные прогнозы, которые помогают компенсировать недостатки каждой из моделей. Однако как и в предыдущем подразделе (подраздел 3.1.5) комбинирование отработало хуже, чем отдельные модели. На рисунке 3.13 отображены полученные комбинированные прогнозы, что помогает визуально оценить их динамику. В таблице 3.12 представлены показатели точности прогнозов, которые подтверждают вывод об ухудшении точности прогнозирования.

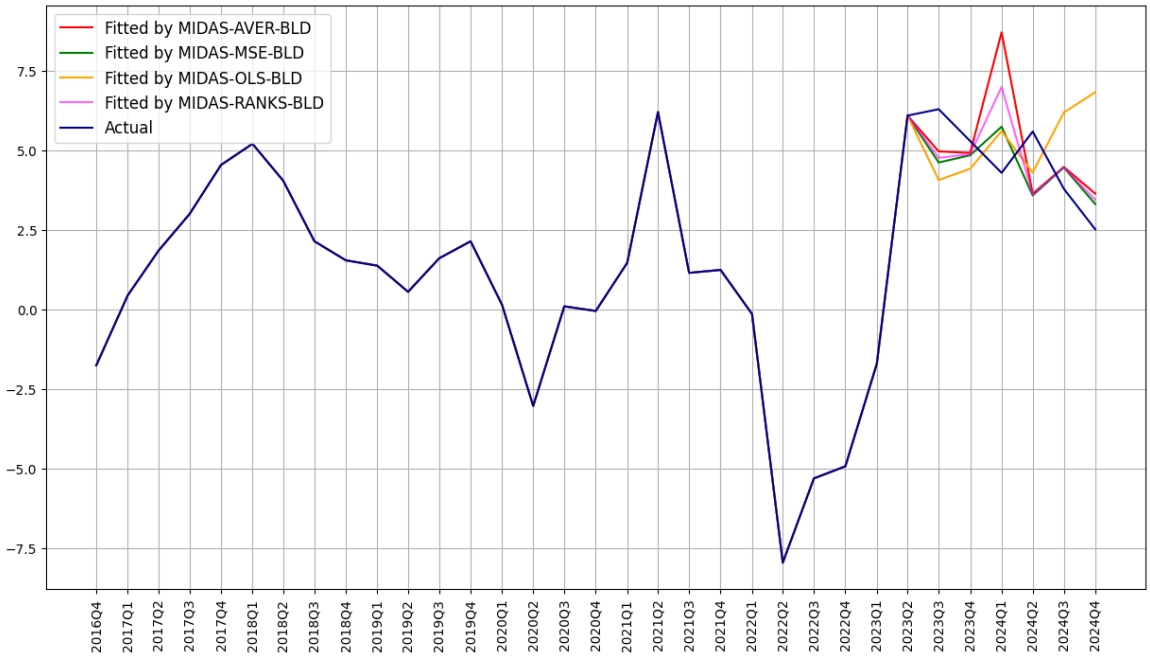


Рисунок 3.13 — Комбинированные прогнозы темпов прироста ВВП

Таблица 3.12 — Результаты оценки точности комбинированных прогнозов моделей MIDAS

Модель	RMSE	MAE
MIDAS-AVER-BLD	2,119604	1,646736
MIDAS-OLS-BLD	2,360006	2,067626
MIDAS-MSE-BLD	1,305781	1,174360
MIDAS-RANKS-BLD	1,583925	1,372195

3.2 Комбинирование прогнозов моделей MIDAS

В данном разделе представлены результаты комбинирования прогнозов, полученных из трёх различных моделей: U-MIDAS, MIDAS-GETS и MIDAS-GETSIS. С помощью каждого из типов MIDAS были построены шесть спецификаций с разными высокочастотными экзогенными переменными. Описание построения данных моделей было приведено в разделе выше.

Комбинирование проводилось отдельно для каждой группы моделей. Это означает, что сначала были сформированы комбинированные прогнозы для моделей U-MIDAS, затем для моделей MIDAS-GETS, и, наконец, для моделей MIDAS-GETSIS. Такой подход позволяет учитывать уникальные характеристики каждой модели и обеспечивать более точные результаты.

Комбинирование было осуществлено путём создания комбинированных прогнозов на основе различных подходов к взвешиванию. В частности, для этой цели были использованы следующие методы: равные веса (MIDAS-AVER), при которых прогнозы объединяются с использованием одинаковых весов для всех моделей; веса, полученные с помощью метода наименьших квадратов (MIDAS-OLS), где веса рассчитываются на основе минимизации суммы квадратов отклонений; обратные значения функции потерь (MIDAS-MSE), при которых веса определяются как обратные значения среднеквадратичной ошибки; и обратные значения рангов функции потерь (MIDAS-RANKS), где веса рассчитываются на основе рангов ошибок прогнозирования [11, с. 21].

Для максимизации точности был выполнен процесс оптимизации, в котором путем систематического анализа были подобраны самые продуктивные сочетания для каждой категории моделей. Это способствовало повышению качества прогнозов, поскольку каждая комбинация задействует преимущества входящих в нее моделей. Выбор оптимальных комбинаций предполагал рассмотрение различных способов усреднения прогнозов, что дало возможность определить наиболее результативные методы для каждой группы.

Лучшие показатели точности прогнозов для модели U-MIDAS обеспечиваются за счёт комбинированного прогноза, построенного на основе прогнозов моделей с высокочастотными переменными, такими как $CESI_SA$, INV_G, PP_G и $AGRO_G$. Для моделей MIDAS-GETS наилучшие результаты достигаются благодаря комбинированию прогнозов, основанных на переменных $CESI_SA$ и INV_G. В случае моделей MIDAS-GETSIS оптимальные показатели обеспечиваются комбинированием прогнозов на основе моделей с экзогенными переменными $CESI_SA$, RET_G, INV_G и BLD_G.

На графиках, представленных на рисунках 3.14-3.16, показаны результаты, полученные с использованием различных методов комбинирования для каждой группы моделей.

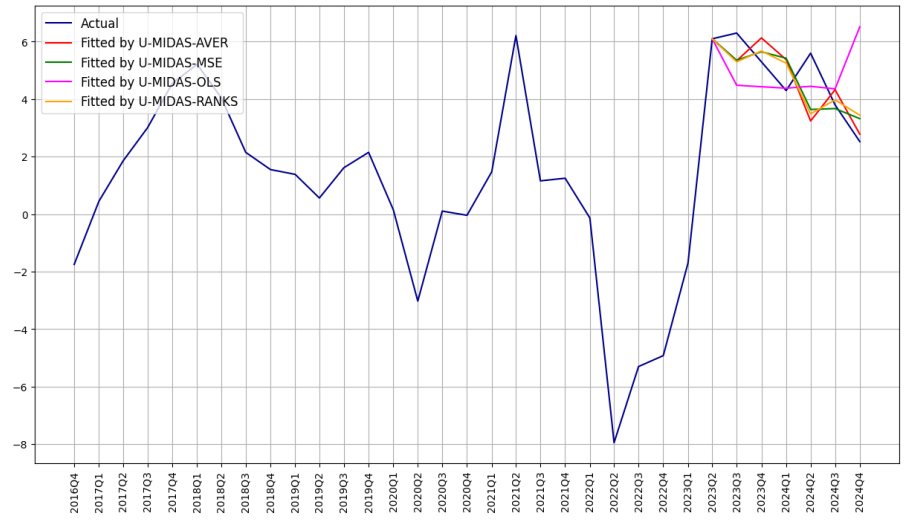


Рисунок 3.14 — Комбинированные прогнозы U-MIDAS моделей

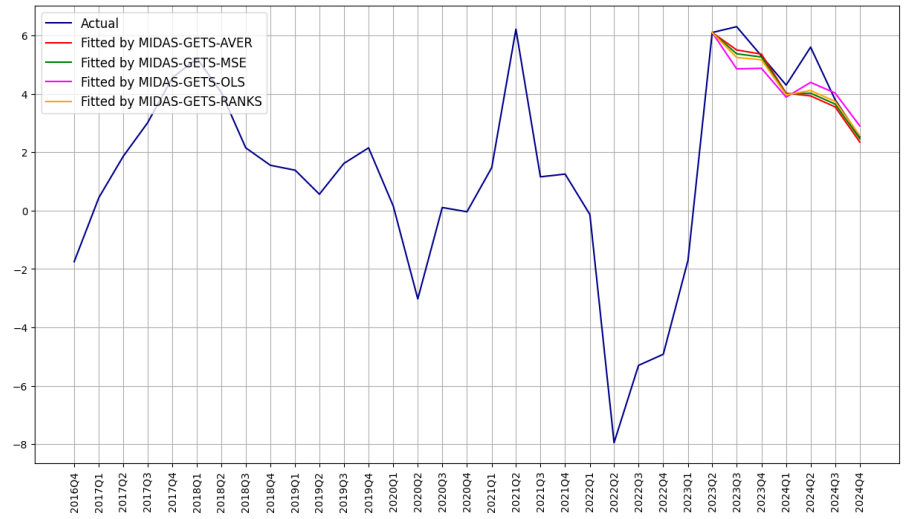


Рисунок 3.15 — Комбинированные прогнозы MIDAS-GETS моделей

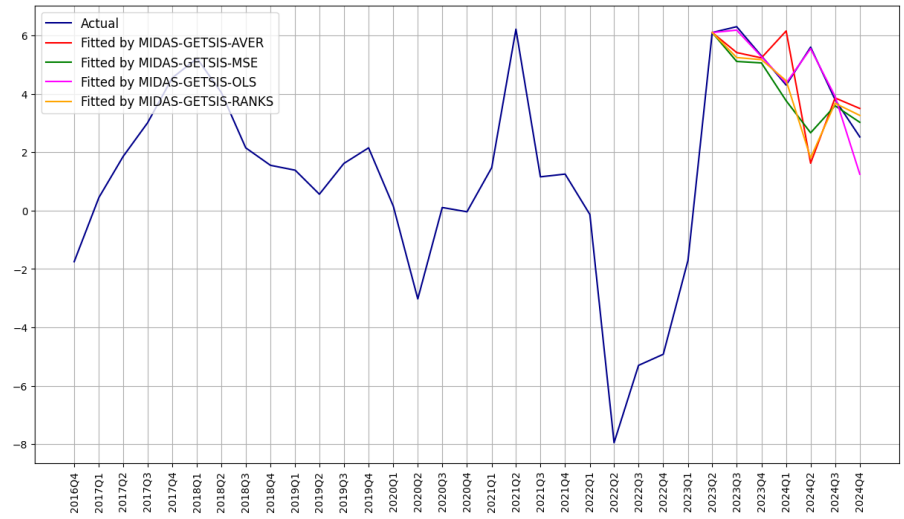


Рисунок 3.16 — Комбинированные прогнозы MIDAS-GETSIS моделей

В таблице 3.13 приведены значения RMSE и MAE для каждого комбинированного прогноза. Эти метрики позволяют оценить качество прогнозов и выявить, насколько успешно каждая модель и комбинированный прогноз справляются с задачей предсказания. RMSE особенно чувствителен к большим ошибкам, что делает его полезным для выявления моделей, которые могут иногда значительно ошибаться. MAE предоставляет более интуитивное представление о средней ошибке, не акцентируя внимание на выбросах.

Таблица 3.13 — Результаты оценки точности комбинированных прогнозов для каждой группы моделей

Модель	RMSE	MAE
U-MIDAS-AVER	1,203146	1,002428
U-MIDAS-OLS	1,900844	1,412285
U-MIDAS-MSE	1,062112	0,883752
U-MIDAS-RANKS	1,108846	0,922419
MIDAS-GETS-AVER	0,774570	0,538316
MIDAS-GETS-OLS	0,825803	0,682538
MIDAS-GETS-MSE	0,760512	0,510534
MIDAS-GETS-RANKS	0,758914	0,517774
MIDAS-GETSIS-AVER	1,870494	1,303127
MIDAS-GETSIS-OLS	0,526760	0,276070
MIDAS-GETSIS-MSE	1,333696	0,935411
MIDAS-GETSIS-RANKS	1,638807	0,998584

В процессе анализа комбинированного прогноза MIDAS-GETSIS-OLS, полученного с использованием взвешивания с помощью МНК, было установлено, что данный прогноз не может быть принят в расчёт. Основной причиной является наличие переобучения модели.

Переобучение происходит, когда модель слишком точно подстраивается под обучающие данные, включая шум и случайные колебания, что приводит к ухудшению её способности обобщать на новых данных. В данном случае, несмотря на высокую точность на обучающем периоде (2023q3-2024q3), в тестовом периоде (2024q4) наблюдается значительное отклонение от реальных значений, что можно наблюдать на рисунке 3.16.

Переобучение происходит по нескольким причинам, такими как [12]:

1. Размер обучающих данных слишком мал и не содержит достаточного количества выборок данных для точного представления всех возможных значений входных данных.
2. Обучающие данные содержат большое количество нерелевантной информации, называемой зашумленными данными.

3. Модель слишком долго обучается на одном наборе выборочных данных.
4. Сложность модели высока, поэтому она изучает шум в обучающих данных.

Таким образом, учитывая указанные аспекты, комбинированный прогноз был исключён из дальнейшего анализа, поскольку его использование могло бы привести к неверным выводам и неправильной интерпретации результатов.

3.3 Построение ARDL моделей

В данном разделе будут представлены результаты построения моделей авторегрессии с распределенными лагами (ARDL). Перед построением моделей было необходимо преобразовать высокочастотные переменные к квартальной частоте. Исходные временные ряды, отражающие объем в среднегодовых ценах, были агрегированы путем суммирования и преобразованы в годовые темпы прироста. Сезонно скорректированные временные ряды, представляющие базисные индексы, были агрегированы с использованием методом усреднения. Это обеспечивает устранение сезонных колебаний и гарантирует, что модели ARDL опираются на стабильные и сопоставимые данные. Далее будут представлены результаты оценки и анализа моделей ARDL, включая оценку их значимости и качество предсказаний.

При построение ARDL моделей для эндогенной переменной *GGDP* было задано число лагов равное 1, где в качестве экзогенных переменных в модели выступают *CESI_SA*, *INV_G*, *RET_G* или *BLD_SA*. Число лагов, которое равняется 2, было задано для зависимой переменной *GGDP* в случае, когда в модели в качестве независимой переменной выступает *PP_G* или *AGRO_G*. Для независимых переменных число лагов приравнивается 4. Данные значения лагов являются оптимальными. Эти значения лагов были определены как оптимальные для данной модели. В дополнение, в модель была включена фиксированная дамми-переменная *DUM2022Q2* для учета специфических эффектов в рассматриваемом периоде.

Для проверки прогностических способностей построенных шести моделей ARDL был использован метод расширяющегося окна, который уже применялся в предыдущих разделах. Этот подход позволяет оценивать эффективность моделей на различных временных интервалах, что особенно важно в контексте экономического анализа. Учитывая динамику экономических процессов, такой метод обеспечивает лучшее понимание поведения моделей в условиях изменяющейся информации. Результаты анализа представлены в виде «волосатых графиков», которые изображены на рисунке 3.17. В таблице 3.14 приведены показатели точности полученных шести точечных прогнозов, включая RMSE и MAE. Эти показатели позволяют сравнить эффективность различных моделей и выбрать наиболее подходящую для дальнейшего использования.

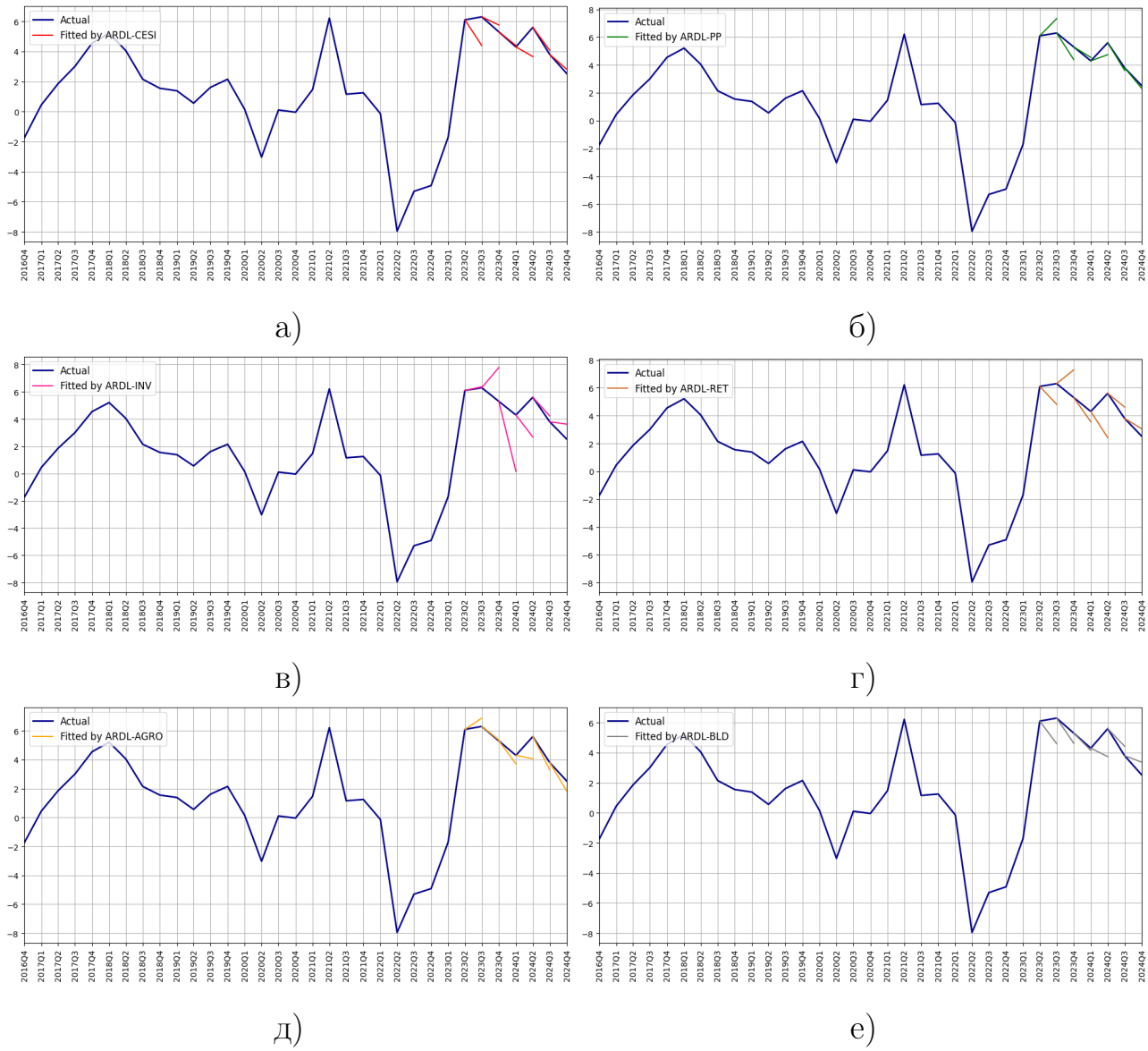


Рисунок 3.17 — Наблюдаемые и прогнозные значение *GGDP* на основе моделей:
а) ARDL-CESI; б) ARDL-PP; в) ARDL-INV; г) ARDL-RET; д) ARDL-AGRO; е) ARDL-BLD

Таблица 3.14 — Результаты оценки точности краткосрочного прогноза годовых темпов прироста ВВП с помощью моделей ARDL

Модель	RMSE	MAE
ARDL-CESI	1,137606	0,827334
ARDL-PP	0,679436	0,572115
ARDL-INV	2,355645	1,859862
ARDL-RET	1,723077	1,459963
ARDL-AGRO	0,791650	0,659611
ARDL-BLD	1,150129	0,976440

На основе построенных моделей ARDL были сформированы комбинированные прогнозы на основе различных подходов к взвешиванию, которые были использованы ранее. Методом перебора была выбрана лучшая комбинация, состоящая из моделей, где в качестве независимых переменных выступают временные ряды *CESI_SA* и *INV_G*. Данные объединённые прогнозы дают возможность учёта различных аспектов и особенностей данных, что позволяет улучшить точность и стабильность предсказаний путём включения в них лучших качеств отдельных прогнозов, полученных с помощью ARDL моделей.

На рисунке 3.18 представлены полученные комбинированные данные. В таблице 3.15 приведены показатели точности, такие как RMSE и MAE, что позволяет оценить эффективность комбинированных прогнозов по сравнению с точечными результатами, полученными от отдельных моделей.

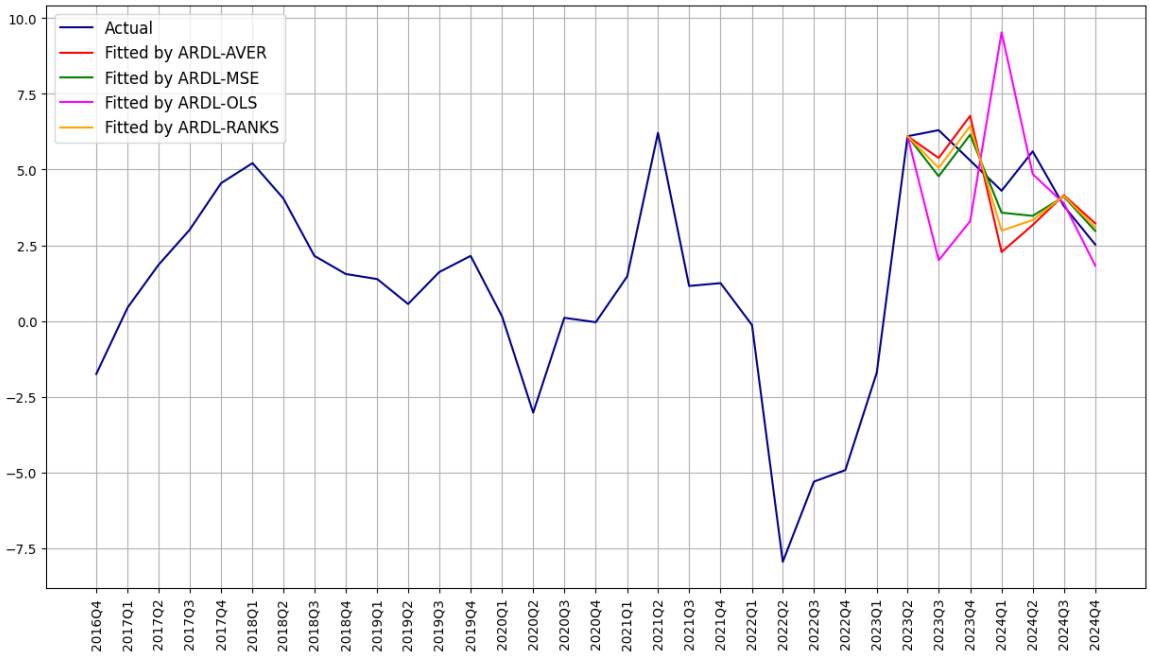


Рисунок 3.18 — Комбинированные прогнозы на основе ARDL моделей

Таблица 3.15 — Результаты оценки точности комбинированных прогнозов на основе моделей ARDL

Модель	RMSE	MAE
ARDL-AVER	1,506768	1,316214
ARDL-OLS	2,912556	2,177425
ARDL-MSE	1,182533	0,996722
ARDL-RANKS	1,299659	1,142657

3.4 Сравнительный анализ точности полученных прогнозов и наукастинг

В данной главе проводится сравнительный анализ точности прогнозов, полученных с помощью различных моделей MIDAS (U-MIDAS, MIDAS-GETS, MIDAS-GETSIS) и ARDL. На основе представленных в предыдущей главе таблиц 3.1-3.15, содержащий метрики RMSE и MAE, была выбрана лучшая модель, представляющая собой комбинированный прогноз MIDAS-GETS-RANKS (р. 3.2), основанный на взвешивании с использованием обратных значений функции потерь. Результаты показали, что данный комбинированный прогноз, основанный на комбинации моделей MIDAS-GETS с высокочастотными переменными $CESI_SA$ и INV_G, продемонстрировал лучшие показатели точности по всем метрикам.

В ходе проведённого исследования были получены прогнозные значения годовых темпов прироста ВВП на 4 квартал 2024 года для данных моделей. При этом в процессе построения моделей в данный момент времени использовалось оценочное значение, рассчитанное Национальным банком Республики Беларусь. В конце марта 2025 года были выложено реальное значение ВВП Беларуси за данный квартал. В таблице 3.16 представлены данные значения годовых темпов прироста ВВП.

Таблица 3.16 — Оценочное, реальное и прогнозные значения годовых темпов прироста ВВП Беларуси в 2024q4

Оценочное значение	Реальное значение	Прогнозное значение, полученное MIDAS-GETS-RANKS
2,5235 %	2,5962 %	2,5672 %

Исходя из представленных данных о годовых темпах прироста ВВП Беларуси, можно сделать несколько выводов. Реальное значение прироста составило 2,5962 %, что немного превышает оценочное значение в 2,5235 %. Это говорит о том, что предварительная оценка НБРБ была более консервативной, возможно, из-за недостатка информации о текущих экономических условиях.

Прогнозное значение, равное 2,5672 %, оказалось достаточно близким к реальному, что свидетельствует о хорошей работе модели. Отклонение между прогнозным и реальным значениями составляет всего 0,029 % в пользу реального значения, что указывает на высокую точность прогноза.

Перейдём к результатам построения прогноза на 1 квартал 2025 года с помощью лучшей модели, где за 4 квартал 2024 года для анализа было использовано реальное значение прироста ВВП. На рисунке 3.19 представлены результаты прогнозирования на 2025q1 с помощью MIDAS-GETS-RANKS.

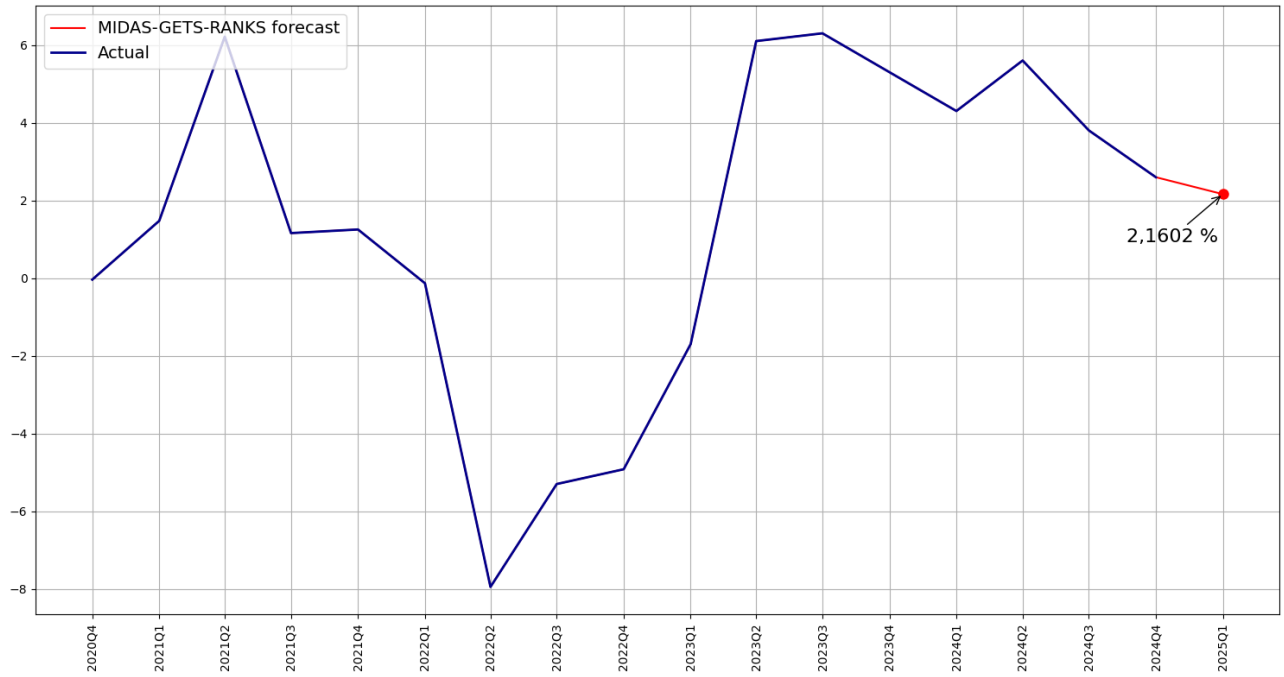


Рисунок 3.19 — Прогноз на 2025q1, построенный с помощью MIDAS-GETS-RANKS

Построенный прогноз (значение 2,1602 %) демонстрирует стабильный рост и отражает благоприятную тенденцию в экономической ситуации в стране.

В результате проведённого сравнительного анализа подтвердилось, что модели на основе смешанных данных MIDAS, превосходят модели ARDL, использующие агрегированные данные. Кроме того, объединение прогнозов является действенным способом для повышения качества предсказаний. Также был построен прогноз годовых темпов прироста ВВП Беларуси на первый квартал 2025 года, что подчёркивает важность применения современных методов прогнозирования в экономическом анализе.

ЗАКЛЮЧЕНИЕ

В данной работе была рассмотрена проблема прогнозирования компонент ВВП на основе регрессионных моделей MIDAS, а также комбинированных прогнозов. В ходе исследования были получены следующие результаты:

1. Подготовлен аналитический обзор эконометрических моделей и методов машинного обучения на основе данных смешанной частоты, а также алгоритмов комбинирования и оценки точности модельных прогнозов.
2. Сформированы временные ряды экономических показателей и проведён их предварительный статистический анализ с целю использования в рассматриваемых эконометрических моделях.
3. Разработаны альтернативные варианты моделей для краткосрочного прогнозирования и наукастинга квартальных темпов роста реального ВВП Республики Беларусь по экономическим переменным с месячной частотой наблюдения на основе регрессионной модели MIDAS по данным смешанной частоты, различающиеся составом используемых экзогенные переменных и методом оценивания.
4. Разработан программный инструментарий, реализующий алгоритмы построения и оценки точности вневыборочных прогнозов на основе одиночных моделей и различных способов комбинирования моделей.
5. Проведён сравнительный анализ точности прогнозов на основе построенных моделей по смешанным и агрегированным данным.

Полученные результаты, включая модели прогнозирования и методы комбинирования прогнозов целевого макроэкономического показателя, целесообразно использовать в качестве одного из инструментов для квартального прогнозирования опережающей оценки целевого показателя до получения его официальных значений.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Интерактивная информационно-аналитическая система распространения официальной статистической информации. [Электронный ресурс]. – https://dataportal.belstat.gov.by/osids/home-page – Дата доступа: 17.02.2025.
2. Макеева, Н.М. Наукастинг элементов использования ВВП России / Н.М. Макеева, И.П. Станкевич – Статья 2022/10, Экономический журнал ВШЭ.
3. Малюгин В. И. Анализ многомерных статистических моделей с неоднородной структурой в случае скрытой марковской зависимости состояний / Малюгин В. И., Новопольцев А. Ю. – Минск: НИЭИ, 2015. – С. 26-36.
4. Малюгин В. И. Краткосрочное прогнозирование и наукастинг темпов роста инфляции на основе моделей по смешанным данным. – Банковский вестник, 2024. – с. 23-36.
5. Национальный банк Республики Беларусь. Мониторинг предприятий реального сектора экономики Республики Беларусь. [Электронный ресурс]. – https://www.nbrb.by/publications/monitoringpredpriyatij/mp_2021_4.pdf – Дата доступа: 05.03.2025.
6. Национальный статистический комитет Республики Беларусь. Валовый внутренний продукт и методы его расчёта. [Электронный ресурс]. – https://www.belstat.gov.by/upload-belstat/upload-belstat-word/Metod_pologenija/VVP_27_02_2018.doc – Дата доступа: 06.03.2025.
7. Национальный статистический комитет Республики Беларусь. Календарь пользователя. [Электронный ресурс]. – https://www.belstat.gov.by/calendar/ – Дата доступа: 17.02.2025.
8. Национальный статистически комитет Республики Беларусь. Основные понятия системы национальных счетов. [Электронный ресурс]. – https://www.belstat.gov.by/upload-belstat/upload-belstat-word/Methodology/m7_sns_29_07_2017.doc – Дата доступа: 06.03.2025.
9. Станкевич И.П. Сравнение методов наукастинга макроэкономических индикаторов на примере российского ВВП – Прикладная эконометрика 2020. – С. 113-127.
10. Харин, Ю. С. Теория вероятностей, математическая и прикладная статистика / Ю. С. Харин, Н. М. Зуев, Е. Е. Жук. – Минск : БГУ, 2011.
11. Цукарев Т. Алгоритмы машинного обучения для краткосрочного прогнозирования темпов прироста реального ВВП / Т. Цукарев, К. Погосян, К. Лемба. – ЕФСР, 2024. – 33 с.
12. Что такое переоснащение? [Электронный ресурс]. – https://aws.amazon.com/ru/what-is/overfitting/ – Дата доступа: 23.02.2025.
13. Aiolfi, M. Persistence in Forecasting Performance and Conditional Combination Strategies. / Aiolfi, M. and A. Timmermann – Journal of Econometrics, Vol. 135, 2006. — pp. 31-53.
14. Bauwens, L. General-to-specific modelling of exchange rate volatility: A forecast evaluation. / Bauwens, L., Sucarrat, G. – International Journal of Forecasting

Elsevier, Vol. 26, Iss. 4, 2010. – pp. 885-907.
15. Claeskens G. The forecast combination puzzle: A simple theoretical explanation. / Claeskens G., Magnus J.R., Vasnev A.L., Wang W. — International Journal of Forecasting, Vol. 32, Iss. 3, 2016. – pp. 754-762.
16. Clemen R.T. Combining forecasts: A review and annotated bibliography. – International Journal of Forecasting, Vol. 5, Iss. 4, 1989. – pp. 559-583.
17. Dickey D. A. Distribution of the Estimators for Autoregressive Time Series With a Unit Root / Dickey D. A., Fuller W. A. – Journal of the American Statistical Association, Vol. 74, No. 366, 1979. – pp. 427-431.
18. Foroni, C. Unrestricted Mixed Data Sampling (U-MIDAS): MIDAS Regressions With Unrestricted Lag Polynomials / C. Foroni, M. Marcellino, C. Schumacher – Journal of the Royal Statistical Society Series A: Statistics in Society, Vol. 178, Iss. 1, January 2015. – pp. 57-82.
19. Ghysels, E. MIDAS Regressions: Further Results and New Directions. / Ghysels, E., Sinko, A., Valkanov, R. – Econometric Reviews, Vol. 26, Iss. 1. – pp. 53-90.
20. Ghysels, E. The MIDAS touch: Mixed data sampling regression models / Ghysels, E., Santa-Clara P., Valkanov R. – Working paper, UNC and UCLA.
21. Kuzin, V. MIDAS vs. Mixed-Frequency VAR: Nowcasting GDP in the Euro Area / V. Kuzin, M. Marcellino, C. Shumacher – EUI Working Paper.
22. Makridakis S. The M5 accuracy competition: Results, findings and conclusions. / Makridakis S., Spiliotis E., Assimakopoulos V. – International Journal of Forecasting, Vol. 38, 2022. – pp. 1346-1364.
23. Perron, P. The Great Crash, the Oil Price Shock and the Unit Root Hypothesis: Erratum / Perron, P., Vogelsang, T.J. – Econometrica, Vol. 61. – pp. 248-249.
24. Soybilgen, B. Nowcasting the New Turkish GDP / B. Soybilgen, E. Yazgan – Economics Bulletin, Vol. 38, Iss. 2. – pp. 1083-1089.
25. Stock, J.H., Forecasting Output and Inflation: The Role of Asset Prices. / Stock, J.H., M.W. Watson – NBER Working Paper No. 8180, 2001.

ПРИЛОЖЕНИЕ А

Характеристики моделей MIDAS-GETS-CESI и MIDAS-GETSIS-CESI

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.947379	0.360125	2.630693	0.0139
Q\DUM2022Q2	-6.398949	1.940342	-3.297846	0.0027
Q\GGDP_SA(-1)	0.501948	0.103360	4.856287	0.0000
Page: M_SA Series: CESI_SA				
LAG3	0.264208	0.083772	3.153899	0.0039
LAG11	-0.164420	0.065019	-2.528807	0.0176

Рисунок А1 — Характеристики модели MIDAS-GETS-CESI

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.555798	0.225804	2.461414	0.0231
Q\GGDP_SA(-1)	0.379045	0.062368	6.077553	0.0000
@ISPERIOD("2020Q...	5.795474	1.712548	3.384124	0.0029
@ISPERIOD("2020Q...	7.709760	1.328033	5.805399	0.0000
@ISPERIOD("2021Q...	3.213689	0.973409	3.301478	0.0036
@ISPERIOD("2022Q...	-4.643942	0.992745	-4.677881	0.0001
@ISPERIOD("2023Q...	2.786079	1.020568	2.729931	0.0129
Page: M_SA Series: CESI_SA				
LAG3	0.415174	0.108673	3.820411	0.0011
LAG4	0.375221	0.116020	3.234112	0.0042
LAG8	-0.657890	0.137127	-4.797675	0.0001
LAG9	0.487825	0.133652	3.649970	0.0016
LAG12	-0.206029	0.040740	-5.057170	0.0001

Рисунок А2 — Характеристики модели MIDAS-GETSIS-CESI

ПРИЛОЖЕНИЕ Б

Метод расширяющегося окна

Данный программный код реализует метод расширяющегося окна для моделей MIDAS с переменной *CESI_SA*.

```
table metrics_table
metrics_table(1,1)="Model"
metrics_table(1,2)="RMSE"
metrics_table(1,3)="MAE"
metrics_table(1,4)="MAPE"
smpl @all

series copy = q\ggdp_sa
series U-MIDAS-CESI = NA
series MIDAS-GETS-CESI = NA
series MIDAS-GETSIS-CESI= NA

!n=0
for !i=0 to 5
 smpl 2016q4 2023q2+!i

 equation cesi_umidas.midas(midwgt=umidas, fixedlag=12)
           q\ggdp_sa c q\dum2022q2 q\ggdp_sa(-1) @ m_sa\cesi_sa
 equation cesi_gets.midas(fixedlag=12, midwgt=autogets)
           q\ggdp_sa c q\dum2022q2 q\ggdp_sa(-1) @ m_sa\cesi_sa
 equation cesi_getsis.midas(fixedlag=12, midwgt=autogets,iis)
           q\ggdp_sa c q\ggdp_sa(-1)  @ m_sa\cesi_sa

 smpl 2023q2+!i+1 2023q2+!i+1

 cesi_umidas.forecast temp1
 cesi_gets.forecast temp2
 cesi_getsis.forecast temp3

 U-MIDAS-CESI = temp1
 MIDAS-GETS-CESI = temp2
 MIDAS-GETSIS-CESI = temp3

 d temp1 temp2 temp3
 !n=!n+1
next !i
```

```
smpl 2023q2 2023q2
U-MIDAS-CESI =q\ggdp_sa
MIDAS-GETS-CESI = q\ggdp_sa
MIDAS-GETSIS-CESI = q\ggdp_sa

smpl 2023q3 2024q4

%eqlist1 = "U-MIDAS-CESI MIDAS-GETS-CESI MIDAS-GETSIS-CESI"

!r=1
for %j {%eqlist1}
 scalar rmse = @rmse(copy, {%j})
    scalar mae = @mae(copy, {%j})
    scalar mape = @mape(copy,{%j})
    metrics_table(!r + 1, 1)=%j
     metrics_table(!r + 1, 2) = rmse
      metrics_table(!r + 1, 3) = mae
      metrics_table(!r + 1, 4) = mape
      !r=!r+1
next

d copy rmse mae mape
metrics_table.display
smpl 2016q4 2025q1
```