

СРАВНЕНИЕ МЕТОДОВ АГРЕГИРОВАНИЯ ПРОГНОЗОВ ДИНАМИКИ КРЕДИТОВ ЮРИДИЧЕСКИХ ЛИЦ

А. А. Мелкумян, О. С. Балаш

*Саратовский национальный исследовательский
государственный университет им. Н. Г. Чернышевского, Россия*
E-mail: melkumyan.ann@yandex.ru, olgalbash@mail.ru

Кредитование играет ключевую роль в развитии экономики, и его исследование является важной задачей для понимания динамики финансовых рынков и их влияния на макроэкономическую стабильность. Представленная статья посвящена анализу динамики кредитов, предоставленных юридическими лицами в РФ с 2019 по 2023 год. Рассматриваются методы агрегирования прогнозов динамики кредитов по регионам и федеральным округам. Оценивается качество моделей ARIMA и ETS для тех или иных регионов. Обсуждается применение различных методов согласования прогнозов и кластеризации регионов по характеристикам динамики кредитов. Исследуется проблема согласования индивидуальных и агрегированных краткосрочных прогнозов, а также вопрос о группировке регионов по особенностям динамики выбранных показателей.

COMPARISON OF FORECAST AGGREGATION METHODS FOR LEGAL ENTITY CREDIT DYNAMICS

A. A. Melkumyan, O. S. Balash

Lending plays a pivotal role in economic development, and its exploration is a significant task for understanding the dynamics of financial markets and their impact on macroeconomic stability. This article is dedicated to analyzing the dynamics of loans provided to legal entities in Russia from 2019 to 2023. The methods of aggregating forecasts for loan dynamics across regions and federal districts are examined. The quality of ARIMA and ETS models for various regions is assessed. The article discusses the application of different forecast reconciliation methods and the clustering of regions based on loan dynamics characteristics. It explores the issue of reconciling individual and aggregated short-term forecasts and the grouping of regions based on the features of the selected indicators' dynamics.

В настоящее время кредиты являются ключевым инструментом в экономике, обеспечивая доступ к необходимым финансовым ресурсам для предприятий, индивидуальных предпринимателей и других хозяйствующих субъектов. Суть кредитования заключается во временном перераспределении финансовых ресурсов, что способствует инвестициям, развитию производства и, в конечном итоге, росту экономики. Важно подчеркнуть, что объем кредитов, предоставленных юридическим лицам, является отражением спроса на финансовые ресурсы для корпоративных нужд. Анализ динамики объема предоставленных кредитов позволяет лучше понять, как экономические и финансовые факторы взаимодействуют с кредитными рынками и могут влиять на структуру финансирования

предприятий и бизнеса в России.

Целью данного исследования является сравнение методов агрегирования прогнозов динамики кредитов юридических лиц и индивидуальных предпринимателей по регионам Российской Федерации за период с 2019 по 2023 гг. Исходные данные представляют собой ежемесячные значения по кредитам в 82 регионах: значения по Ненецкому автономному округу входят в данные Архангельской области, а по Ханты-Мансийскому автономному округу и Ямало-Ненецкому автономному округу – в Тюменскую область, так как отдельно за выбранный период они не рассматриваются [1].

Для достижения поставленной цели были использованы модели ARIMA и ETS, применены методы агрегирования данных на региональном уровне, на уровне федеральных округов и РФ в целом. Также сравнивались методы агрегирования "снизу вверх" (bottom-up), метод наименьших квадратов и (ols) метод минимальных отклонений (min_trace). Прогнозы на более высоком уровне агрегации (по федеральным округам) получались путем суммирования прогнозов на более низком уровне (по областям).

В ходе исследования рассматривается проблема согласования индивидуальных и агрегированных краткосрочных прогнозов. Необходимо отметить, что сумма индивидуальных прогнозов может не совпадать с прогнозом по агрегированным данным и на разных уровнях наибольшую точность прогнозов могут давать разные методы.

В первую очередь проводится анализ прогнозов кредитов по каждому региону с использованием вышеназванных моделей. Для этого строятся графики временных рядов, включающих фактические и прогнозируемые значения. Визуальное сравнение графиков показало, что прогнозы почти у всех моделей по регионам не следуют общему тренду и колебаниям временного ряда. Отметим, что для сопоставления результатов моделирования использовались информационные критерии AIC, AICc, BIC. Известно, что меньшие значения AIC и BIC указывают на лучшую модель.

Так, было выяснено число случаев, когда та или иная модель описывала данные наилучшим образом для конкретного региона. ARIMA в 74 случаях оказалась наилучшей (в 46 случаях информационные критерии принимали отрицательные значения, а в 28 – положительные, но все равно были меньше аналогичных в модели ETS), а ETS – в 8 (информационные критерии хоть и принимали положительное значение, все равно оказывались меньше, чем соответствующие в модели ARIMA).

Заметим, что различия в значениях информационных критериев между моделями ARIMA и ETS могут быть связаны с особенностями данных и характеристиками временного ряда в конкретных регионах.

Далее проведено агрегирование прогнозов по округам с использованием подхода, реализованного в моделях группового прогнозирования временных рядов. Прогнозирование объединенных показателей позволяет получить общую картину и прогноз для группы или региона, основываясь на агрегированных данных.

Для осуществления данного прогнозирования создается иерархический временной ряд с кредитами по регионам на нижнем уровне иерархии, которые агрегированы по федеральным округам на среднем уровне, и далее – в национальный итог на верхнем [2].

Такого рода моделирование полезно, когда интересующие исследователя показатели имеют общие тенденции и паттерны на более высоком уровне агрегации, которые могут быть проще и точнее уловлены моделями. Подобный анализ позволяет принимать более обобщенные решения и строить стратегии на уровне регионов или групп, а не на уровне отдельных наблюдений.

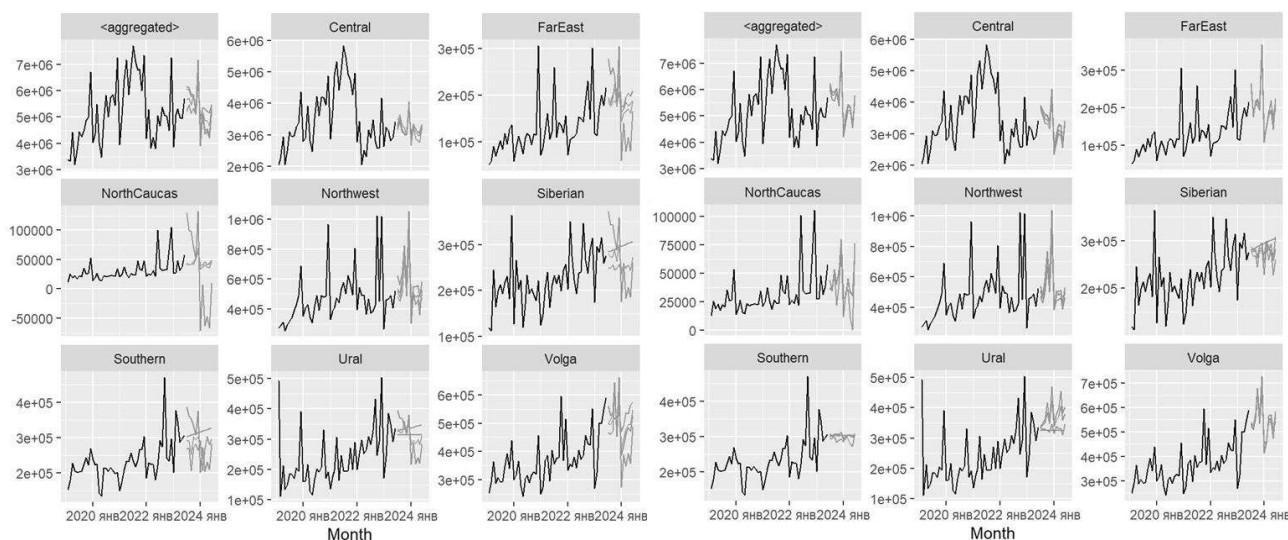


Рис. 1. Агрегированный прогноз по округам ARIMA и ETS

Графики совокупных кредитов на верхнем и среднем уровнях иерархии отразили следующую динамику: в целом по всем федеральным округам зарегистрирован положительный тренд изменения объемов кредитов, предоставленных юридическим лицам. Однако на протяжении всего исследуемого периода динамика характеризовалась резкими периодическими колебаниями, включая временные пики и спады. Эти колебания могут быть связаны с различными факторами, такими как экономическая ситуация внутри страны, геополитические риски, включая кризисы, меры монетарной политики, спрос на кредиты и подобное.

Следует упомянуть, что агрегированный по всем федеральным округам прогноз почти в точности повторяет агрегированный прогноз по Центральному федеральному округу. Вероятно, данная тенденция может быть результатом либо весовых коэффициентов, либо зависимости данных – наиболее очевидных причин. Соответственно, если весовые коэффициенты сильно предпочитают центральный федеральный округ, то агрегированный прогноз будет тесно следовать прогнозу этого округа.

Более того, отраженные на графиках высокие скачкообразные движения исследуемого показателя в основном выпадают на начало каждого последую-

щего года, что в большей степени, на наш взгляд, объясняется сезонными факторами, планированием временных сделок, изменением стратегий на старте нового финансового года, а также бухгалтерской отчетностью.

Далее к агрегированным данным применяются модели ARIMA и ETS, отраженные на рис. 1. Выполняется согласование прогнозов с использованием трех методов согласования: *bottom-up*, *ols* (метод наименьших квадратов) и *mint* (метод наименьших отклонений с учетом тренда). Сравнение формы прогнозов и трендов между моделями позволяет сделать вывод о том, что каждая из них визуально соответствует физическим значениям и достаточно четко захватывает паттерны временного ряда. Графики агрегированных фактических значений совпадают, но прогнозы немного разнятся, хоть и сохраняют схожие тренды.

В случае модели ARIMA с восходящими методами сглаживания (*bottom-up*, *mint*, *ols*), различия в прогнозах могут быть связаны с разными подходами к учету и прогнозированию данных. Каждый из методов может иметь свои особенности и ставить упор на определенные аспекты данных. В модели ETS же различия в прогнозах также могут быть обусловлены различными подходами к учету тренда, сезонности и ошибок модели.

Соответственно, сравнение исследуемых методов позволяет выбрать тот или иной, в зависимости от цели исследования. Например, *bottom-up* метод может быть полезен, если интересует прогнозирование на уровне отдельных регионов, в то время как *mint* и *ols* методы могут учитывать структуру и связь между регионами при агрегации. Более того, если в данных присутствует ярко выраженная сезонность, методы сглаживания, учитывающие сезонные компоненты, могут быть более предпочтительными.

Так, в данном случае мы делаем вывод о том, что методы *ols* и *mint* в модели ETS дают более точные агрегированные прогнозы для кредитов юридических лиц.

Помимо всего вышеперечисленного, мы рассматривали вопрос о группировке регионов по особенностям динамики выбранных показателей. Для этого мы использовали подход, реализованный в библиотеке *tsfeatures*.

В этом подходе кластеризацию временных рядов предлагается проводить с использованием большого числа предварительно вычисленных показателей, отражающих те или иные характеристики динамики, такие как изменчивость и стабильность, наличие точек перелома тенденций, среднее и максимальное значение цепных темпов прироста, нелинейность и т.д.

Так, производится визуализация собственных значений и информации об объясненной дисперсии в результате анализа главных компонент (PCA). Исходя из полученных графиков, можно отметить, что изучаемые данные имеют 10 главных компонент, каждая из которых вносит свой вклад в объяснение дисперсии данных. Выяснено, что первые три компоненты объясняют наибольшую часть дисперсии данных (в сумме около 58.5%), а остальные компоненты объясняют меньший процент дисперсии.

Далее визуализируем индивидуальные наблюдения на двумерной плоскости, которая создана с использованием результатов анализа главных компонент

(PCA). Полученный график показывает индивидуальные наблюдения (регионы) на двумерной плоскости, где оси представляют первые две главные компоненты (PC1 и PC2).

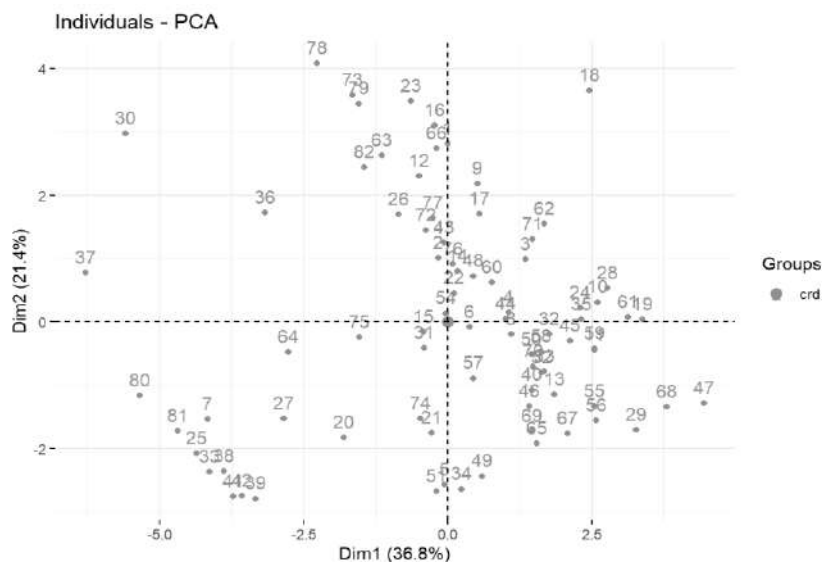


Рис. 2. Индивидуальные наблюдения с использованием PCA

Как видно на графике индивидуальных наблюдений, регионы скорее разбросаны по диапазону, хоть и относительно небольшому. Это может указывать на то, что те, регионы, которые сконцентрированы на одном небольшом участке, могут иметь схожие или связанные характеристики между ними, а вот регионы, находящиеся в отдалении – имеют уникальные характеристики или особенности в своей динамике кредитов.

Дополнительно отметим, что также был проведен анализ графиков вложения данных по методам UMAP и снижения размерности данных по технике t-SNE, который засвидетельствовал о том, что могут существовать группы или кластеры регионов с общими чертами или схожим поведением в отношении динамики кредитов юридических лиц.

Таким образом, можно сказать, что анализ динамики кредитов юридических лиц позволяет делать необходимые выводы для принятия некоторых решений, например, в области инвестиционной активности – увеличение объемов кредитов может указывать на рост бизнеса и инвестиционные возможности в данном регионе, в то время как снижение объема кредитов может свидетельствовать о финансовых трудностях. Более того, информация о динамике данного показателя может позволить оценить ситуацию в мониторинге банковской системы, оценки экономического состояния и рисков, политики монетарного управления и т.д. А также агрегирование данных по федеральным округам позволяет получить более общую картину и сравнить различные регионы между собой: определить успешные стратегии и факторы, которые могут быть применены в других областях.

СПИСОК ЛИТЕРАТУРЫ

1. Официальный сайт Банка России. [Электронный ресурс]. URL: <https://www.cbr.ru/statistics/bbs/> (дата обращения: 21.09.2023).
2. *Hyndman R. J., Athanasopoulos G.* Forecasting: principles and practice, 3rd edition. OTexts: Melbourne, Australia, 2021.