

Наукастинг элементов использования ВВП России¹

Макеева Н.М., Станкевич И.П.

В статье рассматривается вопрос оперативной оценки (наукастинга) текущих темпов роста ВВП России и его компонентов по использованию на квартальных данных. Проводится сравнение качества работы следующих моделей: ограниченные и неограниченные MIDAS-модели (модели со смешанными данными), MIDAS-модель с L1-регуляризацией и MFVAR-модель (байесовская векторная авторегрессия смешанной частоты). Результаты сравниваются с классической авторегрессией для обоснования необходимости использования моделей наукастинга для оперативной оценки макроэкономических показателей. В качестве объясняющих переменных использованы индексы производства по разным отраслям и макропоказатели, характеризующие ВВП России по использованию и его компоненты. В работе предложен способ оперативной оценки текущего состояния экономики, предложен метод наукастинга на основе данных только за первый или за первые два месяца рассматриваемого квартала. В результате для каждой зависимой переменной выбирается лучшая для построения наукаста модель по последним 12 точкам на основе критерия средней абсолютной ошибки (MAE) и среднеквадратичной ошибки прогноза (RMSE).

Ключевые слова: наукастинг; ВВП России; временные ряды; прогнозирование; модели со смешанной частотой; MIDAS-модели.

DOI: 10.17323/1813-8691-2022-26-4-598-622

¹ В работе использованы результаты проекта «Национальный центр научно-технологического и социально-экономического прогнозирования», выполненного в рамках Программы фундаментальных исследований НИУ ВШЭ в 2022 году.

Макеева Наталья Михайловна – магистр экономики, преподаватель Департамента прикладной экономики Факультета экономических наук; стажер-исследователь Научно-учебной лаборатории макроструктурного моделирования экономики России Факультета экономических наук Национального исследовательского университета «Высшая школа экономики». E-mail: nmakeeva@hse.ru
Станкевич Иван Павлович – к.э.н., старший преподаватель Департамента прикладной экономики Факультета экономических наук; стажер-исследователь Научно-учебной лаборатории макроструктурного моделирования экономики России Факультета экономических наук Национального исследовательского университета «Высшая школа экономики». E-mail: istankevich@hse.ru

Статья поступила: 07.02.2022/Статья принята: 27.10.2022

Для цитирования: Макеева Н.М., Станкевич И.П. Наукастинг элементов использования ВВП России. *Экономический журнал ВШЭ*. 2022; 26(4): 598–622.

For citation: Makeeva V., Stankevich I. Nowcasting of the Components of Russian GDP. *HSE Economic Journal*. 2022; 26(4): 598–622. (In Russ.)

1. Введение

В истории современной России можно выделить ряд основных периодов с разными предпосылками для экономического роста. С начала двухтысячных годов и до 2008 г. наблюдались достаточно высокие и стабильные темпы роста ВВП России на фоне восходящего тренда на рост цен на энергоносители на мировых рынках (основные для Российской Федерации – нефть и газ). Вместе с этим в стране наблюдался рост реальных денежных доходов граждан, значительный приток иностранных инвестиций и рост капитализации банковского сектора. Восходящий тренд был нарушен мировым финансовым кризисом 2007–2008 гг.

После кризиса 2007–2008 гг. начался процесс восстановления, но уже в 2014–2015 гг. произошедший в России валютный кризис привел к существенной корректировке денежно-кредитной макроэкономической политики Банка России: был осуществлен переход к свободно плавающему валютному курсу, а основной целью политики Центрального банка стало таргетирование инфляции. Более того, пандемия коронавируса 2020 г. наложила свой отпечаток на динамику темпов экономического роста по всему миру. Сегодня, как и в предыдущие периоды, перед государственными органами и частным сектором стоит вопрос оперативной оценки текущего состояния экономики. Официальная статистика, например, по ВВП и его компонентам, выходит с квартальной задержкой, а далее еще и корректируется, что повышает важность проблемы оперативной оценки макроэкономических показателей по уже имеющимся данным. В современной исследовательской практике для решения данной проблемы используется наукастинг – методика оценки текущих уровней показателей на основе более оперативных данных.

Данная работа дополняет библиотеку отечественных работ, посвященных сравнению методов наукастинга, являясь при этом первым исследованием по построению оценочных моделей для элементов использования ВВП России. Целью данной работы является формирование системы оценки элементов использования ВВП России на основе данных более высокой частоты, публикующихся на ежемесячной основе. Полученная модель для построения наукастов позволит более оперативно по сравнению с официальной статистикой получать оценку интересующих показателей по итогам квартала, что важно как для государственных органов при формировании кратко- и среднесрочного плана действий, так и для крупных коммерческих организаций при оценке результатов рассматриваемого периода.

Данное исследование организовано следующим образом. Раздел 1 – Введение. В разделе 2 содержится полноценный обзор литературы по работам схожей тематики. Раздел 3 посвящен этапам отбора показателей, рассматриваемым в качестве регрессоров к элементам использования ВВП России. Также в этой части работы описана процедура корректировки временных рядов на сезонность и методы оценки качества получаемых наукастов. В разделе 4 рассматривается математический аппарат, применяемый для по-

строения краткосрочных прогнозов. Раздел 5 содержит в себе практические результаты: для каждого из зависимых показателей выбрана лучшая спецификация модели для построения наукастов. В Заключение (раздел 6) содержатся основные результаты работы и полученные выводы, а также рекомендации для дальнейших исследований.

2. Обзор литературы

Существует несколько подходов к наукастингу. Например, в работе [Ingenito et al., 1996] рассматривается эконометрический подход к наукастингу: авторы используют связующие уравнения для оценки интересующих макроэкономических показателей. Исследователи используют ряд более высокой частоты с последующим преобразованием в ряд более низкой частоты и включением в VAR-модель (Vector AutoRegression Model). При этом выбор лучшей спецификации модели осуществлялся с помощью критерия RMSE (Root Mean Square Error). Второй популярный класс эконометрических моделей – это MIDAS-модели (Mixed Data Sampling Models) или модели с данными смешанной частоты. В работе [Schumacher, 2016] сравнивается качество работы двух связующих уравнений и MIDAS-моделей и показывается, что с точки зрения точности наукастов качество моделей является сопоставимым, однако MIDAS-модели менее требовательны к массиву данных, подаваемому на вход. Третий тип моделей, традиционно используемых для наукастинга, – динамические факторные модели (DFM-модели), представленные в исследовании [Soybilgen, Yazgan, 2018]. В данной работе рассматриваются три уже описанные модели, причем в качестве DFM-модели берется динамическая факторная модель с двумя шагами: для определения первоначальных общих факторов используется метод главных компонент, параметры самой модели оцениваются через метод наименьших квадратов, при этом к полному набору данных применяется метод сглаживания Калмана. Динамические факторные модели также используются в работе [Proietti et al., 2021] для наукастинга шестнадцати компонентов ВВП Италии в месячных темпах роста с помощью одновременного суммирования добавленной стоимости различных отраслей экономики. Простота использования и эффективность DFM-модели в более спокойные с точки зрения макроэкономической ситуации времена рассматривается в работе [Зубарев, Рыбак, 2021]: авторы тестируют модель на отрезке с 2016 г. по 2021 г., отмечая преимущество используемого метода для наукастинга ВВП России перед прогнозами Минэкономразвития России. При этом DFM-модель оказывается менее эффективна в период пандемии, но, тем не менее, авторам удалось достаточно точно предсказать восстановление выпуска в III квартале 2020 г.

В работе [Clements, Galvao, 2009] оценивается сила прогноза опережающих показателей в среднесрочной перспективе (до года). Показывается, что одним из преимуществ MIDAS-модели является возможность более точного построения прогнозов за счет комбинирования опережающих переменных и показателей, выходящих на ежемесячной основе. В работах [Ghysels et al., 2006; 2007] MIDAS-модели с линейным регрессором используются для прогнозирования движения рынка (волатильности) в разрезе интервалов в пять минут. В работах сравниваются дневная, недельная, двухнедельная, трехнедельная и месячная волатильности на основе дневных данных, далее с помощью критерия MSE (Mean Square Error) выбирается наиболее точная модель из рассмотренных. Результаты, полученные в данных исследованиях, находят свое подтверждение в работе [Chernis, Sek-

kel, 2017]: авторы строят наукаст для квартальных данных по ВВП Канады по 23 индикаторам, включая три индекса цен на неэнергетические товары Банка Канады, цену на нефть WTI и Индекс деловой активности (PMI). Эффективность использования высокочастотных данных рассматривается в работе [Третьяков, Фокин, 2021]. С помощью рядов недельной частоты номинального обменного курса рубля к доллару, ставки межбанковского кредитования MIACR, цены на нефть и нескольких спецификаций MIDAS-моделей в статье строится наукастинг российской инфляции. Авторы доказывают, что использование высокочастотных данных повышает точность краткосрочного прогноза по сравнению с использованием данных только месячной периодичности.

Существует большое число модификаций MIDAS-моделей. В работе [Ferrara, Marsilli, 2019] используется FA-MIDAS-модель (Factor-Augmented MIXed DATA Sampling Model) для оценки динамики роста мирового ВВП. Суть данных моделей заключается в объединении MIDAS-модели, преимуществом которой является возможность прогнозирования низкочастотных данных через данные более высокой частоты, и факторного анализа, что позволяет сократить размерность рассматриваемой модели. В работе [Marcellino, Schumacher, 2010] также используется FA-MIDAS-модель, позволяющая использовать несбалансированные наборы данных. В работе [Jardet, Meunier, 2022] авторы исследуют вопрос улучшения качества краткосрочного прогноза мирового ВВП в период коронакризиса с помощью 718 месячных, 255 еженедельных рядов и FA-MIDAS-модели. Разработанный исследователями подход позволяет получить годовую модель роста мирового ВВП, близкую по результатам к прогнозам МВФ и ОЭСР, но с опережением на 1–3 месяца.

Более продвинутыми с технической точки зрения среди моделей, используемых для наукастинга, являются MFVAR-модели (Mixed Frequency Vector Autoregression). Данный класс моделей фактически является способом описания динамики переменных с более высокой частотой наблюдений и одновременно ненаблюдаемого разложения данных с низкой частотой в высокочастотные данные. Такой способ оценки макропоказателей рассматривается в статье [Kuzin et al., 2011]: авторы оценивают темпы роста ВВП зоны евро с помощью 20 регрессоров. Результатом исследования становится сопоставление прогнозной силы MIDAS и MFVAR-моделей через критерий MSE [Kuzin et al., 2011]. В работе [Schorfheide, Song, 2015] изучается прирост точности краткосрочного прогноза в момент выхода новых данных и уточнения уже включенных в модель внутри рассматриваемого временного интервала. Показывается, что такая новая информация значительно улучшает качество прогнозов.

Помимо стандартных VAR-моделей, для наукастинга достаточно часто используются и байесовские VAR-модели, в том числе модели смешанной частоты, такие как MFBVAR-модель (Mixed-Frequency Bayesian VAR Model) [McCracken et al., 2015], изначально предложенные в работах [Litterman, 1979; 1986].

Из популярных сегодня методов машинного обучения для наукастинга достаточно часто применяются методы снижения размерности и автоматического отбора переменных. Так, например, в работе [Tiffin, 2016] сравниваются результаты, полученные двумя достаточно популярными методами – регрессией с упругой сетью и алгоритмом случайных лесов. В работе [Serpi et al., 2019] проводится анализ качества методов машинного обучения для краткосрочного прогнозирования ВВП для стран с развивающейся экономикой: L1- и L2-регуляризации, метода главных компонент и LARS. Эффективность методов машинного обучения также доказывается и в статье [Richardson et al., 2021]: авторы

строят краткосрочный прогноз роста ВВП Новой Зеландии. Результаты показывают, что методы машинного обучения оказываются более точными по сравнению с авторегрессионными моделями и динамическими факторными моделями, что позволяет улучшить официальные прогнозы Резервного Банка Новой Зеландии. Для повышения точности краткосрочных прогнозов в работе [Ferrara, Simoni, 2022] используются альтернативные наборы данных в моделях с применением элементов машинного обучения. Авторы добавляют данные поисковых запросов Google для прогнозирования ВВП разных стран в различные экономические периоды. Результатом работы являются теоретическое обоснование необходимости включения альтернативных данных в прогнозы, а также замечание об эффективности включения таких данных в периоды рецессий и стабильности в макроэкономике. Исследование [Майорова, Фокин, 2021] сравнивает методы машинного обучения (Elastic Net, Random Forest, XGBoost и SSVS) и классическую ARIMA-модель для наукастинга долларовой стоимости экспорта и импорта по товарным группам. Авторы доказывают превосходство взвешенной модели методов машинного обучения и близость ее результатов к историческим прогнозам Банка России, построенным в сопоставимых условиях.

Исследование [Nakazawa, 2022] рассматривает вопрос использования альтернативных данных для наукастинга макроэкономических показателей. В качестве временных рядов автор использует оценки Elastic Net с использованием еженедельных данных о розничных продажах и ежедневных данных о поисковых запросах в сети Интернет. Лучший наукаст выбирается с помощью метода комбинированного прогноза, направленного на минимизацию ошибок прогноза. Описанный в работе подход показывает большую эффективность по сравнению со стандартными методами в кризисные времена (например, после распространения вируса COVID-19). Для повышения точности краткосрочных прогнозов в работе [Ferrara, Simoni, 2022] используются альтернативные наборы данных в моделях с применением элементов машинного обучения. Авторы добавляют данные поисковых запросов Google для прогнозирования ВВП разных стран в различные экономические периоды. Результатом работы являются теоретическое обоснование необходимости включения альтернативных данных в прогнозы, а также замечание об эффективности включения таких данных в периоды рецессий и стабильности в макроэкономике. Статья [Marcellino, Sivec, 2021] рассматривает вопрос наукастинга ВВП малой открытой экономики в волатильные периоды на примере Люксембурга. Авторы включают как классические для краткосрочного прогнозирования показатели, так и альтернативные (например, данные продаж топлива с АЗС, Google Trends, данные по краткосрочной государственной помощи и о регистрации новых автомобилей). Исследование показывает, что динамические факторные модели со смешанной частотой и нейронные сети работают хорошо как в абсолютном, так и в относительном выражении по отношению к эталонной авторегрессионной модели. Преимущество рассматриваемых моделей особенно заметно в кризисные периоды, например, в период пандемии COVID-19. Для оценки эффективности включения информации о поисковых запросах в дополнение к официальной статистике при наукастинге макроэкономических переменных в работе [Лазарян, Герман, 2018] используется динамическая факторная модель. Авторами показано, что добавление в модель данных о частоте поисковых запросов не оказывает влияния на точность прогнозов DFM-моделей, построенных на основе официальной экономической статистики, однако обе модели показали лучшее качество прогнозов в сравнении с AR(1)-моделью.

В исследовательской практике вопрос оценки макроэкономических показателей России изучен довольно слабо. Например, в статье [Микош, Соланко, 2019] рассматривается вопрос наукастинга российского ВВП с помощью достаточного разнообразия моделей: связующие уравнения, MIDAS-модели, U-MIDAS-модели. Также авторы рассматривают возможность как включения только данных одной периодичности, так и переменные смешанной частоты, делая акцент на выборе наиболее важных для краткосрочного прогнозирования ВВП переменных. Работа [Поршаков и др., 2016] основана на наукастинге темпов роста ВВП России с помощью DFM-моделей, моделей случайного блуждания, связующих уравнений и разделения факторов на четыре основные группы: «опережающие показатели», «показатели реального сектора», «финансовые показатели» и «показатели внешнего сектора». Основным результатом работы является следующий вывод: качество прогноза моделей улучшается по мере выхода промежуточной статистики и переоценки уже полученных данных в рамках рассматриваемого периода (в данном случае – квартала), что соотносится с выводом, полученным в работе [Schorfheide, Song, 2015]. Интересной с точки зрения техники является статья [Яковлева, 2018]: автор строит наукаст на основе отражающих динамику отечественной экономической активности высокочастотных переменных через применение методов текстового анализа и машинного обучения, направленных на поиск наиболее релевантных регрессоров в новостных источниках. В работе [Станкевич, 2020] сравниваются методы для наукастинга макропоказателей на примере ВВП России: различные вариации MIDAS-моделей и MFBVAR-модели с априорным распределением Миннесоты. Исследователь приходит к выводу о том, что MFBVAR-модели превосходят по точности другие рассматриваемые модели, а также предлагается на основе данного типа моделей построить месячный прогноз квартального показателя – темпов роста российского ВВП. Работа [Zhemkov, 2021] была опубликована Банком России в рамках повышения доверия к публикуемым прогнозам и оценкам: используется метод комбинирования прогнозов, который заключается в усреднении полученных несколькими способами результатов, для прогнозирования темпов экономического роста с помощью нескольких типов моделей: ARMA-модели, DFM-модели, DMA-модели, DMS-модели и FAVAR-модели. Автор приходит к выводу о том, что использование комбинированного прогноза оказывается незначительно лучше по точности по сравнению с использованием только одной модели, поэтому он видит большое пространство для доработки своей методики получения наукастов.

3. Данные и методы оценки качества моделей

Для оценки качества наукастов будут использованы два наиболее популярных критерия: средняя абсолютная ошибка (MAE) и корень из среднеквадратической ошибки (RMSE). Они рассчитываются по формулам

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

где Y_t – фактическая наблюдаемая зависимая переменная (в рамках данной работы – прирост к соответствующему периоду прошлого года); \hat{Y}_t – аналогичное значение, полученное с помощью модели; T – количество периодов, на которых тестируется модель.

При построении наукастов не учитывается информация о последнем доступном квартале – перед оцениванием модели из выборки удаляются значения зависимых переменных и соответствующие данному кварталу месячные значения объясняющих переменных. Далее в выборку возвращаются удаленные значения регрессоров и для них рассчитывается прогнозное значение зависимой переменной (наукаст). Рассматриваемые модели сравниваются по последним 12 точкам. Это вызвано тем, что последние 12 точек соответствуют периоду с III квартала 2018 г. и по начало 2022 г. Таким образом, модели тестируются на достаточно разнородных данных – в тестирование попадают как относительно спокойный период 2018–2019 гг., так и кризисные периоды, что особенно актуально в связи с коронакризисом 2020 г.

Построение моделей наукастинга и дальнейший выбор лучшей спецификации модели осуществляется на основе квартальных данных ВВП по использованию, публикуемых Федеральной службой государственной статистики Российской Федерации. Модели оцениваются за период с I квартала 2001 г. по I квартал 2022 г. Для получения полных временных рядов имеющийся набор данных с базовыми ценами 2016 г. дополнительно были продлены данными за более ранние периоды с другими базовыми ценами с необходимым пересчетом. Зависимые переменные имеют квартальную частотность, а объясняющие – месячную. Все переменные были переведены в темпы роста к соответствующему периоду прошлого года для исключения сезонной компоненты и обеспечения стационарности показателей.

ВВП и его компоненты по использованию рассматриваются как зависимые переменные в рамках данного исследования (табл. 1).

Таблица 1.

Зависимые переменные и календарь их выхода

Переменная	Источник	Описание	Календарь выхода
Real GDP: YoY	Росстат	ВВП России по использованию	Задержка 3 месяца (за I квартал первая оценка выходит в начале июля) + корректировка. Предварительная справка выходит с задержкой в 1,5 месяца (во второй половине ноября выходит за III квартал)
Households Consumption: YoY	Росстат	Расходы на конечное потребление домашних хозяйств	Задержка 3 месяца (за I квартал первая оценка выходит в начале июля) + корректировка
Government Consumption: YoY	Росстат	Расходы на конечное потребление государственного управления	Задержка 3 месяца (за I квартал первая оценка выходит в начале июля) + корректировка
Non-commercial Consumption: YoY	Росстат	Расходы на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства	Задержка 3 месяца (за I квартал первая оценка выходит в начале июля) + корректировка

Окончание табл. 1.

Переменная	Источник	Описание	Календарь выхода
Gross Capital Formation: YoY	Росстат	Валовое накопление основного капитала	Задержка 3 месяца (за I квартал первая оценка выходит в начале июля) + корректировка
Exports: YoY	Росстат	Экспорт	Задержка 3 месяца (за I квартал первая оценка выходит в начале июля) + корректировка
Imports: YoY	Росстат	Импорт	Задержка 3 месяца (за I квартал первая оценка выходит в начале июля) + корректировка

Объясняющие переменные были подобраны как основные показатели, отражающие макроэкономическую ситуацию в стране (табл. 2).

Таблица 2.

Объясняющие переменные и календарь их выхода

Переменная	Источник	Описание	Календарь выхода
Industry: YoY	Доклад «Социально-экономическое положение России», Росстат	Индекс промышленного производства	Задержка 1 месяц (в конце августа выходит за июль)
Mining: YoY	Доклад «Социально-экономическое положение России», Росстат	Индекс производства по добыче полезных ископаемых	Задержка 1 месяц (в конце августа выходит за июль)
Production: YoY	Доклад «Социально-экономическое положение России», Росстат	Индекс обрабатывающего производства	Задержка 1 месяц (в конце августа выходит за июль)
Energy: YoY	Доклад «Социально-экономическое положение России», Росстат	Индекс производства в секторе «Обеспечения электрической энергией, газом и паром; кондиционирования воздуха»	Задержка 1 месяц (в конце августа выходит за июль)
Agriculture: YoY	Доклад «Социально-экономическое положение России», Росстат	Индекс реального объема сельскохозяйственного производства	Задержка 1 месяц (в конце августа выходит за июль)

Продолжение табл. 2.

Переменная	Источник	Описание	Календарь выхода
Construction: YoY	Доклад «Социально-экономическое положение России», Росстат	Индекс реального объема работ, выполненных по виду деятельности «Строительство»	Задержка 1 месяц (в конце августа выходит за июль)
House: YoY	Доклад «Социально-экономическое положение России», Росстат	Ввод в действие жилых домов	Задержка 1 месяц (в конце августа выходит за июль)
Transport: YoY	Доклад «Социально-экономическое положение России», Росстат	Коммерческий грузооборот транспорта	Задержка 1 месяц (в конце августа выходит за июль)
Trade: YoY	Доклад «Социально-экономическое положение России», Росстат	Индекс реального оборота розничной торговли	Задержка 1 месяц (в конце августа выходит за июль)
Unempl	Доклад «Социально-экономическое положение России», Росстат	Уровень безработицы	Задержка 1 месяц (в конце августа выходит за июль)
Inftation	Росстат	Инфляция, рассчитанная по индексу потребительских цен	10 число каждого месяца, без задержек
Oil Price: YoY	Reuters	Цена на нефть марки Brent на последнюю дату месяца	1–2 число каждого месяца, без задержек
ExRate USD: YoY	Банк России	Официальный курс Доллар США – Российский рубль на последнюю дату месяца	5–10 число каждого месяца, без задержек
ExRate EUR: YoY	Банк России	Официальный курс Евро – Российский рубль на последнюю дату месяца	5–10 число каждого месяца, без задержек
IMOEX: YoY	Московская биржа	Значение индекса IMOEX на последнюю дату месяца	1–2 число каждого месяца, без задержек
IRTS: YoY	Московская биржа	Значение индекса IRTS на последнюю дату месяца	1–2 число каждого месяца, без задержек
Gas Price: YoY	Reuters	Цена на газ на последнюю дату месяца	1–2 число каждого месяца, без задержек
MIACR	Банк России	Среднемесячные фактические ставки по кредитам в рублях, предоставленным московскими банками, срок 1 день	5–10 число каждого месяца, без задержек

Окончание табл. 2.

Переменная	Источник	Описание	Календарь выхода
MonBase: YoY	Банк России	Денежная база (в широком определении)	10 число каждого месяца, без задержек
M2: YoY	Банк России	Денежная масса M2	10 число каждого месяца, задержка 1 месяц
Reserve: YoY	Банк России	Международные резервы Российской Федерации	5–10 число каждого месяца, без задержек

4. Методология наукастинга

Целью данной работы фактически является сравнение вневыборочной точности моделей, используемых для наукастинга ВВП России по использованию.

Одной из базовых моделей наукастинга является MIDAS-модель:

$$y_t = \sum_{j=1}^p \alpha_j y_{t-j} + \sum_{i=0}^k \sum_{j=0}^{m_i} \beta_j^{(i)} x_{tm_i-j}^{(i)} + u_t,$$

где y_t – зависимая переменная, выраженная в темпах роста к аналогичному периоду прошлого года, низкочастотный ряд (в данной работе – квартальной периодичности); $x_t^{(i)}$ – объясняющие переменные, выраженные в темпах роста к аналогичному периоду прошлого года, высокочастотный ряд (в данной работе – месячной периодичности); m_i – количество наблюдений переменной $x_t^{(i)}$ при одном периоде для зависимой переменной.

При этом на практике MIDAS-модель без ограничений на какие-либо параметры встречается не часто (Unrestricted MIDAS или U-MIDAS-модели). Больше распространение получили MIDAS-модели с ограничениями на значения параметра $\beta_j^{(i)}$, что фактически является способом агрегации данных высокой частоты в ряд более низкой частоты (в рамках данной работы – месячные данные агрегируются в квартальные). В рамках данного исследования рассматривается MIDAS-модель с неэкспоненциальными полиномами Алмона [Clements, Galvao, 2009]:

$$\beta_j = \frac{\exp(\lambda_1(j+1) + \dots + \lambda_r(j+1)^r)}{\sum_{s=0}^d \exp(\lambda_1(s+1) + \dots + \lambda_r(s+1)^r)}.$$

Следует отметить, что различные ограничения для MIDAS-модели используются с целью снижения размерности данных вследствие наложения ограничения на некоторые параметры рассматриваемой модели. Снижение размерности позволяет, с одной стороны,

пользуясь терминами машинного обучения, предотвратить переобучение модели и, с другой стороны, не потерять имеющуюся вариативность для правильной подстройки под имеющийся набор данных.

В рамках данной работы сравнивается точность MIDAS-моделей, используемых для наукастинга компонентов ВВП по использованию, без авторегрессионной части.

Еще один класс моделей, используемых для получения оперативных оценок макроэкономических показателей, – линейные модели с регуляризацией, позволяющие сначала отобрать наиболее релевантные переменные, а затем вывести за рамки модели «лишние» переменные, что близко к использованию методов машинного обучения для построения краткосрочных прогнозов. Данное исследование включает в себя регуляризацию LASSO, в основе которой заложена U-MIDAS-модель:

$$y_t = \beta^0 + \sum_{i=1}^k \sum_{j=0}^{m_i} \beta_j^{(i)} x_{tm_i-j}^{(i)} + u_t$$

при целевой функции $\sum_{t=1}^T (y_t - \hat{y}_t)^2 + \lambda \sum_{i=1}^k \sum_{j=0}^{m_i} |\beta_j^{(i)}| \rightarrow \min$, где λ – параметр, поиск которого осуществляется кросс-валидацией. Следует отметить, что практически параметр λ носит смысл штрафа. В работе используется λ_{min} , соответствующее минимальной кросс-валидации *RSS*.

В процессе минимизации целевой функции часть коэффициентов принимают значение ноль, что фактически соответствует процессу отбора переменных и выявлению «лишних» из них. При этом, если значение параметра λ оказывается слишком маленьким, модель переходит на этап переобучения. Важным является тот факт, что при больших значениях параметра λ выявить истинные причинно-следственные связи между переменными оказывается затруднительно, а зачастую и вообще невозможно.

Использование MIDAS-модели для прогнозирования накладывает определенные ограничения на количество регрессоров: на практике в модель включается только одна объясняющая переменная, что вызвано необходимостью добавлять дополнительный набор регрессоров при включении еще одной объясняющей переменной в модель. Например, если в модели с данными смешанной частоты в качестве объясняющих переменных рассматриваются месячные данные, а зависимая переменная является квартальной, то при расширении списка объясняющих переменных на один необходимо будет добавить еще три дополнительных регрессора. Линейные модели с регуляризацией снимают данный вопрос для исследователя-практика и позволяют рассматривать желаемый набор данных в качестве объясняющих переменных.

Mixed-Frequency Bayesian VAR – Байесовская векторная авторегрессия смешанной частоты (MFBVAR-модель) является усовершенствованной для использования данных разной частоты формой классической векторной авторегрессии (в рамках данного исследования используются квартальные и месячные переменные). MFBVAR-модель работает по следующему принципу: набор объясняющих переменных является данными более высокой частоты (наиболее частый случай – месячная периодичность), на который накладывается классическая конфигурация VAR-модели, а зависимые переменные являются

данными низкой частоты (обычно квартальной периодичности). Таким образом, квартальные показатели есть среднее из ненаблюдаемых месячных показателей.

Рассмотрим пример, позволяющий понять, как работает простая MFVAR-модель с данными месячной частоты. Обозначим за x_t с размерностью $n \times 1$ описание процесса месячной частоты. Тогда декомпозиция $x_t = (x'_{m,t}; x'_{q,t})'$ в n_m – для переменных месячной частоты, n_q – для переменных квартальной частоты. Пусть $y_t = (y'_{m,t}; y'_{q,t})'$, тогда $y_{m,t} = x_{m,t}$ является компонентой месячной частоты. Для оставшихся квартальных переменных наблюдается x_q – средневзвешенное значение. Далее происходит агрегация квартальных значений с использованием треугольной схемы взвешивания [Mariano, Murasawa, 2013]:

$$y_{q,t} = \begin{cases} \frac{1}{9}(x_{q,t} + 2x_{q,t-1} + 3x_{q,t-2} + 2x_{q,t-3} + x_{q,t-4}), t \in \{Mar, Jun, Sep, Dec\} \\ \emptyset, otherwise \end{cases}$$

Следует отметить, что треугольная схема взвешивания используется в большей степени в случае, если все переменные представлены в темпах роста к соответствующему периоду прошлого года. При этом на практике существенного различия между среднеквартальным и треугольным способами агрегации не наблюдается.

Так как, как было упомянуто выше, для объясняющих переменных месячной частоты используется VAR-модель, покажем ее формально. Введем $Var(p)$ для x_t :

$$(1) \quad x_t = \phi + \Phi_1 x_{t-1} + \dots + \Phi_p x_{t-p} + \varepsilon_t, \varepsilon_t \sim N(0, \Sigma).$$

Перепишем $Var(p)$ через $z_t = (x'_t, x'_{t-1}, \dots, x'_{t-p+1})'$:

$$(2) \quad z_t = \pi + \Pi z_{t-1} + u_t, u_t \sim N(0, \Omega),$$

где π, Π и Ω – матрицы из форм-компаньонов, построенных на основе $(\phi, \Phi_1, \dots, \Phi_p, \Sigma)$ [Hamilton, 1994].

С учетом данных вводных запишем уравнение относительно y_t в виде

$$(3) \quad y_t = M_t \Lambda z_t,$$

где M_t – детерминированная матрица выбора, показывающая вектор наблюдений с помощью выбора необходимых строк, изменяется с изменением времени и с агрегацией Λ ; Λ – матрица агрегирования (скрытый процесс), выбранная в зависимости от типа данных – среднеквартальная или треугольная [Ankargren et al., 2019].

В рамках данного исследования для выбора лучшей функциональной формы модели для краткосрочного прогнозирования компонентов ВВП по использованию рассматривается MFBVAR-модель с априорным распределением Миннесоты и корректировкой, позволяющей использовать разночастотные данные. Общий принцип и механизм работы данного априорного распределения описывается в статьях [Litterman, 1979; 1986]. Технически данный тип моделей реализован в программной среде R [Ankargren, Yang, 2019]. Априорное распределение Миннесоты полагает, что каждый ряд из рассматриваемых представляет из себя процесс случайного блуждания с независимыми ошибками, что дает возможность включения большого количества переменных в модель, а также добавлять достаточно лагов в случае недостатка данных. Следует отметить, что применение к MFBVAR-модели априорного распределения другого класса обязывает исследователя вводить дополнительные более сложные предпосылки, что напрямую влияет на качество получаемого прогноза.

Наконец, последний тип моделей, используемый в рамках данной работы для построения наукаста ВВП России и его компонентов по использованию, – DFM-модель. Данные модели схожи с наукастами на основе метода главных компонент. При этом в качестве недостатка DFM-моделей можно выделить ограниченный набор объясняющих переменных, но данный факт не является ограничением, поскольку цель нашего исследования – сравнить точность отобранных методов наукастинга. В связи с этим в исследование включено минимальное количество регрессоров, напрямую связанных с зависимой переменной.

5. Результаты

В рамках данной работы оценивались результаты работы моделей на данных за последние 3 года, т.е. тестирование проходило по последним 12 точкам. Сопоставлялись результаты нескольких типов моделей: U-MIDAS, MIDAS неэкспоненциальными полиномами Алмона, MIDAS-модели с L1-регуляризацией, MFBVAR и DFM-модели, а также простейшей AR(1)-модели в качестве бенчмарка. Для исследования зависимости точности наукастов от объема использованных данных для каждой зависимой переменной строился наукаст по данным только за первый месяц квартала, при этом второй и третий месяц квартала принимался равным первому месяцу, затем по данным только за первые два месяца квартала, значение за третий месяц принималось равным значению за второй месяц, а также по данным за весь квартал. Благодаря этому можно изучить чувствительность точности наукастов к объему имеющихся данных и оценить точность оценок, которые можно получить еще до завершения квартала.

Сводные результаты по используемым моделям представлены в табл. 3–9. Каждая из таблиц показывает результаты для ВВП и его компонентов по использованию для каждой из рассматриваемых моделей. Для MIDAS-моделей также указана переменная, показавшая лучшую точность наукаста. Результат для предпочтительной модели выделен жирным шрифтом.

Обратимся к результатам для переменной ВВП России по использованию. Наиболее точный наукаст рассматриваемой переменной для целевого квартала рассчитывается при использовании только первого доступного месяца – лучший результат по MAE пока-

зывает MFBVAR-модель. Вторыми по точности оказываются неограниченная MIDAS-модель и MIDAS-R-модель с неэкспоненциальными полиномами Алмона. Если же оценивать качество моделей через критерий RMSE, то лучшей оказывается неограниченная MIDAS-модель по данным за два месяца текущего квартала, немного выигрывая у MFBVAR-модели по данным за один месяц. В целом же результаты по MAE и по RMSE согласуются. Результат сравнения наукаста наиболее точной модели, прогноза, полученного с помощью AR(1)-модели, а также фактический темп роста рассматриваемой переменной представлен на рис. 1. Следует отметить, что использованные модели несколько недооценивают глубину кризиса, однако факт снижения темпов роста ВВП предсказывают точно. При этом посткризисное восстановление предсказывается точно и в полном объеме.

Стоит также отдельно подчеркнуть очень необычный эффект: удаление части данных (последнего или даже последних двух месяцев квартала) повышает точность наукаста, что, вообще-то, является результатом контринтуитивным (как правило, чем больше данных мы используем – тем точнее оценка), и этот результат является достаточно устойчивым для разных типов моделей. Возможно, причина его кроется в том, что последний месяц квартала не очень репрезентативен с точки зрения квартала в среднем в силу определенной инерции, свойственной многим экономическим процессам, и исключение его из рассмотрения позволяет получить более «чистую» картину текущего квартала.

Таблица 3.

**Средняя абсолютная ошибка (MAE) и среднеквадратичная ошибка (RMSE)
для моделей наукастинга по последним 12 точкам,
переменная Real GDP: YoY**

Real GDP: YoY	MAE			RMSE		
	1 мес.	2 мес.	3 мес.	1 мес.	2 мес.	3 мес.
U-MIDAS (переменная)	2,06 <i>Trade:</i> <i>YoY</i>	1,36 <i>Trade:</i> <i>YoY</i>	1,76 <i>Industry:</i> <i>YoY</i>	2,72 <i>Industry:</i> <i>YoY</i>	1,83 <i>Trade:</i> <i>YoY</i>	2,17 <i>Industry:</i> <i>YoY</i>
MIDAS-R-nealmon (переменная)	2,11 <i>Industry:</i> <i>YoY</i>	1,46 <i>Trade:</i> <i>YoY</i>	1,67 <i>Trade:</i> <i>YoY</i>	2,72 <i>Industry:</i> <i>YoY</i>	1,93 <i>Trade:</i> <i>YoY</i>	2,17 <i>Trade:</i> <i>YoY</i>
MIDAS_L1	4,65	4,73	4,68	5,68	5,70	5,69
MFBVAR	1,28	1,51	1,85	1,84	1,99	2,63
DFM	2,16	2,40	2,10	3,03	3,11	2,76
AR (1)	2,90	2,90	2,90	4,42	4,42	4,42

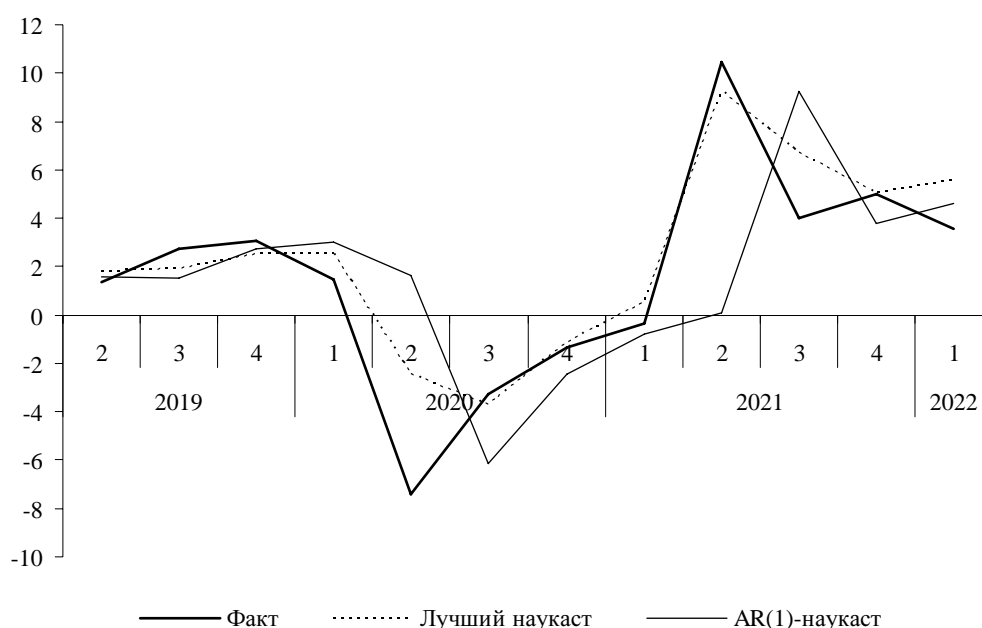


Рис. 1. Темпы роста ВВП: факт, лучший наукаст и AR(1)-наукаст

Оценку динамики расходов на конечное потребление домашних хозяйств разумнее всего с точки зрения точности наукастов проводить с помощью MIDAS-модели без ограничений с использованием данных за первые два месяца рассматриваемого квартала, о чем свидетельствуют результаты табл. 4.

Таблица 4.

Средняя абсолютная ошибка (MAE) и среднеквадратичная ошибка (RMSE) для моделей наукастинга по последним 12 точкам, переменная Households Consumption: YoY

Households Consumption: YoY	MAE			RMSE		
	1 мес.	2 мес.	3 мес.	1 мес.	2 мес.	3 мес.
U-MIDAS (переменная)	3,15 <i>Trade:</i> YoY	2,69 <i>Trade:</i> YoY	3,22 <i>Trade:</i> YoY	3,77 <i>Trade:</i> YoY	3,41 <i>Trade:</i> YoY	4,35 <i>Trade:</i> YoY
MIDAS-R-nealmon (переменная)	3,15 <i>Trade:</i> YoY	2,75 <i>Trade:</i> YoY	3,28 <i>Trade:</i> YoY	3,77 <i>Trade:</i> YoY	3,40 <i>Trade:</i> YoY	4,35 <i>Trade:</i> YoY
MIDAS_L1	10,02	10,15	10,04	12,61	12,62	12,46
MFBVAR	4,14	4,30	4,41	7,27	7,67	7,84
DFM	5,67	6,03	5,61	8,95	9,02	8,62
AR (1)	6,85	6,85	6,85	11,94	11,94	11,94

Интересным также оказывается тот факт, что наилучшую точность среди рассмотренных MIDAS-моделей обеспечивает переменная, отражающая Индекс реального оборота розничной торговли при разных наборах первоначальных данных. Результаты при сравнении MAE и RMSE также согласуются. На рис. 2 представлено сравнение лучшего наукаста, прогноза, полученного с помощью AR(1)-модели, и фактическое значение целевой переменной. Как и в случае с ВВП, модель несколько недооценивает глубину спада в кризис 2020 г., но при этом очень точно оценивает масштаб восстановления после кризиса.

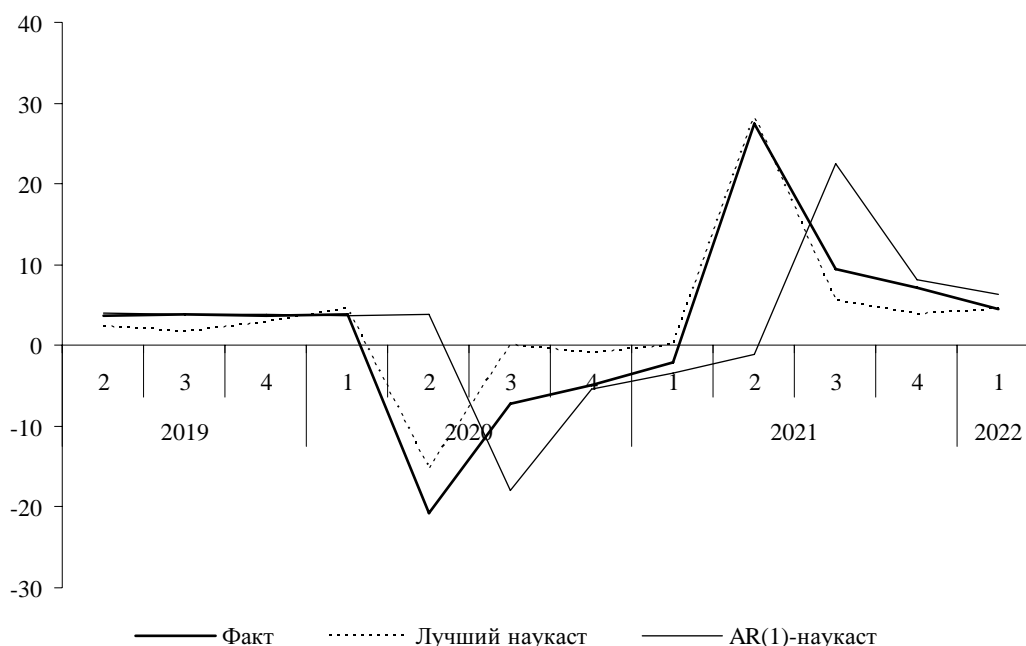


Рис. 2. Темпы роста потребления домашних хозяйств: факт, лучший наукаст и AR(1)-наукаст

В табл. 5 представлены результаты для расходов на конечное потребление государственного управления. Наиболее эффективной для наукастинга как по критерию MAE, так и по RMSE оказалась авторегрессия, что в целом связано со сложным моделированием данной переменной и практически невозможным объяснением влияния различных данных на точность прогноза. При этом схожей по точности оперативной оценки оказывается MIDAS-R-модель с неэкспоненциальными полиномами Алмона.

Расходы на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства (табл. 6), наилучшим образом оцениваются с помощью MFBVAR-модели на наборе данных за первый месяц при сравнении результатов с помощью MAE. При этом при сопоставлении точности через RMSE лучше все же использовать данные за первые два квартала, хотя изменения в точности практически не наблюдается (RMSE = 2,14 и 2,13 при использовании данных за первый и за первые два месяца соответственно). Результаты по переменной, показавшей наибольшую точность в MIDAS-моделях, также можно считать стабильными.

Таблица 5.

Средняя абсолютная ошибка (MAE) и среднеквадратичная ошибка (RMSE)
для моделей наукастинга по последним 12 точкам,
переменная Government Consumption: YoY

Government Consumption: YoY	MAE			RMSE		
	1 мес.	2 мес.	3 мес.	1 мес.	2 мес.	3 мес.
U-MIDAS (переменная) <i>Inftation</i>	0,67	0,66	0,68	0,80	0,79	0,80
MIDAS-R-nealmon (переменная) <i>Inftation</i>	0,64	0,63	0,63	0,75	0,75	0,74
MIDAS_L1	2,86	2,89	3,01	3,32	3,31	3,16
MFBVAR	1,05	0,92	1,01	1,29	1,18	1,22
DFM	1,44	1,39	1,25	1,62	1,58	1,45
AR (1)	0,61	0,61	0,61	0,70	0,70	0,70

Таблица 6.

Средняя абсолютная ошибка (MAE) и среднеквадратичная ошибка (RMSE)
для моделей наукастинга по последним 12 точкам,
переменная Non-commercial Consumption: YoY

Non-commercial Consumption: YoY	MAE			RMSE		
	1 мес.	2 мес.	3 мес.	1 мес.	2 мес.	3 мес.
U-MIDAS (переменная) <i>Unempl</i>	2,65	2,51	2,48	3,45	3,35	3,30
MIDAS-R-nealmon (переменная) <i>Unempl</i>	2,62	2,51	2,49	3,41	3,35	3,29
MIDAS_L1	3,40	3,77	3,82	4,03	3,93	4,14
MFBVAR	1,26	1,40	1,59	2,14	2,13	2,62
DFM	3,47	3,46	3,51	4,25	4,24	4,31
AR (1)	1,82	1,82	1,82	2,73	2,73	2,73

Оперативная оценка валового накопления основного капитала (табл. 7) оказывается наиболее точной при использовании любой из MIDAS-моделей (неограниченной или с экспоненциальными полиномами Алмона). В качестве набора данных можно брать как полный квартал, так и первые два месяца – точность практически не меняется. Результаты

при сравнении с помощью MAE и RMSE идентичны. Модели наукастинга для темпов роста валового накопления основного капитала, как и в случае с темпом роста ВВП России, не всегда точно оценивают глубину кризиса, но безошибочно определяют сам факт спада и очень точно предсказывают период и масштаб восстановления после кризиса (рис. 3).

Таблица 7.

Средняя абсолютная ошибка (MAE) и среднеквадратичная ошибка (RMSE) для моделей наукастинга по последним 12 точкам, переменная Gross Capital Formation: YoY

Gross Capital Formation: YoY	MAE			RMSE		
	1 мес.	2 мес.	3 мес.	1 мес.	2 мес.	3 мес.
U-MIDAS (переменная)	2,50	2,20	2,19	2,89	2,69	2,68
	<i>Transport: YoY</i>	<i>Transport: YoY</i>	<i>Transport: YoY</i>	<i>Transport: YoY</i>	<i>Transport: YoY</i>	<i>Transport: YoY</i>
MIDAS-R-nealmon (переменная)	2,50	2,20	2,19	2,89	2,69	2,68
	<i>Transport: YoY</i>	<i>Transport: YoY</i>	<i>Transport: YoY</i>	<i>Transport: YoY</i>	<i>Transport: YoY</i>	<i>Transport: YoY</i>
MIDAS_L1	4,37	4,33	4,42	5,80	5,67	5,73
MFBVAR	2,70	2,43	3,19	3,45	2,94	4,02
DFM	3,39	3,69	2,77	4,23	4,65	3,99
AR (1)	4,48	4,48	4,48	5,31	5,31	5,31

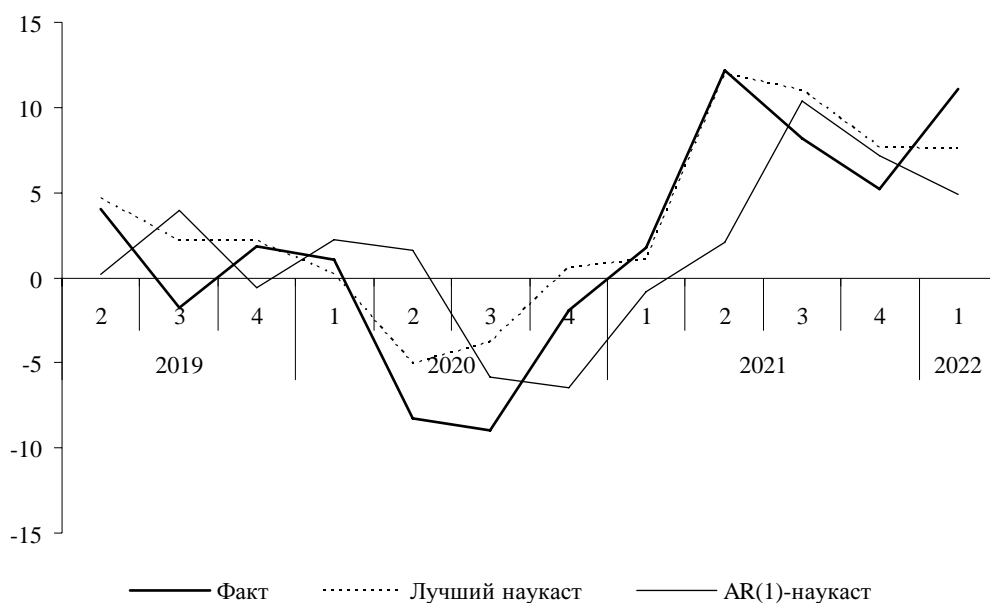


Рис. 3. Темпы роста валового накопления основного капитала: факт, лучший наукаст и AR(1)-наукаст

Окончание табл. 9.

Imports: YoY	MAE			RMSE		
	1 мес.	2 мес.	3 мес.	1 мес.	2 мес.	3 мес.
MIDAS_L1	15,11	14,74	14,86	18,35	16,94	17,24
MFBVAR	6,02	5,28	5,49	7,35	5,74	6,76
DFM	7,92	8,06	6,28	10,54	10,80	8,17
AR (1)	8,44	8,44	8,44	12,64	12,64	12,64

6. Заключение

В работе сравнивается точность моделей для наукастинга нескольких типов: MIDAS-модели без ограничений, MIDAS-модели с неэкспоненциальными полиномами Алмона, MIDAS-модели с регуляризацией L1, DFM и MFBVAR-модели. В качестве обоснования преимущества моделей наукастинга перед классическими моделями прогнозирования результаты сравниваются с авторегрессией. Точность наукаста рассматривается на ВВП России и его компонентах по использованию и оценивалась с помощью средней абсолютной ошибки (MAE) и среднеквадратичной ошибки прогноза (RMSE) по последним 12 точкам. Различий в выводах при сравнении точности по двум критериям не наблюдается. Также в качестве способа оперативной оценки текущего состояния экономики предложен метод наукастинга на основе данных только за первый или за первые два месяца рассматриваемого квартала. Описываемый в работе подход доказывает свою эффективность, что объясняется слабой репрезентативностью третьего месяца в каждом квартале ввиду инерции, свойственной многим экономическим процессам.

Результаты, полученные в рамках данной работы, показывают, что для наукастинга ВВП России и его компонентов по использованию следует использовать разные модели для получения наиболее точного результата. Оперативная оценка валового внутреннего продукта, расходов на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства, а также импорта получается наиболее точной при использовании MFBVAR-модели. При этом расходы на конечное потребление домашних хозяйств и валовое накопление основного капитала оцениваются с помощью неограниченной MIDAS-модели. Интересным оказался тот факт, что переменная, отражающая расходы на конечное потребление государственного управления, которая в целом достаточно плохо подлежит прогнозированию и объяснению через регрессоры, лучше всего прогнозируется с помощью классической авторегрессионной модели. Наконец, для наукастинга экспорта для получения наиболее точной оценки квартальной динамики следует использовать DFM-модель.

* *
*

СПИСОК ЛИТЕРАТУРЫ

- Зубарев А., Рыбак К.* Наукакстинг ВВП: динамическая факторная модель и официальные прогнозы // Экономическое развитие России. 2021. Т. 28. № 12. С. 34–40.
- Лазарян С., Герман Н.* Прогнозирование текущей динамики ВВП на основе данных поисковых запросов // Финансовый журнал. 2018. № 6. С. 83–94.
- Майорова К., Фокин Н.* Наукакстинг темпов роста стоимостных объемов экспорта и импорта России по товарным группам // Деньги и кредит. 2021. Т. 80. № 3. С. 34–48.
- Микош Х., Соланко Л.* Прогнозирование роста российского ВВП с использованием данных со смешанной периодичностью // Деньги и кредит. 2019. Т. 78. № 1. С. 19–35.
- Поршаков А.С., Пономаренко А.А., Синяков А.А.* Оценка и прогнозирование ВВП России с помощью динамической факторной модели // Журнал НЭА. 2016. Т. 2. № 30. С. 60–76.
- Станкевич И.П.* Сравнение методов наукакстинга макроэкономических индикаторов на примере российского ВВП // Прикладная эконометрика. 2020. Т. 59. С. 113–127.
- Третьяков Д., Фокин Н.* Помогают ли высокочастотные данные в прогнозировании российской инфляции? // Вестник Санкт-Петербургского университета. Экономика. 2021. Т. 37. № 2. С. 318–343.
- Яковлева К.* Оценка экономической активности на основе текстового анализа // Деньги и кредит. 2018. Т. 77. № 4. С. 26–41.
- Ankargren S., Yang Y.* Mixed-Frequency Bayesian VAR Models in R: The mfbvar Package. 2019.
- Ankargren S., Unosson M., Yang Y.* A Flexible Mixed-Frequency Vector Autoregression with a Steady-State Prior. 2019.
- Cepni O., Güney I.E., Swanson N.R.* Nowcasting and Forecasting GDP in Emerging Markets Using Global Financial and Macroeconomic Diffusion Indexes // International Journal of Forecasting. 2019. Vol. 35. № 2. P. 555–572.
- Chernis T., Sekkel R.* A Dynamic Factor Model for Nowcasting Canadian GDP Growth // Empirical Economics. 2017. Vol. 53. № 1. P. 217–234.
- Clements M.P., Galvão A.B.* Forecasting US Output Growth Using Leading Indicators: An Appraisal Using MIDAS Models // Journal of Applied Econometrics. 2009. Vol. 24. № 7. P. 1187–1206.
- Ferrara L., Marsilli C.* Nowcasting Global Economic Growth: A Factor-augmented Mixed-frequency Approach // The World Economy. 2019. Vol. 42. № 3. P. 846–875.
- Ferrara L., Simoni A.* When Are Google Data Useful to Nowcast GDP? An Approach via Pre-selection and Shrinkage // Journal of Business and Economic Statistics. 2022.
- Ghysels E., Santa-Clara P., Valkanov R.* Predicting Volatility: Getting the Most Out of Return Data Sampled at Different Frequencies // Journal of Econometrics. 2006. Vol. 131. № 1–2. P. 59–95.
- Ghysels E., Sinko A., Valkanov R.* MIDAS Regressions: Further Results and New Directions // Econometric Reviews. 2007. Vol. 26. № 1. P. 53–90.
- Hamilton J.* Time Series Analysis. Princeton University Press, 1994.
- Ingenito R., Trehan B.* Using Monthly Data to Predict Quarterly Output // Economic Review. Federal Reserve Bank of San Francisco. 1996. 3. P. 3–11.
- Jardet C., Meunier B.* Nowcasting World GDP Growth with High-frequency Data // Journal of Forecasting. 2022. P. 1181–1200.
- Kuzin V., Marcellino M., Schumacher C.* MIDAS vs. Mixed-frequency VAR: Nowcasting GDP in the Euro Area // International Journal of Forecasting. 2011. Vol. 27. № 2. P. 529–542.
- Litterman R.B.* Techniques of Forecasting Using Vector Autoregressions. Working Paper № 115. Federal Reserve Bank of Minneapolis, 1979.

Litterman R.B. Forecasting with Bayesian Vector Autoregressions-Five Years of Experience // Journal of Business & Economic Statistics. 1986. Vol. 4. № 1. P. 25–38.

Marcellino M., Schumacher C. Factor MIDAS for Nowcasting and Forecasting with Ragged-edge Data: A Model Comparison for German GDP // Oxford Bulletin of Economics and Statistics. 2010. Vol. 72. № 4. P. 518–550.

Marcellino M., Sivec V. Nowcasting GDP Growth in a Small Open Economy // National Institute Economic Review. 2021. № 256. P. 127–161.

Mariano RS., Murasawa Y. A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series // Journal of Applied Econometrics. 2013. № 18 (4). P. 427–443.

Nakazawa T. Constructing GDP Nowcasting Models Using Alternative Data. Bank of Japan Working Paper Series. 2022. № 22 (9).

Proiettia T., Giovannelli A., Ricchi O., Citton A., Tegami C., Tinti C. Nowcasting GDP and its Components in a Data-rich Environment: The Merits of the Indirect Approach // International Journal of Forecasting. 2021. № 37 (4). P. 1376–1398.

Richardson A., Florenstein Muldera T., Vehbi T. Nowcasting GDP Using Machine-learning Algorithms: A Real-time Assessment // International Journal of Forecasting. 2021. № 37 (2). P. 941–948.

Schorfheide F., Song D. Real-time Forecasting with a Mixed-frequency VAR // Journal of Business & Economic Statistics. 2015. Vol. 33. № 3. P. 366–380.

Schumacher C. A Comparison of MIDAS and Bridge Equations // International Journal of Forecasting. 2016. Vol. 32. № 2. P. 257–270.

Soybilgen B., Yazgan E. Evaluating Nowcasts of Bridge Equations with Advanced Combination Schemes for the Turkish Unemployment Rate // Economic Modelling. 2018. Vol. 72. P. 99–108.

Soybilgen B., Yazgan E. Nowcasting US GDP Using Tree-Based Ensemble Models and Dynamic Factors // Computational Economics. 2021. № 57. P. 387–417.

Tiffin A. Seeing in the Dark: A Machine-learning Approach to Nowcasting in Lebanon. Working Paper № 115. Federal Reserve Bank of Minneapolis. 2016.

Zhemkov M. Nowcasting Russian GDP Using Forecast Combination Approach // International Economics. 2021. № 168. P. 10–24.

Nowcasting of the Components of Russian GDP

Natalia Makeeva¹, Ivan Stankevich²

¹ National Research University Higher School of Economics,
20, Myasnitskaya st., Moscow, 101000, Russian Federation.
E-mail: nmakeeva@hse.ru

² National Research University Higher School of Economics,
20, Myasnitskaya st., Moscow, 101000, Russian Federation.
E-mail: istankevich@hse.ru

The paper discusses the problem of nowcasting the current growth rates of Russian GDP and its components using quarterly data. The quality of restricted and unrestricted MIDAS models (models with mixed data), MIDAS model with L1 regularisation and MFBVAR model (Bayesian vector autoregression of mixed frequency) are compared. The results are compared with classical autoregression to justify the need to use nowcasting models for the rapid assessment of macroeconomic indicators. Production indices for various industries and macro indicators characterising Russian GDP and its components were used as explanatory variables. The paper proposes a way to quickly assess the current state of the economy and proposes a nowcasting method based on data only for the first or first two months of the quarter under consideration. As a result, for each dependent variable, the best model for building a nowcast based on the last 12 points is selected based on the criterion of mean absolute error (MAE) and root mean square prediction error (RMSE).

Key words: nowcasting; Russian GDP; time series; forecasting; mixed frequency models; MIDAS models.

JEL Classification: C53, E37.

* *
*

References

- Ankargren S., Yang Y. (2019) *Mixed-Frequency Bayesian VAR Models in R: The mfbvar Package*.
Ankargren S., Unosson M., Yang Y. (2019) *A Flexible Mixed-Frequency Vector Autoregression with a Steady-State Prior*.
Cepni O., Güney I.E., Swanson N.R. (2019) Nowcasting and Forecasting GDP in Emerging Markets Using Global Financial and Macroeconomic Diffusion Indexes. *International Journal of Forecasting*, 35, 2, pp. 555–572.

- Chernis T., Sekkel R. (2017) A Dynamic Factor Model for Nowcasting Canadian GDP Growth. *Empirical Economics*, 53, 1, pp. 217–234.
- Clements M.P., Galvão A.B. (2009) Forecasting US Output Growth Using Leading Indicators: An Appraisal Using MIDAS Models. *Journal of Applied Econometrics*, 24, 7, pp. 1187–1206.
- Ferrara L., Marsilli C. (2019) Nowcasting Global Economic Growth: A Factor-augmented Mixed-frequency Approach. *The World Economy*, 42, 3, pp. 846–875.
- Ferrara L., Simoni A. (2022) When Are Google Data Useful to Nowcast GDP? An Approach via Pre-selection and Shrinkage. *Journal of Business and Economic Statistics*. Forthcoming.
- Ghysels E., Santa-Clara P., Valkanov R. (2006) Predicting Volatility: Getting the Most Out of Return Data Sampled at Different Frequencies. *Journal of Econometrics*, 131, 1–2, pp. 59–95.
- Ghysels E., Sinko A., Valkanov R. (2007) MIDAS Regressions: Further Results and New Directions. *Econometric Reviews*, 26, 1, pp. 53–90.
- Hamilton J. (1994) *Time Series Analysis*. Princeton University Press.
- Ingenito R., Trehan B. (1996) Using Monthly Data to Predict Quarterly Output. *Economic Review*, Federal Reserve Bank of San Francisco, 3, pp. 3–11.
- Jardet C., Meunier B. (2022) Nowcasting World GDP Growth with High-frequency Data. *Journal of Forecasting*, pp. 1181–1200.
- Kuzin V., Marcellino M., Schumacher C. (2011) MIDAS vs. Mixed-frequency VAR: Nowcasting GDP in the Euro Area. *International Journal of Forecasting*, 27, 2, pp. 529–542.
- Lazaryan S., German N. (2018) Forecasting Current GDP Dynamics with Google Search Data. *Financial Journal*, 6, pp. 83–94. (In Russian)
- Litterman R.B. (1979) *Techniques of Forecasting Using Vector Autoregressions*. Working Paper no 115. Federal Reserve Bank of Minneapolis.
- Litterman R.B. (1986) Forecasting with Bayesian Vector Autoregressions-Five Years of Experience. *Journal of Business & Economic Statistics*, 4, 1, pp. 25–38.
- Marcellino M., Schumacher C. (2010) Factor MIDAS for Nowcasting and Forecasting with Ragged-edge Data: A Model Comparison for German GDP. *Oxford Bulletin of Economics and Statistics*, 72, 4, pp. 518–550.
- Marcellino M., Sivec V. (2021) Nowcasting GDP Growth in a Small Open Economy. *National Institute Economic Review*, 256, pp. 127–161.
- Mariano R.S., Murasawa Y. (2013) A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series. *Journal of Applied Econometrics*, 18, 4, pp. 427–443.
- Mayorova K., Fokin N. (2021) Nowcasting Growth Rates of Russia's Export and Import by Commodity Group. *Russian Journal of Money and Finance*, 80, 3, pp. 34–48. (In Russian)
- Mikosch H., Solanko L. (2019) Forecasting Quarterly Russian GDP Growth with Mixed-frequency data. *Russian Journal of Money and Finance*, 78, 1, pp. 19–35. (In Russian)
- Nakazawa T. (2022) *Constructing GDP Nowcasting Models Using Alternative Data*. Bank of Japan Working Paper Series, 22, 9.
- Porshakov A.S., Ponomarenko A.A., Sinyakov A.A. (2016) Nowcasting and Short-term Forecasting of Russian GDP with a Dynamic Factor Model. *Journal of the New Economic Association*, 2, 30, pp. 60–76. (In Russian)
- Proietta T., Giovannelli A., Ricchi O., Citton A., Tegami C., Tinti C. (2021) Nowcasting GDP and its Components in a Data-rich Environment: The Merits of the Indirect Approach. *International Journal of Forecasting*, 37, 4, pp. 1376–1398.
- Richardson A., Florenstein Muldera T., Vehbi T. (2021) Nowcasting GDP Using Machine-learning Algorithms: A Real-time Assessment. *International Journal of Forecasting*, 37, 2, pp. 941–948.
- Schorfheide F., Song D. (2015) Real-time Forecasting With a Mixed-frequency VAR. *Journal of Business & Economic Statistics*, 33, 3, pp. 366–380.
- Schumacher C. (2016) A Comparison of MIDAS and Bridge Equations. *International Journal of Forecasting*, 32, 2, pp. 257–270.

- Soybilgen B., Yazgan E. (2018) Evaluating Nowcasts of Bridge Equations with Advanced Combination Schemes for the Turkish Unemployment Rate. *Economic Modelling*, 72, pp. 99–108.
- Soybilgen B., Yazgan E. (2021) Nowcasting US GDP Using Tree-Based Ensemble Models and Dynamic Factors. *Computational Economics*, 57, pp. 387–417.
- Stankevich I. (2020) Comparison of Macroeconomic Indicators Nowcasting Methods: Russian GDP Case. *Applied Econometrics*, 59, pp. 113–127. (In Russian)
- Tiffin A. (2016) *Seeing in the Dark: A Machine-learning Approach to Nowcasting in Lebanon*. Working Paper no 115. Federal Reserve Bank of Minneapolis.
- Tretyakov D., Fokin N. (2021) Does the High-frequency Data Is Helpful for Forecasting Russian Inflation? *St. Petersburg University Journal of Economic Studies*, 37, 2, pp. 318–343. (In Russian)
- Yakovleva K. (2018) Text Mining-based Economic Activity Estimation. *Russian Journal of Money and Finance*, 77, 4, pp. 26–41. (In Russian)
- Zhemkov M. (2021) Nowcasting Russian GDP Using Forecast Combination Approach. *International Economics*, 168, pp. 10–24.
- Zubarev A., Rybak K. (2021) GDP Nowcasting: Dynamic Factor Model vs. Official Forecasts. *Economic Development of Russia*, 28, 12, pp. 34–40. (In Russian)