**Применение MIDAS-моделей с марковским переключением для наукастинга ВВП и его компонент**

В работе рассматривается вопрос наукастинга (оценки текущего уровня) ВВП России и его компонент по использованию при помощи MIDAS-моделей с марковским переключением. Недооценка глубины спада в кризисные периоды, особенно в начале кризиса, является достаточно распространённой проблемой для большинства стандартных наукастинговых моделей. Модели с марковским переключением позволяют более корректно описывать временные ряды с неоднородной структурой, что может положительно сказаться в том числе на качестве наукастинга. В работе рассматривается модификация одной из самых популярных в данной области моделей – MIDAS – как MIDAS-модели с марковскими переключениями для случая двух режимов.

Рассматриваются разные способы получения наукастов на основе полученных результатов: взвешенные по вероятностям нахождения в том или ином режиме в следующий период времени, по наиболее вероятному режиму, в условиях правильного предсказания режима. Проводится сравнение качества её работы со стандартными моделями наукастинга: ограниченными и неограниченными MIDAS-моделями, MIDAS-моделью с L1-регуляризацией и MFBVAR-моделью (байесовской векторной авторегрессией смешанной частоты). Изучается поведение модельных прогнозов в зависимости от объёма доступной при построении наукаста информации.

Показано, что среди всех протестированных моделей максимальная точность обеспечивается MIDAS-моделями с марковским переключением с правильным предсказанием режима. MIDAS-модели с марковским переключением без правильного предсказания режима также стабильно работают лучше стандартных MIDAS-моделей и MFBVAR-моделей для большинства рассмотренных рядов данных.

***Ключевые слова:*** Наукастинг; ВВП России; временные ряды; прогнозирование; модели со смешанной частотой; модели с марковским переключением; MIDAS модели

**Классификация JEL:** C53; E37

**Введение**

Прогнозирование текущих значений публикуемых со значительной задержкой макроэкономических показателей (Nowcasting, в литературе на русском языке часто оставляемый без перевода) на основе значений более оперативных индикаторов является важной задачей с точки зрения оперативной подстройки действий экономических агентов под изменяющиеся макроэкономические условия. Возможность получить первые оценки показателя не через несколько месяцев после окончания периода, а через несколько дней (лаг публикации наиболее оперативно выходящих показателей) очень ценна как для государственных органов, так и для частных компаний. Особенно возрастает актуальность задач наукастинга в кризисные периоды – что, во многом, и обусловило рост интереса к этой задаче в последние несколько лет, после начала кризиса, связанного с распространением COVID-19, и на фоне событий 2022 года. Из всех макроэкономических показателей одним из самых главных и при этом наименее оперативных является ВВП, что обуславливает высокую актуальность задачи наукастинга ВВП и его компонент, публикующихся с ещё более значительной задержкой.

Существует несколько подходов к наукастингу. Например, в работе [Ingenito et al., 1996] используются связующие уравнения для оценки интересующих макроэкономических показателей. Исследователи используют объясняющие переменные 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 оценивается через метод наименьших квадратов, при этом к полному набору данных применяется фильтр 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 в более спокойные с точки зрения макроэкономической ситуации времена DFM-моделей демонстрируется в работе [Зубарев, Рыбак, 2021]: авторы тестируют модель на отрезке с 2016 по 2021 год, 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 третьем квартале 2020 г.

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 темпов 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 [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; Litterman, 1986].

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 [Schorfheide, Song, 2015]. Интересной с точки зрения техники является [Яковлева, 2018]: автор строит наукаст на основе отражающих динамику отечественной экономической активности высокочастотных переменных через применение методов текстового анализа и машинного обучения, направленные 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, а в работе [Макеева, Станкевич, 2022] схожий набор методов используется для наукастинга компонент ВВП. В работе [Гареев, Полбин, 2022] для наукастинга основных макроэкономических индикаторов (включая ВВП и некоторые его компоненты) используется ряд стандартных моделей машинного обучения. 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 показывает, что использование комбинированного прогноза оказывается не значительно лучше по точности по сравнению с использованием только одной модели, поэтому он видит большое пространство для доработки методики получения наукастов.

Одной из значительных проблем наукастинга в кризисные периоды является недооценка глубины спада многими моделями. К примеру, в [Hopp, 2022] рассматриваются наукасты ВВП США широкого спектра моделей от простых МНК-регрессий до нейронных сетей с LSTM архитектурой в кризисные периоды: начало 1980-ых, в кризис 2008 и в кризис 2020. Большая часть моделей (включая MIDAS и VAR смешанной частоты) смогла идентифицировать падение ВВП в 2008 году только с использованием данных за 2 месяца после окончания квартала, когда состоялось падение ВВП. Часть моделей (авторегрессионные и DFM) «перенесли» падение ВВП на 1 квартал позднее. Одним из возможных способов решения проблемы могут быть модели с переключением, в которых разные состояния экономики описываются разными уравнениями.

Модификация MIDAS-моделей до MIDAS-моделей с марковским переключением была предложена в [Guérin, Marcellino, 2013], но использовалась преимущественно для задач прогнозирования волатильности: на фондовых рынках в [Segnon et al., 2023], на рынках товарных фьючерсов в [Lu et al., 2022], или цен на криптовалюты в [Ma et al., 2020]. В приложении к макроэкономическому наукастингу модели с несколькими режимами чаще сводятся просто к проверке на наличие структурных сдвигов и оценке разных моделей для периода до и после сдвига, см. к примеру [Barnett W. A., Tang, 2016].

Данная работа одной из первых, насколько известно автору, рассматривает применение MIDAS-моделей с марковским переключением для наукастинга темпов роста ВВП и его компонент, в том числе с использованием метода главным компонент для снижения размерности пространства объясняющих переменных.

**Методы**

В качестве базовой эконометрической модели в работе используется модель с данными смешанной частоты MIDAS, которая в наиболее общей форме может быть записана в виде:

,

Где  - ряд данных более низкой частоты (в нашем случае – квартальный ряд темпов роста ВВП);  - ряды объясняющих факторов более высокой частоты;  показывает количество наблюдений объясняющей переменной , соответствующих одному значению (одному периоду) в зависимой переменной. При этом в большинстве формулировок MIDAS-моделей предполагается наличие ограничений на значения параметров  общего вида:

,

Где конкретный вид функции  зависит от типа используемой MIDAS модели. Если ограничения на значения параметров не вводятся, полученная модель называется неограниченной MIDAS моделью (Unrestricted MIDAS, U-MIDAS).

В качестве модели с ограничениями в работе рассматривается MIDAS модель с неэкспоненциальными полиномами Алмона:

,

Где  - параметры модели. Разные формулировки MIDAS моделей с ограничениями, несмотря на то что обладают несколько разнящимися свойствами, имеют общую цель ввода ряда ограничений на параметры с целью снижения размерности задачи и, пользуясь терминами из машинного обучения, недопущения переобучения модели, сохраняя при этом достаточную гибкость для правильной подстройки под данные.

Данная работа использует базовую формулировку MIDAS-модели без авторегрессионной части и с технической точки зрения основана на реализации MIDAS моделей в пакете *midasr* языка *R* [Ghysels et al., 2016].

Модели с марковским переключением, предложенные в [Hamilton, 1989], предполагают существование нескольких (как минимум – двух) режимов, в которых временной ряд описывается разными уравнениями. В простейшем случае AR(1) модели предполагается, что

,

Где  - значение переменной в период ,  - коэффициенты модели,  с нулевым математическим ожиданием и постоянной дисперсией, а  - переменная состояния. Таким образом, в рассматриваемой модели в зависимости от состояния меняется математическое ожидание процесса.

Переход из состояния в состояние при этом – марковский процесс с (неизвестными) вероятностями в матрице перехода



Эта модель может быть обобщена до MIDAS-модели с марковским переключением, которая для случая двух состояний (который и рассматривается в данной работе) принимает вид:



В данной работе используются модели без ограничений и авторегрессионной компоненты и с одной объясняющей переменной. Такой выбор обусловлен, с одной стороны, необходимостью сохранить максимальную сопоставимость с использующимися MIDAS-моделями, с другой – небольшим объёмом выборки и необходимостью оценки 6 коэффициентов (по 3 коэффициента – 3 месяца в одном квартале – в каждом из двух состояний) для каждой новой переменной.

Помимо моделей со стандартным набором объясняющих переменных, рассматриваются также MIDAS-модели с марковским переключением с главными компонентами в качестве объясняющих переменных. В каждой модели по-прежнему используется одна переменная, тестируются главные компоненты от 1-ой до 7-ой, в качестве результирующей модели выбирается лучшая.

Оценка моделей проводилась в R при помощи пакета *MSwM* с использованием авторских доработок для интеграции MIDAS-моделей и расчёта наукастов.

Байесовские векторные авторегрессии смешанной частоты (Mixed-Frequency Bayesian VAR, MFBVAR) представляют собой версию стандартных векторных авторегрессий, модифицированную для использования с данными разной частоты. Предполагается, что при исходном высокочастотном (в нашем случае месячном) процессе, описываемом стандартной VAR-моделью, мы наблюдаем часть переменных только на более низкой частоте. При этом, к примеру, наблюдаемые квартальные значения являются средними из ненаблюдаемых месячных значений.

В данной работе используется версия байесовской векторной авторегрессии с априорным распределением Миннесоты, модифицированная для данных со смешанной частотой. Напомним, что общая идея априорного распределения Миннесоты, предложенная в работах Литтермана (Litterman, 1979; Litterman, 1986), заключается в том, что все рассматриваемые ряды предполагаются процессами случайного блуждания с независимыми ошибками. Такой подход позволяет оценивать модели с большим числом переменных и лагов на данных сравнительно небольшого объёма за счёт исключения «лишних» переменных.

Расчеты проведены с использованием пакета *mfbvar* языка *R* (Ankargren, Yang, 2019).

В качестве подхода, сохраняющего идею исключения «лишних» переменных, однако более свойственного машинному обучению, рассматриваются линейные модели с регуляризацией. В работе представлены модели с L1 регуляризацией (LASSO регрессия), построенные на базе U-MIDAS модели. Модель может быть записана как



При целевой функции



Параметр регуляризации  подбирается кросс-валидацией.

**Данные и оценка точности 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:

,

где  – фактическая наблюдаемая зависимая переменная (в рамках данный работы – прирост к соответствующему периоду прошлого года),

– аналогичное значение, полученное с помощью модели,

 – 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 удалённые значения регрессоров и для них рассчитывается прогнозное значение зависимой переменной (наукаст). Рассматриваемые модели сравниваются по последним 12 точкам – это период с третьего квартала 2019 года по первый квартал 2022 года. Таким образом, модели тестируются на достаточно разнородных данных – в тестирование попадают как относительно спокойный период 2019 г., так и кризисные периоды, что особенно актуально в связи с коронакризисом 2020 года.

Для проверки устойчивости к добавлению новых данных все модели тестируются трижды: с использованием данных по объясняющим переменным за все три месяца квартала, для которого рассчитывается наукаст; без данных за последний месяц и без данных за два последних месяца. В случае удаления части данных в объясняющих переменных, «пустым» месяцам в объясняющих переменных проставляется последнее доступное значение показателя (за второй или за первый месяц квартала в зависимости от метода тестирования). Такая проверка позволяет определить, насколько методы устойчивы к объёму используемых данных, и смоделировать встречающиеся в реальной жизни условия, когда наукаст показателя за текущий квартал рассчитывается ещё до окончания квартала.

Отдельно стоит осветить вопрос наукастинга при помощи моделей с марковским переключением. В силу структуры моделей с марковским переключением, в каждый момент времени доступно два наукаста (для каждого из двух режимов), и выбор между ними является задачей нетривиальной. В данной работе сравниваются три подхода для расчёта наукастов:

1. Взвешенный прогноз. Используя вероятности нахождения в каждом из двух режимов и вероятности в матрице переходов (все эти величины рассчитываются моделью), можно рассчитать вероятность нахождения в каждом из режимов в следующий период (для которого рассчитывается наукаст). Результирующий прогноз – взвешенное среднее из двух прогнозов с весами, равными вероятностям нахождения в каждом из состояний.
2. Прогноз по наиболее вероятному режиму. После расчёта вероятностей нахождения в том или ином состоянии в следующий момент времени, в качестве результирующего прогноз берётся прогноз для наиболее вероятного режима
3. Прогноз с правильным предсказанием режима. В качестве наукаста берётся прогноз для того режима, который фактически реализовался. Этот вариант нельзя считать «честным» наукастингом, потому что на момент расчёта наукаста, строго говоря, ещё неизвестно, в каком режиме экономика будет находиться в данный период, но его можно использовать как своего рода бенчмарк максимально достижимой точности. С другой стороны, ситуацию наличия двух прогнозов можно интепретировать как «сценарный наукастинг» с двумя сценариями (оптимистичным и пессимистичным), и с такой точки зрения выбор правильного прогноза можно интерпретировать просто как хорошую работу аналитика, пользующегося моделью, и тогда оценка точности на основе правильного режима не выглядит столь уж «нечестной».

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 1 квартала 2001 года. Для получения полных временных рядов имеющийся набор данных с базовыми ценами 2016 года 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 (см. таблицу 1)

*Таблица 1. Зависимые переменные и календарь их выхода*

| **Переменная** | **Описание** | **Календарь выхода** |
| --- | --- | --- |
| Real GDP: YoY | ВВП России по использованию | Задержка 2 месяца (за второй квартал вышел в начале сентября) + корректировка |
| Households Consumption: YoY | Расходы на конечное потребление домашних хозяйств | Задержка 3 месяца (за второй квартал вышел в конце сентября) + корректировка |
| Government Consumption: YoY | Расходы на конечное потребление государственного управления | Задержка 3 месяца (за второй квартал вышел в конце сентября) + корректировка |
| Non-commercial Consumption: YoY | Расходы на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства | Задержка 3 месяца (за второй квартал вышел в конце сентября) + корректировка |
| Gross Capital Formation: YoY | Валовое накопление основного капитала | Задержка 3 месяца (за второй квартал вышел в конце сентября) + корректировка |
| Exports: YoY | Экспорт | Задержка 3 месяца (за второй квартал вышел в конце сентября) + корректировка |
| Imports: YoY | Импорт | Задержка 3 месяца (за второй квартал вышел в конце сентября) + корректировка |

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

*Таблица 2. Объясняющие переменные и календарь их выхода*

| **Переменная** | **Описание** | **Календарь выхода** |
| --- | --- | --- |
| Industry: YoY | Индекс промышленного производства | Задержка 1 месяц (в конце августа выходит за июль) |
| Mining: YoY | Индекс производства по добыче полезных ископаемых | Задержка 1 месяц (в конце августа выходит за июль) |
| Production: YoY | Индекс обрабатывающего производства | Задержка 1 месяц (в конце августа выходит за июль) |
| Energy: YoY | Индекс производства в секторе "Обеспечения электрической энергией, газом и паром; кондиционирования воздуха" | Задержка 1 месяц (в конце августа выходит за июль) |
| Agriculture: YoY | Индекс реального объема сельскохозяйственного производства | Задержка 1 месяц (в конце августа выходит за июль) |
| Construction: YoY | Индекс реального объема работ, выполненных по виду деятельности "Строительство" | Задержка 1 месяц (в конце августа выходит за июль) |
| House: YoY | Ввод в действие жилых домов | Задержка 1 месяц (в конце августа выходит за июль) |
| Transport: YoY | Коммерческий грузооборот транспорта | Задержка 1 месяц (в конце августа выходит за июль) |
| Trade: YoY | Индекс реального оборота розничной торговли | Задержка 1 месяц (в конце августа выходит за июль) |
| Unempl | Уровень безработицы | Задержка 1 месяц (в конце августа выходит за июль) |
| Inftation | Инфляция, рассчитанная по индексу потребительских цен | Задержка 1 месяц (в конце августа выходит за июль) |
| Oil Price: YoY | Цена на нефть марки Brent на последнюю дату месяца | 1-2 число каждого месяца, без задержек |
| ExRate USD: YoY | Официальный курс Доллар США – Российский рубль на последнюю дату месяца | 5-10 число каждого месяца, без задержек |
| ExRate EUR: YoY | Официальный курс Евро – Российский рубль на последнюю дату месяца | 5-10 число каждого месяца, без задержек |
| IMOEX: YoY | Значение индекса IMOEX на последнюю дату месяца | 1-2 число каждого месяца, без задержек |
| IRTS: YoY | Значение индекса IRTS на последнюю дату месяца | 1-2 число каждого месяца, без задержек |
| Gas Price: YoY | Цена на газ на последнюю дату месяца | 1-2 число каждого месяца, без задержек |
| MIACR | Среднемесячные фактические ставки по кредитам в рублях, предоставленным московскими банками, срок 1 день | 5-10 число каждого месяца, без задержек |
| MonBase: YoY | Денежная база (в широком определении) | 10 число каждого месяца, без задержек |
| M2: YoY | Денежная масса М2 | 10 число каждого месяца, задержка 1 месяц |
| Reserve: YoY | Международные резервы Российской Федерации | 5-10 число каждого месяца, без задержек |

**Результаты**

В таблицах 3-5 приведены средние абсолютные ошибки наукастов по всем рассматриваемым моделям и переменным. Для MIDAS моделей всех модификаций приведён лучший результат из всех объясняющих переменных. Используются следующие обозначения для моделей:

* MFBVAR – байесовская векторная авторегрессия смешанной частоты
* DFM – динамическая факторная модель
* MIDAS L1 – MIDAS модель с полным набором объясняющих переменных и L1-регуляризацией
* MIDAS R nealmon – MIDAS-модель с неэкспоненциальными лагами Алмона
* MIDAS U – MIDAS-модель без ограничений
* MS MIDAS U weighted – MIDAS-модель с марковским переключением: взвешенный прогноз
* MS MIDAS U most probable – MIDAS-модель с марковским переключением: прогноз по наиболее вероятному режиму
* MS MIDAS U best – MIDAS-модель с марковским переключением: прогноз по фактическому режиму
* MS PC MIDAS U weighted – MIDAS-модель с марковским переключением с главными компонентами: взвешенный прогноз
* MS PC MIDAS U most probable – MIDAS-модель с марковским переключением с главными компонентами: прогноз по наиболее вероятному режиму
* MS PC MIDAS U best – MIDAS-модель с марковским переключением с главными компонентами: прогноз по фактическому режиму
* AR(1) – прогноз по AR(1) модели в качестве бенчмарка

В таблице 3 приведены результаты для случая использования всех доступных данных за квартал, жирным выделены минимальные ошибки, курсивом – минимальные среди моделей без правильного угадывания фактического режима:

*Таблица 3. MAE наукастов для полной выборки*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Real GDP | Households Consumption | Government Consumption | Non commercial Consumption | Gross Capital Formation | Exports | Imports |
| MFBVAR | 1.850 | 4.408 | 1.009 | 1.590 | 3.192 | 4.498 | 5.493 |
| DFM | 2.096 | 5.611 | 1.250 | 3.509 | 2.772 | 3.954 | 6.278 |
| MIDAS L1 | 4.678 | 10.038 | 3.014 | 3.815 | 4.421 | 4.274 | 14.864 |
| MIDAS R nealmon | 1.667 | 3.276 | 0.628 | 2.492 | *2.186* | 4.119 | 6.114 |
| MIDAS U | 1.764 | 3.219 | 0.678 | 2.485 | 2.189 | 4.298 | 6.146 |
| MS MIDAS U weighted | 1.825 | 2.754 | *0.521* | ***0.564*** | 2.201 | 4.046 | 7.493 |
| MS MIDAS U most probable | 1.973 | *2.627* | 0.536 | 0.635 | 2.204 | *3.733* | 8.845 |
| MS MIDAS U best | **0.770** | **1.983** | **0.492** | 0.635 | **1.954** | **2.703** | 4.816 |
| MS PC MIDAS U weighted | *1.274* | 5.006 | 0.621 | 2.606 | 2.754 | 4.803 | *5.115* |
| MS PC MIDAS U most probable | 1.341 | 4.995 | 0.619 | 2.509 | 2.453 | 5.206 | 5.876 |
| MS PC MIDAS U best | 0.871 | 4.141 | 0.587 | 2.321 | 2.453 | 3.255 | **4.202** |
| AR(1) | 2.901 | 6.848 | 0.611 | 1.822 | 4.481 | 3.436 | 8.438 |

Для всех показателей, кроме расходов на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства, максимальная точность обеспечивается MIDAS-моделями с марковским переключением и правильным предсказанием фактического режима. Если рассматривать «реалистичные» модели, не требующие правильного угадывания режима, то наибольшая точность также достигается MIDAS-моделями с марковским переключением: с главными компонентами и взвешиванием для темпов роста ВВП и импорта, со взвешиванием для государственного потребления и потребления НКО, по наиболее вероятному режиму – для потребления домашних хозяйств и экспорта. Единственным показатель, для которого лучший наукаст обеспечивает модель без переключения – это валовое накопление основного капитала, где обычная MIDAS R показывает MAE 2.186 против 2.201 у лучшей из моделей с переключением.

MIDAS-модели с переключением и правильным предсказанием фактического режима оказываются лучше базовой AR(1) модели для всех изучаемых показателей, модели без правильного предсказания режима уступают AR(1) только для экспорта.

В таблицах 4-5 приведены MAE для случаев расчёта наукастов без последнего месяца квартала и без последних двух месяцев квартала соответственно:

*Таблица 4. MAE наукастов без последнего месяца квартала*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Real GDP | Households Consumption | Government Consumption | Non commercial Consumption | Gross Capital Formation | Exports | Imports |
| MFBVAR | 1.506 | 4.297 | 0.918 | 1.397 | 2.426 | 4.531 | *5.284* |
| DFM | 2.404 | 6.030 | 1.389 | 3.455 | 3.693 | 3.820 | 8.060 |
| MIDAS L1 | 4.734 | 10.147 | 2.894 | 3.770 | 4.331 | 4.291 | 14.744 |
| MIDAS R nealmon | 1.457 | 2.752 | 0.634 | 2.515 | *2.204* | 4.250 | 6.116 |
| MIDAS U | 1.364 | 2.691 | 0.656 | 2.508 | *2.204* | 4.364 | 6.139 |
| MS MIDAS U weighted | 1.655 | 2.557 | *0.515* | ***0.611*** | 2.225 | 4.027 | 7.701 |
| MS MIDAS U most probable | 1.808 | *2.400* | 0.523 | 0.677 | 2.175 | *3.754* | 8.221 |
| MS MIDAS U best | **0.784** | **1.847** | **0.496** | 0.677 | **1.853** | **2.810** | 5.352 |
| MS PC MIDAS U weighted | *1.328* | 4.836 | 0.618 | 2.600 | 2.475 | 4.749 | 5.296 |
| MS PC MIDAS U most probable | 1.337 | 5.043 | 0.618 | 2.449 | 2.150 | 4.756 | 5.613 |
| MS PC MIDAS U best | 0.912 | 3.831 | 0.618 | 2.449 | 2.150 | 3.332 | **4.490** |
| AR(1) | 2.685 | 5.993 | 0.635 | 2.155 | 3.673 | 4.224 | 8.312 |

*Таблица 5. MAE наукастов без последних двух месяцев квартала*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Real GDP | Households Consumption | Government Consumption | Non commercial Consumption | Gross Capital Formation | Exports | Imports |
| MFBVAR | 1.281 | 4.139 | 1.047 | 1.262 | 2.698 | 4.171 | 6.022 |
| DFM | 2.159 | 5.674 | 1.436 | 3.473 | 3.392 | 3.627 | 7.919 |
| MIDAS L1 | 4.655 | 10.023 | 2.858 | 3.398 | 4.365 | 4.411 | 15.115 |
| MIDAS R nealmon | 2.110 | 3.147 | 0.636 | 2.616 | 2.502 | 3.581 | 7.420 |
| MIDAS U | 2.063 | 3.149 | 0.669 | 2.647 | 2.503 | *3.579* | 7.420 |
| MS MIDAS U weighted | 1.633 | 2.667 | ***0.444*** | ***0.638*** | 2.394 | 3.756 | 8.912 |
| MS MIDAS U most probable | 1.524 | *2.508* | 0.564 | 0.704 | 2.499 | 3.810 | 9.122 |
| MS MIDAS U best | 1.197 | **1.770** | 0.496 | 0.673 | **1.935** | **2.426** | 5.457 |
| MS PC MIDAS U weighted | *1.209* | 4.602 | 0.599 | 2.620 | 2.469 | 4.530 | *4.595* |
| MS PC MIDAS U most probable | 1.230 | 4.665 | 0.605 | 2.568 | *2.077* | 4.529 | 4.665 |
| MS PC MIDAS U best | **0.799** | 3.426 | 0.571 | 2.492 | 1.968 | 3.190 | **3.999** |
| AR(1) | 2.448 | 5.692 | 0.647 | 2.122 | 4.262 | 4.409 | 8.219 |

Анализируя полученные результаты, можно сделать ряд выводов:

1. Как и для случая использования полной выборки, для большинства показателей максимальная точность достигается при использовании той или иной версии наукастинга на основе MIDAS-моделей с марковским переключением. Единственное исключение – это импорт по данным за первые два месяца квартала, когда максимальную точность среди моделей без правильного угадывания режима показывает MFBVAR.
2. Для всех случаев расчёта наукастов на основе неполных данных модели с переключением демонстрируют точность выше базовых AR(1) моделей
3. Использование главных компонент даёт наибольший прирост точности для показателей, на которые (потенциально) может влиять большее количество качественно разных переменных – ВВП и импорт. При этом по мере уменьшения объёма данных важность использования главных компонент возрастает: отношение MAE моделей с главными компонентами и исходными показателями для прогнозов с правильным выбором режимов для ВВП изменяется от 1.13 для наукастов по полным данным до 0.67 по данным за 1 месяц квартала; для импорта – от 0.87 до 0.73
4. Для моделей с марковским переключением, как и для остальных рассмотренных моделей, характерен отмеченный в [Макеева, Станкевич, 2022] эффект роста точности наукастов по мере уменьшения объёма данных, по которым этот наукаст рассчитывается, для многих показателей.
5. Отдельно стоит отметить особенно заметный в некоторых рядах для наукастов без последнего месяца эффект одинаковой точности моделей по наиболее вероятному режиму и по фактическому режиму. Государственное потребление, потребление НКО и ВНОК во все вошедшие в тестовую выборку периоды времени меняют свои состояния в строгом соответствии с модельными вероятностями перехода

Интересно также изучить поведение моделей с пересечением с точки зрения точности и глубины идентификации кризисных периодов (см. рис. 1)

*Рисунок 1: наукасты темпов роста ВВП для моделей с переключением и без переключения*

Модели с переключением гораздо точнее улавливают эффекты смены тренда: спад во 2 квартале 2020 предсказан более точно и минимальная точка спада верно спрогнозирована во 2 квартале, а не в 3, как для лучшей модели без переключения (MIDAS R модель). Аналогично и для 3 квартала 2021, когда завершилось пост-кризисное восстановление, а модель без переключения не смогла идентифицировать этот переход.

**Заключение**

Предложенная в работе методика использования MIDAS-моделей с марковским переключением для задач наукастинга продемонстрировала высокое качество работы на данных по ВВП России и его компонентам. При использовании фактического режима при расчёте наукастов точность MS MIDAS модели оказывается наибольшей среди всех рассмотренных в исследовании моделей, при использовании моделей без фактического режима точность для большинства рядов также оказывается выше, чем у других моделей.

Полученные результаты позволяют сделать вывод о перспективности использования моделей с переключением для наукастинга и необходимости дальнейшего развития этого модельного аппарата: как в сторону усложнения структуры переключения (моделирования вероятности перехода из состояния в состояние как отдельного процесса), так и в сторону развития базовых моделей – к примеру, использования хорошо зарекомендовавших себя в задачах наукастинга VAR-моделей смешанной частоты в модификации с марковским переключением.

**Литература**

Гареев М. Ю., Полбин А. В. Наукастинг: оценка изменения ключевых макроэкономических показателей с использованием методов машинного обучения // Вопросы экономики. – 2022. – №. 8. – С. 133-157.

Зубарев А., Рыбак К. Наукастинг ВВП: динамическая факторная модель и официальные прогнозы // Экономическое развитие России. – 2021. – Т.28. – №12. – С. 34-40

Макеева Н. М., Станкевич И. П. Наукастинг элементов использования ВВП России // Экономический журнал Высшей школы экономики. – 2022. – Т. 26. – №. 4. – С. 598-622.

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. – С. 60.

Станкевич И.П. Сравнение методов наукастинга макроэкономических индикаторов на примере российского ВВП // Прикладная эконометрика. – 2020. – Т. 59. – С. 113-127

Barnett W. A., Tang B. Chinese divisia monetary index and GDP nowcasting //Open Economies Review. – 2016. – Т. 27. – №. 5. – С. 825-849.

Ferrara L., Marsilli C. Nowcasting global economic growth: A factor‐augmented mixed‐frequency approach // The World Economy. – 2019. – Т. 42. – №. 3. – С. 846-875

Ghysels E., Kvedaras V., Zemlys V. Mixed frequency data sampling regression models: the R package midasr // Journal of statistical software. – 2016. – С. 1-35

Guérin P., Marcellino M. Markov-switching MIDAS models //Journal of Business & Economic Statistics. – 2013. – Т. 31. – №. 1. – С. 45-56.

Hamilton J. D. A new approach to the economic analysis of nonstationary time series and the business cycle // Econometrica: Journal of the econometric society. – 1989. – С. 357-384.

Hopp, D. (2022). Benchmarking Econometric and Machine Learning Methodologies in Nowcasting. UNCTAD Research Paper, (83).

Ingenito R. et al (1996). Using monthly data to predict quarterly output. Economic Review (Federal Reserve Bank of San Francisco), 3, 3–11.

Jardet C., Meunier B. Nowcasting world GDP growth with high-frequency data // Journal of Forecasting. – 2022. – С. 1181-1200

Kuzin V., Marcellino M., Schumacher C. MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area // International Journal of Forecasting. – 2011. – Т. 27. – №. 2. – С. 529-542

Litterman, Robert B. Techniques of forecasting using vector autoregressions. No. 115. 1979.

Litterman, Robert B. Forecasting with Bayesian vector autoregressions—five years of experience // Journal of Business & Economic Statistics – 1986. – Т. 4 – №. 1. – С. 25-38

Lu X. et al. Forecasting oil futures realized range‐based volatility with jumps, leverage effect, and regime switching: New evidence from MIDAS models // Journal of Forecasting. – 2022. – Т. 41. – №. 4. – С. 853-868.

Ma F. et al. Cryptocurrency volatility forecasting: A Markov regime‐switching MIDAS approach //Journal of Forecasting. – 2020. – Т. 39. – №. 8. – С. 1277-1290.

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. – Т. 72. – №. 4. – С. 518-550

McCracken M. W., Owyang M., Sekhposyan T. Real-time forecasting with a large, mixed frequency, Bayesian VAR //FRB St. Louis Working Paper. – 2015. – №. 2015-30.

Proiettia 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), 1376-1398

Schorfheide F., Song D. Real-time forecasting with a mixed-frequency VAR // Journal of Business & Economic Statistics. – 2015. – Т. 33. – №. 3. – С. 366-380

Schumacher C. (2016). A comparison of MIDAS and bridge equations. International Journal of Forecasting, 32 (2), 257–270

Segnon M., Gupta R., Wilfling B. Forecasting stock market volatility with regime-switching GARCH-MIDAS: the role of geopolitical risks // International Journal of Forecasting. – 2023.

Soybilgen B., Yazgan E. (2018). Evaluating nowcasts of bridge equations with advanced combination schemes for the Turkish unemployment rate. Economic Modelling, 72, 99–108

*Zhemkov M.* Nowcasting Russian GDP using forecast combination approach // International Economics. – 2021. – №168. – С. 10-24