

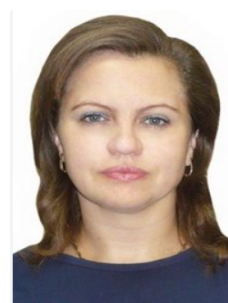
УДК 330.45

ВЫРАВНИВАНИЕ ДИНАМИЧЕСКИХ РЯДОВ В АНАЛИТИКЕ БОЛЬШИХ ДАННЫХ В ЦИФРОВОЙ ЭКОНОМИКЕ



Э. Е. Дадашев

Бакалавр направления подготовки «Экономика»,
профиль «Экономика и организация ИТ бизнеса»,
гр. БЭЭ2401, МТУСИ
ericandeev2016@yandex.ru



М. Д. Самакаева

Доцент кафедры цифровой экономики,
управления и бизнес-технологий, МТУСИ
кандидат экономических наук, доцент
samakaeva-m@yandex.ru

Э. Е. Дадашев

Студент направления подготовки «Экономика», профиль «Экономика и организация ИТ бизнеса», гр. БЭЭ2401, Московского технического университета связи и информатики. Круг научных интересов: экономика, бизнес-технологии, аналитика, экономическая статистика.

М. Д. Самакаева

Окончила Оренбургский государственный педагогический университет. Область научных интересов: цифровая экономика, бизнес-технологии, аналитика, экономическая статистика, маркетинг, управление проектами.

Аннотация. В условиях ускоряющейся цифровой трансформации экономики ключевые бизнес-процессы, взаимодействие участников рынков и поведение потребителей всё чаще описываются большими массивами данных, формируемыми в режиме близком к реальному времени. Показатели электронной коммерции, использования облачных сервисов, цифровых финансовых услуг, телеком-инфраструктуры и онлайн-активности пользователей фиксируются в виде протяженных временных рядов, которые становятся основой аналитики и поддержки управленческих решений [3,6,7]. В данной статье рассматриваются модели и методы анализа и выравнивания динамических рядов при анализе больших данных в экономике. Применение динамических рядов при анализе данных позволяет выстраивать более точные прогнозы и формировать стратегии развития с более предсказуемым результатом.

Ключевые слова: цифровая трансформация, динамические ряды, Big Data, модели, динамические ряды, статистическое выравнивание, методы, модели, управленческие, градиентный бустинг, автоматическое прогнозирование, платформы, искусственный интеллект

Введение. Вопросы анализа и выравнивания динамических рядов в экономической статистике подробно рассматриваются в работах отечественных и зарубежных исследователей, посвящённых трендовым моделям, методам скользящего среднего, экспоненциального сглаживания и аналитическому выравниванию временных рядов [2,3,4,5,6,11]. Значительный вклад в развитие экономико-статистического аппарата внесли исследования, в которых обосновываются подходы к выбору функциональной формы тренда (линейной, полиномиальной, экспоненциальной) и оценивается их точность при анализе социально-экономических процессов.

Традиционные динамические ряды рассматриваются как результат совместного действия нескольких компонент: долгосрочного тренда, сезонной и циклической составляющей, а также случайных (нерегулярных) колебаний. Однако такая система имеет ряд преимуществ в классической модели, но несет в себе критические недостатки, в случае с Big Data что подробно описано в учебной и монографической литературе по экономической

статистике и эконометрике [2,3,4,5]. Эти и другие проблемы обусловили актуальность нашего исследования.

Цель – исследовать теоретические проблемы и разработать практические рекомендации по повышению эффективности использования выравнивания динамических рядов в Больших данных.

Из цели вытекают следующие задачи:

1. Исследовать теоретические аспекты использования динамических рядов в бизнес аналитике.
2. Выделить проблемы использования данных моделей.
3. Разработать практические рекомендации по выравниванию динамических моделей при использовании больших данных в цифровой статистике.

Данными проблемами занимались российские ученые, такие как: Капица С.П. [2], Курдюмов С.П., Малинецкий Г.Г., Ильясов Р.Х., Куразова Д.А. Кузовкова Т.А., Шаравов И. М. , Курицын Н. С. [3], Кумратова А.М., Плотников В.А. [4], Платунина Г.П. [6], Салюткина Т.Ю. [6], М. Д. Самакаева [11], Дадашев Э. Е. [11], Минаков В.Ф. [4] и др.

Зарубежные авторы, занимающиеся проблемами математического моделирования и прогнозированием на основе Искусственного интеллекта: Синьей Ху, Олчай Чирит, Танмай Бинайкия, Рамит Хора [13], Си Чен, Мин Го, Уюэ Шангуань [12] и др.

Объектом исследования выступают процессы цифровой экономики, описываемые временными рядами показателей электронной коммерции, цифровых услуг и онлайн-активности пользователей [3, 7, 9]. Предмет исследования - методы анализа, выравнивания и прогнозирования динамических рядов в условиях использования больших данных, а также их влияние на качество управленческих решений в цифровой среде.

В работе применены общенаучные методы исследования: анализ и синтез, индукция и дедукция, сравнение и обобщение, а также логический и системный подход к изучению социально-экономических процессов. Методологическую основу составляет экономико-статистический и эконометрический анализ динамических рядов, включающий аналитическое выравнивание тренда (линейные и нелинейные зависимости), методы скользящего среднего, экспоненциальное сглаживание и модели класса ARIMA/ARIMAX [12]. Для учёта особенностей больших данных и сложной структуры цифровых временных рядов дополнительно рассматриваются модели со структурными сдвигами и сменой режимов, а также методы машинного обучения, основанные на деревьях решений и градиентном бустинге [13].

Следует отметить, что процессы цифровой экономики, описываемые временными рядами показателей электронной коммерции, цифровых услуг и онлайн-активности пользователей являются современной основой использования больших данных в цифровой среде. Дополнительно значимым направлением цифровизации является развитие фриланс-платформ, формирующих новые модели занятости и требующих анализа временных рядов спроса и предложения на цифровом рынке труда [2,9,11] Предмет исследования - методы анализа, выравнивания и прогнозирования динамических рядов в условиях использования больших данных, а также их влияние на качество управленческих решений в цифровой среде.

Вопросы анализа и выравнивания динамических рядов в экономической статистике подробно рассматриваются в работах отечественных и зарубежных исследователей, посвящённых трендовым моделям, методам скользящего среднего, экспоненциального сглаживания и аналитическому выравниванию временных рядов [5]. Значительный вклад в развитие экономико-статистического аппарата внесли исследования, в которых обосновываются подходы к выбору функциональной формы тренда (линейной, полиномиальной, экспоненциальной) и оценивается их точность при анализе социально-экономических процессов [6,7]. В современных работах по цифровой экономике и Big Data акцент переносится на интеграцию классических методов временных рядов с моделями

машинного обучения и интеллектуального анализа данных, что позволяет обрабатывать высокочастотные и высокоразмерные массивы цифровой информации [8].

В современных работах по цифровой экономике и Big Data акцент переносится на интеграцию классических методов временных рядов с моделями машинного обучения и интеллектуального анализа данных, что позволяет обрабатывать высокочастотные и высокоразмерные массивы цифровой информации [8] (рисунок 1).

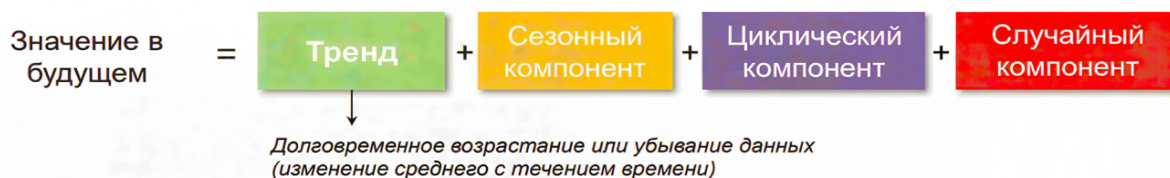


Рисунок 1. Классическая аддитивная модель временного ряда [13]

Классический подход опирается на явную структурную модель: тренд, сезонность и случайная компонента. В итоге малое число параметров позволяет четко интерпретировать результаты, но плохо масштабируется на тысячи и миллионы параллельных рядов, вынуждая строить модель под каждый ряд значений. Стационарность, фиксированная и регулярная сезонность, ограничивается относительно коротким временным горизонтом анализа. Сезонность и тренд предполагаются относительно стабильными, однако современные паттерны быстро меняются под влиянием алгоритмических рекомендаций и конкуренции, представлены в статистических исследованиях [1].

Классические модели с регрессорами (ARIMAX/SARIMAX и т.п.) умеют включать ограниченный набор внешних переменных, но когда мы имеем дело с большими данными мы говорим об интеграции в систему анализа факторов о сотне и тысяче поведенческих и контекстных факторов: клики, глубина просмотра, источники трафика, свойства пользователя, конкуренты, погода и т.п. [12]. Важно отметить, что вышеуказанные особенности классических моделей не предполагают отказа от них, а стимулируют создавать новые методы анализа данных, для их адаптации или совместного применения с более современными подходами [13]. Среди ключевых современных методик анализа динамических рядов, можно выделить: машинное обучение, глубинное обучение, Break-point и Regime-switching методы, градиентный бустинг и дерево решений [12].

Можно выделить следующие положительные и отрицательные стороны использования больших данных для формирования статистических моделей, представленных в таблице 1.

Таблица 1. Положительные и отрицательные стороны использования больших данных

| Положительные стороны | Отрицательные стороны |
|---|--|
| 1 | 2 |
| 1. Улучшение принятия решений: анализ больших объемов данных позволяет организациям принимать более обоснованные решения, основанные на фактических данных и тенденциях. | 1. Конфиденциальность и безопасность: сбор и хранение больших объемов личных данных вызывает опасения по поводу конфиденциальности и безопасности информации. Утечки данных могут иметь серьезные последствия. |
| 2. Персонализация: компании могут использовать большие данные для создания персонализированных предложений и улучшения клиентского опыта, что может привести к повышению лояльности клиентов. | 2. Сложность анализа: обработка и анализ больших данных требуют значительных ресурсов и специализированных навыков, что может быть дорогостоящим для компаний. |

Продолжение таблицы 1

| 1 | 2 |
|--|---|
| 3. Оптимизация процессов: анализ данных помогает выявить узкие места в бизнес-процессах и оптимизировать их, что может привести к снижению затрат и повышению эффективности. | 3. Качество данных: большие объемы данных могут содержать ошибки или нерелевантную информацию, что может привести к неправильным выводам и решениям. |
| 4. Инновации: большие данные могут способствовать разработке новых продуктов и услуг, основанных на потребностях и предпочтениях клиентов. | 4. Зависимость от технологий: организации могут стать зависимыми от технологий и инструментов анализа данных, что может создать риски в случае их сбоя или устаревания. |
| 5. Предсказательная аналитика: использование алгоритмов машинного обучения позволяет прогнозировать будущие тренды и поведение пользователей, что может быть полезно в различных отраслях. | |

По материалам исследований авторов

Использование больших данных (Big Data) имеет как положительные, так и отрицательные стороны, важно, чтобы эти данные были объективными.

В цифровой экономике большие данные становятся основой не только для классических онлайн-сервисов, но и для новых моделей монетизации, таких как криптовалютные платформы, где поведенческие данные пользователей служат ключевым ресурсом для генерации дохода [2,6, 8,11].

В современных задачах анализа массивов транзакционных и поведенческих данных, ограничение классической модели компенсируется переходом к методам машинного обучения [3].

Основным преимуществом использования данного метода является особенность построения регрессии не на «тренде и сезонности», а на большом наборе признаков: лаги, скользящие окна, календарные признаки; а также таких внешних факторов, как маркетинг и поведение пользователей. Такой подход позволяет строить деревья решений, автоматически моделируя нелинейности и взаимодействия признаков, при таком методе структура не задается вручную, а обучается на основе данных, что позволяет модели построить многомерную зависимость и одновременно учитывать множество переменных факторов [10].

Нестабильность структуры ряда является одной из ключевых проблем в e-commerce и онлайн трафике, с которой классический подход не может справиться. Большое количество структурных сдвигов в тренде, изменения в системе рекомендаций и системе аукционов в рекламных кабинетах.

Сезонные паттерны не статичны и постоянно меняются, а динамическое распределение спроса и трафика вызывает медленное или резкое изменение статистических свойств во времени [9].

Для решения проблем со стационарностью, современные методы и подходы предлагают использовать модели со сменой режим Break-point (точки разрыва) модели ищут конкретные даты, в которые произошли резкие изменения трендов, такие как обвал или рост, разделяя тем самым весь динамический ряд значений на части и подбирая под каждый отдельный участок наиболее точные параметры для расчетов. Regime-switching (Смена режимов) модели представляют из себя более совершенную систему, полностью меняя алгоритмы анализа не только на основе временных периодов, но и учитывая помимо них ряд других внутренних и внешних параметров системы, таких как: покупательский спрос, дата, суточное время, текущая занятость сотрудников, погодные условия и прочие [8].

Список параметров изменяется в зависимости от того в какой сфере и отрасли применяется конкретная модель.

Помимо этого, необходимо учитывать изменение параметров и значений в реальном времени для этого в моделях используются онлайн-обновление и адаптивное взвешивание.

Технически это реализуется через одновременное использование двух моделей, где первая строится на исторических данных, а вторая только на недавних событиях. Финальные же прогнозы строятся на основе взвешенных сумм прогнозов обеих моделей, благодаря адаптивной системе весов для каждого параметра модели [12] (рисунок 2).

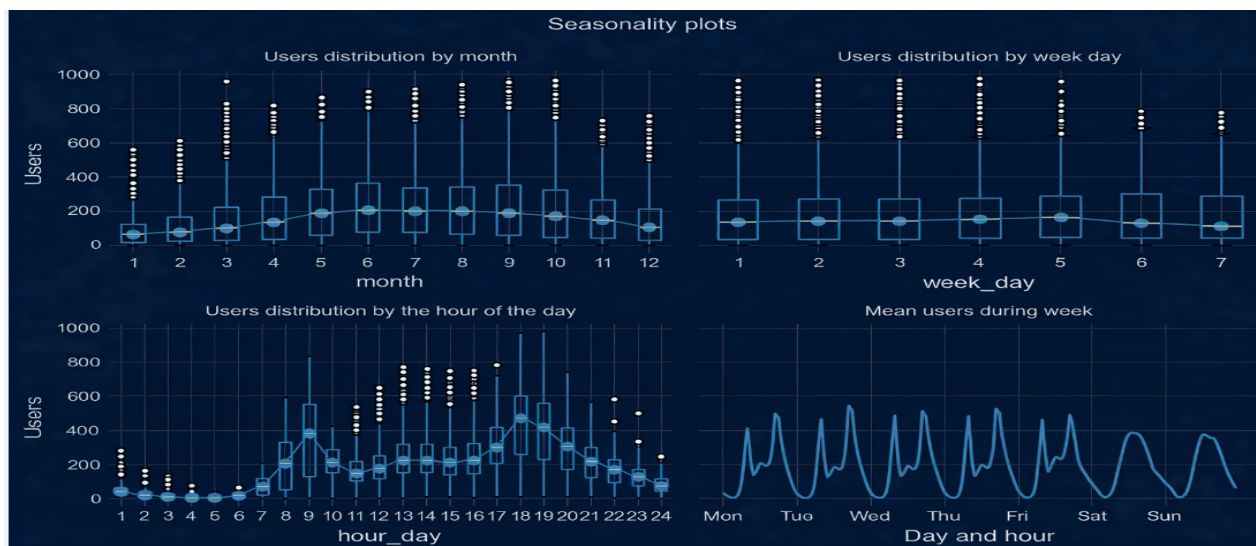


Рисунок 2. Пример сезонных графиков с данными для выявления сезонных закономерностей и трендов во временных рядах [7]

Решение третьей ключевой проблемы классического подхода в виде ограничений линейной регрессии достигается благодаря градиентному бустингу и ансамблям деревьев. В данной ситуации мы говорим об огромном количестве внешних переменных, которые в свою очередь влияют на внутренние данные.

Дерево решений позволяет учитывать множество параметров таких как: количество и периодичность использования приложения пользователем, каким устройством он пользовался, в какое время он это делал, какая была погода в момент пользования приложением, по какой цене он приобрел товар или услугу и ряд других параметров, которые также меняются в зависимости от сферы и отрасли в которой проводится анализ.

Градиентный бустинг же позволяет выявлять нелинейную зависимость между всеми этими параметрами, построив ансамбль деревьев из множества деревьев решений.

Этот метод построен на основе правил по признаку, где модель последовательно задает вопросы по этим признакам, например: если клиент заказывает такси, а сейчас идет дождь, то он готов заплатить за такси большую цену.

Такая цепочка рассуждений и является одним из вариантов дерева решений, разумеется таких вариантов может быть множество и они могут полностью меняться в зависимости от набора признаков и заданных параметров.

Деревья решений обучаются независимо, а результаты работы модели представляют из себя усредненные решения набора деревьев [7,8] (рисунок 3).

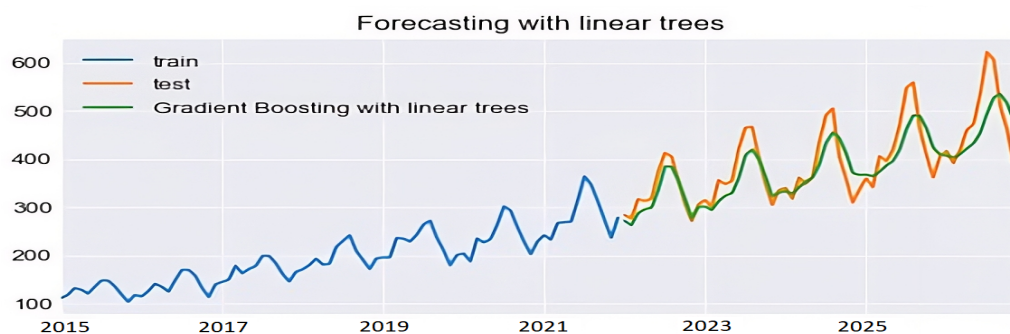


Рисунок 3. График, иллюстрирующий применение градиентного бустинга с линейными деревьями [7]

Заключение. Вышеописанные методы уже активно внедряются во множестве компаний и показывают положительные результаты от их применения [9].

Дочерняя компания Сбера «Сбер Бизнес Софт» разработала ML модель прогнозирования спроса на косметику для дистрибьютора «Градиент»; модель обучалась на данных о продажах, клиентах, складах, промо и остатках и дала +29% к качеству прогнозирования по сравнению с базовыми статистическими методами [9].

Wildberries обрабатывает более 1 млн заказов в день и строит Big Data архитектуру вокруг потоковой обработки событий, где продажи, остатки и поведение пользователей используются для онлайн аналитики и прогнозов спроса и логистики.

Во внешних сервисах для продавцов на Wildberries уже реализовано автоматическое прогнозирование поставок по временным рядам продаж и остатков по каждому товару и складу: система анализирует историю продаж, неснижаемый остаток и логистические задержки и выдает прогноз, какое количество товара и на какой именно склад везти, чтобы избежать дефицита товара на складе [10].

ВТБ протестировал платформу Smart Data Hub для аналитики цифрового маркетинга: система анализировала траектории пользователей на сайте банка по данным Яндекс.Метрики за 11 месяцев (временные ряды визитов, кликов, конверсий по страницам продуктов).

За 3 месяца работы пилотного проекта было создано более 40 гипотез, где более 40-50% из них были признаны практически значимыми, а потенциал роста конверсий оценен более чем в 5% [12].

Стриминговый сервис Netflix применяет модели временных рядов и последовательностей для прогнозирования нагрузки, персонализации и оценки удержания, комбинируя историю просмотров и поведение пользователей с контентными признаками [7].

В Uber использование градиентного бустинга и нейросетей для прогноза спроса на поездки снижало среднеквадратичную ошибку прогноза почти вдвое по сравнению с одиночными деревьями и классическими моделями, что напрямую влияет на динамическое ценообразование и балансировку предложения машин [13].

Сегодня формируется новая парадигма анализа динамических рядов, в которой классические статистические модели рассматриваются не как устаревший инструмент, а как важный базовый слой, дополняемый глобальными, нелинейными и высокоразмерными методами Big Data-аналитики.

Список литературы

[1] Стратегия развития государственной статистики до 2030 года Федеральная служба государственной статистики (Росстат), [Электронный ресурс]. – URL: https://rosstat.gov.ru/storage/mediabank/Strategiia_2030_bol.pdf (Дата обращения 28.02.2026г.)

[2] Капица С.П., Курдюмов С.П., Малинецкий Г.Г. (1997) Синергетика и прогнозы будущего, М.: Наука.1997 г.

[3] Кузовкова Т.А., Особенности и перспективы развития цифровых услуг и сервисов инфокоммуникационных компаний / Т.А. Кузовкова, И. М. Шаравов, Н. С. Курицын // журнал «Экономика и качество систем связи» № 3/2024 [Электронный ресурс]. – URL: <https://journal-ekss.ru/wp-content/uploads/2024/10/27-39-1.pdf>

[4] Кумратова А.М. Применение методов нелинейной динамики и машинного обучения для прогнозирования экономических волатильных процессов / А.М. Кумратова В.А. Плотников // Электронный журнал *π-Economy*. 2024, Том 17, № 3. С. 81–95. [Электронный ресурс]. – URL: https://economy.spbstu.ru/userfiles/files/articles/2024/3-2024/06_Kumratova,-Plotnikov.pdf 25 (Дата обращения 28.02.2026г.)

[5] Перепелица В.А., Попова Е.В. (2001) Многокритериальный подход к моделированию финансово-экономических рисков. Известия вузов. Северо-Кавказский регион. Естественные науки, 4, 37–41.

[6] Платунина Г.П. Монетизация криптосайтов как новый способ заработка в условиях цифровой экономики / Г.П. Платунина, А.С. Старовойтова [Электронный ресурс]. – URL: <https://cyberleninka.ru/article/n/monetizatsiya-kriptosaytov-kak-novyy-sposob-zarabotka-v-usloviyah-tsifrovoy-ekonomiki/viewer> (Дата обращения 26.11.2025)

[7] Прогнозирование товарооборота арендаторов торгово-развлекательного центра // Электронный журнал Cyberleninka [Электронный ресурс]. – URL: <https://cyberleninka.ru/article/n/prognozirovanie-tovarooborota-arendatorov-torgovo-razvlekatelnogo-tsentra> (Дата обращения 28.02.2026г.)

[8] Способы прогнозирования курса валют на основе моделей экспоненциального сглаживания и Хольта // Электронный журнал Cyberleninka [Электронный ресурс]. – URL: <https://cyberleninka.ru/article/n/sposoby-prognozirovaniya-kursa-valyut-na-osnove-modeley-eksponentsialnogo-sglazhivaniya-i-holta> (Дата обращения 28.02.2026г.)

[9] ИИ-модель Сбербанка прогнозирует спрос на косметику на треть лучше статистических моделей // Электронный журнал CNews от 12.09.2024г [Электронный ресурс]. – URL: https://axenix.cnews.ru/news/line/2024-09-12_ii-model_sberbanka_prognoziruet (Дата обращения 28.02.2026г.)

[10] Ильясов Р.Х., Куразова Д.А. (2017) Прогнозирование конъюнктуры финансового рынка и оценка его влияния на развитие промышленности. Финансы и кредит, 23 (43), 2575–2591. DOI: <https://doi.org/10.24891/fc.23.43.2575>

[11] Самакаева, М. Д. Фриланс платформы как основа развития рынка труда в цифровой экономике / М. Д. Самакаева, Е. Д. Критина, Р. Ф. Мартынова // Первый экономический журнал. – 2024. – № 12(354). – С. 69-78. – DOI 10.58551/20728115_2024_12_69. – EDN GPOAXT. (Дата обращения 28.02.2026г.)

[12] Си Чен Оценка влияния облачных вычислений на показатели деятельности компаний: эмпирическое исследование компаний, акции которых котируются на бирже / Си Чен, Мин Го, Уюэ Шангуань // Электронный журнал Journal of Econometrics (S0378720622000155) [Электронный ресурс]. – URL: <https://www.sciencedirect.com/science/article/pii/S0378720622000155> (Дата обращения 28.02.2026г.)

[13] Синьйю Ху ДеерЕТА: Как Uber прогнозирует время прибытия с помощью глубокого обучения / Синьйю Ху, Олчай Чирит, Танмай Бинайкия, Рамит Хора // Блог UBER от 10.02.2022г. [Электронный ресурс]. – URL: <https://www.uber.com/blog/deerpeta-how-uber-predicts-arrival-times/> (дата обращения: 28.02.2026).

ALIGNMENT OF DYNAMIC SERIES IN BIG DATA ANALYTICS IN THE DIGITAL ECONOMY

E. E. Dadashev

*Bachelor's degree in Economics,
Profile «Economics and organization of IT
business», BEE2401, MTUCI
ericandeev2016@yandex.ru*

M. D. Samakaeva

*Associate Professor of the Department of Digital
Economics, Management and Business Technology,
MTUCI
Candidate of Economic Sciences, Associate Professor
samakaeva-m@yandex.ru*

Abstract. In the context of the accelerating digital transformation of the economy, key business processes, the interaction of market participants and consumer behavior are increasingly described by large amounts of data generated in close to real time. Indicators of e-commerce, the use of cloud services, digital financial services, telecom infrastructure, and online user activity are recorded as long time series that become the basis for analytics and management decision support. This article discusses models and methods of dynamic series analysis and alignment in the analysis of big data in economics. The use of dynamic series in data analysis allows for more accurate forecasts and development strategies with more predictable results.

Keywords: digital transformation, dynamic series, Big Data, models, dynamic series, statistical alignment, methods, models, management, gradient boosting, automatic forecasting, platforms, artificial intelligence.