

Мосин Владимир Геннадьевич

**МЕТОДОЛОГИЯ И ИНСТРУМЕНТАРИЙ ДЕТЕКЦИИ АНОМАЛИЙ
В УПРАВЛЕНИИ КАЧЕСТВОМ ПРОЦЕССОВ АВТОМОБИЛЬНОЙ ОТРАСЛИ**

2.5.22. Управление качеством продукции. Стандартизация.
Организация производства

Автореферат
диссертации на соискание учёной степени
доктора технических наук

Работа выполнена в ФГБОУ ВО «Самарский государственный технический университет» на кафедре «Теоретическая и общая электротехника».

Научный консультант: **Козловский Владимир Николаевич**, доктор технических наук, профессор, федеральное государственное бюджетное образовательное учреждение высшего образования «Самарский государственный технический университет», кафедра «Теоретическая и общая электротехника», заведующий кафедрой.

Официальные оппоненты: **Димитров Валерий Петрович**, доктор технических наук, профессор, ФГБОУ ВО «Донской государственный технический университет», кафедра «Управление качеством», заведующий кафедрой;

Киселев Эдуард Валентинович, доктор технических наук, профессор, ФГБОУ ВО «Рыбинский государственный авиационный технический университет имени П.А. Соловьева», кафедра «Организация производства и управление качеством», заведующий кафедрой;

Полякова Марина Андреевна, доктор технических наук, доцент, ФГБОУ ВО «Магнитогорский государственный технический университет им. Г.И. Носова», кафедра обработки материалов давлением им. М.И. Бояршинова, профессор кафедры.

Ведущая организация: ФГБОУ ВО «**Иркутский национальный исследовательский технический университет**», г. Иркутск.

Защита состоится «25» марта 2026 года в 10.00 на заседании диссертационного совета 24.2.379.05, созданного на базе федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева», по адресу: 443086, г. Самара, ул. Московское шоссе, д. 34.

С диссертацией можно ознакомиться в библиотеке федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева» и на сайте: https://www.ssau.ru/resources/dis_protection/mosin.

Автореферат разослан «__» _____

Учёный секретарь
диссертационного совета 24.2.379.05
доктор технических наук, доцент

Я.А. Ерисов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. В деятельности любой компании качество корпоративной системы управления является приоритетным, поскольку напрямую влияет на успешность стратегического развития. Ключевым инструментом для улучшения деятельности является система менеджмента качества (СМК), основанная на международных стандартах ISO 9000, при этом в автомобильной промышленности определяющим стандартом СМК выступает IATF 16949, а национальный стандарт ГОСТ Р ИСО 9001-2015 адаптирует международные требования в области качества к отечественным условиям. Соблюдение положений стандартов в области качества является обязательным требованием, учитывая их важность для признания организации и её возможности работать как на внутреннем, так и на международном рынках.

Управление качеством особенно важно в отраслях промышленности с высокой конкуренцией, где производятся сложные технические изделия, например, современном машиностроении, в частности — автомобилестроении. Современный автопроизводитель рассматривает свой продукт не только как совокупность технических характеристик, но и как сложную систему, включающую в себя весь жизненный цикл автомобиля, от маркетинговых исследований до организации сервисного обслуживания и формирования потребительской лояльности к бренду. Автомобилестроение характеризуется разветвлённой цепочкой поставок компонентов и предприятий, выполняющих кооперационные технологические процессы, при этом стремительная цифровизация процессов управления качеством предполагает освоение новых компетенций в области применения цифровых методов планирования, организации, мониторинга и управления качеством продукции. В контексте маркетинговых инициатив, сервисного обслуживания, а также в условиях аддитивного производства, процессы, могут быть лишены стабильности, и при этом именно сегодня, в условиях тотальной цифровизации автопроизводители располагают практически неограниченными массивами данных, отражающих динамику этих процессов: количество заявок на гарантийное обслуживание в различных сервисных центрах; статистика просмотров контента на онлайн-платформах бренда; данные с активных датчиков контроля технологических производственных линий. Однако применение традиционных статистических методов, таких как контрольные карты Шухарта, разработанных для анализа стабильных процессов, оказывается либо принципиально невозможным, либо приводит к получению результатов, характеризующихся неприемлемой степенью погрешности, что исключает их практическую ценность.

В машиностроении накапливается колоссальный объем данных, генерируемых на различных этапах жизненного цикла продукции. Датчики, станки с числовым программным управлением, автоматизированные комплексы с системой предиктивного управления, а также производственные системы диагностирования непрерывно поставляют информацию о параметрах технологических процессов, состоянии узлов и агрегатов, качестве выпускаемой продукции. Однако, как это ни парадоксально, основная проблема заключается не в дефиците данных, а в неспособности эффективно их анализировать. Существующие методы мониторинга и аналитики, зачастую, не справляются с обработкой таких массивов разнородной и сложной информации. Данные могут быть зашумлены, неполны, противоречивы и представлены в различных форматах. Традиционные статистические подходы к управлению качеством оказываются недостаточно мощными для выявления скрытых закономерностей и зависимостей. Кроме того, данные отражающие ход машиностроительного производства характеризуются высокой многофакторностью и нелинейностью. Сложно выявить взаимосвязи между множеством параметров, влияющих на качество и надёжность продукции.

Таким образом, можно сделать существенный вывод о наличии в современной науке и практике управления качеством системного пробела и противоречия, которые образуют отраслевую проблему.

Пробел заключается в том, что широкоизвестные традиционные инструменты контроля и мониторинга показателей качества в условиях взрывного роста объемов электронных данных отражающих характеристики протекания процессов системы менеджмента машиностроительного (автомобильного) производства не всегда в состоянии полноценно и адекватно дать оценку функционированию, результативности и эффективности деятельности.

Противоречие заключается в том, что в актуальных стандартах менеджмента качества, таких как IATF 16949:2016, на уровне требований декларируется обязательность применения инженерных методик управления качеством, таких, например, как SPC, MSA и т.д., в то время как фундамент указанных методов выстроен, исходя из статистических инструментов, подразумевающих высокий уровень стабильности процессов.

Для устранения пробела и преодоления противоречия необходимы новые подходы к мониторингу данных, сочетающие в себе статистические методы, машинное обучение и экспертные знания специалистов. Только тогда можно будет превратить поток данных в процессах системы менеджмента качества автосборочного предприятия в ценный ресурс для повышения конкурентоспособности. Решение выделенной отраслевой научно-технической проблемы является ключевым фактором необходимым для дальнейшего развития отрасли и внедрения концепции «Индустрии 4.0». Сегодня недостаёт 1) целостной концепции, 2) методологии и 3) инструментария, позволяющих эффективно использовать имеющиеся информационные ресурсы управления качеством. Настоящая диссертационная работа восполняет этот пробел, предлагая концептуальные, методологические и алгоритмические основы для управления качеством процессов в машиностроении (автомобилестроении), с учётом гарантированности управленческих решений вне зависимости от уровня стабильности процессов. Таким образом, в диссертации предлагается переход от базового понятия стабильности процесса к понятию его предсказуемости в управлении качеством. Успешное применение аналитики данных позволяет перейти от реактивного управления к проактивному в системе менеджмента, предсказывая и предотвращая проблемы до их возникновения. Это, в свою очередь, приведёт к сокращению издержек, повышению эффективности производства и улучшению качества выпускаемой продукции.

Актуальность темы диссертационного исследования обусловлена необходимостью повышения результативности и эффективности управления качеством в условиях современной машиностроительной отрасли в общем и автомобилестроении в частности за счёт системного развития комплексных инструментов детекции аномалий в управлении качеством, то есть, специализированной системы статистического управления, предполагающей использование алгоритмов машинного обучения (таких как алгоритмы классификации и анализа временных рядов), которые обучаются на исторических данных, представляющих предсказуемое поведение как стабильных, так и нестабильных процессов, разрабатываются с учётом специфики автомобильного производства и способны эффективно распознавать отклонения от прогнозируемого поведения процессов системы менеджмента даже в условиях высокой вариативности данных. Анализ больших данных, поступающих из систем мониторинга качества, позволяет выявлять скрытые зависимости и нелинейные корреляции, зачастую недоступные для традиционных методов статистического анализа, при этом алгоритмы машинного обучения способны эффективно обнаруживать аномальные повторяющиеся участки (паттерны) в этих данных, что способствует оперативному вмешательству и предотвращению развития аномальных состояний. С практической точки зрения, актуальность темы диссертационного исследования подтверждается тем, что применение детекции аномалий в управлении качеством в автомобилестроении повышает оперативность выявления отклонений (так как автоматизированный анализ данных обеспечивает быстрое обнаружение аномалий (несоответствий), минимизируя время реакции на проблемы) и снижает затраты на контроль качества (так как автоматизация процессов детекции аномальных состояний снижает потребность в ручном мониторинге и инспекциях, сокращая затраты на рабочую силу и оборудование).

Степень разработанности темы. Различные подходы к вопросам управления качеством, в частности — концепции, в той или иной степени апеллирующие к статистическим методам широко представлены в исследованиях как зарубежных, так и отечественных авторов.

Основополагающие научные работы таких авторов, как Э. Деминг, Дж. Джуран, П. Друкер, К. Исикава, Н. Кано, Р. Каплан, Ф. Котлер, Ф. Кросби, Г. Тагути, Н. Талеб, Ф. Тейлор, А. Фейгенбаум и В. Шухарт, оказали значительное влияние на формирование теоретической базы данного диссертационного исследования.

Научно-практическая направленность работы определена на основе исследований ведущих отечественных ученых, среди которых Ю. П. Адлер, В. Н. Азаров, Г. Г. Азгальдов, И. З. Аронов, В. А. Барвинок, В. Я. Белобрагин, Б. В. Бойцов, В. В. Бойцов, В. А. Васильев, В. Г. Версан, Г. П. Воронин, А. В. Гличев, В. А. Лapidус, В. В. Окрепилов, И. И. Чайка и другие.

Важнейшие научно-прикладные аспекты исследования рассматриваются в работах Д. В. Антипова, В. Ф. Безъязычного, С. А. Васина, В. Е. Годлевского, О. А. Горленко, С. Я. Гродзенского, В. П. Димитрова, А. Я. Дмитриева, В. В. Ефимова, А. В. Зажигалкина, А. Г. Ивахненко, В. А. Качалова, В. Я. Кершенбаума, Ю. С. Ключкова, В. Н. Клячкина, В. Н. Козловского, П. А. Лончиха, С. В. Мищенко, С. Н. Николаева, И. Н. Омельченко, К. Г. Пивоварова, Е. В. Плахотниковой, М. А. Поляковой, С. В. Пономарева, В. Б. Протасьева, С. В. Пугачева, М. И. Розно, Т. А. Салимовой, Е. Г. Семеновой, Л. Е. Скрипко, А. Г. Суслова, Х. А. Фасхиева, А. И. Хаймовича, И. Н. Хаймович, Ю. К. Чернова, А. Д. Шадрина, А. П. Шалаева, В. Л. Шпера, В. В. Щипанова, Г. Л. Юнака и других российских исследователей.

Цель диссертационной работы заключается в повышении результативности и эффективности системы менеджмента качества автопроизводителя в условиях отраслевого развития информатизации и цифровизации посредством разработки нового методологического инструментария статистического управления качеством при помощи детекции аномалий в процессах автомобильной отрасли.

Научно-прикладные задачи. Для достижения поставленной цели в настоящей диссертационной работе решаются следующие научно-прикладные задачи:

1. Научно-технический анализ проблемы, связанной с выявлением фактов оценки стабильности процессов в системе менеджмента качества автопроизводителя и их контролем, а также выделение группы процессов, обладающих свойствами нестабильности.
2. Формирование и реализация научной концепции методологии детекции аномалий в управлении качеством процессов, дающей универсальное представление о процессах и подходах к определению нормы и аномалии на основании измеримых характеристик в автомобильной отрасли.
3. Разработка методологии детекции аномалий процессов системы менеджмента качества автопроизводителя, позволяющей давать оценку «норма-аномалия» любому его состоянию на основании наблюдаемой характеристики без апелляции к экспертной оценке.
4. Разработка и реализация научно-программного комплекса моделей и алгоритмов детекции аномалий в процессах автомобильной отрасли, входящих в систему менеджмента качества.
5. Апробация и практическая реализация полученного инструментария детекции аномалий в процессах системы менеджмента качества автопроизводителя.

Область исследования. Область исследования соответствует направлениям исследований паспорта научной специальности 2.5.22. Управление качеством продукции. Стандартизация. Организация производства в части пунктов:

4. Инновации при разработке, развитии, цифровизации систем менеджмента качества (СМК) предприятий и организаций;

8. Разработка научно-практического статистического инструментария управления качеством;

9. Разработка и совершенствование научных инструментов оценки, мониторинга и прогнозирования качества продукции и процессов;

14. Развитие основных положений и содержания Всеобщего Управления Качеством (TQM), и других концепций управления качеством;

25. Разработка моделей описания, методов и алгоритмов решения задач проектирования производственных систем, организации производства и принятия управленческих решений в цифровой экономике.

Объект исследования. Комплексные системы статистического мониторинга качества процессов менеджмента предприятий автомобильной отрасли, в условиях информатизации и цифровизации.

Предмет исследования. Статистические методы организации мониторинга качества процессов менеджмента предприятий автомобильной отрасли.

Научная новизна работы. Заключается в разработке системных научно-практических инструментов и комплексов моделей, составляющих методологию детекции аномалий в управлении качеством автопроизводителя, направленных на повышение результативности и эффективности системы менеджмента качества, в условиях отраслевого развития информатизации и цифровизации и включает следующее:

1. Разработана концепция методологии детекции аномалий в управлении качеством процессов менеджмента автопроизводителя, которая в отличие от заложенных норм и требований в стандарте IATF 16949:2016 по применению инструментов SPC определяет новый подход к анализу данных, основанный на предиктивном описании процессов системы менеджмента качества. Пункт 14 научной специальности 2.5.22.
2. Разработан комплекс научно-программных инструментов по обработке данных о процессах СМК автомобильной отрасли, отличающийся возможностью генерировать синтетические данные о процессах и их экспертных оценках, направленный на решение задачи тестирования инструментов детекции аномалий в условиях различных искусственно смоделированных экспертиз. Пункт 4 научной специальности 2.5.22.
3. Разработан комплекс методов и инструментов выявления аномалий в процессах, который позволяет детектировать аномальные состояния в характеристиках измеримых процессов системы менеджмента качества автомобильной отрасли. Пункт 8 научной специальности 2.5.22. Комплекс включает в себя:
 - 3.1. Метод моделирования временных рядов, отличающийся возможностью одновременного исследования семейства процессов, протекающих в совокупности локаций (метод направлен на детекцию аномалий процесса по отношению к его состояниям в предыдущие моменты времени).
 - 3.2. Метод моделирования состояния семейства процессов, протекающих в совокупности локаций, в фиксированные моменты времени (метод направлен на детекцию аномалий процесса по отношению к его состояниям в других локациях).
 - 3.3. Статистический критерий оценки выраженности аномалий в характеристиках измеримых процессов, основанный на определении локтевой точки, который, во-первых, применим к любым законам статистических распределений, в отличие от критерия 6 сигм, применимого только к нормальным распределениям, а во-вторых, не апеллирует к экспертной оценке.
4. Разработан комплекс научно-программных инструментов, позволяющий реализовать детекцию аномалий в управлении качеством процессов системы менеджмента предприятий автомобильной отрасли, отличающийся универсальностью применения к

анализу процессов, имеющих как стабильные, так и нестабильные характеристики протекания. Пункт 9 научной специальности 2.5.22.

5. Предложен метод принятия решений построенный на основе рационального подхода, отличающийся от инструментария управления, реализованного на основе выборочного контроля, полным анализом всей совокупности статистических данных о процессах и направленный на исключение человеческого фактора. Пункт 25 научной специальности 2.5.22.

Теоретическая значимость. Теоретическая значимость диссертационного исследования заключается в создании методологии детекции аномалий в управлении качеством процессов предприятий автомобильной отрасли, направленной на повышение результативности и эффективности системы менеджмента качества и обеспечивающей системное развитие положений теории управления качеством, а также установлением и углублением связей между базовыми положениями, образующими принципы системы менеджмента качества — процессным подходом и принятием решений основанных на свидетельствах. Предложенный методологический инструментарий расширяет арсенал статистических инструментов мониторинга качества процессов системы менеджмента качества автомобильной отрасли в условиях развития информатизации и цифровизации.

Разработанная методология обладает высокой степенью универсальности. В рамках автомобильной отрасли она позволяет изучать с точки зрения соответствия или аномалий норме самые разнообразные процессы: маркетинг, проектирование, производство, сервисное техническое обслуживание, что обеспечивает широкий охват характеристик, обеспечивающих качество автомобильной продукции. Вместе с тем, область применения комплекса разработанных методов не ограничивается автомобилестроением и может быть расширена на другие отрасли машиностроения.

Существенную значимость работы составляет аспект, связанный с развитием отраслевого стандарта IATF 16949:2016 и его российского аналога ГОСТ Р 58139-2018 в части обязательного применения инженерных инструментов статистического управления процессами системы менеджмента (SPC), за счёт включения предложенного в работе инструментария, образующего методологию детекции аномалий в процессы СМК предприятий автомобильной отрасли.

Практическая значимость. Практическая значимость результатов диссертации состоит в разработке комплекса научно-прикладных решений, повышающих результативность и эффективность мониторинга процессов системы менеджмента качества в автомобильной отрасли.

Полученные в диссертационной работе результаты вошли в устойчивую отраслевую практику крупных автомобилестроительных предприятий. В частности, внедрены инструменты автоматизированного поиска и устранения нарушений, связанных с необоснованным ремонтом автомобильной техники на предприятиях фирменной сервисной сети автопроизводителя, а также инструментарий поиска позиций дефектов автомобилей в эксплуатации, имеющих наибольший прирост за отчётный анализируемый период времени. Также в работе предложены инструменты анализа стабильности ключевых параметров качества автомобилей, для серийной и перспективной продукции автопроизводителя. В практику процесса испытаний при проектировании автомобильной техники внедрён инструмент детекции аномалий, позволяющий выявлять несоответствия продукции по ключевым параметрам качества. Предложенный в работе метод принятия решения о качестве автомобилей по ключевым параметрам обеспечивает рационализацию выводов по результатам статистического контроля при проектировании, производстве и эксплуатации продукции.

Комплексные научно-технические решения внедрены в организациях: ПАО «КАМАЗ», г. Набережные Челны; АО «АВТОВАЗ», г. Тольятти.

Предложенные научно-прикладные инструменты доказали свою результативность и эффективность при проведении мониторинга качества автомобилей по ключевым параметрам. При применении инструментов мониторинга достигнут уровень снижения трудоёмкости анализа данных о качестве продукции на 40% по сравнению с ранее применяемыми методами и подходами. Применение разработанных технических решений существенно расширяет базу инструментов управления качеством в рамках действующей системы менеджмента автопроизводителя.

Методология и методы исследования. Для решения поставленных в работе задач использовались принципы всеобщего управления качеством (TQM), принципы Деминга и Шухарта, методы теории систем и системного анализа, статистические методы управления качеством, методы стохастического прогнозирования. Базы данных о процессах проектирования, производства, сервиса и маркетингового сопровождения автомобильного бренда реализованы в среде SQL. Программный инструментарий для решения задачи алгоритмической детекции аномалий реализован в среде Jupyter Notebook на языке Python при помощи библиотек numpy, pandas, sklearn. Визуализации выполнены с использованием библиотек matplotlib и seaborn.

На защиту выносятся:

1. Концепция методологии детекции аномалий в управлении качеством процессов менеджмента предприятий автомобильной отрасли.
2. Комплекс научно-программных инструментов и алгоритмов генерации синтетических данных для тестирования инструментов, составляющих содержательную часть методологии детекции аномалий в управлении качеством процессов менеджмента.
3. Методы моделирования временных рядов, реализующие задачи детекции аномалий процессов менеджмента качества по отношению к их состояниям в предыдущие моменты времени и по отношению к их состояниям в других локациях.
4. Критерий определения аномальных состояний в процессах, входящих в СМК предприятий автомобильной отрасли, позволяющий детектировать аномалии в условиях отсутствия экспертной оценки, и исключающий, тем самым, человеческий фактор при принятии управленческих решений.
5. Комплекс научно-программных инструментов для детекции аномальных состояний в процессах, входящих в систему менеджмента предприятий автомобильной отрасли.
6. Метод принятия управленческих решений на основе полного анализа всей совокупности статистических данных о процессах, входящих в СМК предприятий автомобильной отрасли, и алгоритмически выявленных в них аномалиях.
7. Результаты комплексной апробации методологии детекции аномалий в процессах, входящих в СМК автомобильной отрасли.

Личный вклад автора. Все результаты диссертационной работы получены автором самостоятельно. Под научным руководством автора (или при его преобладающем участии) разработана методология и инструментарий детекции аномалий в процессах автомобильного производства, сервисного обслуживания и маркетингового сопровождения автомобильного бренда. Все направления исследований диссертационной работы, формулировки проблем и постановки задач обсуждались с научным консультантом — д. т. н., профессором В. Н. Козловским, что отражено в совместных публикациях, в которых основные результаты принадлежат диссертанту.

Связь работы с научными программами, темами, грантами. Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (тема №FSSE-2023-0003) в рамках государственного задания Самарского государственного технического университета.

Степень достоверности и апробация результатов. Основные положения настоящей диссертационной работы в виде научных работ и тезисов докладывались на международных и всероссийских конференциях и семинарах. Результаты диссертационного исследования нашли своё отражение в сборниках научных трудов и были доложены автором на международных и всероссийских научно-практических конференциях, в том числе: Московский авиационно-космический салон МАКС-2021. II Международная конференция. Москва, 2021 г.; VIII Международная научно-практическая конференция «Инновации в информационных технологиях, машиностроении и автотранспорте». Кемерово. 2024 г.; IV всероссийская научно-практическая конференция «Энергетика, управление и автоматизация: инновационные решения проблем», Санкт-Петербург, 2024 г.; X Международная научно-практическая конференция «Актуальные вопросы экономики и управления: Теоретические и прикладные аспекты», Донецк, 2025 г.

Публикации. Основные результаты диссертации представлены в 50 научных трудах, из них — 1 монография; 33 статьи, опубликованные в рецензируемых периодических изданиях, рекомендованных ВАК, 5 статей в иностранных научных изданиях, 11 статей, опубликованных в других научных изданиях.

Структура и объем работы. Диссертация состоит из введения, семи глав, заключения, списка литературы и пяти приложений. Общий объем диссертации составляет 491 страницу, включая 122 рисунка, 76 таблиц, список литературы содержит 249 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** показана актуальность диссертационного исследования, сформулированы цель и задачи, дана общая характеристика работы, проведён выбор объекта и предмета исследования, указаны научная новизна и практическая значимость результатов работы, определены выносимые на защиту научные положения и результаты.

В **первой главе** решаются две взаимосвязанные научно-прикладные задачи. Во-первых, проводится научно-технический анализ проблемы стабильности/ нестабильности процессов в СМК автомобильной отрасли (см. п. 1 в списке научно-прикладных задач), а во-вторых, реализуется научная концепция методологии детекции аномалий в процессах, специфичных для автомобильного производства (см. п. 2 в списке научно-прикладных задач).

В начале главы рассматриваются известные концепции управления качеством, такие как:

1. концепция Деминга,
2. концепция Джурана,
3. концепция 6σ,
4. концепция Фейгенбаума,
5. концепция Тачучи.

Отмечается, что все классические концепции опирались на методы статистического анализа (либо вообще ставили его во главу угла, как это было, например, в концепции Деминга или, в ещё более выраженной форме, в концепции 6σ), и, в той или иной формулировке, в них всегда присутствуют обязательные к исполнению положения:

1. должен производиться непрерывный мониторинг показателей качества;
2. результаты мониторинга должны подвергаться статистическому анализу;
3. управляющие решения должны приниматься на основании этого анализа (см. рисунок 1).

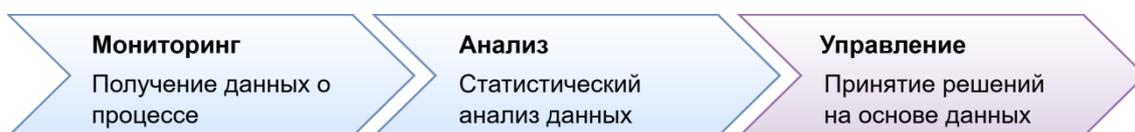


Рисунок 1 — Элементы статистического управления качеством

Эти положения являются неотъемлемой частью любой теории управления качеством, базирующейся на статистическом анализе данных о процессах. Однако настоящая диссертационная работа обладает определённой спецификой. Во-первых (и об этом подробно говорится во второй главе и далее по тексту диссертации), в ней используются принципиально новые методы статистического анализа, основанные на машинном обучении с применением сложных прогнозирующих моделей, то есть алгоритмов, способных обучаться на основе имеющихся исторических данных. А во-вторых, сами процессы, которые подвергаются статистическому анализу, существенно расширяются, захватывая классы процессов, ранее не рассматривавшиеся системой управления качеством автомобильного производства как процессы, к которым применимы статистические методы управления. Для демонстрации этой специфики в первой главе проводится типология процессов автомобильного производства, включая собственно производственные, сервисные и маркетинговые процессы.

А именно, все процессы сначала делятся на измеримые (то есть, процессы, которые можно описать какими-либо числовыми характеристиками) и неизмеримые (когда такое описание невозможно). Например, эволюция корпоративной культуры на производстве является неизмеримым процессом (она оценивается косвенно по нечисловым показателям), а производительность труда — измеримым (она измеряется в единицах продукции на одного сотрудника в единицу времени). Восприятие автомобильного бренда — это неизмеримый процесс (так как оно отражается в мнениях и ассоциациях, которые не описываются наборами чисел), а трафик на сайте автопроизводителя — измеримый (так как оценивается количеством посетителей). Перечень таких примеров может быть неограниченно продолжен. Затем все измеримые процессы делятся на:

1. стабильные процессы — то есть, измеримые процессы, удовлетворяющие условию стационарности (когда математическое ожидание, дисперсия и функция автокорреляции являются константами относительно времени);
2. нестабильные процессы — то есть, процессы, для которых условия стационарности нарушены (см. рисунок 2).

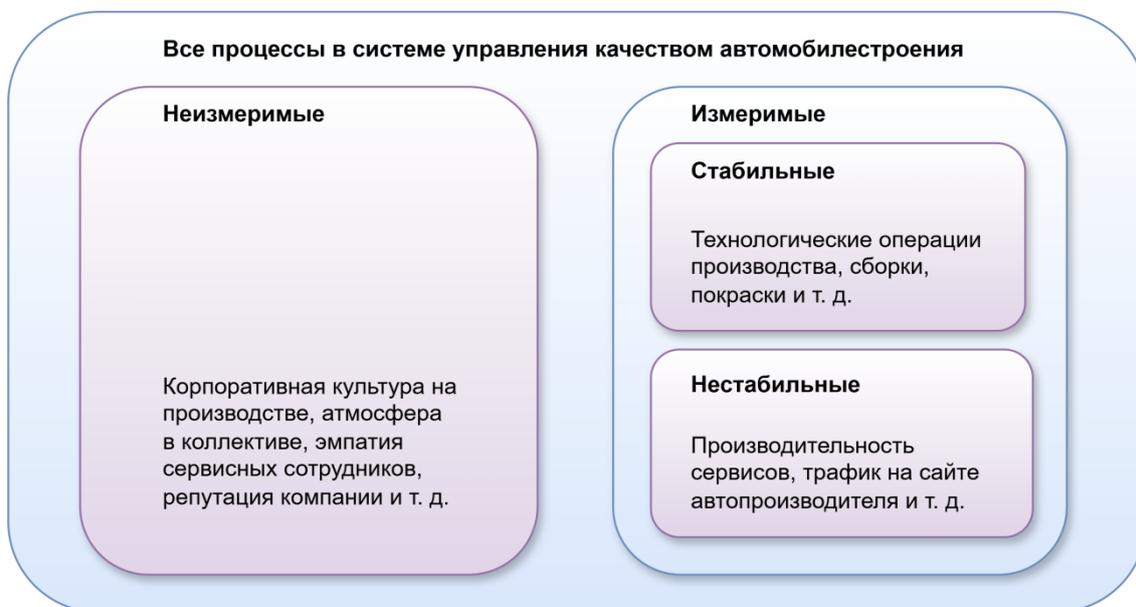


Рисунок 2 — Типология процессов в системе управления качеством автомобильной отрасли

Например, любая технологическая операция (обработка поверхностей деталей вращения, нарезание резьбы, обработка плоских поверхностей и пазов, сверление отверстий и т. д.), по определению, должна быть стабильной, а нарушение стабильности согласно традиционным методикам статистического управления качеством является маркером аномального состояния процесса, которое требует немедленного вмешательства и устранения. Вместе с

тем, производительность центра сервисного обслуживания автомобилей даёт пример процесса, нестабильного по своей природе, так как она зависит от множества факторов, неподвластных производителю (сезонность, маркетинговые акции конкурентов, опоздания и неявки клиентов, погодные условия, дорожная обстановка и т. д.), и, тем не менее, несмотря на свою существенную нестабильность, этот процесс точно так же как и стабильные по своей природе процессы требует управления и оперативной реакции на возникающие аномалии. Именно нестабильные процессы являются главными действующими лицами настоящей диссертационной работы, и именно возможность анализировать нестабильные процессы отличает методологию и инструментарий заявленной в диссертации парадигмы детекции аномалий от традиционных статистических парадигм управления качеством.

Так как в настоящей диссертационной работе ключевую роль играет стохастическое прогнозирование и моделирование процессов на основе большого объёма предварительно накопленных знаний, в первой главе проводится чёткое разделение двух подходов к прогнозированию, которое впоследствии используется на протяжении всего диссертационного исследования:

1. прогнозирование процесса по времени, когда в качестве накопленных знаний о процессе используются значения наблюдаемой характеристики, зафиксированные в предыдущие моменты времени;
2. прогнозирование процесса по локациям, когда в качестве накопленных знаний о процессе используются значения наблюдаемой характеристики, зафиксированные в других локациях в тот же момент времени.

Для прогнозирования (как по времени, так и по локациям) в рамках настоящей диссертации используются модели, обучающиеся на выборке типичных данных о процессе и предсказывающих его будущее поведение (в новый момент времени или в новой локации), причём, в работе применяются как простые модели (такие, как, например, линейная регрессия), так и более сложные (случайный лес, нейронные сети или градиентный бустинг), о чем подробно говорится в последующих главах. Но уже в первой главе заявленный модельный подход позволяет выдвинуть одно из главных концептуальных положений диссертации, а именно — предложение рассматривать стохастическую прогнозирующую модель нестабильного процесса как альтернативную версию константного стандарта. Дело в том, что константный стандарт подразумевает наличие некоего фиксированного и заранее известного значения, отклонение от которого сверх установленных границ интерпретируется наблюдателем как аномалия и требует вмешательства в процесс с целью установления её причин и последующего устранения (см. рисунок 3(a), где по горизонтали откладываются моменты наблюдений, а по вертикали — зафиксированные значения наблюдаемой характеристики процесса). Однако применение константного подхода к нестабильным процессам может привести к абсурдным выводам (см. рисунок 3(b)) и, как следствие, к неверным управленческим решениям. Если же в качестве стандарта рассматривать не константу, а стохастическую модель нестабильного процесса, и если при этом для детекции аномалий использовать отклонение не от стандартной константы, а от модели (см. рисунок 3(c)), то эта проблема снимается.

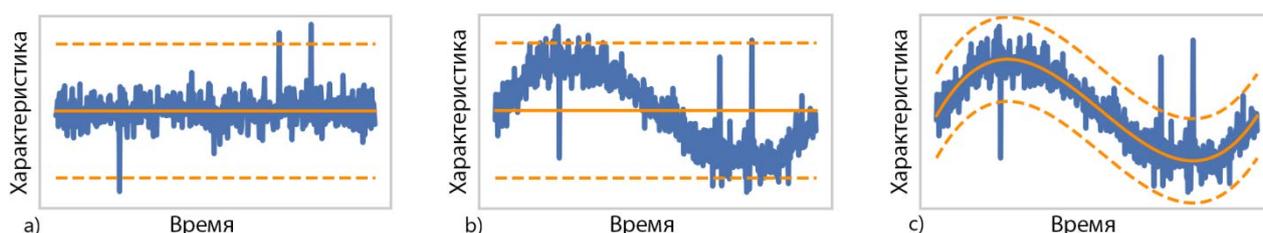


Рисунок 3 — а) стабильный процесс с отклонением от среднего, б) нестабильный процесс с отклонением от среднего, в) нестабильный процесс с отклонением от модели

Ещё одним концептуальным аспектом настоящей диссертации является формат данных, пригодных для реализации её методов и алгоритмов: это формат ML (от англ. — Moment, Location), который является основой для моделей машинного обучения, позволяющих прогнозировать состояние процесса по времени (см. таблицу 1, где символ c_{ij} означает значение измеряемой характеристики процесса в i -й момент времени в j -й локации).

Таблица 1. Процесс в формате ML (Moment, Location)

	Локация 1	Локация 2	Локация 3	...	Локация m
Момент 1	c_{11}	c_{12}	c_{13}	...	c_{1m}
Момент 2	c_{21}	c_{22}	c_{23}	...	c_{2m}
Момент 3	c_{31}	c_{32}	c_{33}	...	c_{3m}
...
Момент n	c_{n1}	c_{n2}	c_{n3}	...	c_{nm}

С точки зрения данной выше типологии, возможность описания процесса в формате ML означает, что процесс является измеримым (так как он представлен в виде числовых значений), причём, предполагается, что аналитик системы управления качеством в любой момент времени располагает сведениями о процессе, который протекает одновременно сразу в нескольких локациях (например, в нескольких производственных подразделениях предприятия автомобильной отрасли, в нескольких центрах по сервисному обслуживанию автомобилей и т. д.). Для прогнозирования процесса по локациям используются данные в формате LM (Location, Moment) — они получаются из данных в формате ML при помощи обычной операции транспонирования, и, по сути, если у корпоративной службы качества есть данные в одном из форматов, то они есть также и в другом. Если же процесс описывается не одной, а несколькими характеристиками (а это весьма распространённая на практике ситуация, так как сложный процесс нельзя описать одной числовой характеристикой), то обобщением формата ML на набор из нескольких характеристик служит объединение нескольких таблиц формата ML (или LM), каждая из которых описывает поведение какой-то одной его характеристики.

Итак, на концептуальном уровне в настоящей диссертационной работе заявляется существенное расширение пространства статистического контроля и управления качеством в автомобильной отрасли, захватывающее, помимо изучавшихся ранее стабильных процессов, процессы, являющиеся существенно нестабильными, но, тем не менее, предсказуемыми. При этом заменой традиционного статистического инструментария, основанного, в первую очередь на контрольных картах Шухарта, служит разработанная в рамках диссертации методика алгоритмической детекции аномалий, работающая на кардинально более сложных математических принципах, но, тем не менее, вполне реализуемая в условиях современных вычислительных возможностей и алгоритмики (см. рисунок 4).



Рисунок 4 — Обобщающие свойства алгоритмической детекции аномалий

В заключение первой главы отмечается, что методология и инструментарий детекции аномалий не отрицает традиционные подходы к статистическим методам управления качеством в автомобильной отрасли, а обобщает их (см. рисунок 4), так как стабильность

является частным случаем предсказуемости (если процесс стабилен, то он предсказуем, обратное неверно), а константа — это частный случай модели (например, можно считать, что это полиномиальная регрессионная модель нулевой степени).

Целью **второй главы** является решение основной научно-прикладной задачи, заявленной в рамках настоящей диссертации, а именно — разработка методологии детекции аномалий в процессах автомобильной отрасли (см. п. 3 списка научно-прикладных задач). Для решения этой задачи разработан и описан комплекс методов и инструментов выявления аномальных состояний, объединённых в методику МССР (Modeling, Challenge, Calibration, Production), где аббревиатуру МССР можно перевести на русский язык, в соответствии с входящими в неё процедурами, как последовательное: 1) моделирование, 2) сопоставление, 3) калибровка и 4) эксплуатация; причём, ключевая роль в этой схеме отводится именно моделированию нестабильных процессов средствами машинного обучения.

В целом, схема функционирования инструментов и методов МССР представлена в виде диаграммы (см. рисунок 5), её основные этапы разобраны ниже.

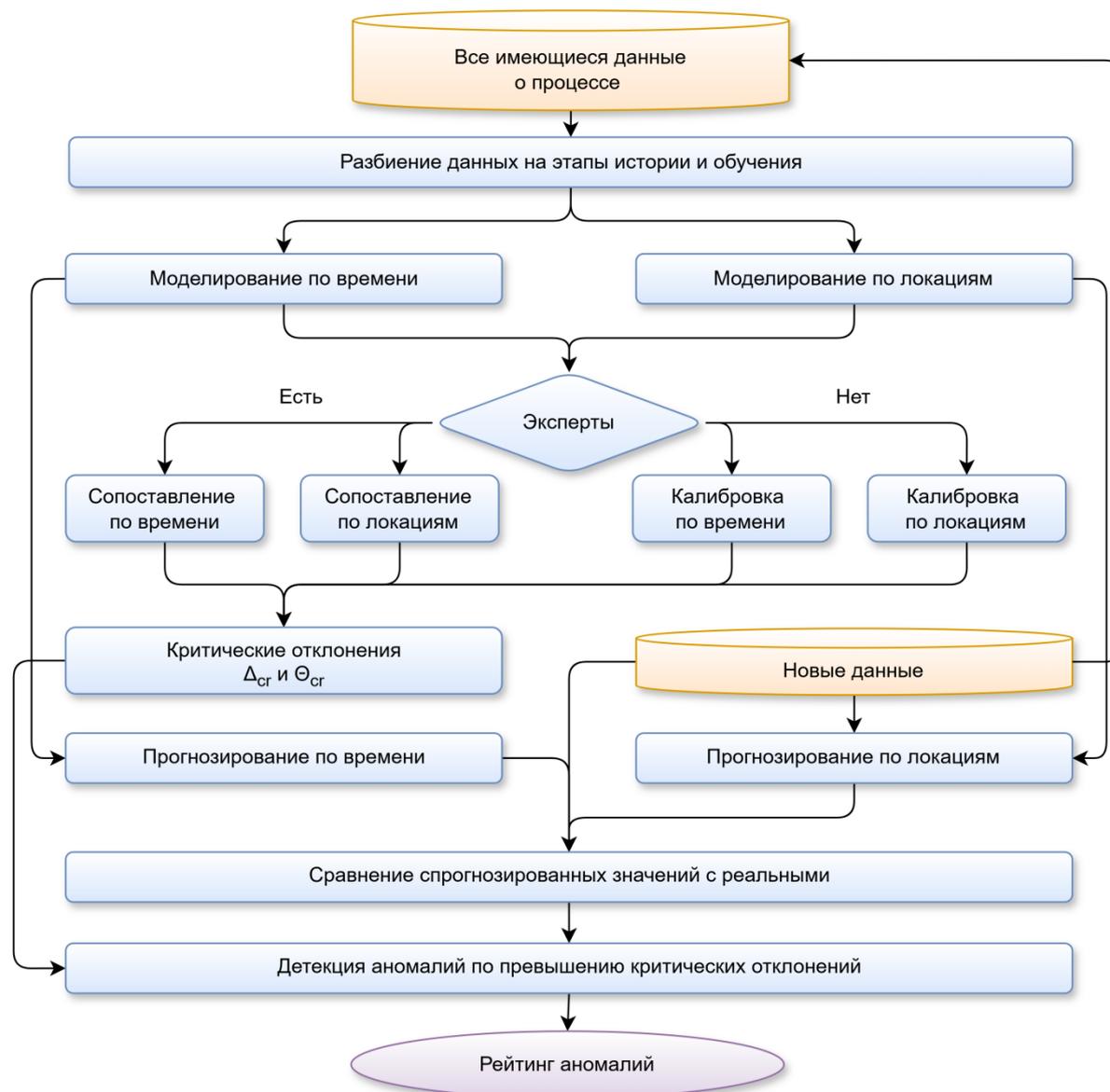


Рисунок 5 — Общая схема методики МССР

Прежде всего, на предварительном этапе, предшествующем основным этапам методики МССР, данные о процессе разбиваются на три последовательных блока (см. таблицу 2): 1) блок исторических наблюдений, который необходим для обучения моделей (в таблице 2

этот блок маркирован синим цветом), 2) блок обучения, необходимый для выбора наилучшей схемы моделирования и определения критической величины отклонения реальных наблюдений от значений, предсказанных лучшей моделью (маркирован жёлтым цветом), 3) блок пилотной эксплуатации детектора аномалий, необходимый для верификации двух предыдущих блоков (маркирован серым цветом).

Таблица 2. Разбиение данных в формате ML на блоки истории, обучения и эксплуатации

	Локация 1	Локация 2	Локация 3	...	Локация m
Момент 1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$...	$C_{1,m}$
...
Момент n1	$C_{n1,1}$	$C_{n1,2}$	$C_{n1,3}$...	$C_{n1,m}$
Момент n1+1	$C_{n1+1,1}$	$C_{n1+1,2}$	$C_{n1+1,3}$...	$C_{n1+1,m}$
...
Момент n2	$C_{n2,1}$	$C_{n2,2}$	$C_{n2,3}$...	$C_{n2,m}$
Момент n2+1	$C_{n2+1,1}$	$C_{n2+1,2}$	$C_{n2+1,3}$...	$C_{n2+1,m}$
...
Момент n	$C_{n,1}$	$C_{n,2}$	$C_{n,3}$...	$C_{n,m}$

Затем формируется набор из моделей машинного обучения, относящихся к разным семействам моделей (линейным, метрическим, древесным и т. д.) среди которых будет выбираться наилучшая схема моделирования процесса. В рамках настоящей диссертации применяется пул из семи моделей (обзор различных моделей машинного обучения и принципов их функционирования приведён в приложении А настоящей диссертации):

1. `LinearRegression` (линейная регрессия);
2. `DecisionTreeRegressor` (регрессия на решающих деревьях);
3. `RandomForestRegressor` (регрессия на случайных лесах);
4. `GradientBoostingRegressor` (градиентный бустинг над случайными лесами);
5. `SVR` (регрессия на методе опорных векторов);
6. `KNeighborsRegressor` (регрессия на ближайших соседях);
7. `MLPRegressor` (нейронная сеть многослойного перцептрона).

Этот список, разумеется, не является единственно возможным, он используется для демонстрации возможностей методики МССР (другие аналитики в других исследованиях могут формировать собственные модельные наборы по своему усмотрению). Однако, выбор именно такого набора моделей продиктован как минимум двумя нетривиальными соображениями. Во-первых, методика МССР предполагает не только временное, но и локационное прогнозирование, поэтому модели, специально разработанные для прогнозирования временных рядов, такие, например, как модели семейства ARIMA, в данном случае применять нельзя. А во-вторых, модели 1–7 относятся к разным семействам регрессионных моделей, в силу чего по-разному учитывают скрытые закономерности в данных о процессе и, следовательно, прогнозируют его состояние с разных позиций, что повышает шансы обнаружить качественный прогноз.

Задача процедуры моделирования (как по времени, так и по локациям) состоит в том, чтобы выяснить:

1. какая именно модель из числа используемых способная обеспечить наилучший прогноз для данных периода сопоставления,
2. при использовании какого числа переменных это происходит.

Так как для прогнозирования по времени применяется приём авторегрессии, ответ на этот вопрос, может быть, например, таким: «Наилучший прогноз обеспечивает модель регрессии на случайных лесах `RandomForestRegressor` при глубине авторегрессии, равной 4 наблюдениям (при этом используются 4 переменные, так как четыре раза происходит

последовательное смещение вектора наблюдаемой характеристики каждый раз на одну позицию в прошлое)». Или таким: «Наилучший прогноз даёт модель регрессии на ближайших соседях `KNeighborsRegressor` при глубине авторегрессии, равной 7 моментам (используются 7 переменных, так как смещение происходит семь раз)». И т. д.

Для получения конкретного результата при осуществлении процедуры моделирования выполняется вложенный двухуровневый цикл (см. рисунок 6) с полным перебором всех используемых моделей (то есть, моделей 1–7, хотя, список базовых моделей может быть и другим) и глубин авторегрессии (в рамках настоящей диссертации это были глубины от 1 до 10, хотя, разумеется, этот диапазон можно было бы уменьшить или, наоборот, увеличить).

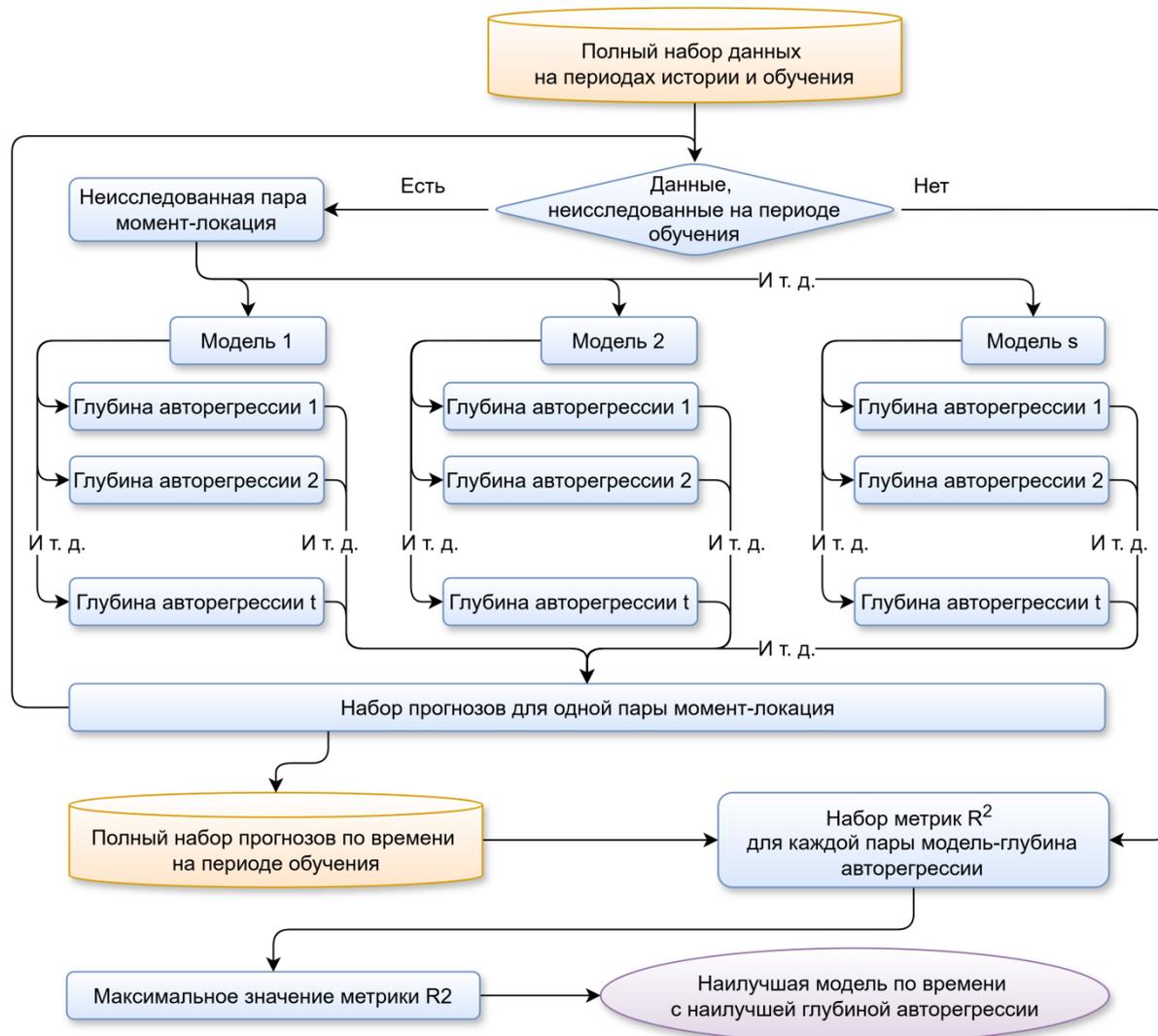


Рисунок 6 — Схема процедуры моделирования по времени

При моделировании по локациям приём авторегрессии невозможен (так как локации, в отличие от моментов времени, образуют неупорядоченное множество), поэтому в качестве переменных для регрессионных задач используются несколько предыдущих наблюдений, число которых задаёт глубину истории наблюдений (в рамках настоящей диссертации от 1 до 10), но в целом принципиальная схема локационного моделирования точно такая же, как и схема моделирования по времени (см. рисунок 6), с той лишь разницей, что вместо глубины авторегрессии следует применять глубину истории наблюдений. Так же, как и в случае моделирования по времени (см. выше), в результате отработки алгоритма моделирования по локациям возвращаемый ответ представляет собой пару модель-глубина, то есть, он может быть, например, таким: «Наилучшей моделью является линейная модель `LinearRegression`, при этом наилучшая глубина истории наблюдений равна 2», или таким: «Наилучшая

модель — это регрессия на ближайших соседях K NeighborsRegressor с глубиной истории наблюдений, равной 5» и т. д.

Модели позволяют алгоритму дать серию прогнозов относительно значений измеряемой характеристики процесса, среди которых какие-то прогнозы точны, какие-то менее точны, а какие-то совсем неточны. После того, как прогнозы получены, оценка их качества осуществляется при помощи коэффициента детерминации R^2 (подробно о метриках регрессионных моделей говорится в приложении А настоящей диссертационной работы), который принимает значения от минус бесконечности до 1, и, если не принимать во внимание возможный эффект переобучения, прогноз тем качественнее, чем выше значение коэффициента детерминации. Таким образом, при проведении этапа моделирования вычисляется серия детерминаций, наибольшая из которых указывает на наилучшую пару модель-глубина, где под глубиной понимается глубина авторегрессии для моделирования по времени или глубина истории наблюдений для моделирования по локациям. Такая пара называется наилучшей схемой моделирования, и на дальнейших этапах алгоритмической детекции аномалий алгоритм использует именно эту, отобранную на этапе моделирования схему. Следует отметить, что этап моделирования является наиболее затратным этапом методики МССР с алгоритмической точки зрения: использование сложных моделей во вложенном двумерном цикле занимает наибольшее время работы алгоритма. Но именно этот этап отличает предложенную для защиты методику МССР от традиционных статистических методов управления качеством процессов в автомобилестроении, так как именно именно здесь проявляют себя преимущества современных статистических методов, способных прогнозировать существенно нестабильные процессы с неконстантными и даже нелинейными трендами.

Применяя наилучшие схемы моделирования (как по времени, так и по локациям), алгоритм вычисляет прогнозируемые величины наблюдаемой характеристики процесса во всех парах момент-локация для периода обучения и анализа результатов, после чего каждой такой паре соответствует три числа: c_{ij} — истинное значение характеристики процесса в i -й момент времени в j -й локации, p_{ij} — значение характеристики, спрогнозированное по времени, и q_{ij} — ее значение, спрогнозированное по локациям (см. таблицу 3).

Таблица 3. Матрица истинных и спрогнозированных значений характеристики процесса

	Локация 1	Локация 2	...	Локация m
Момент $n1+1$	$c_{n1+1,1}, p_{n1+1,1}, q_{n1+1,1}$	$c_{n1+1,2}, p_{n1+1,2}, q_{n1+1,2}$...	$c_{n1+1,m}, p_{n1+1,m}, q_{n1+1,m}$
...
Момент $n2$	$c_{n2,1}, p_{n2,1}, q_{n2,1}$	$c_{n2,2}, p_{n2,2}, q_{n2,2}$...	$c_{n2,m}, p_{n2,m}, q_{n2,m}$

Кроме того, вычисляя разность между истинными и спрогнозированными значениями, алгоритм получает отклонения прогнозов от реальности, причём, это отклонения двух типов:

1. отклонение по времени: $\Delta_{ij} = c_{ij} - p_{ij}$;
2. отклонение по локациям: $\Theta_{ij} = c_{ij} - q_{ij}$.

Тем самым, каждая пара момент-локация характеризуется двумя числами, показывающими то, насколько реальное наблюдение, зафиксированное в этой локации в этот момент времени, отличается от того, которое прогнозировалось (см. таблицу 4).

Таблица 4. Матрица отклонений

	Локация 1	Локация 2	...	Локация m
Момент $n1+1$	$\Delta_{n1+1,1}, \Theta_{n1+1,1}$	$\Delta_{n1+1,2}, \Theta_{n1+1,2}$...	$\Delta_{n1+1,m}, \Theta_{n1+1,m}$
...
Момент $n2$	$\Delta_{n2,1}, \Theta_{n2,1}$	$\Delta_{n2,2}, \Theta_{n2,2}$...	$\Delta_{n2,m}, \Theta_{n2,m}$

Понятно, что чем больше абсолютная величина такого отклонения, тем более достоверно, что истинное наблюдаемое значение аномально, но в данный момент своей работы алгоритм ещё не располагает сведениями о том, каковы те пороговые значения, каждое преодоление

которых о должен интерпретировать как аномалии. Для вычисления пороговых отклонений методика МССР использует либо сопоставление с экспертным мнением (если таковое имеется), либо калибровку по локтевой точке (если мнение экспертов отсутствует).

При наличии экспертного мнения алгоритм сопоставляет найденные им отклонения с бинарным набором экспертных оценок пар момент-локация для каждой такой пары из периода обучения, а именно: если эксперт считает, что в i -й момент времени в j -й локации наблюдалось нормальное значение характеристики процесса, то он снабжает соответствующей наблюдение меткой 0; если же он считает, что наблюдалась аномалия, он ставит 1. Точно так же должен вести себя и алгоритм, но в его распоряжении на данном этапе работы имеются лишь числовые данные Δ_{ij} (или, если речь идёт о прогнозах по локациям, Θ_{ij}), которые являющиеся бинарными, их нужно предварительно бинаризовать для того, чтобы иметь возможность сравнить их с экспертной разметкой, которая изначально является бинарной. Алгоритм в цикле выставляет порог Δ_{cr} (или, если речь идёт о прогнозах по локациям, Θ_{cr}), то есть, критическое значение отклонения, превышение которого он считает аномалией, и все пары момент-локация, заменяет на 0 и 1 по очень простому правилу: если $\Delta_{ij} > \Delta_{cr}$, то ставится 1, в противном случае ставится 0 (аналогично он поступает и с прогнозами по локациям: если $\Theta_{ij} > \Theta_{cr}$, то ставится 1, в противном случае ставится 0). Критические пороги Δ_{cr} и Θ_{cr} прогоняются в цикле от минимально возможных значений (например, от нуля) до максимально возможных (например, до максимальных зафиксированных в процессе наблюдений отклонений Δ_{max} и Θ_{max}) с небольшим шагом, и на каждой итерации этого цикла производится сравнение алгоритмически полученной бинарной разметки с экспертной (см. рисунок 7).

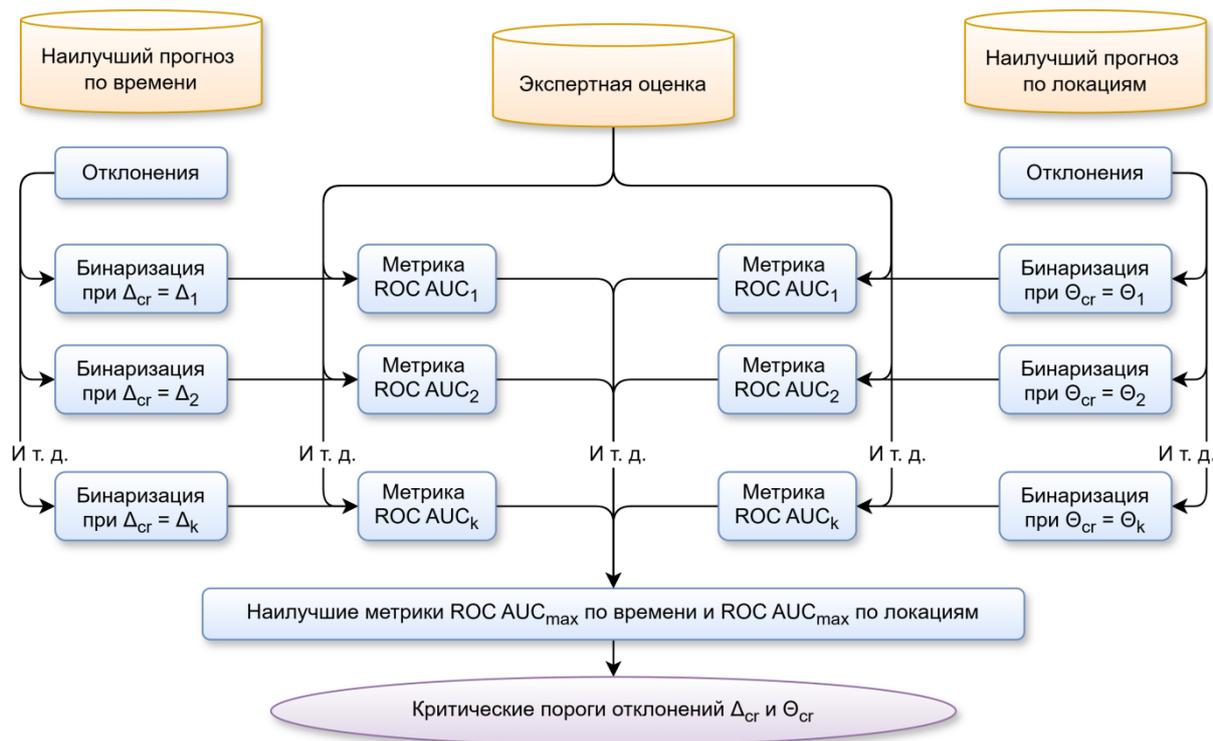


Рисунок 7 — Схема процедуры сопоставления

В качестве метрики такого сравнения двух бинарных наборов данных используется хорошо известный в теории бинарных классификаторов показатель: метрика ROC AUC, которая, по определению, изменяется от 0 до 1, и, если не принимать во внимание возможного эффекта переобучения, классификация считается тем более качественной, чем выше значение этой метрики (об этой и других метриках бинарной классификации подробно говорится в приложении А настоящей диссертации). Таким образом, после исполнения алгоритмом этой процедуры, алгоритм располагает следующими сведениями: 1) он знает, какие именно модели является наилучшими для осуществления временных и локационных

прогнозов; 2) какие именно схемы прогнозирования ему следует применять (какую давать глубину авторегрессии для прогнозов по времени, и какую именно давать глубину истории наблюдений для прогнозов по локациям); 3) какие именно отклонения от прогнозов считать критическими, а, значит — какие наблюдаемые значения признавать аномальными. Другими словами, после этапа сопоставления у алгоритма есть все сведения для того, чтобы переходить к этапу эксплуатации.

Однако, на практике экспертная разметка данных, как правило, недоступна. Это обусловлено многими причинами, одна из которых состоит в том, что квалифицированный эксперт — это всегда высоко дефицитный сотрудник, задействовать которого для предварительной разметки большого объема данных является, как минимум, нерентабельным решением (такой сотрудник гораздо более полезен на своём рабочем месте: конструктора, технолога, менеджера и т. д.). Кроме того, объем данных для разметки может быть настолько большим, а период обновления данных — настолько малым, что экспертная разметка может оказаться (и часто оказывается) физически невозможной. Поэтому для определения критических порогов отклонений в методике МССР предусмотрен ещё один метод: калибровка по локтевой точке. Этот метод состоит в том, что все зафиксированные выше отклонения упорядочиваются по абсолютной величине и располагаются в двумерной координатной системе, в результате чего возникает так называемая калибровочная кривая причём, в рамках методики МССР таких кривых две: одна из них визуализирует отклонения реальных значений от прогнозов по времени, а вторая — по локациям. Калибровочные кривые всегда являются неубывающими и, как правило, обладают тем свойством, что на стартовом отрезке возрастающий тренд таких кривых является слабым, а на финишном — сильным. Точка смены тренда со слабого на сильный и называется локтевой точкой. Она не является точкой экстремума (что вполне очевидно), она не является точкой перегиба (хотя в литературе приходится иногда встречать и такие выражения, являющиеся грубейшей математической ошибкой), она не обязательно является точкой максимальной кривизны и т. д., поэтому правильнее всего её определять именно так, как это было сделано выше: как точку смены тренда. Есть несколько алгоритмических методов вычисления локтевой точки: метод максимального удаления, метод максимального ускорения (метод второй производной) и т. д., а в рамках настоящей диссертации автор применял собственный метод, основанный на построения регрессионной сплит-модели (см. [16]), но так или иначе, если алгоритм располагает упорядоченным набором абсолютных отклонений, то он может вычислить локтевую точку на соответствующей калибровочной кривой. Ордината локтевой точки и есть критическая величина отклонения, превышение которой говорит о том, что отклонение от прогноза слишком большое для того, чтобы считать такое состояние процесса нормальным: если отклонение по абсолютной величине превышает ординату локтевой точки, то это явная аномалия. Схема калибровки по локтевой точке очень проста (см. рисунок 8), примеры калибровочных кривых приведены ниже (см. рисунки 17(а), 17(б) и 20(б)).

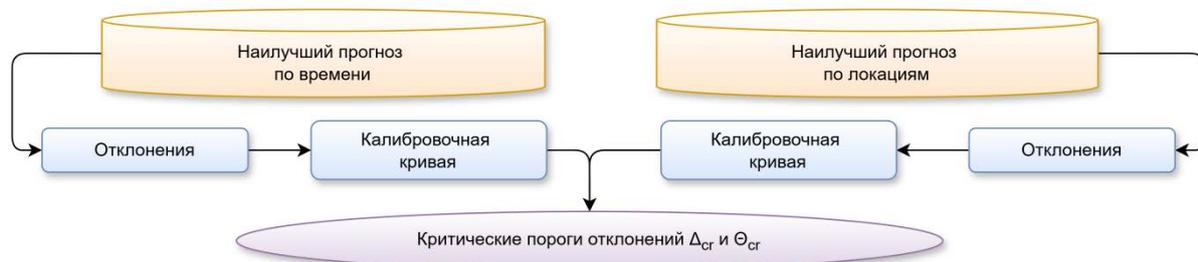


Рисунок 8 — Схема процедуры калибровки

Итак, если аналитик системы управления качеством располагает экспертным мнением, то критические величины отклонений получаются алгоритмом за счёт процедуры сопоставления, если же экспертная оценка отсутствует, то критические отклонения получаются за счёт процедуры калибровки по локтевой точке. В любом случае, после выполнения этого этапа алгоритм располагает необходимыми сведениями для того, чтобы

осуществлять процедуру эксплуатации: он знает 1) наилучшие модели для прогнозирования и по времени, и по локациям; 2) наилучшие схемы моделирования; 3) критические величины отклонений реально наблюдаемых значений от прогнозируемых.

Процедура пилотной эксплуатации состоит в том, что, прежде всего, данные, относящиеся к этапам вплоть до эксплуатационного (см. таблицу 2), объединяются в один блок (см. таблицу 5).

Таблица 5. Разбиение данных в формате ML на этапы истории и эксплуатации

	Локация 1	Локация 2	Локация 3	...	Локация m
Момент 1	C_{11}	C_{12}	C_{13}	...	C_{1m}
...
Момент n2	$C_{n2,1}$	$C_{n2,2}$	$C_{n2,3}$...	$C_{n2,m}$
Момент n2+1	$C_{n2+1,1}$	$C_{n2+1,2}$	$C_{n2+1,3}$...	$C_{n2+1,m}$
...
Момент n	$C_{n,1}$	$C_{n,2}$	$C_{n,3}$...	$C_{n,m}$

После этого реализация алгоритма осуществляется гораздо проще, чем это было на этапе моделирования (см. рисунок 9).

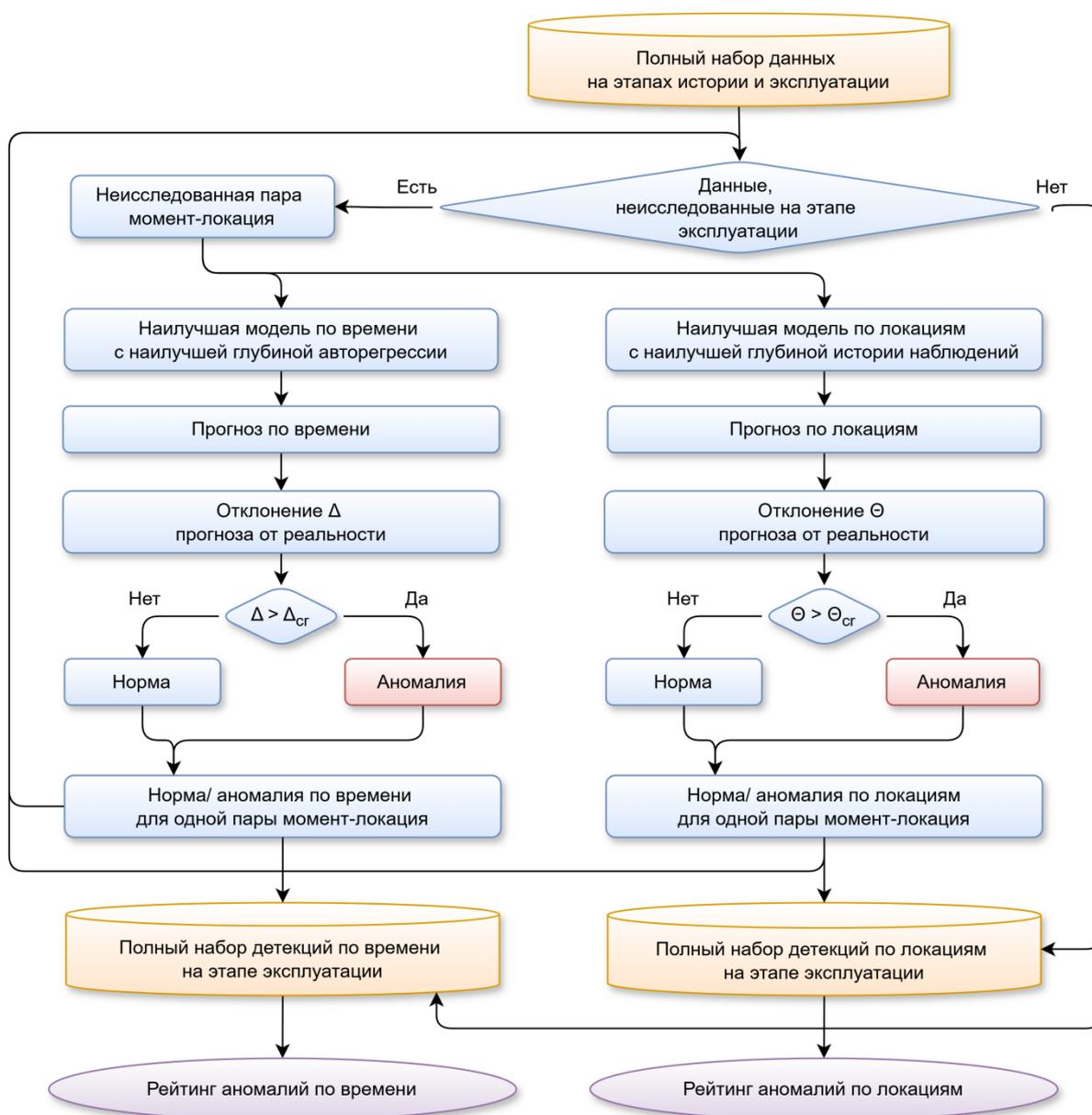


Рисунок 9 — Схема процедуры эксплуатации

Алгоритм берет наилучшую модель (известную заранее, так как она уже найдена выше), включает её в наилучшую схему с наилучшей глубиной (тоже найденную выше), осуществляет прогноз и сравнивает его с реальным значением наблюдаемой характеристики. Если отклонение превышает критический порог (тоже найденный выше), алгоритм помечает соответствующую пару модель-локация как аномалию и отправляет её в список аномалий. В таком режиме алгоритм анализирует каждую строку из блока эксплуатации таблицы 5. Каждая строка этого блока представляет собой очередной момент эксплуатации алгоритма, и по итогам анализа каждой строки алгоритм возвращает два списка аномалий: 1) список локаций, аномальных по отношению к прогнозу по времени, и 2) список локаций, аномальных по отношению к прогнозу по локациям. Таким образом, в каждый момент эксплуатации алгоритмического детектора аномалий аналитик службы качества имеет прямое указание к проведению проверочных мероприятий: следует провести внеплановый аудит именно в тех локациях, которые вернул алгоритм. Более того, алгоритм возвращает не только бинарную разметку норма-аномалия, но и выраженность аномальных состояний, которая позволяет упорядочить списки аномальных локаций в порядке убывания выраженности. Следовательно, в каждый момент времени у аналитика службы качества есть не только список аномальных локаций, но и предпочтительный порядок проведения проверочных мероприятий: начинать нужно с тех локаций, в которых аномалии выражены наиболее ярко, и заканчивать теми, где выраженность слабее. Это один из наиболее важных практических аспектов алгоритмической детекции аномалий (см. ниже, главу 7).

Третья глава посвящена экспериментальному исследованию возможностей методики МССР. В этой главе продолжается решение задачи, начатое ранее, а именно — разработка методологии алгоритмической детекции аномалий процессов в системе управления качеством в автомобильной отрасли (см. п. 3 списка научно-прикладных задач), и, кроме того — закладываются основы для решения задачи по разработке научно-программного комплекса моделей и алгоритмов алгоритмической детекции аномалий (см. п. 4 списка научно-прикладных задач), являющегося ключевым техническим элементом предложенной методологии.

Здесь синтезируются искусственные данные о процессе, протекающем в наборе локаций, и, при помощи вариации параметров синтеза, алгоритмы методики МССР применяются в разных ситуациях, другими словами — проводится компьютерный эксперимент для отладки алгоритмов моделирования, сопоставления алгоритмической оценки с экспертной оценкой и пилотной эксплуатации, причём, в первую очередь, это делается ради процедуры сопоставления, так как (и это уже отмечалось выше) получить реальную экспертную оценку на реальных процессах в производстве, сервисе или маркетинге автомобильной отрасли бывает весьма затруднительно, в тот момент как синтетические «эксперты» могут быть получены в неограниченном количестве, и более того: их мнение можно произвольно варьировать, наблюдая при этом подстройку алгоритма под изменяющееся экспертное мнение.

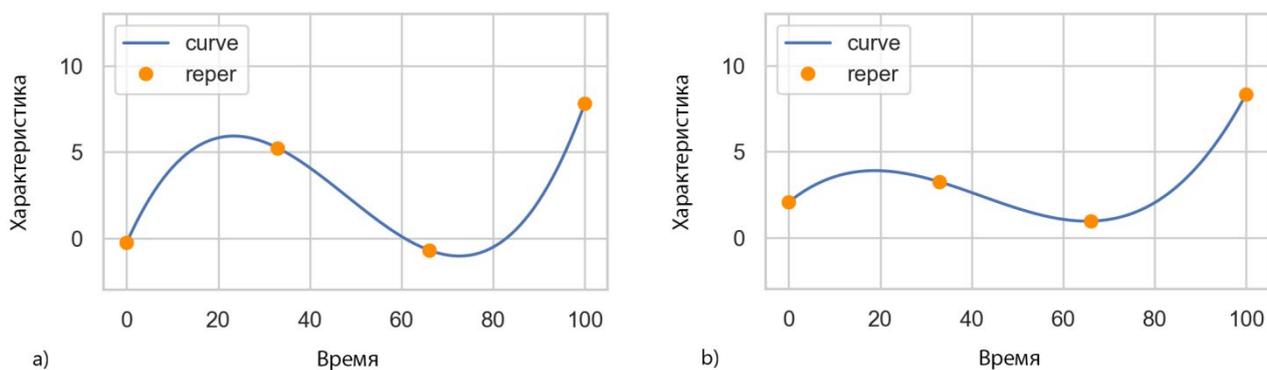


Рисунок 10 — Кубические параболы из одного семейства, построенные по разным реперам

Поскольку методика МССР ориентирована на процессы с существенно нелинейным трендом, в качестве базы синтетического процесса используется семейство кубических парабол, каждая из которых однозначно задаётся репером из четырёх точек (см. рисунок 10). При помощи генератора случайных чисел формируется семейство из девяти трендов (см. рисунок 11(a)), а, после наложения на них случайного шума, — семейство из девяти временных рядов (см. рисунок 11(b)).

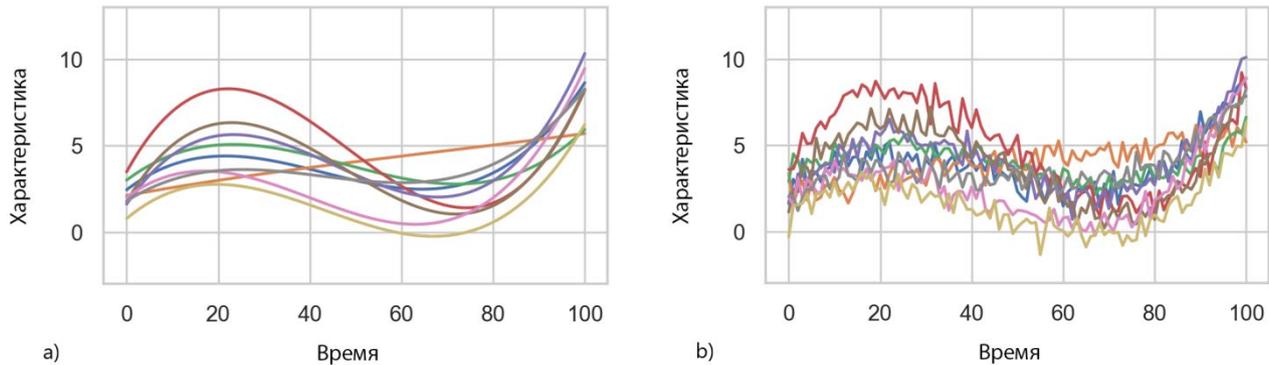


Рисунок 11 — Семейство а) трендов, б) временных рядов, описывающих синтетический процесс из ста наблюдений в девяти локациях

Тем самым, после процедуры генерации на случайных числах, получается описание процесса, который происходит в девяти локациях и наблюдается в течение ста моментов времени (см. таблицу 6, где представлены синтетические данные в формате ML).

Таблица 6. Синтетические данные в формате ML

	Локация 1	Локация 2	Локация 3	Локация 4	...	Локация 9
Момент 1	2.556096	2.852326	1.925826	4.213096	...	0.579310
Момент 2	2.963065	1.956280	3.053146	3.274485	...	1.568918
Момент 3	2.904049	2.634007	2.977713	4.606357	...	1.292128
...
Момент 100	9.230322	5.896291	5.240052	8.533112	...	6.980886

После этого в соответствии с изложенной выше методикой МССР данные разбиваются на три периода: 1) период исторических наблюдений (с 1 по 50), 2) период обучения (с 51 по 75) и 3) период пилотной эксплуатации (с 76 по 100). На этапе моделирования модели из зафиксированного выше набора моделей из семи моделей (см. описание главы 2) обучаются на данных исторического периода и предсказывают поведение процесса в каждой локации для всех моментов периода калибровки. Результаты работы этапа моделирования представлены ниже (см. рисунок 12).

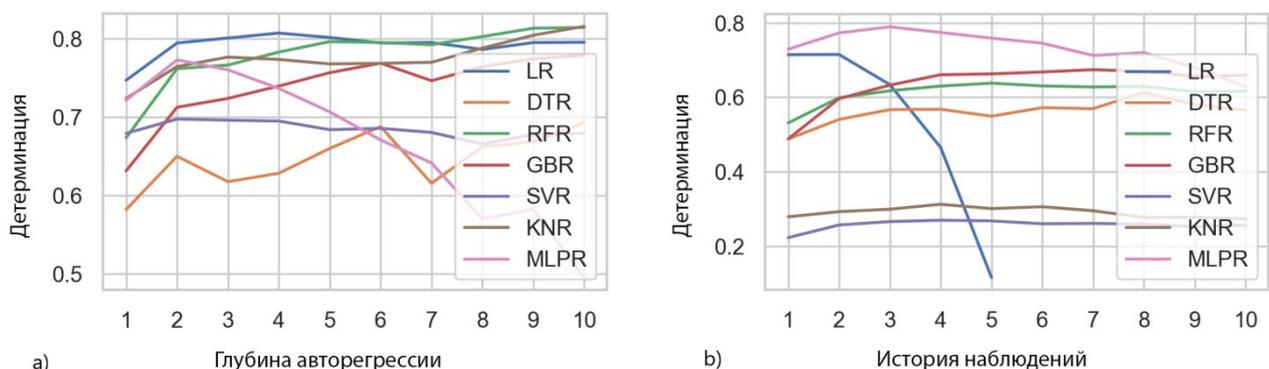


Рисунок 12 — Сравнительный анализ схем моделирования на синтетических данных а) по времени, б) по локациям

На этих иллюстрациях по горизонтали отложены глубина авторегрессии или глубина истории наблюдений (в случае моделирования по локациям), а по вертикали — значение коэффициента детерминации, вычисленного относительно истинных и предсказанных значений наблюдаемой характеристики. При моделировании по времени (см. рисунок 12(a)) наилучшие прогнозирующие возможности демонстрирует метрическая модель `KNeighborsRegressor` (описанная в легенде как `KNR`) с глубиной авторегрессии 10 (так как соответствующий коэффициент детерминации $R^2 = 0.8245$ является наибольшим среди всех рассмотренных схем), а при моделировании по локациям (см. рисунок 12(b)) вне всяких сомнений лидирует нейронная сеть `MLPRegressor` (описанная как `MLPR`) с глубиной исторических наблюдений, равной 3. Таким образом, по итогам этапа моделирования алгоритм возвращает следующие сведения:

1. при моделировании по времени лучшей является метрическая модель регрессии на ближайших соседях `KNeighborsRegressor` с глубиной авторегрессии 10;
2. при моделировании по локациям лучшей моделью является нейросетевая модель `MLPRegressor` с глубиной истории наблюдений 3.

Эти схемы моделирования дают предсказания во всех локациях для всех моментов времени на этапе обучения, что позволяет вычислить матрицы отклонений истинных значений наблюдаемой характеристики от спрогнозированных. Для моделирования по времени это матрица Δ (см. таблицу 7).

Таблица 7. Отклонения реальности от прогноза при прогнозировании пот времени

	Локация 1	Локация 2	Локация 3	Локация 4	...	Локация 9
Момент 51	-1.200534	0.518195	-0.491902	-0.231303	...	-0.118877
Момент 52	0.608003	0.255684	-0.138603	-1.513516	...	-0.708029
Момент 53	-0.101635	-0.382656	-2.023102	-1.341958	...	0.439615
...
Момент 75	-0.839624	0.068506	-0.126449	0.477126	...	-1.051690

Для моделирования по локациям это матрица Θ (см. таблицу 8), причём матрица Θ не получается из матрицы Δ транспонированием, так как они получены за счёт разных моделей и в силу разных принципов: первая из них получается после того, как была проведена процедура прогнозирования по времени, а вторая — после прогнозирования по локациям.

Таблица 8. Отклонения реальности от прогноза при прогнозировании по локациям

	Момент 51	Момент 52	Момент 53	Момент 54	...	Момент 75
Локация 1	-1.019977	1.472237	0.707600	-0.818095	...	-0.507161
Локация 2	0.200560	0.049882	0.562108	-1.258869	...	0.139162
Локация 3	-0.632156	0.376437	-1.678220	-0.811391	...	-0.039805
...
Локация 9	-0.322844	-0.590254	0.298375	-0.044333	...	-1.357558

Для интерпретации каждого такого отклонения как нормального или аномального, их абсолютные величины следует бинаризовать по пороговым значениям Δ_{cr} и Θ_{cr} , преодоление которых будет означать аномалию (то есть, в бинарном варианте матрицы Δ и Θ будут состоять из нулей и единиц, при этом, чем ниже пороги, тем больше в них будет единиц, и чем выше пороги — тем больше будет нулей). В третьей главе настоящей диссертации реализуется идея определения пороговых величин Δ_{cr} и Θ_{cr} , исходя из наличия экспертного мнения, поэтому, дополнительно к синтетическим данным о протекании процесса, синтезированным выше, в этой главе синтезируется сообщество в количестве 1000 «экспертов» (по сути, каждый такой «эксперт» — это просто бинарный вектор, содержащий

нули и единицы). В качестве управляемого параметра синтетического «экспертного сообщества» используется степень толерантности «эксперта», а именно:

1. если «эксперт» любое, даже незначительное отклонение процесса от прогнозируемого состояния признает аномалией, то относительно такого «эксперта» говорится, что он занимает жёсткую позицию;
2. наоборот, если «эксперт» даже весьма значительные отклонения признает, тем не менее, нормой, то считается, что такой «эксперт» занимает мягкую позицию;
3. наконец, «эксперты», с промежуточными значениями мнений считаются умеренными.

Если большинство «экспертов» занимает жёсткую позицию, то аномальными будут признаваться незначительные отклонения, и, значит, критический порог, который вычисляет алгоритм, подстраиваясь под такое экспертное мнение, должен быть меньше, чем в ситуации, когда большинство «экспертов» занимает мягкую позицию. При формировании жёсткого «экспертного мнения» доля жёстких «экспертов» составляет 90%, при этом порог бинаризации для прогнозов по времени составляет $\Delta_{cr} = 0.9625$, а для прогнозов по локациям $\Theta_{cr} = 0.9916$ (см. рисунок 13).

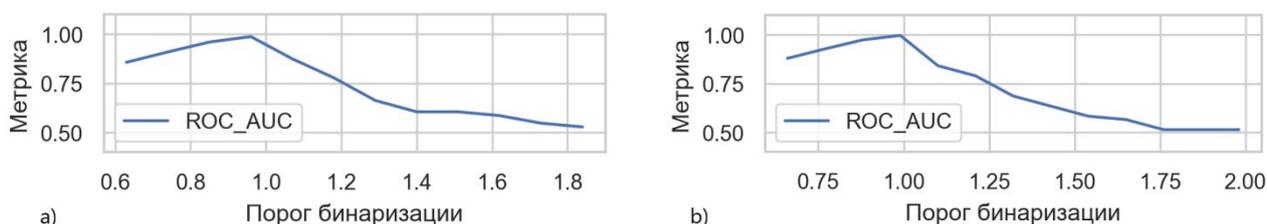


Рисунок 13 — Алгоритмический порог бинаризации относительно жёсткого «экспертного мнения»: а) отклонения от прогнозов по времени, б) отклонения от прогнозов по локациям

В случае умеренного «экспертного мнения» жёсткие и мягкие «эксперты» имеют по 50%, при этом для прогнозов по времени порог бинаризации $\Delta_{cr} = 1.1825$, а для прогнозов по локациям $\Theta_{cr} = 1.2205$ (см. рисунок 14).

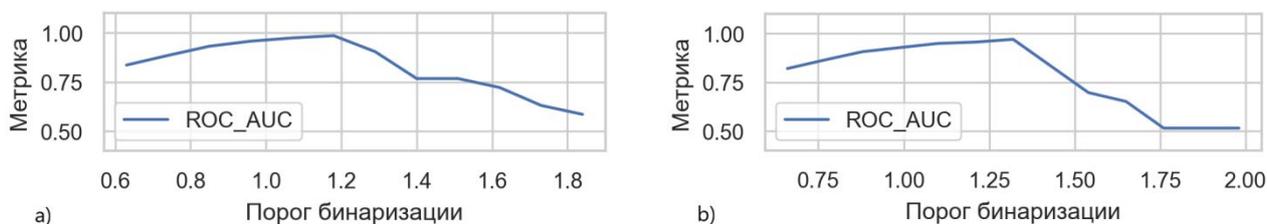


Рисунок 14 — Алгоритмический порог бинаризации относительно умеренного «экспертного мнения»: а) отклонения от прогнозов по времени, б) отклонения от прогнозов по локациям

Наконец, при преобладании мягких «экспертных мнений» доля жёстких «экспертов» составляет всего 10%, при этом пороги бинаризации отклонений по времени и по локациям соответственно равны $\Delta_{cr} = 1.4921$ и $\Theta_{cr} = 1.5415$ (см. рисунок 15).

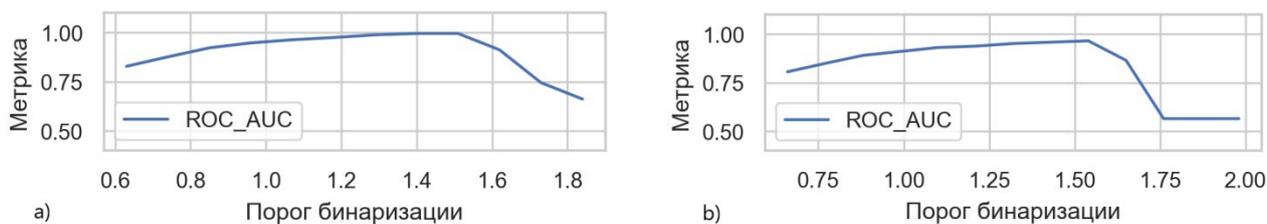


Рисунок 15 — Алгоритмический порог бинаризации относительно мягкого «экспертного мнения»: а) отклонения от прогнозов по времени, б) отклонения от прогнозов по локациям

Проведённый эксперимент на синтетических «экспертах» показывает, что алгоритм хорошо подстраивается под экспертное мнение и успешно проходит процедуру сопоставления. По итогам процедуры сопоставления алгоритм детекции аномалий располагает сведениями, необходимыми для осуществления пилотной эксплуатации (ему известны наилучшие схемы моделирования и критические пороги отклонений как по времени, так и по локациям). Процедура эксплуатации также описывается в конце третьей главы, но основной её результат состоит именно в демонстрации высокой адаптивности алгоритма к мнению экспертного сообщества (разумеется, при его наличии), а эксплуатация более подробно рассматривается в следующих главах, где средствами методики МССР исследуются реальные, а не синтетические процессы.

В **четвёртой главе** начинается решение научно-прикладной задачи апробации и практической реализации инструментария разработанной методологии детекции аномалий по отношению к процессам различной природы, включая производственные, сервисные и маркетинговые процессы (см. п. 5 списка научно-прикладных задач), и, наряду с этим, продолжается решение задачи разработки и реализации научно-программного комплекса моделей и алгоритмов детекции аномалий (см. п. 4 списка научно-прикладных задач), с акцентом на его интеграцию в практическую систему мониторинга и контроля качества в автомобильной отрасли.

В этой главе методика МССР применяется к производственному процессу установки уплотнителей на стекла автомобиля. В качестве наблюдаемой характеристики используется число выполненных за смену операций, а в качестве локаций процесса — когорта работников, принятых на конвейер в один и тот же период (см. таблицу 9).

Таблица 9. Данные о числе исполненных операций в формате ML

	Локация 1	Локация 2	Локация 3	...	Локация 24
Момент 1	98	106	102	...	105
Момент 2	95	102	108	...	110
Момент 3	97	103	107	...	111
...
Момент 515	194	191	213	...	186

Затем под множество исторических наблюдений отводятся первые 100 наблюдений, для этапа обучения — наблюдения со 101 по 150, а оставшиеся наблюдения используются для пилотной эксплуатации детектора аномалий и формирования календаря аномалий с указанием рейтингов их выраженности (см. ниже). Моделирования процесса осуществляются при помощи того же пула из семи моделей, что и при анализе синтетических данных, в результате чего возникают две серии детерминаций (см. рисунок 16), из которых выбираются две лучшие схемы моделирования.

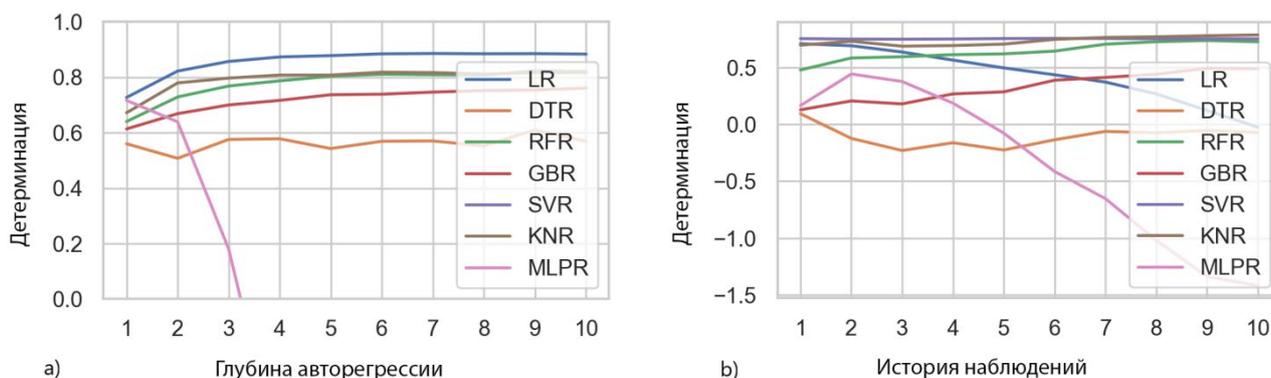


Рисунок 16 — Сравнительный анализ моделей в процессе установки уплотнителей на стекла автомобиля: а) по времени, б) по локациям

По итогам этапа моделирования алгоритм располагает следующими сведениями:

1. при моделировании по времени следует использовать модель линейной регрессии `LinearRegression` с глубиной авторегрессии 7 (описана в легенде как LR), так как именно её коэффициент детерминации $R^2 = 0.8401$ является максимальным среди всех рассмотренных схем;
2. при моделировании по локациям — тоже модель регрессии на ближайших соседях `KNeighborsRegressor` с глубиной истории наблюдений 10 (описана в легенде как KNR), так как именно её коэффициент детерминации $R^2 = 0.6022$ является максимальным.

В силу того, что на этих данных экспертное мнение отсутствует, для калибровки отклонений применяется метод локтевой точки, после чего алгоритм детекции аномалий получает сведения о порогах возможных отклонений наблюдаемой характеристики от модельных прогнозов (см. рисунок 17), а именно:

1. отклонения от прогнозов по времени не должны превышать $\Delta_{cr} = 55$;
2. отклонения от локационных прогнозов не должны превышать $\Theta_{cr} = 44$.

Эти сведения составляют основу для дальнейшей работы алгоритма.

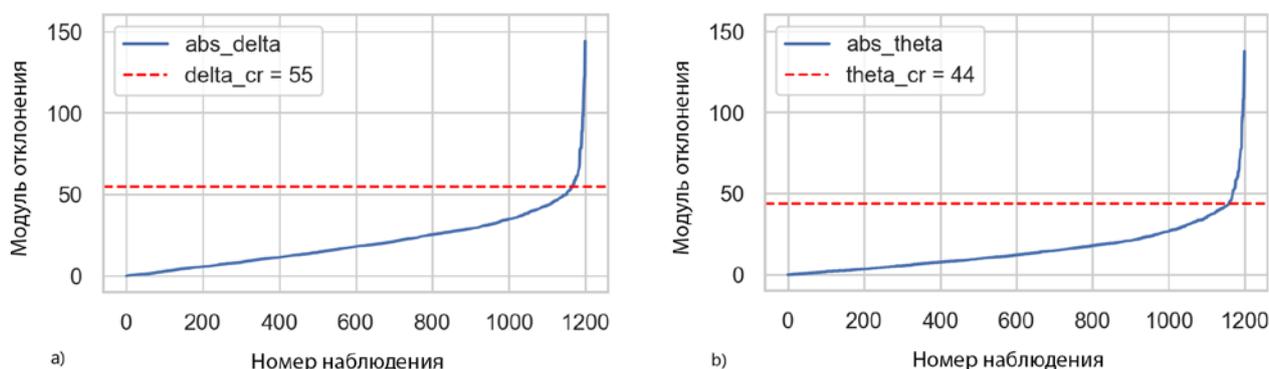
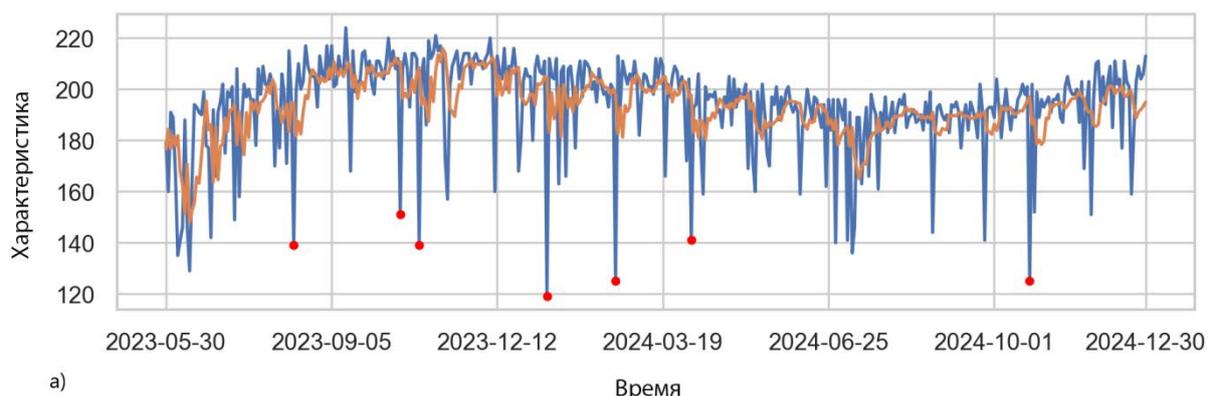


Рисунок 17 — Калибровка по локтевой точке отклонений в процессе установки уплотнителей на стекла автомобиля: а) по времени, б) по локациям

На последних наблюдениях, начиная со 151 момента времени и до конца наблюдений, проводится процедура пилотной эксплуатации детектора, для чего поведение каждой локации (синяя линия на рисунке 18) прогнозируется при помощи найденных выше наилучших схем моделирования (жёлтая линия на рисунке 18), и отклонения, превышающие критические значения, помечаются как аномальные (красные точки на рисунке 18). Например, в локации AZ (это код сотрудника) на этапе пилотной эксплуатации с 2023-05-30 по 2024-12-30 было зафиксировано 7 аномальных состояний относительно прогнозов по времени (см. рисунок 18(a)) и 12 аномальных состояний относительно прогнозов по локациям (см. рисунок 18(b)).



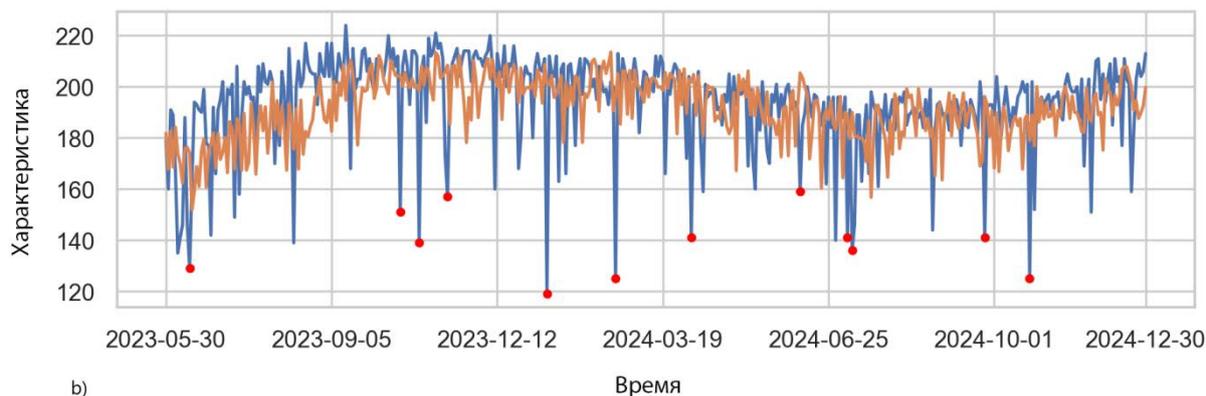


Рисунок 18 — Процесс установки уплотнителей на стекла автомобиля в локации AZ, его прогноз и детектированные аномалии: а) по времени, б) по локациям

Всего таких (точнее, подобных) пар графиков получается 24 (в каждой паре один график используется для временного прогнозирования и один — для локационного), по числу локаций, в которых наблюдается процесс. На каждом из таких графиков алгоритм отмечает моменты времени, которые он посчитал аномальными, и полная информация о аномальных состояниях процесса на этапе пилотной эксплуатации может рассматриваться как совокупность таких графиков.

Однако с практической точки зрения гораздо важнее сопряжённый срез результатов, то есть, важны не аномалии, зафиксированные в одной отдельно взятой локации во все моменты времени (как это сделано на рисунке 18), а наоборот: аномалии в отдельный момент времени во всех локациях сразу. Алгоритмически это совсем несложно и означает всего лишь другой формат вывода (см. рисунок 19, где для примера изображена ситуация в момент времени 2024-07-03).

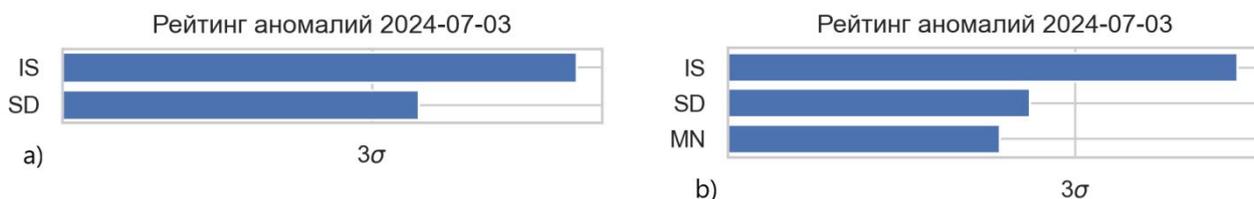


Рисунок 19 — Аномалии в процессе установки уплотнителей и их выраженность в момент времени 2024-07-03: а) по времени, б) по локациям

На этой иллюстрации аномалии, зафиксированные в локациях в момент времени 2024-07-03, упорядочены по степени их выраженности (в качестве которой используется абсолютная величина отклонения наблюдаемого значения от спрогнозированного, а для удобства их сравнения в качестве единицы масштаба использована психологически важная величина 3σ , где σ берётся относительно всех зафиксированных отклонений). Рейтинг выявленных аномалий непосредственно указывает, в какой последовательности необходимо проводить мероприятия по проверке и выяснению причин аномальных результатов: в первую очередь надо разобрать ситуацию с сотрудником IS, после чего — с сотрудником SD, и лишь в последнюю очередь — с сотрудником MN (а остальных сотрудников вообще не трогать).

В **пятой главе** продолжается решение научно-прикладных задач апробации и практической реализации инструментария разработанной методологии детекции аномалий (см. пп. 4 и 5 списка научно-прикладных задач). Здесь с помощью методики МССР аналогичным образом анализируется процесс сервисного обслуживания автомобилей. В качестве наблюдаемой характеристики используются число оформленных в течение рабочего дня заявок на сервисное обслуживание, а в качестве локаций — сервисные центры

одного из ведущих отечественных автопроизводителей. В результате отработки алгоритма, так же как и выше, в каждый момент пилотной эксплуатации формируется рейтинг аномалий, демонстрирующий специалисту по качеству сервисного обслуживания тот порядок, в котором следует проводить проверочные мероприятия, начиная с локаций, где аномалии процесса выражены наиболее сильно, и далее в порядке убывания выраженности аномалий.

Наконец, в **шестой главе** завершается решение научно-прикладных задач апробации и практической реализации инструментария методологии детекции аномалий (см. пп. 4 и 5 списка научно-прикладных задач). Здесь изучению средствами методики МССР подвергается процесс маркетингового сопровождения автомобильного бренда. В качестве наблюдаемой характеристики используется число просмотров, зарегистрированное в течение недели на официальном канале автопроизводителя, а в качестве локаций процесса — источники трафика. Результатом отработки алгоритма, так же как и выше, служит серия рейтингов аномалий, которая формируется в каждый момент пилотной эксплуатации и указывает порядок реагирования на выявленные аномалии.

В **седьмой главе** происходит возвращение к задаче научно-технического анализа проблемы контроля процессов в системе менеджмента качества автопроизводителя (см. п. 1 списка научно-прикладных задач). Данная глава охватывает вопросы интеграции разработанной методологии анализа процессов в существующую теорию и практику управления качеством, определяя её потенциальную роль и место в общей системе управления качеством предприятия.

В первой части этой главы рассматривается вопрос практического эффекта от внедрения алгоритмической детекции аномалий, состоящий в том, что такая методика существенно (как минимум — кратно) повышает результативность аудиторских проверок. Это объясняется тем, что методика МССР строится на моделях с высокой прогнозирующей способностью, в силу чего процент действительных аномалий среди состояний процесса, помеченных алгоритмом как аномалии, достаточно высок, в тот момент как доля аномалий среди всех наблюдений сравнительно низка. Например, в процессе сервисного обслуживания автомобилей доля аномальных наблюдений составляет порядка 5% (см. главу 5), и это означает, что при традиционном аудите сервисных центров результативной окажется только каждая двадцатая проверка. При этом доля действительных аномалий среди состояний, отмеченных алгоритмом как аномальные, составляет примерно 64%, и это значит, что как минимум каждая вторая проверка окажется результативной. То же справедливо и для других процессов, рассмотренных в настоящей диссертации: производственных (глава 4) и маркетинговых (глава 6).

Кроме того, во второй части этой главы поднимается вопрос о реализации в методике МССР принципа непрерывного улучшения, который основывается не на пассивном наблюдении, а на активном вмешательстве в процесс и устранении причин возникновения аномалий. При устранении одной из причин общая характеристика отклонений (для отклонений по времени это матрица Δ , а для отклонений по локациям — матрица Θ , описанные выше) становится меньше, калибровочная кривая приобретает новый вид, а критические величины отклонений (Δ_{cr} и Θ_{cr} соответственно) понижаются. Так как экспериментальное вмешательство в реальный процесс невозможно (это естественным образом сопряжено с необходимостью соответствующих допусков и полномочий), в седьмой главе, как и в третьей, эксперименты проведены на синтетических данных, и результаты экспериментов убедительно демонстрируют уменьшение критических величин допустимых отклонений с течением времени при устранении причин той или иной аномалии. Например, на иллюстрации ниже процесс до пятисотого наблюдения подвержен аномалиям двух типов, которые отмечены как зелёные и красные точки (см. рисунок 20(a)), но алгоритм не считает зелёные точки аномалиями. В пятисотый момент причина приводящая к возникновению красных точек устраняется, и, так как алгоритм в каждый

момент заново пересчитывает критические пороги отклонений, новый порог оказывается ниже (см. рисунок 20(b)), поэтому то, что раньше считалось допустимой нормой, начиная с пятисотого наблюдения считается аномалией (зелёные точки становятся красными).

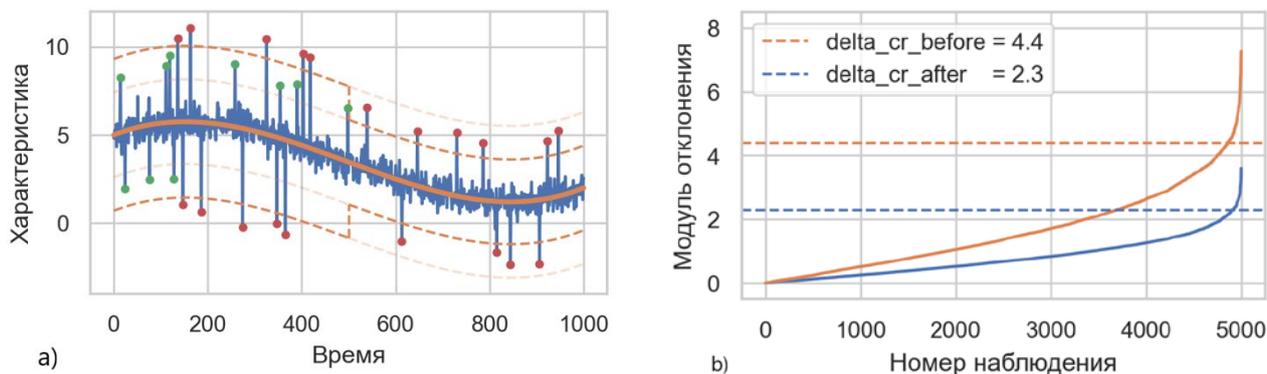


Рисунок 20 — а) временной ряд синтетического процесса, б) калибровочные кривые до и после устранения причин одной из аномалий

Таким образом, в самой логике алгоритма детекции аномалий заложен один из важнейших принципов управления качеством — принцип непрерывного улучшения, состоящий в постоянном совершенствовании абсолютно всех аспектов деятельности автопроизводителя, включая нестабильные по своей природе процессы.

В **заключении** формулируется сжатое резюме работы и определяются направления дальнейших исследований, такие как: повышение эффективности прогнозирующих моделей, индивидуальный подход к каждой локации на этапе калибровки, детекция аномалий по совокупности характеристик и т. д.

В **приложениях** приводятся краткие сведения об основах метода стохастического прогнозирования, основные методы моделирования и анализа данных на языке Python, коды основных функций, составляющих алгоритмическое ядро методики МССР, и основные практические результаты настоящей диссертационной работы, представленные в виде календарей аномальных состояний процессов, изучавшихся в главах 4–6.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ ПО РАБОТЕ И ВЫВОДЫ

В ходе работы над диссертацией решена важная научно-отраслевая проблема, предложена совокупность комплексных технических решений, в результате чего достигнута цель диссертационного исследования, состоящая в разработке нового методологического инструментария статистического управления качеством при помощи детекции аномалий в процессах СМК предприятий автомобильной отрасли.

Получены следующие научно-прикладные результаты:

1. По итогам научно-прикладного анализа проблемы, связанной с анализом процессов в системе менеджмента качества автопроизводителей и необходимостью обеспечения принципа принятия решений, основанных на свидетельствах, выделено существующее в отраслевой науке и практике управления качеством противоречие, заключающееся в отсутствии адекватных статистических методов и инструментов, направленных на повышение результативности и эффективности процессов, являющихся существенно нестабильными. Также выявлено, что действующие стандарты системы менеджмента качества ISO 9001:2015, и его отраслевой аналог IATF 16949:2016 не содержат требований по применению статистических инструментов управления по отношению к нестабильным процессам. В условиях информатизации и цифровизации процессов системы менеджмента, взрывного характера роста объемов электронных данных, предназначенных для контроля, мониторинга и управления качеством, требуется

опережающая разработка комплексных научно-технических решений, дающих адекватный ответ на современные вызовы. В качестве такого ответа следует рассматривать предлагаемую методологию и инструментарий детекции аномалий в управлении качеством процессов предприятий автомобильной отрасли.

2. Сформирована научная концепция методологии детекции аномалий в процессах, входящих в СМК предприятий автомобильной отрасли. Концепция является обобщением известных концепций статистического управления качеством, таких как концепция Деминга и тесно связанная с ней концепция бσ, и позволяет анализировать как стабильные, так и существенно нестабильные по своей природе процессы. Разработанная концепция представляет собой комплексный подход, обеспечивающий универсальное представление о различных процессах и подходах к определению нормы и аномалии на основании измеримых характеристик, и может быть использована для автоматизации управления качеством в таких процессах как маркетинг, проектирование, производство автомобильной продукции, сервисное обслуживание автомобилей в сети сервисных центров автопроизводителя, сопровождение автомобильного бренда и т. д.
3. Предложен комплекс научно-программных инструментов по обработке данных о процессах СМК автомобильной отрасли, позволяющий решать задачи по тестированию инструментов детекции аномалий (несоответствий) при мониторинге процессов. В рамках предложенного комплекса разработан инструментарий генерирования синтетических данных о протекании процессов системы менеджмента предприятий автомобильной отрасли и их экспертных оценках. В производственной практике применение предложенного комплекса обеспечивает создание виртуального инструментария объективного контроля качества мониторинга процессов СМК посредством тестирования инструментов детекции аномалий в условиях различных искусственно смоделированных экспертиз.
4. Разработаны методы моделирования состояния семейства процессов системы менеджмента, которые используются для описания производственных, сервисных и маркетинговых процессов, позволяющие одновременно анализировать группу из множества взаимосвязанных характеристик процессов СМК, протекающих на различных участках автомобильного производства — от штамповки кузовных деталей до сборки продукта, его последующего обслуживания в сервисных центрах и маркетингового сопровождения. Эти методы особенно ценны для крупных автоконцернов, имеющих разветвлённую сеть производственных и сервисных площадок (семейство локаций). Предложенные методы предназначены для оперативной детекции аномальных состояний (несоответствий) процессов, например, отклонений в параметрах технологических характеристик, объёме выполненных операций, суммарной стоимости затрат на устранение дефектов и т. д. Использование предложенных инструментов обеспечивает значительное улучшение показателей качества на предприятиях автомобильной отрасли (в части применения к процессам статистического мониторинга качества исполнения технологических операций), укреплению репутации производителя (в части его применения к процессам сервисного обслуживания автомобилей) и повышению конкурентоспособности на рынке (в части его применения к процессам маркетингового сопровождения автомобильного бренда). А именно:
 - 4.1. Для группы процессов, относящихся к эффективности выполнения технологических операций автомобильного производства, процент выявленных аномалий вырос с 4,1% до 56,4%.
 - 4.2. Для группы процессов сервисного обслуживания автомобилей в фирменных сервисных центрах автопроизводителя процент выявленных аномалий вырос с 5,6% до 64,2%.
 - 4.3. Для группы процессов, относящихся к маркетинговому сопровождению автомобильного бренда, процент выявленных аномалий вырос с 9,3% до 61,7%.

5. Разработан критерий детекции аномальных состояний процесса, основанный на методе локтевой точки, который позволяет автоматизировать выявление аномалий (несоответствий) процессов в системе менеджмента качества. Математический алгоритм, автоматически определяющий величину допустимых отклонений, позволяет мгновенно выявлять проблемы, такие как негативные изменения в клиентском сервисе или снижение заинтересованности потребителей в продукции автопроизводителя. Ключевое преимущество этого критерия заключается в исключении экспертной оценки, которая становится острой проблемой автомобильной отрасли в условиях ресурсных ограничений, с которыми столкнулись многие предприятия в настоящее время. Кроме того, исключение человеческого фактора повышает объективность, способствуя обоснованным управленческим и организационным решениям и как следствие — повышению качества продукции и услуг в автомобильной отрасли.
6. Разработан и реализован научно-программный комплекс моделей и алгоритмов, основанный на предложенных методологии и инструментарии детекции аномальных состояний. Комплекс, реализованный на языке Python для обеспечения гибкости и масштабируемости, включает в себя 6 ключевых алгоритмов, специально разработанных для обнаружения отклонений измеряемых характеристик процессов системы менеджмента от моделируемых трендов в режиме реального времени. Это критически важно для своевременного обнаружения потенциально опасных отклонений процессов, входящих в СМК автомобильной отрасли, от нормальных режимов их протекания. Разработанный научно-программный комплекс позволяет оперативно проводить количественный анализ выявленных аномалий (несоответствий), ранжируя их по степени выраженности, что позволяет оперативно принимать решения об уровне необходимого реагирования для каждой аномалии. Испытания комплекса на синтетических и реальных данных о процессах автомобильного производства, сервисного обслуживания автомобилей и маркетингового сопровождения автомобильного бренда подтвердили его высокую эффективность в задачах оперативной детекции аномальных состояний, позволяя сократить время реакции на возникающие критические отклонения как минимумкратно.
7. Ключевым практическим результатом диссертации является разработка и обоснование метода принятия управленческих решений по отношению к процессам, входящим в СМК автомобильной отрасли. Метод основан на рациональном анализе данных и, в отличие от традиционного статистического выборочного контроля, часто применяемого в автомобилестроении, использует полный анализ всей доступной статистической информации о производственных, сервисных и маркетинговых процессах. Это обеспечивает более точную и объективную картину для принятия обоснованных управленческих решений, причём, основным преимуществом разработанного метода является минимизация влияния человеческого фактора и субъективных оценок, часто приводящих к ошибкам в организационных и управленческих решениях. Достоверность принимаемых решений обеспечивается за счёт использования автоматизированных алгоритмов анализа данных и чётко определённых, объективных критериев оценки. Это позволяет избежать предвзятости и обеспечивает последовательность решений, что критически важно для поддержания стабильно высокого качества выпускаемой продукции и предоставляемых услуг. Реализованный метод позволяет руководителям принимать обоснованные управленческие решения по устранению выявленных отклонений, что, в конечном итоге, приводит к снижению затрат на контроль качества.
8. Предложенные научно-технические решения нашли своё применение и внедрены в устойчивую практику таких автомобильных концернов как АО «АВТОВАЗ» и ПАО «КАМАЗ». В ПАО «КАМАЗ» на основании данных о гарантийном обслуживании реализованы решения по автоматизированному выявлению и предотвращению необоснованных ремонтов. Также внедрён комплекс для обнаружения аномалий в базах

данных системы менеджмента качества, снижающий риски принятия неэффективных управленческих решений и способствующий развитию системы управления качеством предприятия. В АО «АВТОВАЗ» в целях повышения эффективности контроля качества автомобилей по параметрам электромагнитной совместимости (ЭМС) внедрены методы моделирования временных рядов и состояния семейства процессов, обеспечивающие анализ параметров качества автомобилей с позиции ЭМС как для перспективных, так и для серийных моделей. Используется научно-программный инструмент для детекции аномалий (несоответствий) в управлении качеством автомобильного производства на основе анализа данных испытаний автомобилей по параметрам электромагнитной совместимости и помехоустойчивости. Кроме того, реализован метод принятия решений о качестве автомобилей по параметрам ЭМС, обеспечивающий рационализацию выводов и исключение субъективного фактора при оценке. Все предложенные инструменты прошли апробацию и доказали свою производственную эффективность. В настоящее время они используются для анализа данных о качестве автомобилей с позиции ЭМС, что позволило снизить трудоёмкость процессов исследования и анализа данных на 40% при одновременном повышении достоверности и точности полученных результатов.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ АВТОРОМ ПО ТЕМЕ ДИССЕРТАЦИИ

Монографии

1. **Мосин В. Г.** Математические модели и методы медиаметрии / В. Г. Мосин. Самара: издательство СГАСУ, 2007. 112 с.

В изданиях, рекомендованных ВАК

2. **Мосин В. Г.** Предиктивная детекция аномалий в процессе гарантийного обслуживания автомобилей / В. Г. Мосин // Известия Тульского государственного университета. Технические науки. 2024. № 12. С. 506–513.
3. **Мосин В. Г.** Методология поиска и оценки предиктивной модели обращений в пункты гарантийного обслуживания автомобилей / В. Г. Мосин // Известия Тульского государственного университета. Технические науки. 2024. № 10. С. 126–132.
4. **Мосин В. Г.** О влиянии востребованности контента на некоторые характеристики вовлеченности в его потребление / В. Г. Мосин, В. Н. Козловский, О. В. Пантюхин // Известия Тульского государственного университета. Технические науки. 2024. № 1. — С. 250–258.
5. **Мосин В. Г.** Обоснование методологии детекции аномалий в качестве нового цифрового инструмента управления качеством / В. Г. Мосин, В. Н. Козловский // Известия Самарского научного центра Российской академии наук, т. 26, № 3, 2024, С. 16-22.
6. **Мосин В. Г.** Анализ качества электронного контента. Разбиение данных о потреблении контента на оптимальное число кластеров при помощи алгоритма KMEANS / В. Г. Мосин, В. Н. Козловский // Известия Тульского государственного университета. Технические науки. 2024. № 2. С. 604–609.
7. **Мосин В. Г.** Анализ стабильности процессов с сильно коррелирующими признаками / В. Г. Мосин, В. Н. Козловский, О. И. Антипова, С. А. Васин // Известия Тульского государственного университета. Технические науки. 2024. № 5. С. 11–16.
8. **Мосин В. Г.** Бинарная модель оценки качества и востребованности электронного контента и ее вероятностный прогноз / В. Г. Мосин, В. Н. Козловский, Д. И. Благовещенский // Известия Тульского государственного университета. Технические науки. 2023. № 12. С. 79–84.

9. **Мосин В. Г.** Востребованность контента как один из факторов качества информационного канала / В. Г. Мосин, В. Н. Козловский // *Качество и жизнь*. 2023. № 4 (40). С. 6–11.
10. **Мосин В. Г.** Детекция аномалий информационного канала на основе прогнозирующих моделей в решении задач анализа качества контента / В. Г. Мосин, В. Н. Козловский, О. В. Пантюхин // *Известия Тульского государственного университета. Технические науки*. 2024. № 3. С. 421–425.
11. **Мосин В. Г.** Инструменты управления качеством. Поиск калибровочных границ методами машинного обучения / В. Г. Мосин, В. Н. Козловский, О. И. Антипова, Р. Р. Гафаров // *Известия Тульского государственного университета. Технические науки*. 2024. № 7. С. 10–15.
12. **Мосин В. Г.** Мониторинг данных о результативности процессов системы менеджмента. Настройка параметров кластеризации DBSCAN по подходящим значениям метрики силуэта / В. Г. Мосин, В. Н. Козловский, С. А. Васин, Г. С. Стрелков // *Известия Тульского государственного университета. Технические науки*. 2024. № 4. С. 28–33.
13. **Мосин В. Г.** Мониторинг качества. Сравнительный анализ двух методик поиска значимых предикторов регрессионной модели / В. Г. Мосин, В. Н. Козловский, О. В. Пантюхин, Д. В. Антипов // *Известия Тульского государственного университета. Технические науки*. 2024. № 5. С. 39–45.
14. **Мосин В. Г.** Новые инструменты мониторинга качества. Анализ потребления контента на разреженных данных / В. Г. Мосин, В. Н. Козловский, О. В. Пантюхин // *Известия Тульского государственного университета. Технические науки*. 2024. № 1. С. 208–214.
15. **Мосин В. Г.** Новые инструменты мониторинга качества. Мониторинг данных, временные ряды потребления контента и усиление их прогнозирующих свойств ансамблированием слабых моделей / В. Г. Мосин, В. Н. Козловский, Д. И. Благовещенский // *Известия Тульского государственного университета. Технические науки*. 2024. № 1. С. 237–246.
16. **Мосин В. Г.** Новые инструменты управления качеством. Метод поиска локтевой точки для определения приемлемых значений метрик машинного обучения / В. Г. Мосин, В. Н. Козловский, А. П. Новикова // *Известия Тульского государственного университета. Технические науки*. 2024. № 4. С. 53–56.
17. **Мосин В. Г.** О влиянии востребованности контента на некоторые характеристики вовлеченности в его потребление / В. Г. Мосин, В. Н. Козловский, О. В. Пантюхин // *Известия Тульского государственного университета. Технические науки*. 2024. № 1. С. 250–258.
18. **Мосин В. Г.** Детекция аномалий в процессе гарантийного обслуживания автомобилей по методике МССР (Modeling, Calibration, Challenge, Production) / В. Г. Мосин, В. Н. Козловский, С. А. Васин, О. В. Пантюхин // *Известия Тульского государственного университета. Технические науки*. 2024. № 10. С. 132–141.
19. **Мосин В. Г.** О некоторых особенностях определения качества продукции на примере данных одного из ведущих отечественных автопроизводителей / В. Г. Мосин, В. Н. Козловский, Н. А. Антонова // *Известия Самарского научного центра Российской академии наук*. 2023. Т. 25. № 4 (114). С. 47–55.
20. **Мосин В. Г.** Повышение производительности регрессионных моделей при оценке качества потребления электронного контента / В. Г. Мосин, В. Н. Козловский // *Известия Тульского государственного университета. Технические науки*. 2024. № 2. С. 593–599.
21. **Мосин В. Г.** Понижение размерности данных методом SVD на примере данных о потреблении контента / В. Г. Мосин, В. Н. Козловский // *Известия Тульского государственного университета. Технические науки*. 2024. № 2. С. 599–603.
22. **Мосин В. Г.** Применение сингулярных разложений к обнаружению выбросов в данных о потреблении информационного контента / В. Г. Мосин, В. Н. Козловский, С. А. Васин //

Известия Тульского государственного университета. Технические науки. 2024. № 3. С. 464–470.

23. **Мосин В. Г.** Сезонная структура потребления контента / В. Г. Мосин, В. Н. Козловский, Д. И. Благовещенский // Известия Тульского государственного университета. Технические науки. 2024. № 3. С. 490–495.
24. **Мосин В. Г.** Сравнительный анализ двух метрик кластеризации данных о потреблении контента / В. Г. Мосин, В. Н. Козловский // Известия Тульского государственного университета. Технические науки. 2024. № 2. С. 609–615.
25. **Мосин В. Г.** Мониторинг данных о качестве процессов. Теорема Байеса в приложении к бинарной классификации пользователей на основе анализа их транзакций / В. Г. Мосин, В. Н. Козловский, А. С. Клентак, С. А. Васин // Известия Тульского государственного университета. Технические науки. 2024. № 9. С. 115–120.
26. **Мосин В. Г.** Методика МССР (Modeling, Calibration, Challenge, Production) в сравнительном анализе пунктов гарантийного обслуживания автомобилей / В. Г. Мосин, В. Н. Козловский, А. С. Клентак, О. В. Пантюхин // Известия Тульского государственного университета. Технические науки. 2025. № 1. С. 272–289.
27. **Мосин В. Г.** Прогнозирование качества продукции. Моделирование коротких временных рядов с выраженной периодичностью / В. Г. Мосин, В. Н. Козловский, А. С. Клентак, О. В. Пантюхин // Известия Тульского государственного университета. Технические науки. 2024. № 9. С. 120–128.
28. **Мосин В. Г.** Детекция аномалий в процессе гарантийного обслуживания автомобилей по методике МССР (Modeling, Calibration, Challenge, Production) / В. Г. Мосин, В. Н. Козловский, С. А. Васин, О. В. Пантюхин // Известия Тульского государственного университета. Технические науки. 2024, № 10. С. 132–141.
29. **Мосин В. Г.** Парето-анализ качества работы сервисных центров автопроизводителей / В. Г. Мосин, К. А. Брагина, В. Н. Козловский, А. В. Гусев // Известия Самарского научного центра Российской академии наук, т. 27, № 3, 2025. С. 92–98.
30. Благовещенский Д. И. Комплексная программа улучшений как инструмент развития организации: Стратегический уровень организации и управления / Д. И. Благовещенский, В. Н. Козловский, **В. Г. Мосин**, О. И. Антипова // Известия Тульского государственного университета. Технические науки. 2024, № 7. — С. 3–6.
31. Крицкий А. И. Контроль качества электрокомпонентов в состоянии поставки в автосборочное производство / А. В. Крицкий, В. Н. Козловский, Е. В. Стрижакова, **В. Г. Мосин** // Известия Тульского государственного университета. Технические науки. 2023. № 7. С. 326–331.
32. Беляева И. А. Организация деятельности операционного комитета по продуктам в машиностроительном производстве / И. А. Беляева, А. С. Подгорный, **В. Г. Мосин**, В. Н. Козловский // Известия Самарского научного центра Российской академии наук. 2024. Т. 26. № 6 (122). С. 38–46.
33. Гусев А. В. Концепция, логика, инструментарий, информационные связи при взаимодействии автопроизводителя и поставщиков автомобильных компонентов / А. В. Гусев, В. Н. Козловский, **В. Г. Мосин**, И. А. Беляева // Известия Самарского научного центра Российской академии наук. 2025. Т. 27, № 5. С. 94–101.
34. Демкин А. С. Развитие методов мониторинга качества автомобилей в период гарантийной эксплуатации / А. С. Демкин, А. С. Саксонов, В. Н. Козловский, **В. Г. Мосин** // Известия Самарского научного центра Российской академии наук. 2025. Т. 27, № 5. С. 166–171.

В иностранных научных изданиях

35. **Mosin V. G.** Predictive analysis of two bijectively related families of functions which are expressed as tuple pairs / V. G. Mosin, A. A. Abashkin // В сборнике: E3S Web of

Conferences. Сер. «Topical Problems of Agriculture, Civil and Environmental Engineering, TPACEE 2020», 2020.

36. **Mosin V. G.** Mathematical model of bank scoring in conditions of insufficient data / V. G. Mosin, A. A. Abashkin, O. V. Yusupova // В сборнике: E3S Web of Conferences. Сер. «Topical Problems of Agriculture, Civil and Environmental Engineering, TPACEE 2021», 2021.
37. **Mosin V. G.** Novel Tools for Monitoring Process Quality: Detection of Anomalies in Digital Data on Content Consumption Using the DBSCAN Clustering Method / V. G. Mosin, V. N. Kozlovsky, M. V. Nenashev // Russian Engineering Research, 2024, Vol. 44, No. 9. pp. 1370–1374.
38. **Mosin V. G.** Comprehensive Algorithmic Detection of Abnormal States in the Process of Automotive Warranty Services / V. G. Mosin, K. A. Bragina, V. N. Kozlovsky, A. S. Klentak // Russian Engineering Research, 2025, Vol. 45, No. 11. pp. 1594–1600.
39. Deryugina V. A. Detecting Anomalous Data in the Marketing of Automobile Brands / V. A. Deryugina, **V. G. Mosin**, V. N. Kozlovskii, U. V. Brachunova // Russian Engineering Research, 2025, Vol. 45, No. 9, pp. 1318–1321.

Научные публикации в других научных изданиях

40. **Мосин В. Г.** Новые инструменты мониторинга качества процессов: Детекция аномалий в электронных данных о потреблении контента методом кластеризации dbSCAN / В. Г. Мосин, В. Н. Козловский, Дм. В. Ненасhev // СТИН. 2024. № 8. С. 49–52.
41. **Мосин В. Г.** Комплексная алгоритмическая детекция аномальных состояний в процессе гарантийного обслуживания автомобилей / В. Г. Мосин, К. А. Брагина, В. Н. Козловский, А. С. Клентак // СТИН, № 10, 2025. С. 23–28.
42. Егорова И. П. Об одном аспекте математической модели конвективного атмосферного переноса загрязняющей аэрозольной субстанции / И. П. Егорова, **В. Г. Мосин**, А. И. Насретдинова // Мониторинг. Наука и технологии. 2018. № 2 (35). С. 39–43.
43. **Мосин В. Г.** Кодирование категориальных данных в задачах регрессионного анализа / В. Г. Мосин, И. П. Егорова // В сборнике: Инновационное развитие современной науки: Теория, методология, практика. Сборник статей II Всероссийской научно-практической конференции. Петрозаводск, 2021. С. 163–169.
44. **Мосин В. Г.** Линеаризация целевой функции в регрессионных задачах методом сингулярных разложений / В. Г. Мосин // В книге: Математическое моделирование. Тезисы II Международной конференции. Москва, 2021. С. 66–67.
45. **Мосин В. Г.** О некоторых проблемах моделирования измеряемых социально-психологических переменных / В. Г. Мосин // В сборнике: Математическое образование в современном мире: теория и практика. Материалы Всероссийской научно-методической конференции с международным участием. Самара, 2022. С. 165–174.
46. **Мосин В. Г.** Временные ряды данных о пунктах гарантийного обслуживания автомобилей и их актуальная глубина / В. Г. Мосин // В сборнике: Актуальные вопросы экономики и управления: Теоретические и прикладные аспекты. Материалы X Международной научно-практической конференции, г. Горловка, 21 марта 2025 г.
47. Дерюгина В. А. Детекция аномальных состояний в маркетинговых процессах автомобильного бренда / В. А. Дерюгина, **В. Г. Мосин**, В. Н. Козловский, У. В. Барчунова // СТИН, № 8, 2025. С. 38–41.
48. Бахарева В. Б. Регрессионный анализ вовлеченности пользователя в контент / В. Б. Бахарева, **В. Г. Мосин** // Тенденции развития науки и образования. 2024. № 106-11. — С. 17–21.
49. Караваев А. В. Оценка важности категориальных признаков с использованием one-hot-кодирования для модели линейной и гребневой регрессии / А. В. Караваев, **В. Г. Мосин** // В сборнике: Прикладная математика и информатика: современные исследования в области естественных и технических наук. VIII Международная научно-практическая

конференция (школа-семинар) молодых ученых: сборник материалов. Тольятти, 2022. С. 174–179.

50. Хрисанфова К. С. Применение полиномиальных признаков в области линейного регрессионного анализа / К. С. Хрисанфова, **В. Г. Мосин**, В. П. Корнеев // В сборнике: Математика и математическое моделирование. Материалы Всероссийского конкурса научно-исследовательских работ (проектов) обучающихся. Самара, 2023. С. 67–73.