

Костанайский региональный университет имени Ахмет Байтұрсынұлы

УДК 004.891.3

На правах рукописи

**МАДИН ВЛАДИМИР АНАТОЛЬЕВИЧ**

**Разработка и исследование моделей интеллектуального анализа больших  
данных в задачах планирования**

8D06102 – Информационные технологии и робототехника

Диссертация на соискание степени  
доктора философии (PhD)

Научный консультант  
доктор PhD,  
ассоциированный профессор  
О.С. Салыкова

Зарубежный научный консультант  
доктор технических наук,  
профессор  
И.М. Ячиков  
(ФГАОУ ВО «ЮУрГУ (НИУ)»)

Республика Казахстан  
Костанай, 2026

## СОДЕРЖАНИЕ

|          |                                                                                                                               |    |
|----------|-------------------------------------------------------------------------------------------------------------------------------|----|
|          | <b>НОРМАТИВНЫЕ ССЫЛКИ.....</b>                                                                                                | 4  |
|          | <b>ОПРЕДЕЛЕНИЯ.....</b>                                                                                                       | 5  |
|          | <b>ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ.....</b>                                                                                          | 7  |
|          | <b>ВВЕДЕНИЕ.....</b>                                                                                                          | 8  |
| <b>1</b> | <b>АНАЛИТИЧЕСКИЙ ОБЗОР МЕТОДОВ ПРОГНОЗИРОВАНИЯ И ПЛАНИРОВАНИЯ ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ.....</b> | 15 |
| 1.1      | Неопределенность в производственном планировании: постановки и источники вариативности.....                                   | 15 |
| 1.2      | Производственные данные и методы машинного обучения для прогнозирования параметров.....                                       | 17 |
| 1.3      | Представление неопределенности и риск-ориентированные подходы в планировании.....                                             | 23 |
| 1.4      | Интеграция прогнозирования и календарного планирования: существующие подходы и разрыв.....                                    | 27 |
|          | Выводы по разделу.....                                                                                                        | 30 |
| <b>2</b> | <b>МОДЕЛИ И МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОГО КАЛЕНДАРНОГО ПЛАНИРОВАНИЯ.....</b>                                                       | 32 |
| 2.1      | Формализация задачи календарного планирования.....                                                                            | 32 |
| 2.2      | Учет неопределенности параметров в задачах планирования.....                                                                  | 36 |
| 2.3      | Методы оптимизации календарного планирования.....                                                                             | 39 |
| 2.4      | Интеграция прогнозных моделей и оптимизации.....                                                                              | 42 |
|          | Выводы по разделу.....                                                                                                        | 45 |
| <b>3</b> | <b>РАЗРАБОТКА И РЕАЛИЗАЦИЯ ГИБРИДНЫХ МОДЕЛЕЙ ИНТЕЛЛЕКТУАЛЬНОГО ПЛАНИРОВАНИЯ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ.....</b>             | 46 |
| 3.1      | Формализация задач прогнозно-оптимизационного планирования...                                                                 | 46 |
| 3.2      | Данные и экспериментальный контур прогнозирования и планирования.....                                                         | 51 |
| 3.3      | Гибридная модель прогнозирования длительности операций.....                                                                   | 62 |
| 3.4      | Интеграция прогнозов машинного обучения в оптимизационную модель.....                                                         | 66 |
| 3.5      | Робастное планирование на основе квантильных и риск-ориентированных оценок.....                                               | 71 |
| 3.6      | Сравнительный анализ сценариев и интегральная оценка эффективности.....                                                       | 76 |
|          | Выводы по разделу.....                                                                                                        | 79 |
| <b>4</b> | <b>ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ И АНАЛИЗ ЭФФЕКТИВНОСТИ ГИБРИДНОЙ СИСТЕМЫ ПРОИЗВОДСТВЕННОГО ПЛАНИРОВАНИЯ.....</b>               | 81 |
| 4.1      | Экспериментальная постановка и сценарии вычислительных экспериментов.....                                                     | 81 |

|          |                                                                                           |            |
|----------|-------------------------------------------------------------------------------------------|------------|
| 4.2      | Анализ результатов вычислительных экспериментов для различных сценариев планирования..... | 84         |
| 4.3      | Сравнительный и статистический анализ сценариев планирования                              | 86         |
| 4.4      | Обсуждение результатов.....                                                               | 89         |
| 4.5      | Валидация, устойчивость и обобщаемость результатов вычислительных экспериментов.....      | 93         |
|          | Выводы по разделу.....                                                                    | 95         |
| <b>5</b> | <b>ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ПЛАНИРОВАНИЯ.....</b>                 | <b>96</b>  |
| 5.1      | Архитектура интеллектуальной системы планирования.....                                    | 96         |
| 5.2      | Реализация ключевых программных модулей.....                                              | 101        |
| 5.3      | Интеграция прогнозирования, оптимизации и аналитики.....                                  | 104        |
| 5.4      | Экспериментальная апробация и воспроизводимость.....                                      | 110        |
|          | Выводы по разделу.....                                                                    | 113        |
|          | <b>ЗАКЛЮЧЕНИЕ.....</b>                                                                    | <b>115</b> |
|          | <b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....</b>                                              | <b>118</b> |
|          | <b>ПРИЛОЖЕНИЕ А – Свидетельства об авторском праве.....</b>                               | <b>133</b> |
|          | <b>ПРИЛОЖЕНИЕ Б – Акт внедрения.....</b>                                                  | <b>135</b> |

## НОРМАТИВНЫЕ ССЫЛКИ

В настоящей диссертации использованы ссылки на следующие стандарты:  
Инструкция по оформлению диссертации и автореферата, ВАК МОН РК, №377-3ж.

ГОСТ 7.32-2017. Система стандартов по информации, библиотечному и издательскому делу. Отчет о научно-исследовательской работе. Структура и правила оформления.

СТ РК 34.005-2002. Информационная технология. Основные термины и определения.

СТ РК 34.006-2002. Информационная технология. Базы данных. Основные термины и определения.

СТ РК 34.014-2002. Информационная технология. Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Термины и определения.

СТ РК 34.019-2005. (ISO/IEC 12207:1995, MOD) Информационная технология. Процессы жизненного цикла программных средств.

Указ Президента Республики Казахстан. О Государственной программе «Информационный Казахстан-2020» и внесении дополнения в Указ Президента Республики Казахстан от 19 марта 2010 года, №957 «Об утверждении Перечня государственных программ»: утв. 8 января 2013 года, №464.

Закон Республики Казахстан. Об информатизации: принят 24 ноября 2015 года, №418.

Постановление Республики Казахстан. Об утверждении Государственной программы «Цифровой Казахстан»: утв. 12 декабря 2017 года, №827.

Закон Республики Казахстан. О науке: принят 18 февраля 2011 года, №407-IVЗРК.

ГОСТ 7.1-2003. Библиографическая запись. Библиографическое описание. Общие требования и правила составления.

## ОПРЕДЕЛЕНИЯ

В настоящей диссертации применяют следующие термины с соответствующими определениями:

**Ансамблевый метод обучения** – метод машинного обучения, объединяющий несколько моделей для повышения точности прогнозирования.

**Гибридная прогнозно-оптимизационная система планирования** – система календарного планирования, объединяющая прогнозные модели, данные предиктивного обслуживания и оптимизационные методы.

**Задача календарного планирования типа Job Shop Scheduling Problem (JSSP)** – задача распределения операций работ по оборудованию при заданных технологических маршрутах и ограничениях ресурсов.

**Интерпретируемость модели** – возможность объяснения влияния входных параметров на результат прогнозирования.

**Календарно-производственное планирование** – процесс формирования расписания выполнения производственных операций с учетом технологических ограничений и доступности ресурсов.

**Квантильный прогноз** – прогноз значения случайной величины, соответствующего заданной вероятности распределения.

**Машинное обучение** – методы построения моделей на основе анализа данных и выявления закономерностей.

**Нейронная сеть LSTM** – рекуррентная нейронная сеть с механизмом долговременной памяти для обработки временных последовательностей.

**Неопределенность производственных параметров** – вариабельность длительностей операций, состояния оборудования и условий выполнения работ.

**Ограниченное программирование** – метод решения задач оптимизации, основанный на описании задачи системой ограничений.

**Оптимизационная модель планирования** – математическая модель формирования расписания на основе ограничений и целевой функции.

**Остаточный ресурс (Remaining Useful Life, RUL)** – прогнозируемое время функционирования оборудования до отказа.

**Предиктивное обслуживание (Predictive Maintenance, PdM)** – подход к техническому обслуживанию оборудования, основанный на прогнозировании его технического состояния.

**Прогнозная модель** – вычислительная модель, предназначенная для оценки будущих значений параметров процесса.

**Расписание производства** – упорядоченное во времени распределение операций по оборудованию.

**Робастное планирование** – метод планирования, ориентированный на получение решений, устойчивых к неопределенности параметров.

**Сценарий планирования** – вариант оптимизационной постановки задачи, отражающий определенные предположения о параметрах системы и уровне неопределенности.

**Сценарные ограничения** – ограничения оптимизационной модели, сформированные на основе альтернативных значений параметров системы.

**CP-SAT-решатель** – программный решатель задач дискретной оптимизации, использующий методы ограниченного программирования и булевой оптимизации.

**Makespan** – момент завершения выполнения всех операций расписания.

**SHAP-оценки** – метод количественной оценки вклада признаков в результат модели машинного обучения.

**XGBoost** – алгоритм градиентного бустинга на решающих деревьях.

## ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

|         |                                                      |
|---------|------------------------------------------------------|
| AI      | – Artificial Intelligence                            |
| CNN     | – Convolutional Neural Network                       |
| CP      | – Constraint Programming                             |
| CP-SAT  | – Constraint Programming with Boolean Satisfiability |
| DL      | – Deep Learning                                      |
| IoT     | – Internet of Things                                 |
| IPS     | – Integrated Planning and Scheduling                 |
| JSSP    | – Job-Shop Scheduling Problem                        |
| LSTM    | – Long Short-Term Memory                             |
| ML      | – Machine Learning                                   |
| NP      | – Non-Polynomial complexity class                    |
| PdM     | – Predictive Maintenance                             |
| RUL     | – Remaining Useful Life                              |
| SHAP    | – SHapley Additive exPlanations                      |
| STLF    | – Short-Term Load Forecasting                        |
| XGBoost | – eXtreme Gradient Boosting                          |
| ТОиР    | – техническое обслуживание и ремонт.                 |

## ВВЕДЕНИЕ

Современный этап развития промышленности характеризуется активным внедрением цифровых технологий, методов анализа больших данных и интеллектуальных систем поддержки принятия решений [1, 2]. Концепции *Industry 4.0*, цифровых производственных платформ и предиктивной аналитики формируют новые требования к системам управления производством, ориентированным на использование данных мониторинга оборудования, интеллектуальных методов анализа и адаптивных механизмов планирования [3, 4]. В этих условиях производственное планирование рассматривается не как изолированная оптимизационная задача, а как элемент интегрированной информационно-аналитической среды, включающей мониторинг технологических процессов, диагностику оборудования, прогнозирование параметров производства и адаптивное управление ресурсами [5].

В Республике Казахстан вопросы цифровизации промышленности и внедрения интеллектуальных производственных систем рассматриваются как приоритетное направление развития экономики, отраженное в государственных программах цифровой трансформации и индустриально-инновационного развития [6]. В этих условиях особое значение приобретает разработка методов интеллектуального планирования производственных процессов, ориентированных на использование данных мониторинга оборудования и производственных систем.

Эволюция цифровых производственных систем характеризуется переходом от традиционных систем управления производством к интеллектуальным сервисам поддержки управленческих решений, основанным на анализе больших данных и прогнозной аналитике. В таких системах календарное планирование интегрируется в замкнутый управленческий цикл, объединяющий этапы мониторинга, предиктивной оценки технологических параметров и динамической корректировки графиков [7].

Индустриальная цифровая трансформация приводит к формированию киберфизических производственных систем и увеличению объема доступных данных о режимах работы оборудования, текущей операционной динамике и загрузке производственных ресурсов [8, 9]. Обозначенные факторы формируют фундамент для замещения статических алгоритмов планирования прогнозно-ориентированными методологиями и определяют направление настоящего диссертационного исследования [10, 11].

Одной из ключевых задач управления производственными системами является построение эффективных производственных расписаний, обеспечивающих рациональное использование ресурсов и соблюдение технологических ограничений [12]. Классической математической моделью таких задач является задача календарного планирования типа *Job Shop Scheduling Problem (JSSP)*, широко применяемая для описания процессов обработки заказов на многооперационном оборудовании. Несмотря на значительный прогресс в области комбинаторной оптимизации и теории расписаний, большинство существующих моделей календарного планирования



опирается на детерминированные значения длительностей операций и предполагает относительную стабильность производственной среды [13, 14].

Однако в реальных производственных системах длительности операций, состояние оборудования и условия выполнения работ подвержены неопределенности [15, 16]. Вариабельность технологических параметров, износ оборудования и вероятность отказов приводят к отклонениям от запланированных сроков выполнения операций, что снижает устойчивость расписаний и требует разработки методов планирования, способных учитывать подобные факторы. В связи с этим возрастает интерес к робастным и вероятностным подходам к планированию, ориентированным на обеспечение устойчивости решений в условиях неопределенности [17].

Параллельно с развитием методов оптимизации активно развиваются методы машинного обучения и предиктивной аналитики, позволяющие прогнозировать длительности операций, состояние оборудования и вероятность его деградации [18]. Особое значение в производственных системах приобретает направление *Predictive Maintenance (PdM)*, связанное с прогнозированием остаточного ресурса оборудования на основе сенсорных данных. Несмотря на значительное количество исследований в области машинного обучения и *PdM*, прогнозные модели чаще всего используются изолированно от задач календарного планирования, а их формализованная интеграция в оптимизационный контур остается недостаточно проработанной [19].

Существующие подходы к применению машинного обучения в задачах планирования, как правило, ориентированы либо на замену оптимизационных методов эвристическими или обучаемыми моделями, либо на ускорение поиска решений [20]. При этом вопрос формализованной интеграции прогнозных оценок неопределенности в оптимизационные модели календарного планирования остается недостаточно проработанным [21]. Особенно актуальной является задача использования интервальных и квантильных прогнозов длительностей операций, позволяющих количественно учитывать неопределенность входных данных и управлять компромиссом между эффективностью и устойчивостью расписаний [22].

Дополнительную сложность представляет необходимость совместного учета неопределенности технологических параметров и рисков деградации оборудования [23]. В большинстве существующих работ задачи производственного планирования и предиктивного обслуживания рассматриваются отдельно, что ограничивает возможности построения интеллектуальных систем управления производственными процессами. Интеграция прогнозных моделей технического состояния оборудования в контур календарного планирования представляет собой важное научное и практическое направление, связанное с развитием интеллектуальных производственных систем [24].

Задачи календарно-производственного планирования, включая задачу типа *JSSP*, на протяжении нескольких десятилетий являются одним из центральных направлений исследований в области комбинаторной оптимизации и теории расписаний [25]. Сформированы фундаментальные подходы к решению

подобных задач, основанные на методах дискретной оптимизации, ограниченного программирования, смешанного целочисленного программирования, а также на различных классах эвристических и метаэвристических алгоритмов. Современные решатели ограничений, включая *CP-SAT*-решатели, демонстрируют высокую эффективность при решении задач планирования средней и большой размерности и широко применяются как в научных исследованиях, так и в инженерной практике.

Вместе с тем, большинство классических моделей календарного планирования рассматривает длительности операций как детерминированные параметры, что ограничивает применимость таких моделей в условиях реального производства. В ответ на это сформировалось направление робастного и стохастического планирования, ориентированное на учет неопределенности параметров производственных процессов [26]. В рамках этих исследований разрабатываются методы интервального планирования, сценарного анализа, вероятностных ограничений и робастной оптимизации. Однако подобные подходы часто требуют существенного усложнения модели и увеличения вычислительных затрат [27].

Несмотря на активное развитие робастных и стохастических методов планирования, их практическое применение часто ограничивается высокой вычислительной сложностью и трудностями параметризации моделей неопределенности. В реальных производственных системах параметры неопределенности редко известны заранее и могут изменяться во времени [28]. В этой связи особый интерес представляют подходы, в которых источником информации о неопределенности выступают прогнозные модели машинного обучения, формируемые на основе исторических и сенсорных данных [29].

Использование методов интеллектуального анализа данных позволяет перейти от априорных предположений о неопределенности к ее количественной оценке на основе наблюдаемых данных производственного процесса. Это открывает возможность формализованной интеграции прогнозных оценок в оптимизационные модели календарного планирования [30].

В последние годы активно развивается направление применения машинного обучения в задачах управления производственными системами. Методы регрессии, нейронных сетей и ансамблевых алгоритмов используются для прогнозирования длительностей операций, загрузки оборудования и параметров технологических процессов [31]. В последние годы особое внимание уделяется гибридным подходам, в которых методы машинного обучения используются совместно с оптимизационными алгоритмами. Тем не менее, в значительной части работ машинное обучение рассматривается либо как инструмент аппроксимации оптимизационных решений, либо как средство ускорения поиска, тогда как задачи формальной интеграции прогнозной информации в оптимизационную модель остаются недостаточно исследованными [32].

Отдельное направление исследований связано с предиктивным обслуживанием оборудования, основанным на анализе сенсорных данных и прогнозировании остаточного ресурса оборудования. В рамках данного

направления разработаны методы оценки технического состояния и вероятности отказов, включая модели прогнозирования *Remaining Useful Life (RUL)* [33]. Несмотря на значительное количество работ в области *PdM*, вопросы интеграции результатов диагностики оборудования в задачи производственного планирования остаются ограниченно проработанными и чаще рассматриваются на концептуальном уровне [34].

Таким образом, анализ современного состояния исследований позволяет выделить три относительно самостоятельных направления: методы календарного планирования, методы машинного обучения для прогнозирования параметров производственных процессов и методы предиктивного обслуживания оборудования. Их интеграция в рамках единой вычислительной архитектуры интеллектуального планирования остается недостаточно разработанной и представляет собой актуальную научную задачу.

Совокупное развитие методов календарного планирования, машинного обучения и предиктивного обслуживания оборудования формирует предпосылки для создания гибридных прогнозно-оптимизационных систем управления производством. Однако в существующих исследованиях указанные направления чаще всего развиваются независимо друг от друга, что ограничивает возможности построения интеллектуальных систем планирования, способных учитывать неопределенность производственных процессов и риски деградации оборудования в единой вычислительной архитектуре.

На предприятиях обрабатывающей промышленности Республики Казахстан, как и в мировой практике, сохраняется проблема износа оборудования. Это делает модернизацию производственных систем необходимой. Традиционные подходы к планированию не справляются с реальной нестабильностью производства – износом оборудования, непредсказуемостью операций, постоянно меняющимися условиями [35]. Чтобы решить эту проблему, в диссертации разрабатываются методы и модели интеллектуального планирования.

**Актуальность темы исследования.** Сегодня производство требует совсем другого подхода к составлению расписания. Мы предлагаем решение, которое строится на трех основных вещах: прогнозах с помощью машинного обучения, оценке неопределенности длительности операций и учете рисков деградации оборудования. Вместе это дает расписания, которые не разваливаются при первом же сбое, остаются вполне реалистичными по вычислениям и могут гибко подстраиваться под изменения прямо во время работы.

Тема хорошо ложится в общую картину цифровой трансформации промышленности – там, где пересекается цифровое производство, интеллектуальное управление системами и работа с большими данными. Разработка гибридных прогнозно-оптимизационных систем планирования представляет собой важный шаг к созданию интеллектуальных систем поддержки принятия решений в производственных и инженерных приложениях.

**Научная новизна:** состоит в разработке и обосновании гибридного прогнозно-оптимизационного подхода к календарному планированию

производственной деятельности, основанного на интеграции моделей машинного обучения и методов оптимизации на ограничениях. Разработана методика формирования сценарных ограничений оптимизационной модели на основе квантильных прогнозов длительности операций для учета неопределенности параметров производственных процессов. Разработан способ интеграции оценок технического состояния оборудования, полученных методами предиктивного обслуживания, в контур календарного планирования. В рамках предложенного подхода реализован единый вычислительный контур «прогнозирование – синтез ограничений – оптимизация – анализ устойчивости», обеспечивающий воспроизводимость вычислительных экспериментов и интерпретируемость плановых решений.

**Цель исследования:** разработка и апробация комплекса моделей интеллектуального анализа больших данных для прогнозного планирования производственной деятельности, направленного на повышение обоснованности и эффективности управленческих решений в условиях неопределенности параметров производственных процессов.

**Задачи исследования:**

- проанализировать современные методы интеллектуального анализа данных и машинного обучения, применяемые в задачах прогнозирования и календарного планирования в производственных системах;
- разработать концептуальную модель интеграции разнородных данных производственного и эксплуатационного характера, а также внешних факторов, рассматриваемых как расширяемые источники информации в задачах прогнозного планирования;
- построить гибридные модели прогнозирования с формализованной процедурой оценки неопределенности прогнозов;
- разработать оптимизационную модель формирования производственного плана с учетом сценарных ограничений по рискам и квантильным прогнозным оценкам параметров операций;
- интегрировать прогнозные модели и данные предиктивного обслуживания в единый контур планирования;
- провести экспериментальную апробацию методики на открытых наборах данных с сравнением результатов с базовыми методами по метрикам точности прогнозирования, устойчивости решений и вычислительной эффективности.

**Объект исследования:** процессы календарного планирования производственной деятельности в условиях неопределенности параметров выполнения операций и состояния оборудования.

**Предмет исследования:** гибридные прогнозно-оптимизационные модели календарного планирования производственной деятельности, основанные на интеграции прогнозных моделей машинного обучения и оптимизационных методов посредством формирования сценарных ограничений, учитывающих неопределенность параметров производственных процессов.

**Методы исследований:** в диссертационной работе на различных этапах использовались методы системного анализа и математического моделирования сложных производственных систем, методы машинного обучения и

интеллектуального анализа данных, методы оптимизационного моделирования и комбинаторной оптимизации, а также методы математической статистики для оценки точности прогнозирования, робастности получаемых решений и верификации результатов вычислительных экспериментов.

**Теоретическая значимость** заключается в развитии методов интеграции моделей машинного обучения и оптимизационных процедур календарного планирования производственных процессов в условиях неопределенности. Предложена формализованная схема интеграции прогнозных моделей и оценок технического состояния оборудования в оптимизационную постановку задачи планирования, основанная на использовании квантильных оценок параметров операций и расширяющая методологию робастного планирования производственных процессов.

**Практическая значимость** заключается в разработке программного прототипа интеллектуальной системы календарного планирования, интегрирующей методы машинного обучения, предиктивного обслуживания оборудования и оптимизационного моделирования. Предложенные методы гибридного прогнозно-оптимизационного планирования могут быть использованы при создании и модернизации систем поддержки принятия решений в области календарного планирования производственных процессов, а также в образовательной практике при подготовке специалистов в области анализа данных и интеллектуальных информационных систем.

**Социальная значимость** заключается в возможности использования разработанной интеллектуальной системы планирования при создании систем поддержки принятия решений в сфере промышленного производства и цифрового управления технологическими процессами. Предложенный подход к интеграции прогнозных моделей и оптимизационных методов может применяться инженерами и специалистами по производственному планированию, способствуя повышению устойчивости производственных процессов, эффективности использования оборудования и внедрению технологий интеллектуального анализа данных в практику цифровизации промышленности.

**Положения, выносимые на защиту:**

- модель гибридного прогнозно-оптимизационного календарного планирования, основанная на использовании квантильных прогнозов параметров производственных процессов и обеспечивающая интеграцию результатов прогнозных моделей машинного обучения в оптимизационные модели календарного планирования для учета неопределенности длительностей операций;

- математическая модель календарного планирования с учетом рисков технического состояния оборудования, основанная на использовании сценарных ограничений, формируемых по результатам анализа показателей предиктивного обслуживания оборудования;

- алгоритм интеграции прогнозных моделей машинного обучения и оптимизационного планирования, реализующий вычислительный контур «прогнозирование – синтез ограничений – оптимизация – анализ устойчивости»

и обеспечивающий повышение устойчивости формируемых производственных расписаний.

**Апробация работы.** Основные результаты диссертационного исследования представлены на международных научных конференциях, опубликованы в научных изданиях, а разработанные программные продукты зарегистрированы в государственном реестре объектов, охраняемых авторским правом. По теме диссертации опубликовано 9 научных работ, в том числе 6 статей в журналах, рекомендованных Комитетом по обеспечению качества в сфере науки и высшего образования МОН РК, 1 статья в международном научном издании, индексируемом в базе данных Scopus, а также получено 2 свидетельства об авторском праве (Приложение А).

**Объем и структура диссертационной работы.** Диссертация состоит из введения, пяти разделов, заключения, списка использованной литературы и приложений. Объем диссертационной работы составляет 132 страницы (38 рисунков и 17 таблиц).

Диссертационная работа выполнена в рамках проекта программно-целевого финансирования по теме BR24992785 «Организация и проведение комплексных исследований по обеспечению устойчивого развития агропромышленного комплекса Костанайской области с созданием научно-исследовательского технологического центра», реализуемого в соответствии с приоритетными направлениями научно-технологического развития Республики Казахстан.

# 1 АНАЛИТИЧЕСКИЙ ОБЗОР МЕТОДОВ ПРОГНОЗИРОВАНИЯ И ПЛАНИРОВАНИЯ ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

## 1.1 Неопределенность в производственном планировании: постановки и источники вариативности

Производственное планирование и его календарная составляющая составляют центральную часть исследований в области оптимизации операций. Одной из наиболее распространенных формальных моделей является задача планирования на предприятии (*job shop*), где множество операций различных работ должно быть упорядочено на ограниченном множестве ресурсов с целью минимизации целевых критериев, таких как общая длительность выполнения всех работ (*makespan*) или задержка завершения заданий [36]. В классической формулировке задача планирования последовательностей работ (*JSSP*) предполагает фиксированные длительности операций и последовательность их обработки на машинах с жесткими технологическими ограничениями, что делает ее задачей комбинаторной оптимизации класса *NP*-трудных задач [38, 40]. Аналогично, ее частные случаи – *flow shop* и *flexible job shop* – рассматриваются в литературе как варианты, в которых последовательность операций на машинах либо фиксирована для всех работ, либо допускает множественность альтернативных маршрутов [39].

Кроме календарного планирования, в операционной практике выделяют стратегическое, тактическое и оперативное планирование. В свою очередь, оперативное планирование ориентировано на краткосрочные задачи формирования расписаний при изменяющихся входных данных, в то время как тактическое планирование формирует более общие планы загрузки ресурсов на среднесрочные горизонты [37, р. 3-345; 38, р. 3-216]. Отдельным направлением является диспетчеризация, рассматривающая детерминированную или динамическую перестановку уже сформированных планов в реальном времени.

Типичные модели календарного планирования, рассматриваемые в литературе, основаны на предположении, что параметры системы известны заранее и не изменяются в ходе исполнения плана. Это представление удобно для математических формулировок, поскольку позволяет свести планирование к решению задачи оптимизации с фиксированными параметрами и жесткими ограничениями, например, в виде смешано-целочисленных линейных программ или графовых моделей с дисъюнктивными графами [39, р. 1239-1262; 41].

Однако в реальных производственных системах параметры процессов подвержены вариативности, что существенно ограничивает применимость детерминированных моделей. Игнорирование изменчивости длительностей операций и простоев оборудования может приводить к формированию расписаний, оптимальных в математическом смысле, но неустойчивых при практической реализации [42, 43].

Одной из основных причин является вариативность длительностей операций из-за отличий в материалах, квалификации операторов, условиях эксплуатации оборудования и переналадке между заданиями. При этом

распределение длительностей часто моделируется как случайная величина с неизвестной или сложно поддающейся априорному описанию вероятностной функцией, что делает фиксированные значения недостаточными для адекватного описания реального производственного процесса [38, p. 3-216; 44, 45].

Не менее значимой является неоднородная доступность оборудования, обусловленная как плановыми действиями по техническому обслуживанию (*preventive maintenance*), так и внезапными отказами, и ремонтом (*corrective maintenance*). В литературе по устойчивому планированию рассматриваются расширенные формулировки задач, включающие в себя как регулярные, так и стохастические простои машин, что приводит к моделям робастного или стохастического планирования [46].

Совокупность перечисленных источников неопределенности обуславливает необходимость разработки методов планирования, способных учитывать вариативность параметров и обеспечивать устойчивость расписаний к изменениям среды исполнения задачи.

Наличие неопределенности в параметрах планирования может существенно влиять на качество полученных расписаний, снижая эффективность и. Так, при игнорировании изменчивости длительностей операций и внезапных простоев машин реальные планы могут приводить к заметным отклонениям от целевых показателей, таким как суммарное время выполнения всех работ (*makespan*), задержки завершения операций (*tardiness*), а также к увеличению числа случаев пересмотра планов (*rescheduling*) в ходе исполнения. В литературе по оптимизации с неопределенностью отмечается, что формулы, основанные на детерминированных прогнозах, зачастую дают слишком оптимистичные оценки ожидаемых показателей, тогда как учет вариативности позволяет формировать более устойчивые расписания.

Для учета неопределенности предложен ряд направлений развития моделей календарного планирования. К ним относятся стохастическое планирование, в котором неопределенные параметры рассматриваются в виде вероятностных распределений для получения решений, оптимальных по ожидаемому значению целевой функции; робастное планирование, призванное находить решения, устойчивые к вариативности параметров; а также сценарные и гибридные подходы, основанные на анализе возможных реализаций параметров.

В последнее время также наблюдается рост интереса к обучаемым методам, которые могут использовать данные о прошлых реализациях производства для прогнозирования будущих неопределенностей и использования этих прогнозов внутри оптимизационных моделей [45, p. 1025-1043; 47].

Эволюция подходов к календарному планированию отражает постепенный переход от детерминированных моделей с фиксированными параметрами к использованию данных, прогнозов и риск-ориентированных формализаций. Однако, как показано на рисунке 1, даже современные стохастические и робастные модели не обеспечивают системной интеграции прогнозирования и



оптимизации, что ограничивает их практическую применимость в сложных производственных системах [43, р. 3-284; 45, р. 1025-1043].

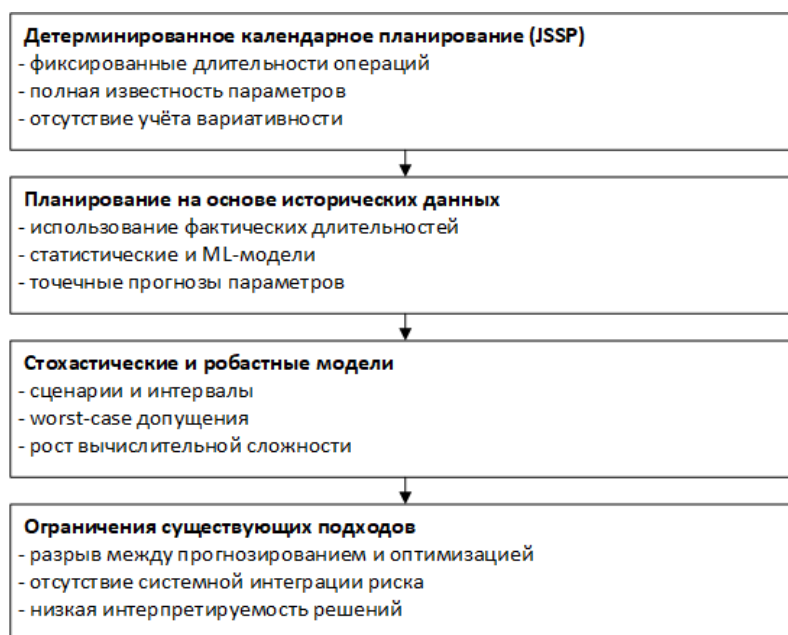


Рисунок 1 – Эволюция подходов к календарному планированию производственных процессов

Таким образом, существующие модели календарного планирования формируют теоретическую основу оптимизации, однако их практическая применимость в условиях неопределенности остается ограниченной и требует развития формализованных методов учета вариативности параметров и интеграции неопределенности в оптимизационные модели. Для производственных систем с дискретной структурой операций и высокой вариативностью режимов работы оборудования (например, машиностроительные предприятия) данные ограничения приобретают принципиальный характер.

## 1.2 Производственные данные и методы машинного обучения для прогнозирования параметров

В современных производственных системах прогнозирование и принятие управленческих решений опираются не только на экспертные оценки, но и на обработку больших объемов данных, генерируемых в ходе функционирования предприятий. Источниками таких данных выступают системы автоматизации производственных процессов (*MES – Manufacturing Execution Systems*, *ERP – Enterprise Resource Planning*, *SCADA – Supervisory Control and Data Acquisition*), а также датчики оборудования и журналирование выполнения операций [48]. Данные из *MES* обычно содержат информацию о фактически выполненных операциях, длительностях, загрузке ресурсов и взаимодействии между операциями; *ERP-системы* предоставляют сведения о заказах, сроках и приоритетах; *SCADA-данные* дают телеметрические параметры оборудования, включая измерения температур, скоростей, вибраций и состояний рабочих

режимов, что имеет ключевое значение в контексте интеграции данных *MES/SCADA* для аналитических приложений [49]. Источники данных в промышленности также включают потоковые данные *IoT*, траектории сенсорных измерений и *ERP-метрики* времени выполнения задач, которые интегрируются в аналитику производительности и оптимизации планирования процессов [50].

Производственные данные обладают рядом характерных особенностей, которые усложняют их анализ [51]. Во-первых, данные часто являются шумными и неполными вследствие технических сбоев, пропусков записей или ошибочного ввода, что требует применения методов предобработки и очистки. Во-вторых, данные могут быть неоднородными – они представлены в виде временных рядов с различной частотой измерений, категориальных признаков (например, тип операции, номер оборудования) и числовых параметров (длительности, загрузки). В-третьих, данные в значительной степени зависят от контекста исполнения, поскольку длительности операций могут изменяться в зависимости от конкретного материала, смены персонала или состояния оборудования, что затрудняет применение классических статистических моделей без их предварительной адаптации к контексту выполнения операций [52]. В результате требуется использование специализированных процедур предобработки данных и отбора признаков, а также адаптивных моделей анализа временных рядов.

В таблице 1 приведены основные источники данных, используемые в современных производственных системах, и их характеристики с точки зрения частоты поступления, типов параметров и применимости в задачах прогнозирования и календарного планирования.

Таблица 1 – Источники данных в современных производственных системах и их использование в задачах прогнозирования и календарного планирования

| Источник данных | Тип данных               | Частота поступления | Примеры параметров                                     | Использование                                       |
|-----------------|--------------------------|---------------------|--------------------------------------------------------|-----------------------------------------------------|
| MES             | события, временные ряды  | средняя             | длительности операций, маршруты, загрузка оборудования | прогноз длительностей операций, анализ выполнения   |
| ERP             | транзакционные данные    | низкая              | заказы, сроки выполнения, приоритеты                   | формирование ограничений и приоритетов планирования |
| SCADA / IoT     | сенсорные временные ряды | высокая             | вибрации, температура, токи, состояния                 | оценка технического состояния, PdM                  |

В исследованиях, посвященных анализу временных рядов в производственных системах, отмечается важность использования качественных наборов данных для прогнозирования будущих событий, таких как длительности операций и простои оборудования [53]. Показано, что точность прогнозов в промышленных приложениях существенно зависит от предобработки данных,

выбора признаков и применения адаптивных алгоритмов, способных учитывать сложные паттерны в многомерных временных рядах [54].

Стоит отметить, что данные о состоянии оборудования, включая диагностические признаки и телеметрию, могут быть использованы не только для прогноза длительностей операций, но и для предварительной оценки риска отказов оборудования, что создает предпосылки для интеграции прогнозирования и задач предиктивного обслуживания. Их интеграция в прогнозные модели позволяет более полно описывать характеристики выполнения операций и предсказывать поведение системы в будущем, улучшая устойчивость процессов и качество планирования [55, 56].

Таким образом, в задачах промышленного прогнозирования данные представляют собой сложные, разнородные временные ряды, требующие специализированных подходов обработки и анализа. Это создает условия для применения современных методов машинного обучения, которые способны выявлять скрытые зависимости и сезонные эффекты в многомерных производственных данных [57-60].

Для прогнозирования производственных параметров чаще всего используются регрессионные модели. Они предназначены для оценки числовых величин. Например, длительностей операций, времени простоя оборудования или показателей загрузки ресурсов. К традиционным методам можно отнести линейную и полиномиальную регрессию, а также регуляризованные модели *Ridge* и *Lasso*. При этом в реальных производственных условиях их применение ограничено сложной структурой зависимостей и выраженной нелинейностью данных [61].

Для повышения точности прогнозирования применяют ансамблевые методы машинного обучения *Random Forest* и градиентный бустинг над решающими деревьями (*XGBoost*, *LightGBM*, *CatBoost*). Эти методы работают с неоднородными и шумными данными, могут учитывать нелинейные зависимости и взаимодействия признаков, а также являются устойчивыми к выбросам. Результаты исследований показывают, что методы градиентного бустинга повышают точность прогнозирования длительностей операций и производственных показателей в условиях реального производства [62-64].

Широкое распространение получили рекуррентные нейронные сети (*RNN*) и их модификации, такие как *LSTM* и *GRU*. Они позволяют учитывать долгосрочные временные зависимости и также применяются для прогнозирования длительностей операций, риска отказа оборудования и динамики загрузки ресурсов [65]. Помимо этого, для обработки многомерных временных рядов в производственных системах рассматриваются архитектуры *CNN* и *Transformer*. Их использование позволяет учитывать сложные временные и межпризнаковые зависимости при прогнозировании длительностей операций и состояний оборудования [66].

Несмотря на высокую точность прогнозов, получаемых с использованием перечисленных методов, большинство моделей машинного обучения ориентировано на формирование точечных оценок прогнозируемых параметров, что ограничивает их применимость в задачах производственного планирования,

особенно в условиях вариативности и неопределенности параметров процессов [67].

Как правило, такие модели обучаются путем минимизации среднеквадратичной или абсолютной ошибки, что приводит к получению прогнозов, соответствующих ожидаемому или среднему значению случайной величины [68]. Однако идентичные точечные прогнозы могут соответствовать принципиально различным распределениям длительностей операций [69], что делает невозможной корректную оценку риска при использовании только средних значений (рисунок 2).

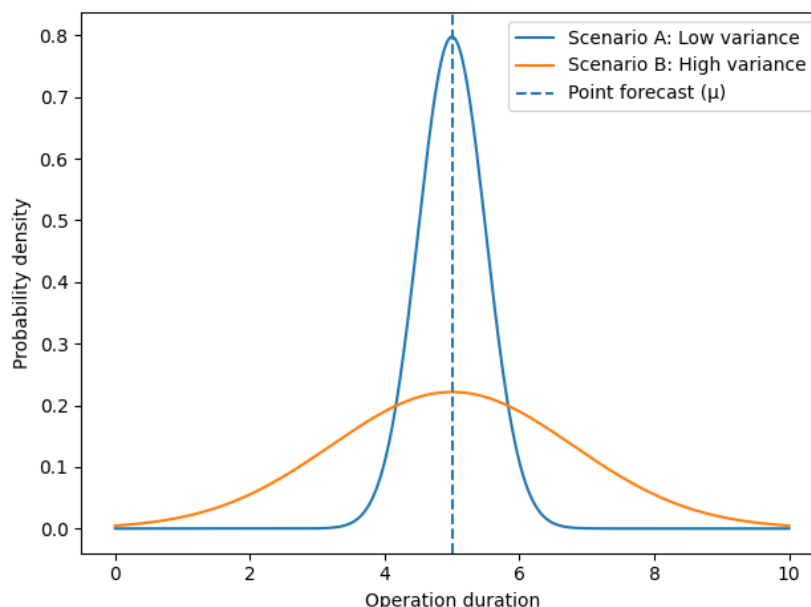


Рисунок 2 – Пример одинакового точечного прогноза при различной вариативности длительностей операций

Основной проблемой точечных прогнозов является их неспособность отражать неопределенность и вариативность производственных процессов: фактические длительности операций, времена переналадки и периоды простоя оборудования подвержены колебаниям, обусловленным внутренними (износ оборудования, человеческий фактор) и внешними факторами (изменение входных материалов, организационные сбои).

Использование единственного прогнозного значения не позволяет учесть вероятность отклонений от ожидаемого сценария и может приводить к систематическому недооцениванию рисков срыва сроков выполнения заданий [70, 71], что в задачах календарного планирования проявляется в формировании расписаний, устойчивых лишь при совпадении фактических параметров с ожидаемыми, и в каскадном накоплении задержек с ростом фактического *makespan* [72, 73].

В литературе по оптимизации уже давно отмечают проблему с точечными прогнозами – это *optimistic bias*. Модель просто выдает самый хороший результат, а все редкие, но очень неприятные отклонения отбрасывает и не смотрит на них. [74]. На производстве это реально выходит боком. Один

серьезный сбой – и потери могут быть очень большими. Поэтому сейчас все чаще переходят на вероятностные и интервальные прогнозы. Они хотя бы показывают реальные риски и дают возможность заложить резервы именно там, где нужно [70, p. 195-200; 75].

Отдельная история – как понять, что модель вообще делает. В производственном планировании одной цифры мало. Хочется видеть, на каких данных она основывалась и в каких случаях ей можно верить [76]. В промышленности решения всегда нужно обосновывать и потом проверять. А когда все переплетено между собой, без нормальной прозрачности модели просто не обойтись [77].

Многие современные алгоритмы машинного обучения, такие как градиентный бустинг или глубокие нейронные сети, обладают высокой предсказательной способностью, но одновременно рассматриваются как модели, поведение которых трудно интерпретировать непосредственно [78]. В то же время в ряде практических сценариев интерпретация преимущественно используется на этапе анализа и верификации моделей, а не как непосредственный компонент оптимизационного алгоритма. Однако в задачах прогнозирования отказов оборудования или длительностей операций недостаток интерпретируемости может затруднить выявление факторов, приведших к повышенному прогнозу риска, и усложнить корректировку плана на основе прогноза [79].

В ответ на эти вызовы активно развиваются методы объяснимого машинного обучения (*Explainable AI, XAI*), направленные на повышение прозрачности моделей. Например, постхок-аналитика, которая оценивает влияние входных признаков на прогнозы модели. Современные исследования рекомендуют использовать *SHAP* (*SHapley Additive exPlanations*) или *LIME* (*Local Interpretable Model-agnostic Explanations*). Указанные методы позволяют получать локальные и глобальные объяснения работы модели [80]. Это повышает доверие к результатам прогнозирования и делает возможным их использование для аналитической оценки на этапе разработки и экспериментальной апробации [81].

Развитие объяснимости также связано с построением интерпретируемых моделей прогнозирования. К ним относятся деревья решений с ограниченной глубиной, обобщенные аддитивные модели (*GAM*) и правила деревьев. В отличие от более сложных моделей, они сами позволяют выявлять понятные зависимости между входными данными и прогнозом без необходимости дополнительной интерпретации. Такие модели применяются в задачах, где особенно важна прозрачность. Например, при прогнозировании отказов оборудования, оценке бизнес-рисков и в других прикладных задачах [82].

Отдельного внимания требует корректное представление результатов прогнозирования в оптимизационных моделях планирования. В данном случае важно не только объяснить, что именно было спрогнозировано, но и показать, как неопределенность прогноза влияет на решение оптимизационной задачи [77]. В задачах календарного планирования это означает, что прогнозы длительностей должны быть представлены таким образом, чтобы

оптимизационный модуль мог учитывать распределение прогнозируемых значений, а не только их точечное значение. Для этого используются такие методы, как представление прогнозов в виде распределений, квантилей или сценариев, что позволяет включить риски и неопределенность в модель планирования [78, р. 82-114].

Таким образом, методы объяснимости выступают практическим механизмом повышения доверия к прогнозным моделям и обеспечивают возможность их корректного использования при формировании ограничений и сценариев в задачах производственного планирования [79, р. 206-214].

Ограничения точечных прогнозов и проблемы интерпретируемости моделей машинного обучения привели к активному развитию подходов, ориентированных на учет неопределенности в прогнозировании. В отличие от традиционных методов, формирующих единственное ожидаемое значение параметра, такие подходы позволяют описывать диапазон возможных значений и вероятность их реализации. Это особенно важно для производственных систем, функционирующих в условиях изменчивости и стохастических воздействий, где фактические параметры процессов могут существенно отклоняться от средних оценок. В литературе описываются методы, которые формируют распределения прогнозов, а не только точечные оценки, и концептуально связывают это с задачами риск-ориентированного принятия решений и последующей оптимизации [83].

Одно из направлений развития связано с вероятностным прогнозированием, при котором результатом работы модели становится оценка распределения прогнозируемой величины. Однако использование полных вероятностных распределений в оптимизационных задачах планирования часто связано с высокой вычислительной сложностью и дополнительными допущениями. В частности, методы на основе ансамблей и агрегирования распределений прогнозов показывают, что объединение прогноза нескольких моделей или нескольких реализаций одной модели может повышать надежность прогнозов, особенно при анализе временных рядов в условиях неопределенности [84].

Квантильные методы прогнозирования представляют собой практически значимый и вычислительно приемлемый компромисс между точечными и полными вероятностными прогнозами. Для получения асимметричной оценки неопределенности и учета стоимости ошибок, связанных с занижением или завышением прогнозируемых параметров, модель обучается предсказывать заданные квантили распределения целевой переменной. Различные вариации квантильного прогнозирования широко применяются в задачах временных рядов, где важно учитывать распределение возможных исходов [85].

Использование квантильных прогнозов позволяет формировать риск-ориентированные оценки параметров производственных процессов. Например, для построения расписаний с низкой вероятностью нарушения сроков могут использоваться верхние квантили длительностей операций. Тогда как для сценариев с более высокой вероятностью нарушения сроков подходят медианные или нижние квантили. Такой подход повышает гибкость управления

и позволяет адаптировать планирование под разные уровни допустимого риска. При этом квантильные оценки используются как ограничивающие значения в оптимизационной задаче [86].

Для повышения точности прогноза и оценки неопределенности также применяются ансамблевые подходы, когда прогнозы нескольких моделей комбинируются для моделирования распределения возможных исходов. Такие ансамбли, в том числе основанные на методах конформного прогнозирования (*conformal prediction*) или агрегировании квантильных прогнозов, позволяют повысить качество прогнозов и применяются преимущественно на этапе аналитической оценки и верификации моделей [87].

Таким образом, переход от точечных прогнозов к методам, учитывающим неопределенность и риск, является естественным этапом развития интеллектуальных систем планирования. Использование вероятностных, квантильных и ансамблевых прогнозов создает методологическую основу для интеграции машинного обучения и оптимизационных моделей, в которых неопределенность прогнозов учитывается через формализованные ограничения и сценарии планирования, и позволяет формировать более устойчивые и интерпретируемые решения в задачах календарного планирования производственных процессов [88].

### **1.3 Представление неопределенности и риск-ориентированные подходы в планировании**

В практических задачах прогнозирования параметров производственных процессов принципиальное значение приобретает не только оценка центрального тренда (среднего значения), но и количественная характеристика неопределенности прогнозов. В отличие от точечных моделей, вероятностные модели формируют распределения прогнозируемых величин, что позволяет учитывать вариативность будущих исходов и формировать количественную основу для интеграции риска в последующие этапы планирования и оптимизации [89].

Такие подходы широко рассматриваются в экономике, статистике и *ML*-литературе как инструменты для более надежного прогнозирования в условиях стохастических воздействий и неопределенных данных [90]. В частности, выбор формы представления прогноза определяет доступные способы формализации риска в оптимизационных моделях [91].

Вероятностное прогнозирование базируется на построении распределения вероятностей будущих значений целевой переменной. В отличие от классических моделей, которые минимизируют *MSE/MAE* для получения центральной оценки, вероятностные методы моделируют функцию плотности распределения исходов, что позволяет строить доверительные интервалы, оценивать вероятность выхода параметра за заданный предел и учитывать асимметрию рисков в различных частях распределения [83, р. 99-122]. Вместе с тем использование полного распределения в задачах оптимизации планирования часто сопровождается ростом вычислительной сложности и требует дополнительных допущений о форме распределений [92].

В случаях, когда построение полного распределения затруднено из-за ограниченности данных, применяются интервальные методы, задающие диапазон возможных значений параметра с заданным уровнем доверия [93]. Такие подходы обеспечивают более осмотрительные решения по сравнению с точечными оценками и используются в условиях повышенных рисков и информационной неопределенности [94].

На рисунке 3 показано концептуальное различие между точечным, интервальным и вероятностным прогнозами.

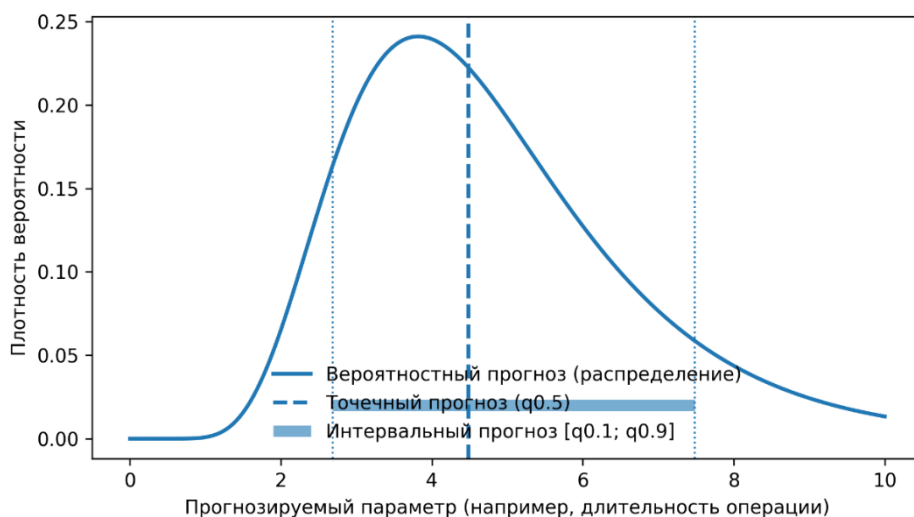


Рисунок 3 – Сопоставление точечного, интервального и вероятностного подходов к прогнозированию параметров производственного процесса

Точечный прогноз представлен единственным значением (например, медианой распределения), интервальный прогноз задает диапазон возможных значений с заданным уровнем покрытия, а вероятностный прогноз описывает полное распределение прогнозируемого параметра [95]. Такое различие принципиально влияет на способы включения неопределенности в задачи календарного планирования. Вероятностные прогнозы позволяют оценивать вероятность превышения критических порогов, а интервальные – задавать допустимые диапазоны значений и использовать их в качестве ограничений модели [96]. Тем самым становится возможным не только минимизировать ожидаемое значение целевой функции (например, *makespan*), но и контролировать вероятность неблагоприятных исходов, что критично в промышленных условиях с жесткими сроками и дорогими простоями [97].

В условиях неопределенности производственных процессов используются квантильные методы прогнозирования для учета характеристик распределения [98]. Квантильная регрессия позволяет оценивать условные квантили целевой переменной вместо ее математического ожидания [99]. Это позволяет моделировать асимметрию, учитывать хвосты и экстремальные значения для производственных систем с выраженной неоднородностью и рисками отклонений.



В разрезе производства квантильная регрессия может дать не только базовый прогноз длительности операции, но также и пессимистический или консервативный сценарии выполнения. Таким образом, модель может формировать несколько альтернативных прогнозов для определения длительности одной и той же операции в зависимости от уровня надежности и допустимого риска [100].

Для обучения квантильных моделей используется специализированная функция потерь, известная как *pinball loss* (квантильная функция потерь). Эта функция обеспечивает статистическую согласованность оценки соответствующего квантиля. В отличие от симметричных функций потерь, основанных на среднеквадратичной или абсолютной ошибке, *pinball loss* является асимметричной и позволяет по-разному штрафовать недооценку и переоценку прогнозируемого значения в зависимости от выбранного уровня квантиля.

В результате модель обучается так, чтобы прогнозируемое значение соответствовало заданному квантилю распределения, а не его среднему. По этой причине квантильные модели особенно востребованы в задачах, в которых последствия недооценки параметров (например, длительности операции) значительно превышают последствия их переоценки с точки зрения устойчивости и выполнимости производственных планов [101].

Для одной операции может формироваться набор прогнозов, соответствующих различным уровням надежности, что позволяет более гибко управлять производственными процессами. Верхние квантили используются при построении консервативных расписаний, нижние – при анализе оптимистичных сценариев, что позволяет варьировать степень консервативности расписания в зависимости от допустимого уровня риска [102, 103]. Особую ценность квантильные прогнозы приобретают в системах, где распределения длительностей являются асимметричными и характеризуются «тяжелыми хвостами» [104], поскольку использование средних оценок в таких условиях может приводить к систематическому недоучету рисков [105].

Несмотря на методологическую зрелость квантильных и распределенных методов, в большинстве существующих исследований они применяются преимущественно на уровне аналитики и мониторинга и редко интегрируются непосредственно в оптимизационные постановки календарного планирования, что формирует разрыв между возможностями прогнозных моделей и их практическим использованием [106].

Наряду с распределенными и квантильными методами в условиях неопределенности активно развиваются сценарные и риск-ориентированные подходы к формированию расписаний [107]. Сценарное моделирование предполагает формирование конечного множества возможных реализаций неопределенных параметров (длительностей операций, доступности оборудования, ресурсов), по которым осуществляется оптимизация либо строятся решения, допустимые для всех сценариев [108]. Данный подход позволяет анализировать чувствительность расписаний к изменениям входных

параметров. При этом качество решений зависит от полноты и репрезентативности набора сценариев [109].

Среди подходов к учету неопределенности рассматривается робастное планирование. Оно направленно на поиск решений, допустимых при наихудших реализациях параметров в пределах заданных множеств неопределенности [110]. Робастные методы обеспечивают высокую надежность решений. Однако их избыточная консервативность может приводить к ухудшению средних показателей эффективности и увеличению *makespan* при широких множествах неопределенности [111].

Более гибкий вариант представляют риск-ориентированные методы оптимизации. Они используют формализованные меры риска *Conditional Value-at-Risk* (CVaR) [112] и вероятностные ограничения (*chance constraints*) [113]. По сути, эти методы ищут компромисс: дают хорошую эффективность, но при этом не позволяют рискам выходить за допустимые рамки. Они контролируют, чтобы вероятность сильных сбоев или больших потерь не была слишком высокой.

Несмотря на то, что эти подходы довольно хорошо проработаны, на практике в реальном производстве они сталкиваются с серьезными ограничениями. Сценарии чаще всего формируют вручную, опираясь на мнение экспертов. Из-за этого страдает масштабируемость и воспроизводимость результатов [114]. Кроме того, прогнозные данные обычно берут в виде жестких фиксированных интервалов или готовых сценариев, при этом они слабо связаны с моделями машинного обучения, которые обучаются на реальных производственных данных. В итоге прогнозирование и оптимизация плохо интегрируются между собой [115].

Сравнительный анализ методов учета неопределенности представлен в таблице 2.

Таблица 2 – Методы учета неопределенности в задачах планирования

| Подход                           | Учитывает неопределенность | Интерпретируемость | Масштабируемость |
|----------------------------------|----------------------------|--------------------|------------------|
| Точечные прогнозы                | -                          | +                  | +                |
| Стохастическое планирование      | +                          | -                  | -                |
| Робастное планирование           | + (worst-case)             | +                  | -                |
| Квантильные прогнозы+ограничения | +                          | +                  | +                |

Из таблицы видно, что квантильные прогнозы вместе с ограничениями хорошо балансируют три вещи: учет неопределенности, понятность модели и нормальную скорость расчетов [116]. Поэтому такой подход вполне перспективный и его можно внедрять в модели календарного планирования.

Проведенный анализ показывает, что в современной литературе представлены вероятностные, интервальные, квантильные, сценарные и риск-ориентированные методы учета неопределенности. Однако отсутствует единая формализованная методологическая связка, обеспечивающая последовательную интеграцию прогнозирования, количественной оценки риска и оптимизации в рамках единого контура интеллектуального планирования. Данный разрыв

формирует основу для разработки интегрированных подходов, объединяющих машинное обучение и методы оптимизации с формализованным учетом неопределенности и риска.

#### **1.4 Интеграция прогнозирования и календарного планирования: существующие подходы и разрыв**

В задачах календарного планирования производственных процессов, формализуемых в виде *JSSP*, длительности операций традиционно рассматриваются как детерминированные параметры, задаваемые априори [117]. Такое предположение упрощает математическую постановку задачи и позволяет применять классические методы оптимизации, однако слабо соответствует условиям реального производства, где фактические длительности операций подвержены значительной вариативности.

С развитием цифровых производственных систем и накоплением больших массивов исторических данных все большее внимание уделяется использованию методов машинного обучения для прогнозирования длительностей операций [118]. В этом случае прогнозные модели выступают источником параметров для задач календарного планирования, подменяя или уточняя фиксированные нормативные значения длительностей операций. Такой подход позволяет учитывать фактическое поведение производственной системы, влияние состояния оборудования, характеристик материалов и организационных факторов.

На практике интеграция чаще всего осуществляется путем прямой подстановки точечных прогнозов в оптимизационную модель *JSSP* [119]. Таким образом, детерминированная структура задачи сохраняется, но параметры формируются на основе данных, а не экспертных оценок. Такой подход позволяет повысить точность исходных параметров, однако сопровождается рядом методологических ограничений [120]. Прогнозные значения, даже обладая высокой средней точностью, не отражают вариативность и асимметрию распределений, а оптимизационная модель воспринимает их как абсолютно точные величины.

Следствием этого становится снижение устойчивости расписаний: даже небольшие расхождения между прогнозируемыми и фактическими длительностями способны вызывать каскадные задержки последующих операций, перераспределение очередей на оборудовании и рост фактического *makespan* по сравнению с плановым значением [121]. Таким образом, подмена нормативных параметров точечными прогнозами не решает задачу формализованного управления риском [122], поскольку информация о вероятности превышения временных порогов и структуре распределений остается вне оптимизационной модели. В результате планирование по-прежнему ориентируется на средний сценарий, игнорируя хвосты распределений и потенциально неблагоприятные реализации параметров.

Указанные ограничения обуславливают необходимость перехода к моделям, способным учитывать неопределенность параметров более формализованным образом [123]. В литературе для этого активно развиваются

стохастические и робастные подходы к решению задач *JSSP*. Стохастические модели предполагают представление длительностей операций и других параметров в виде случайных величин с заданными распределениями вероятностей, а целевая функция формулируется с учетом математического ожидания, дисперсии или вероятностных ограничений (*chance constraints*) [124]. Это позволяет анализировать ожидаемое поведение системы, вероятность нарушения сроков и чувствительность решений к изменениям параметров [125].

Однако учет распределений и вероятностей существенно увеличивает размерность задачи и усложняет алгоритмы решения, вплоть до необходимости многосценарной или вложенной оптимизации [126]. В условиях *NP*-трудности *JSSP* это приводит к экспоненциальному росту вычислительной сложности и ограничивает масштабируемость стохастических методов на реальные производственные системы с десятками и сотнями операций [127]. Поэтому такие модели чаще применяются к задачам малой размерности или в виде приближенных схем.

Альтернативным направлением является робастное планирование, ориентированное на получение решений, допустимых для всех реализаций параметров в пределах заранее заданных множеств неопределенности [128]. Робастные модели не требуют явного задания распределений вероятностей и опираются на интервальные или многогранные описания неопределенности. Решения, сформированные в рамках *worst-case* подходов, обеспечивают выполнение ограничений даже при наиболее неблагоприятных сочетаниях параметров [129], что делает их привлекательными для производств с высокой стоимостью простоев.

Вместе с тем чрезмерная консервативность робастных моделей может приводить к ухудшению средних показателей эффективности, увеличению *makespan* и недоиспользованию оборудования [130]. Дополнительной проблемой является экспертный характер выбора множеств неопределенности, что снижает воспроизводимость и адаптивность моделей при изменении производственных условий.

Стохастические и робастные подходы позволяют учитывать неопределенность в формализованном виде. Однако их применение связано с рядом ограничений. В одном случае растет вычислительная сложность. В другом – появляется избыточная консервативность и необходимость ручной настройки параметров неопределенности [131, 132]. В результате их практическое использование в промышленных системах остается ограниченным.

Методы прогнозирования и риск-ориентированного планирования активно развиваются. Однако их интеграция лишь частично [133]. Прогнозные модели машинного обучения используются для оценки длительностей операций или состояния оборудования. Затем полученные результаты передаются в оптимизационные процедуры в виде фиксированных параметров без учета распределения неопределенности [134]. Таким образом информация о вариативности и рисках существенно упрощается на этапе формирования расписаний.

Сценарные и робастные модели также имеют свои ограничения. Во многих случаях они требуют ручного формирования сценариев или множеств неопределенности [135]. Также сценарии и множества обычно не обновляются автоматически на основе новых данных и прогнозов. Это ограничивает применение таких методов в динамических производственных средах и препятствует формированию единого контура «данные – прогнозирование – оптимизация».

Дополнительным ограничением является низкая интерпретируемость сложных стохастических и робастных оптимизационных моделей [136]. Пользователь не всегда может определить, какие факторы оказали решающее влияние на формирование расписания, что снижает доверие к результатам оптимизации и ограничивает их внедрение в производственной практике.

Кроме того, лицо, принимающее решения (ЛПР), в большинстве существующих подходов не включено в формализованный контур планирования. Настройка уровня риска и выбор сценариев осуществляются вне оптимизационной модели, что усиливает разрыв между математической постановкой и управленческой практикой.

На рисунке 4 схематически показан разрыв между этапами прогнозирования параметров производственных процессов и календарного планирования в большинстве существующих подходов.

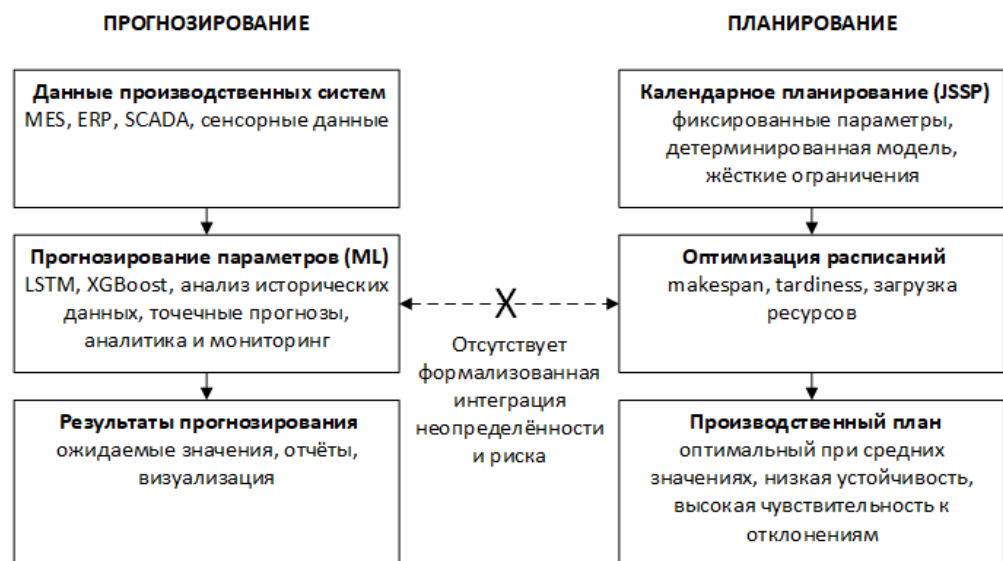


Рисунок 4 – Разрыв между прогнозированием и календарным планированием в существующих подходах

Прогнозные модели, в том числе основанные на методах машинного обучения, используются преимущественно на уровне аналитики и мониторинга и не оказывают системного влияния на формирование расписаний. Оптимизационные модели, в свою очередь, опираются на фиксированные параметры и не используют информацию о неопределенности и рисках, содержащуюся в прогнозах. Отсутствие формализованной связки между прогнозированием, оценкой риска и оптимизацией приводит к снижению

устойчивости планов и ограничивает практическую применимость таких решений.

Таким образом, существующие подходы характеризуются разрозненностью методов, отсутствием формализованной связки между прогнозированием и оптимизацией, недостаточной автоматизацией сценарной логики, ограниченной интерпретируемостью решений и отсутствием формализованной роли ЛПР в контуре планирования. Совокупность указанных факторов свидетельствует о наличии научно-методического разрыва между развитием методов машинного обучения и их системной интеграцией в механизмы календарного планирования. Это формирует научную и практическую мотивацию для разработки интегрированных интеллектуальных систем, объединяющих прогнозирование, количественную оценку риска, оптимизацию и поддержку принятия решений в рамках единого вычислительного контура.

### **Выводы по разделу**

1. Проведен аналитический обзор современных подходов к прогнозированию параметров производственных процессов и их использованию в задачах календарного планирования. Рассмотрены источники производственных данных, методы машинного обучения для прогнозирования длительностей операций, а также вероятностные, квантильные, сценарные, стохастические, робастные и риск-ориентированные подходы к учету неопределенности.

2. Установлено, что большинство существующих методов прогнозирования ориентировано на получение точечных оценок параметров производственных процессов. Несмотря на повышение точности по сравнению с традиционными подходами, такие модели не отражают вариативность и асимметрию распределений, что ограничивает их применимость в задачах календарного планирования и приводит к формированию расписаний, чувствительных к отклонениям фактических условий исполнения, включая каскадный рост задержек и увеличение фактического *makespan*.

3. Показано, что в литературе представлены вероятностные, квантильные, сценарные и робастные методы учета неопределенности. Однако прогнозирование, оценка риска и оптимизация, как правило, реализуются разрозненно, без единого методологического контура. Сценарные и риск-ориентированные модели часто требуют ручной настройки, обладают ограниченной масштабируемостью и не обеспечивают автоматизированной адаптации к изменяющимся данным и прогнозам машинного обучения.

4. Отмечено, что существующие подходы к календарному планированию не обеспечивают достаточной интерпретируемости решений и не предусматривают формализованной роли лица, принимающего решения (ЛПР), что затрудняет внедрение моделей в производственную практику и усиливает разрыв между математической постановкой и управленческими задачами.

5. Таким образом, выявлен исследовательский разрыв между современными методами прогнозирования параметров производственных

процессов и задачами календарного планирования. Это обосновывает необходимость разработки интегрированного подхода, обеспечивающего формализованную связку прогнозирования, количественной оценки риска, оптимизации и поддержки принятия решений в рамках единого интеллектуального контура производственного планирования.

## 2 МОДЕЛИ И МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОГО КАЛЕНДАРНОГО ПЛАНИРОВАНИЯ

### 2.1 Формализация задачи календарного планирования

В рамках настоящего исследования задачи календарного планирования производственных процессов формализуются в виде *Job Shop Scheduling Problem (JSSP)*, которая является одной из базовых моделей для описания дискретных производственных систем с конкуренцией операций за ограниченные ресурсы [137, 138]. Данная постановка используется как базовая формальная модель для последующего введения прогнозных и риск-ориентированных модификаций, связанных с учетом неопределенности параметров выполнения операций.

В классической формулировке *JSSP* рассматривается конечное множество производственных заданий (1):

$$J = \{J_1, J_2, \dots, J_n\} \quad (1)$$

каждое из которых состоит из упорядоченной последовательности операций. Каждая операция должна быть выполнена на строго определенной единице оборудования из множества (2):

$$M = \{M_1, M_2, \dots, M_m\} \quad (2)$$

Пусть операция  $O_{ij}$  обозначает  $j$ -ю операцию задания  $J_i$ . Для каждой операции заранее задана:

- машина  $M(O_{ij})$ , на которой она должна быть выполнена;
- длительность выполнения  $p_{ij}$ .

Операции одного задания должны выполняться в заданном технологическом порядке, а каждая единица оборудования может обрабатывать не более одной операции в единицу времени.

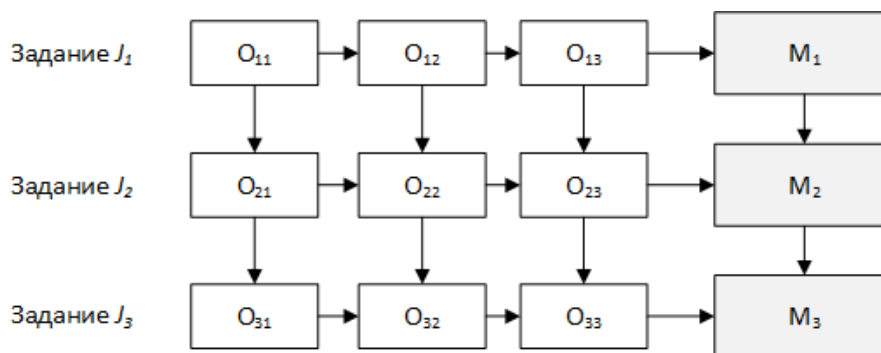


Рисунок 5 – Схема задачи календарного планирования типа JSSP

На схеме (рисунок 5) показано, что каждое задание имеет индивидуальный маршрут обработки, а операции конкурируют за использование оборудования. Данная структура отражает ключевую особенность *JSSP* – пересечение



технологических маршрутов, приводящее к необходимости согласования операций во времени.

Классическая постановка *JSSP* включает следующие типы ограничений:

1. Технологические ограничения (*precedence constraints*). Каждая операция может начаться только после завершения предыдущей операции того же задания (3):

$$S_{i,j+1} \geq S_{ij} + p_{ij} \quad (3)$$

2. Ресурсные ограничения (*machine constraints*). Для любых двух операций  $O_{ij}$  и  $O_{kl}$ , выполняемых на одной машине, должно выполняться одно из условий (4):

$$S_{ij} \geq S_{kl} + p_{kl} \text{ или } S_{kl} \geq S_{ij} + p_{ij} \quad (4)$$

3. Ограничения неотрицательности (5):

$$S_{ij} \geq 0 \quad (5)$$

Указанные ограничения задают детерминированную базовую модель календарного планирования, в которой параметры длительностей операций считаются заранее известными и неизменными.

В качестве базовой целевой функции в задачах *JSSP* в работе рассматривается минимизация *makespan* – общего времени завершения всех заданий (6) [139, 140]:

$$C_{max} = \max_{i,j} (S_{ij} + p_{ij}) \quad (6)$$

Целью планирования является нахождение такого набора времен начала операций  $S_{ij}$ , при котором значение  $C_{max}$  минимально и все ограничения задачи выполняются. Для формального описания задачи календарного планирования используется классическая система обозначений, принятая в постановке *JSSP*.

Вместе с тем в рамках настоящего исследования отдельные элементы модели получают расширенную интерпретацию, не изменяющую структуру базовой постановки, но необходимую для последующего учета неопределенности параметров выполнения операций и формирования риск-ориентированных ограничений (таблица 3). Тем самым базовая детерминированная модель сохраняется как формальный каркас, а прогнозная информация рассматривается как источник параметризации и анализа устойчивости расписаний.

Представленная интерпретация элементов модели позволяет рассматривать задачу календарного планирования не только как оптимизацию временных параметров, но и как задачу обеспечения устойчивости расписания в

условиях неопределенности за счет использования прогнозной информации и риск-ориентированных ограничений в рамках оптимизационной модели.

Таблица 3 – Формальное описание элементов задачи JSSP

| Элемент      | Обозначение | Классическая трактовка      | Интерпретация в рамках настоящего исследования                                                                                         |
|--------------|-------------|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| Задание      | $J_i$       | Последовательность операций | Производственный заказ, реализация которого может сопровождаться отклонениями фактических сроков выполнения                            |
| Операция     | $O_{ij}$    | j-я операция i-го задания   | Элементарный технологический процесс, характеризующийся неопределенностью длительности и влияющий на устойчивость итогового расписания |
| Машина       | $M_k$       | Ресурс выполнения операций  | Производственный ресурс с ограниченной пропускной способностью и возможной вариативностью доступности, влияющей на выполнение графика  |
| Длительность | $p_{ij}$    | Время выполнения операции   | Неопределенный параметр, аппроксимируемый прогнозными оценками и используемый при формировании риск-ориентированных ограничений        |
| Время начала | $S_{ij}$    | Переменная планирования     | Управляемая переменная, определяемая в процессе оптимизации с учетом технологических и параметризованных ограничений                   |
| Цель         | $C_{max}$   | Минимизация makespan        | Базовый критерий эффективности, при котором устойчивость расписания обеспечивается посредством структуры ограничений                   |

Задача *JSSP* относится к классу *NP*-трудных задач, что обуславливает необходимость применения специализированных методов оптимизации и эвристических подходов при решении задач реальной размерности [141, 142]. Несмотря на это, *JSSP* остается одной из наиболее универсальных и выразительных моделей календарного планирования, широко применяемой как в теоретических исследованиях, так и в практических системах управления производством.

Классическая детерминированная постановка задачи *JSSP* опирается на ряд допущений, обеспечивающих формальную строгость и вычислительную управляемость модели [143]. Вместе с тем эти допущения формируют ключевые источники расхождения между плановым расписанием и фактическим ходом производства, что снижает устойчивость решений в реальной среде.

Прежде всего, в детерминированной модели предполагается, что длительности всех операций  $p_{ij}$  известны заранее и остаются неизменными на

всем горизонте планирования [144]. Такое предположение соответствует рассмотрению производственной среды как полностью предсказуемой и исключает влияние случайных факторов на ход выполнения операций на уровне планирования. На практике длительности операций варьируются под воздействием состояния оборудования, человеческого фактора и внешних условий, что приводит к отклонениям от плановых параметров.

Вторым важным допущением является отсутствие сбоев и отказов оборудования. В классической *JSSP* предполагается, что все машины доступны на протяжении всего интервала планирования и способны выполнять операции без перерывов. Такое предположение упрощает модель, однако исключает из рассмотрения один из ключевых источников неопределенности в промышленности – простои и деградацию оборудования, оказывающие непосредственное влияние на фактическое выполнение производственных графиков [145].

Также детерминированная постановка исходит из фиксированности технологических маршрутов. Для каждого задания заранее задана последовательность операций и соответствующих машин, и возможность изменения маршрута в процессе выполнения не рассматривается [146]. Это ограничивает адаптивность планов и не позволяет учитывать альтернативные пути обработки на этапе формирования расписания, которые на практике могут использоваться для компенсации задержек или перераспределения нагрузки.

Существенным ограничением является и ориентация детерминированной модели на оптимизацию одной целевой функции, чаще всего *makespan*. Хотя минимизация общего времени завершения заданий является важным критерием эффективности, в реальных условиях принятие решений на практике связано с учетом дополнительных аспектов, таких как соблюдение сроков и управление рисками отклонений, которые неявно или вовсе не отражаются в детерминированной постановке. Детерминированная постановка, как правило, не предусматривает явного учета подобных аспектов [147].

Наконец, классическая модель предполагает, что оптимальное расписание, полученное на этапе планирования, может быть реализовано без существенных отклонений [148]. Это приводит к формированию планов, оптимальных лишь при строгом выполнении исходных предположений. Даже небольшие отклонения фактических параметров от заданных значений могут привести к каскадному накоплению задержек и снижению практической ценности расписания [149].

Таким образом, детерминированная постановка задачи *JSSP* обеспечивает формальную строгость и служит важной теоретической основой для исследований в области календарного планирования. Вместе с тем заложенные в нее допущения существенно ограничивают устойчивость и применимость получаемых решений в условиях неопределенной и динамичной производственной среды. Это обосновывает необходимость развития моделей планирования, в которых вариативность параметров и риски отклонений учитываются посредством интеграции прогнозной информации и риск-ориентированных ограничений в оптимизационный контур [150].

## 2.2 Учет неопределенности параметров в задачах планирования

В реальных производственных системах длительности технологических операций не являются детерминированными величинами и подвержены вариациям, что необходимо учитывать при формализации задач календарного планирования [151]. В связи с этим возникает необходимость явного представления неопределенности параметров операций, прежде всего их длительностей. Переход от фиксированных значений параметров к более общим формам их задания является принципиальным отличием моделей планирования в условиях неопределенности от детерминированной постановки, рассмотренной в разделе 2.1.

В рамках формализации задач календарного планирования используются несколько основных подходов к представлению неопределенных длительностей операций.

Первый подход основан на рассмотрении длительностей операций как случайных величин, характеризуемых вероятностными распределениями. В этом случае каждая операция описывается распределением времени выполнения, что позволяет учитывать вариативность процессов и формально анализировать вероятность неблагоприятных исходов на уровне параметров модели [152]. Однако использование полных распределений в задачах календарного планирования существенно усложняет оптимизационную модель и, как правило, приводит к значительным вычислительным затратам.

Второй подход заключается в интервальном представлении длительностей, при котором время выполнения операции задается в виде диапазона допустимых значений. Интервальная модель позволяет учитывать границы возможных отклонений без перехода к полной стохастической постановке, однако не содержит информации о вероятностной структуре отклонений и не позволяет дифференцировать уровни риска внутри заданного интервала.

Третий подход основан на сценарном представлении неопределенности, при котором формируется конечный набор возможных реализаций длительностей операций. Каждому сценарию соответствует конкретная комбинация параметров, отражающая возможное состояние производственной системы. Сценарные модели позволяют учитывать совместные реализации неопределенных параметров и зависимости между операциями, однако, как правило, требуют ручного или эвристического задания сценариев и плохо масштабируются при увеличении размерности задачи [153].

Наконец, в качестве компромиссного способа представления неопределенности используется квантильное описание длительностей операций, при котором для каждой операции задаются характерные значения распределения (например, медиана и верхние квантили). Например, 90%-й квантиль длительности операции  $Q_{0.9}(p_{ij})$  означает, что с вероятностью 0,9 фактическое время ее выполнения не превысит данного значения [154]. Использование верхних квантилей позволяет формировать более устойчивые («защищенные») расписания за счет контролируемого повышения

консервативности параметров без необходимости задания полного закона распределения.

Квантильная оценка отражает уровень параметра, соответствующий заданной доле распределения, и может интерпретироваться как верхняя граница ожидаемых отклонений [155]. Такое представление позволяет формализовать неблагоприятные отклонения и управлять уровнем консервативности планирования на этапе формирования оптимизационной модели при сохранении вычислительной управляемости задачи.

Таблица 4 – Основные формы представления неопределенности длительностей операций

| Подход к представлению неопределенности | Форма задания параметров                                       | Учитываемые аспекты неопределенности                                 | Основные ограничения подхода                                                                  | Применимость в рамках настоящего исследования                         |
|-----------------------------------------|----------------------------------------------------------------|----------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|
| Стохастический                          | Полное вероятностное распределение длительности операций       | Вероятностная структура отклонений, хвосты распределений             | Высокая вычислительная сложность; необходимость задания распределений                         | Используется на уровне концептуального анализа, без прямой реализации |
| Интервальный                            | Диапазон допустимых значений                                   | Границы возможных отклонений                                         | Отсутствие информации о вероятностной структуре; одинаковая трактовка всех значений интервала | Ограниченно применим, не обеспечивает управления уровнем риска        |
| Сценарный                               | Конечный набор сценариев параметров                            | Совместные реализации неопределенных параметров                      | Плохая масштабируемость; необходимость ручного задания сценариев                              | Используется для анализа подходов, но не является базовым             |
| Квантильный                             | Характерные значения распределения (медиана, верхние квантили) | Уровень неблагоприятных отклонений без задания полного распределения | Требует прогнозной модели для оценки квантилей                                                | Выбран в качестве базового способа параметризации неопределенности    |

В соответствии с таблицей 4, для сопоставления основных форм представления неопределенности длительностей операций, используемых в задачах календарного планирования, приведен их обобщенный сравнительный анализ с точки зрения учитываемых аспектов неопределенности и практической применимости. Следует подчеркнуть, что перечисленные формы задают формат параметров, но сами по себе не определяют способ использования прогнозной информации в оптимизационной задаче [156]. Они формируют основу для последующей модификации целевой функции и ограничений модели. Таким образом, выбор формы представления неопределенных длительностей операций

является ключевым этапом при построении моделей календарного планирования в условиях неопределенности и напрямую влияет на возможность использования риск-ориентированных критериев оптимизации и интеграции прогнозных моделей в оптимизационный контур.

Учет неопределенности параметров операций требует не только выбора формы представления длительностей, но и соответствующей модификации критериев оптимизации и ограничений модели. В отличие от детерминированной постановки, где оптимизация ориентирована на минимизацию значения целевой функции при фиксированных параметрах, в условиях неопределенности возникает необходимость явного управления рисками неблагоприятных отклонений.

Одним из распространенных подходов являются вероятностные ограничения (*chance constraints*), при которых выполнение условий планирования гарантируется с заданной вероятностью [157]. В таком случае оптимизационная модель формулируется таким образом, чтобы вероятность превышения допустимых значений длительностей операций или сроков выполнения заказов не превышала заданного уровня надежности. Практическая реализация таких ограничений явного задания либо аппроксимации распределений неопределенных параметров.

Другим классом методов являются риск-ориентированные критерии оптимизации, основанные на использовании специальных мер риска. Одной из наиболее распространенных мер риска является *CVaR* (*Conditional Value-at-Risk*). Она показывает среднее значение целевой функции в самых плохих сценариях [158]. По сути, *CVaR* сдвигает фокус оптимизации: вместо среднего результата модель старается контролировать именно худшие случаи. Но у этого подхода есть минус: чтобы его использовать, нужно либо задавать много сценариев, либо хорошо знать распределение вероятностей. Из-за этого модель сильно усложняется и медленнее считается [159].

К риск-ориентированным методам относится и робастная оптимизация. Здесь целевую функцию или ограничения специально настраивают под самые неблагоприятные условия [160]. Такие методы дают очень устойчивые расписания, но часто получаются слишком осторожными и работают хуже в обычных условиях.

Хотя вероятностные и риск-ориентированные критерии хорошо обоснованы с теоретической точки зрения, на практике с ними возникает немало проблем. Во-первых, почти все такие модели требуют, чтобы распределения или сценарии неопределенности были заданы заранее. Во-вторых, когда задача становится большой, вычисления сильно усложняются и тормозят. Из-за этого их сложно применять на реальных промышленных системах большого масштаба [161]. В-третьих, риск-ориентированные критерии нередко обладают ограниченной интерпретируемостью для лиц, принимающих управленческие решения.

Дополнительным ограничением является отсутствие системной связки между прогнозированием параметров и оптимизационной моделью [162]. В большинстве случаев неопределенность учитывается на уровне оптимизации, но

не интегрируется напрямую с результатами прогнозных моделей, построенных на фактических данных. Это приводит к разрыву между аналитическими модулями прогнозирования и механизмами формирования расписаний.

Указанные ограничения обуславливают необходимость интегрированных подходов, в которых прогнозные модели выступают источником параметров неопределенности, а управление риском осуществляется через структуру ограничений и сценариев планирования без перехода к полной стохастической или робастной оптимизации.

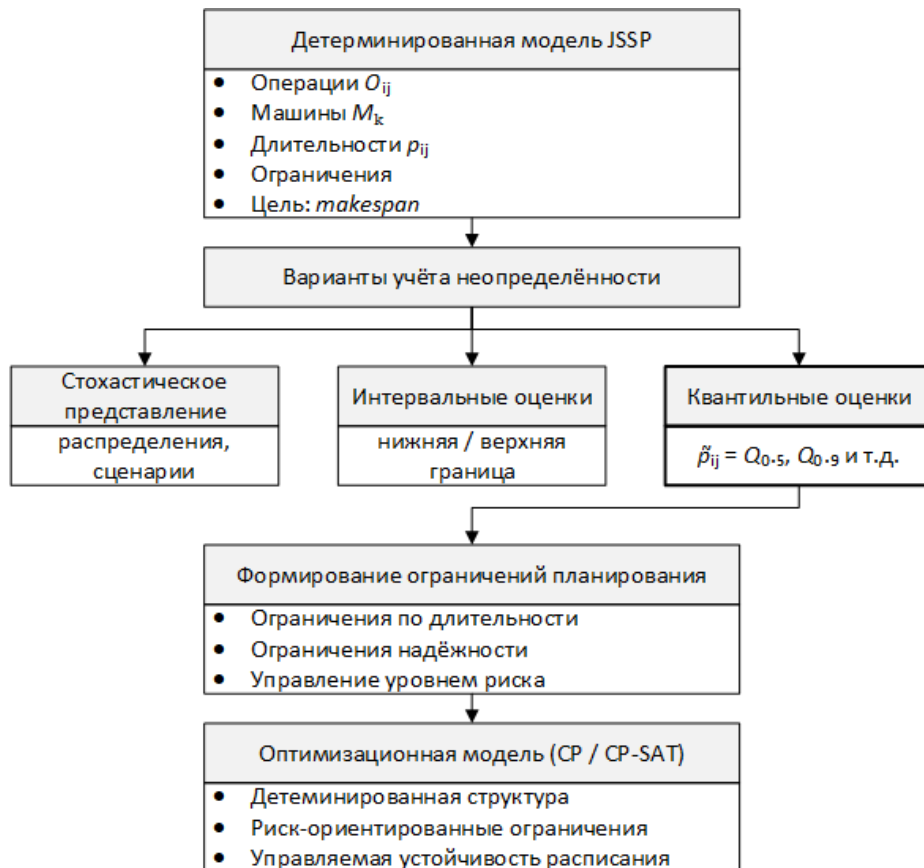


Рисунок 6 – Место учета неопределенности в задаче календарного планирования типа JSSP

На рисунке 6 схематично показано, на каких уровнях постановки задачи календарного планирования может учитываться неопределенность параметров и каким образом это влияет на формирование расписания.

Таким образом, вероятностные и риск-ориентированные критерии оптимизации предоставляют формальный аппарат для учета неопределенности в задачах календарного планирования, однако их практическая реализация требует разработки методов, обеспечивающих связь между прогнозными моделями, представлением риска и оптимизационным контуром.

### 2.3 Методы оптимизации календарного планирования

Задача календарного планирования типа JSSP относится к классу NP-трудных задач комбинаторной оптимизации, что обуславливает значительное

разнообразие методов, применяемых для ее решения. В зависимости от масштаба задачи, требований к оптимальности решения и допустимого времени вычислений традиционно выделяют три основных класса методов: точные методы, эвристические алгоритмы и метаэвристики [163].

Для наглядного представления основных классов методов решения задач календарного планирования типа *JSSP* на рисунке 7 приведена их обобщенная классификация.

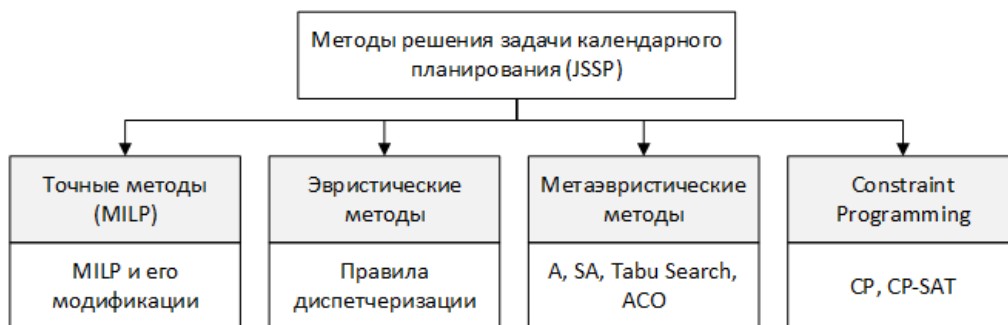


Рисунок 7 – Классификация основных методов решения задачи календарного планирования типа *JSSP*

Точные методы основаны на строгой математической формализации задачи и направлены на нахождение глобально оптимального решения. Наиболее распространенным представителем данного класса являются модели смешанно-целочисленного линейного программирования (*MILP*), в которых технологические и ресурсные ограничения описываются системой линейных неравенств с целочисленными переменными [164]. Такие модели обеспечивают формальную корректность постановки и гарантии оптимальности при условии разрешимости задачи в допустимое вычислительное время.

Однако практическое применение *MILP*-моделей для задач календарного планирования существенно ограничено их вычислительной сложностью. Даже при умеренной размерности число бинарных переменных и ограничений быстро возрастает, что приводит к экспоненциальному росту времени решения [165]. В результате точные методы используются преимущественно для задач малой размерности либо в качестве эталона при оценке приближенных алгоритмов.

В качестве альтернативы точным методам широко применяются эвристические алгоритмы, ориентированные на быстрое получение допустимых, но не обязательно оптимальных расписаний. К таким подходам относятся различные правила диспетчеризации (*dispatching rules*), например, кратчайшей операции (*SPT*), наибольшей оставшейся работы (*LPT*), минимального запаса времени и их модификации [166]. Эвристические алгоритмы отличаются высокой вычислительной эффективностью и сравнительной простотой реализации. Они широко применяются при оперативном формировании расписаний в условиях ограниченных ресурсов.

Главное ограничение эвристических алгоритмов связано с их зависимостью от структуры конкретной задачи. Один и тот же метод может давать допустимый результат для одного класса задач и заметно уступать для



другого. Кроме того, обычные эвристики обычно не дают систематического улучшения решения внутри одной модели. Они также плохо справляются со сложными логическими ограничениями и дополнительными критериями.

Чтобы улучшить качество решений и расширить возможности эвристик, чаще всего используются метаэвристики. К ним относятся генетические алгоритмы, имитации отжига, табу-поиск, муравьиные алгоритмы и их разные гибриды [167]. Метаэвристики работают за счет постепенного итеративного улучшения решений. Благодаря этому они неплохо «прощупывают» большое пространство возможных решений.

Метаэвристики хорошо себя показывают при решении задач *JSSP* среднего и большого размера. Они дают разумный баланс между качеством расписания и временем расчета. Но у них есть свои минусы: требуют тщательной настройки параметров, сильно зависят от структуры задачи и не гарантируют глобально оптимальное решение [168]. Кроме того, добавлять в метаэвристики четкие логические ограничения и понятные механизмы управления рисками обычно бывает довольно сложно.

Классические методы решения *JSSP* включают большой набор подходов – от точных математических моделей до эвристик и метаэвристик. Каждый из них имеет свои слабые места: либо слишком долго считает, либо недостаточно гибкий и прозрачный. Поэтому сейчас активно ищут новые подходы, которые могли бы сочетать понятную модель, хорошую масштабируемость и возможность добавлять логические и рискованные ограничения.

Одним из таких перспективных подходов является *Constraint Programming (CP)* – программирование с ограничениями [169]. В отличие от математического программирования, здесь не нужно приводить все к линейному виду. Можно напрямую описывать логические, временные и ресурсные зависимости, которые типичны для задач календарного планирования.

Для задач типа *JSSP* такой подход очень хорошо подходит. Ведь в *JSSP* есть жесткие технологические последовательности операций и ограничения на использование оборудования. В *CP* это все удобно описывается с помощью специальных типов переменных и глобальных ограничений. Модель получается более понятной и наглядной, чем в *MILP*.

Важным элементом *CP* являются интервальные переменные. Они описывают каждую операцию как временной интервал с определенной длительностью и временем начала [170]. Благодаря этому легко задавать, чтобы операции на одном станке не перекрывались, и соблюдать порядок внутри задания. В итоге модель становится проще, в ней меньше вспомогательных переменных и ограничений, и все выглядит гораздо нагляднее.

*CP-SAT* – это один из современных сильных решателей *Constraint Programming*. Он объединяет в себе программирование с ограничениями, *SAT*-решение и целочисленный поиск [171]. Работает он по гибриднему принципу: ограничения распространяются, идет логический вывод и одновременно поиск решений. За счет этого он неплохо справляется с разными комбинаторными задачами.

В задачах *JSSP* у *CP-SAT* есть важные преимущества. Он строго следит за всеми ограничениями, даже если задача среднего или большого размера [172]. К тому же модель легко менять: можно добавлять новые ограничения или менять цели оптимизации, и не приходится все переделывать с нуля [173].

Особенно полезен *CP-SAT*, когда нужно планировать в условиях неопределенности. Ограничения и целевую функцию можно гибко настраивать, поэтому его удобно использовать для сценарных и риск-ориентированных моделей – если данные о неопределенности берутся из прогнозов [174]. В отличие от эвристических и метаэвристик, здесь можно добавлять новые ограничения, не ломая всю структуру модели.

В общем, анализ показал, что *Constraint Programming* и особенно *CP-SAT* – это действительно удобный и мощный инструмент для задач *JSSP*. Он сочетает формальную строгость, нормальную скорость работы и хорошую гибкость. Именно это и нужно, чтобы в дальнейшем нормально связывать прогнозные модели с оптимизацией в интеллектуальных системах планирования.

## **2.4 Интеграция прогнозных моделей и оптимизации**

Современные подходы к интеллектуальному управлению производственными системами, как правило, основываются на функциональном разделении процессов прогнозирования и оптимизации. В рамках данной парадигмы прогнозные модели используются для оценки параметров производственной среды – длительностей операций, вероятностей отказов оборудования, уровней загрузки ресурсов – после чего полученные прогнозы передаются в оптимизационную модель, формирующую план или расписание. Такой подход в литературе и практических системах принято обозначать как концепцию «*predict* → *optimize*» [175].

В классической постановке данной концепции прогнозирование и оптимизация рассматриваются как последовательно независимые этапы архитектуры системы планирования. На первом этапе строится модель прогнозирования, ориентированная на минимизацию стандартных статистических метрик ошибки, после чего на втором этапе оптимизационная задача решается с использованием прогнозных значений в качестве фиксированных параметров [176]. Подобная схема широко применяется в практических системах планирования благодаря своей модульности и относительной простоте реализации.

Однако раздельное рассмотрение этапов прогнозирования и оптимизации приводит к ряду методологических ограничений. Прогнозные модели, оптимизированные по средним метрикам точности, не учитывают специфики последующей оптимизационной задачи и не ориентированы на управление рисками, значимых с точки зрения планирования. В результате оптимизационная модель получает детерминированные входные параметры, которые не отражают вариативность производственных процессов [177].

В контексте календарного планирования это означает формирование расписаний, оптимальных лишь при реализации ожидаемого сценария, но чувствительных к отклонениям фактических параметров. Даже при высокой

точности прогнозирования небольшие ошибки могут вызывать каскадные эффекты и снижение устойчивости расписаний [178]. Тем самым качество прогноза в статистическом смысле не гарантирует качества итогового управленческого решения.

Дополнительной проблемой является отсутствие обратной связи между этапами прогнозирования и оптимизации. Оптимизационная модель не предоставляет, какие параметры критичны для устойчивости плана, а прогнозная модель не адаптируется к требованиям планирования и не дифференцирует ошибки по степени их влияния на итоговое решение [179]. Это усиливает разрыв между аналитическими модулями прогнозирования и механизмами формирования расписаний.

В современных исследованиях подчеркивается необходимость более тесной интеграции прогнозных и оптимизационных компонентов, при которой прогнозы используются не только как входные значения, но и как элементы, формирующие структуру ограничений и сценариев оптимизационной задачи [180]. Такой подход позволяет учитывать неопределенность непосредственно на этапе планирования и переходить от минимизации средних значений к управлению устойчивостью решений.

Одним из практических механизмов такой интеграции является использование квантильных прогнозов не в качестве точечных параметров, а в виде параметризованных ограничений. В отличие от классического подхода, при котором в оптимизационную модель передается ожидаемое значение длительности операции, квантильное представление позволяет учитывать асимметрию распределений и различный уровень риска, связанный с отклонениями фактических параметров [181].

Квантильный прогноз определяет такое значение параметра, соответствующее заданному уровню распределения, и используется как консервативная оценка при формировании ограничений оптимизационной модели [182]. В контексте производственного планирования это означает возможность формулировать ограничения вида: длительность операции не превышает значения, соответствующего выбранному уровню квантиля. Тем самым прогноз становится инструментом управления допустимым пространством решений, а не просто числовой подстановкой в модель.

Принципиальное отличие данного подхода от детерминированного заключается в том, что оптимизация осуществляется не по средним оценкам, а в рамках множества решений, сформированного с учетом вариативности параметров [183]. Использование квантилей позволяет управлять уровнем консервативности планирования: низкие квантили ориентированы на более агрессивные решения, тогда как высокие формируют более устойчивые, но потенциально менее эффективные расписания с точки зрения целевой функции [184].

Квантильные прогнозы не изменяют структуру оптимизационной задачи. Они задают верхние или нижние границы допустимых значений. Эти границы используются при параметризации ограничений. Такой подход сохраняет математическую целостность модели календарного планирования и

одновременно позволяет учитывать неопределенность. Это позволяет учитывать риск без перехода к полной стохастической или робастной постановке. При этом остается возможность сохранять вычислительную управляемость модели с учетом риска.

Квантильные ограничения также позволяют менять стратегию планирования. В зависимости от требований к надежности графика можно формировать различные сценарии. Они различаются уровнем консервативности ограничений. При этом нет необходимости явно перебирать все возможные реализации случайных параметров.

Использование квантильных прогнозов в виде ограничений создает основу для интеграции машинного обучения и календарного планирования. Данный подход позволяет связать результаты прогнозирования с оптимизационным контуром на уровне управления допустимым пространством решений при сохранении детерминированной структуры модели [185]. Это создает методологические предпосылки для построения интеллектуальных систем планирования, в которых прогнозные и оптимизационные компоненты функционируют как взаимосвязанные элементы единого процесса принятия решений.



Рисунок 8 – Интеграция квантильных прогнозов в оптимизационную модель календарного планирования

На рисунке 8 представлена обобщенная схема интеграции квантильных прогнозов параметров операций в оптимизационную модель календарного планирования, демонстрирующая переход от результатов прогнозирования к формированию риск-ориентированных ограничений и устойчивых расписаний.

### Выводы по разделу

1. В разделе обосновано использование задачи *JSSP* в качестве базовой формальной модели календарного планирования производственных процессов. Показано, что классическая детерминированная постановка *JSSP* обеспечивает строгую математическую основу для описания технологических маршрутов, ресурсных ограничений и целевого критерия оптимизации, однако опирается на допущения, существенно ограничивающих устойчивость получаемых расписаний в условиях реального производства.

2. Проведенный анализ подходов к учету неопределенности параметров операций показал, что использование случайных величин, интервальных оценок, сценарных моделей и квантильного описания длительностей, а также риск-ориентированных критериев оптимизации позволяет формально учитывать вариативность длительностей операций и управлять рисками отклонений. Вместе с тем установлено, что большинство существующих моделей либо требует ручного задания параметров неопределенности, либо сопровождается значительным ростом вычислительной сложности, что затрудняет их применение в задачах промышленного масштаба.

3. Анализ методов оптимизации календарного планирования показал, что классические точные, эвристические и метаэвристические методы обладают существенными ограничениями, связанными либо с вычислительной неэффективностью, либо с недостаточной гибкостью и интерпретируемостью. В этом контексте парадигма *Constraint Programming* и реализация *CP-SAT* обоснованы как методологически и практически перспективный инструмент решения задач *JSSP*, обеспечивающий выразительность формальной модели, масштабируемость и возможность интеграции дополнительных логических и риск-ориентированных ограничений.

4. Методологический анализ раздела обосновывает необходимость интеграции прогнозных моделей и оптимизационного контура календарного планирования. Показано, что использование квантильных прогнозов в качестве параметризованных ограничений, а не точечных параметров, позволяет связать результаты прогнозирования с управлением допустимым пространством решений и уровнем риска при сохранении детерминированной структуры модели и вычислительной эффективности. Предложенный подход формирует концептуальную основу для разработки интегрированных методов интеллектуального планирования, рассматриваемых в последующих главах диссертационного исследования.

### 3 РАЗРАБОТКА И РЕАЛИЗАЦИЯ ГИБРИДНЫХ МОДЕЛЕЙ ИНТЕЛЛЕКТУАЛЬНОГО ПЛАНИРОВАНИЯ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

#### 3.1 Формализация задач прогнозно-оптимизационного планирования

Цифровизация производственных систем и переход к анализу больших массивов разнородных данных усложняют задачи оперативного и стратегического планирования. Управленческие опираются не только на ретроспективную статистику, но и на прогнозы ключевых характеристик: спроса, технического состояния оборудования, энергопотребления и загрузки ресурсов. В этих условиях задачи планирования приобретают выраженный прогнозный характер и требуют интеграции методов интеллектуального анализа данных (ИАД) с модельно-оптимизационными подходами.

В современной научной литературе задачи прогнозирования временных рядов и производственного планирования, как правило, исследуются отдельно. В прогнозировании спроса доминируют методы градиентного бустинга и глубоких нейронных сетей; это подтверждается результатами соревнования *M5*, где лучшие решения базировались на *LightGBM/Gradient Boosting* с тщательно сконструированными признаками и иерархическими моделями спроса [186]. В задачах оценки остаточного ресурса (*RUL*) широко используется датасет *C-MAPSS NASA* и архитектуры глубинного обучения (*CNN*, *LSTM*, *attention*-модели), позволяющие выявлять деградационные паттерны в многомерных временных рядах [187]. В области краткосрочного прогнозирования нагрузки энергосистемы (*STLF*) сформировалось самостоятельное направление, ориентированное на глубокие и гибридные нейросетевые модели [188].

Параллельно развиваются методы оптимизационного и эвристического решения задачи диспетчеризации производства *JSSP* [189], включая классические смешанно-целочисленные модели, эволюционные алгоритмы и подходы «*learning to optimise*» на основе машинного обучения и обучения с подкреплением [190]. Работы по интеграции прогнозов *RUL* и моделей обслуживания с производственным планированием (*PdM-IPS*, совместное планирование ТОиР и производства) показывают, что использование прогностической информации повышает доступность оборудования и производительность [191], однако комплексные фреймворки, объединяющие прогноз спроса, *RUL*, энергопотребления и производственные расписания в единую систему, остаются фрагментарными [192].

В настоящем исследовании указанные направления рассматриваются как взаимосвязанные элементы единого прогнозно-оптимизационного контура. При этом прогнозирование спроса и энергопотребления используется преимущественно как исследовательский домен для анализа свойств и универсальности прогностических архитектур, тогда как в итоговый интегрированный контур интеллектуального планирования включаются задачи прогнозирования остаточного ресурса оборудования и календарного планирования типа *JSSP*.

Формально рассмотрим каждую из задач и их взаимосвязь.

Для набора  $M5$  рассмотрим для каждой позиции временной ряд ежедневных продаж (7):

$$\{y_t\}_{t=1}^T, y_t \in R_+ \quad (7)$$

Дополнительно введем матрицу признаков  $X_t \in R^p$  [193].

Точечное прогнозирование на горизонте  $h$  задается отображением (8):

$$\hat{y}_{t+h} = f(y_t^{(L)}, X_t; \theta) \quad (8)$$

где  $y_t^{(L)} = (y_t, y_{t-1}, \dots, y_{t-L+1})$  – окно из  $L$  последних значений спроса;

$f(\cdot)$  – параметрическая или гибридная прогнозная модель;

$\theta$  – параметры модели.

Обучение модели сводится к минимизации функции потерь на обучающей выборке (9):

$$L_{demand}(\theta) = \frac{1}{N} \sum_{i=1}^N \ell(y_i, \hat{y}_i(\theta)) \quad (9)$$

где  $\ell(\cdot, \cdot)$  – выбранная функция потерь ( $MSE$ ,  $RMSE$ ,  $MAE$ ,  $MAPE$ );

$N$  – число наблюдений в выборке.

Градиентный бустинг является практическим решением задач иерархического прогнозирования спроса. При этом требует подбора временных и категориальных признаков. В то же время гибридные архитектуры, объединяющие бустинг и рекуррентные сети (*DeepAR* и др.), позволяют учитывать более сложные временные зависимости [194].

В контексте настоящей работы прогноз спроса используется в качестве исследовательского домена для анализа свойств и универсальности нейросетевых и гибридных моделей, однако в итоговый интегрированный контур интеллектуального планирования непосредственно не включается [195].

Аналогично формализуется задача прогнозирования остаточного ресурса оборудования ( $RUL$ ). Пусть для каждого объекта (агрегата, двигателя) задан многомерный временной ряд (10):

$$z_t = (z_t^{(1)}, \dots, z_t^{(m)}) \in R^m \quad (10)$$

где компоненты описывают показания сенсоров и соответствующих режимных параметров. Для каждого объекта известен момент отказа  $T_{fail}$ , что позволяет определить истинный остаточный ресурс на момент  $t$  [196]:

$$\tau_t = T_{fail} - t \quad (11)$$

Требуется построить отображение (12):

$$\hat{\tau}_t = g(z_t^{(L)}, \varphi_t, \beta) \quad (12)$$

где  $z_t^{(L)}$  – окно из  $L$  последних многомерных наблюдений;

$\varphi_t$  – дополнительные признаки (скользящие средние, разности, агрегированные индексы деградации);

$g(\cdot)$  – прогностическая модель (например, глубокие *RNN/CNN*, а также гибридные архитектуры *ML* и *DL*);

$\beta$  – параметры модели.

В регрессионной постановке используется функционал ошибки (13):

$$L_{RUL}(\beta) = \frac{1}{N} \sum_{i=1}^N (\tau_i - \hat{\tau}_i(\beta))^2 \quad (13)$$

Датасет *C-MAPSS NASA* стал де-факто стандартом для оценки методов прогнозирования *RUL* авиационных двигателей; на нем исследуются разнообразные архитектуры глубинного обучения (*CNN*, *LSTM*, *attention-модели*), а также методы расширенной обработки признаков.

В ряде работ демонстрируется эффективность гибридных схем, объединяющих нейронные сети и традиционные *ML*-модели, а также методов интерпретации, позволяющих связывать изменения признаков с физическими механизмами деградации [197].

В рамках предлагаемого подхода оценки *RUL* используются как прогностический слой для поддержки решений по обслуживанию и перепланированию, влияя на параметры задач *JSSP* и планирования простоев оборудования.

Краткосрочное прогнозирование нагрузки энергосистемы (*Short-Term Load Forecasting, STL*) играет ключевую роль в балансировании спроса и предложения, управлении мощностями и тарифной политике [198]. Для краткосрочного прогнозирования нагрузки энергосистемы вводится временной ряд почасовой нагрузки  $p_t \in R_+$  и вектор внешних факторов  $W_t$ . Прогноз формализуется как (14):

$$\hat{p}_{t+h} = h(p_t^{(L)}, W_t, \omega) \quad (14)$$

где  $p_t^{(L)} = (p_t, p_{t-1}, \dots, p_{t-L+1})$ ;

$h(\cdot)$  – прогностическая модель временного ряда (например, *LSTM*, *TCN*, гибридные архитектуры);

$\omega$  – параметры.

При обучении используются стандартные метрические функционалы (*MSE*, *RMSE*, *MAPE*) и специальные сценарии валидации (*rolling-origin evaluation*), обеспечивающие корректную оценку качества на горизонтах нескольких часов и дней [199]. Обзор недавних работ показывает, что глубокие



нейронные сети (в том числе гибриды *CNN-LSTM*, *encoder-decoder*-архитектуры) могут существенно превосходить классические статистические модели при наличии сложных нелинейностей и сезонностей [200].

Прогноз нагрузки используется для анализа устойчивости моделей на данных с выраженной сезонностью и может вводиться в оптимизационную постановку в виде ограничений по мощности или штрафных функций.

Классическая задача *JSSP* формализует процесс назначения операций на оборудование при наличии технологических и ресурсных ограничений [201, 202]. Пусть:

- $J = \{1, \dots, n\}$  - множество работ (*jobs*);
- каждая работа  $j$  состоит из упорядоченной последовательности операций  $O_j = (o_{j1}, \dots, o_{j,k_j})$ ;
- $M = \{1, \dots, m\}$  – множество станков (*machines*);
- каждой операции  $o_{jk}$  соответствует станок  $m(o_{jk}) \in M$  и длительность обработки  $d_{jk} > 0$ .

Ищется расписание, определяющее время начала  $s_{jk}$  каждой операции, при котором минимизируется критерий (чаще всего – *makespan*) (15):

$$C_{max} = \max_{j,k} (s_{jk} + d_{jk}), \min C_{max} \quad (15)$$

при ограничениях технологической последовательности (16):

$$s_{j,k+1} \geq s_{jk} + d_{jk}, \forall j, k = 1, \dots, k_j - 1 \quad (16)$$

и ресурсных ограничениях (17):

$$\forall (j, k) \neq (j', k'), m(o_{jk}) = m(o_{j'k'}) \Rightarrow (s_{jk} \geq s_{j'k'} + d_{j'k'} \vee s_{j'k'} \geq s_{jk} + d_{jk}) \quad (17)$$

В рамках диссертации новым аспектом является параметризация длительностей и доступности оборудования на основе прогнозов *RUL*, что позволяет формировать адаптивные и риск-ориентированные расписания.

Представленные выше задачи в традиционных исследованиях часто рассматриваются изолированно: прогноз спроса – в цепях поставок, *RUL* – в системах ТОиР, *JSSP* – в операционном менеджменте. Современные работы по интеграции прогнозов *RUL* и планирования производства (*PdM-IPS*, совместное планирование обслуживания и производства) демонстрируют преимущества объединенного подхода, однако в основном фокусируются на паре «*RUL* + *JSSP*» либо «ТОиР + расписание одного станка» [203].

Отдельные элементы предлагаемой архитектуры прогнозно-оптимизационной системы рассматривались ранее в работах автора и соавторов [204-206], где анализировались задачи прогнозирования, адаптивного управления и интеллектуальной обработки данных в производственных и технических системах.

С учетом изложенного интегрированная многокритериальная постановка прогнозного планирования может быть записана в общем виде (18):

$$\min_{\theta, \beta, \omega, S} [\lambda_1 \cdot Err_{demand}(\theta) + \lambda_2 \cdot Err_{RUL}(\beta) + \lambda_3 \cdot C_{max}(S) + \lambda_4 \cdot PeakLoad(S, \omega)] \quad (18)$$

где  $\theta$ ,  $\beta$ ,  $\omega$  – параметры прогностических моделей спроса,  $RUL$  и нагрузки соответственно;

$S$  – расписание (множество  $S_{jk}$ );

$Err_{demand}$ ,  $Err_{RUL}$  – агрегированные показатели качества прогнозов (например,  $RMSE$  или  $MAPE$ );

$C_{max}(S)$  – *makespan*;

$PeakLoad(S, \omega)$  – функционал, отражающий пиковую нагрузку энергосистемы при реализации расписания  $S$  и профиль нагрузки, прогнозируемый моделью с параметрами  $\omega$ ;

$\lambda_i$  – весовые коэффициенты, задающие относительную важность критериев [207].

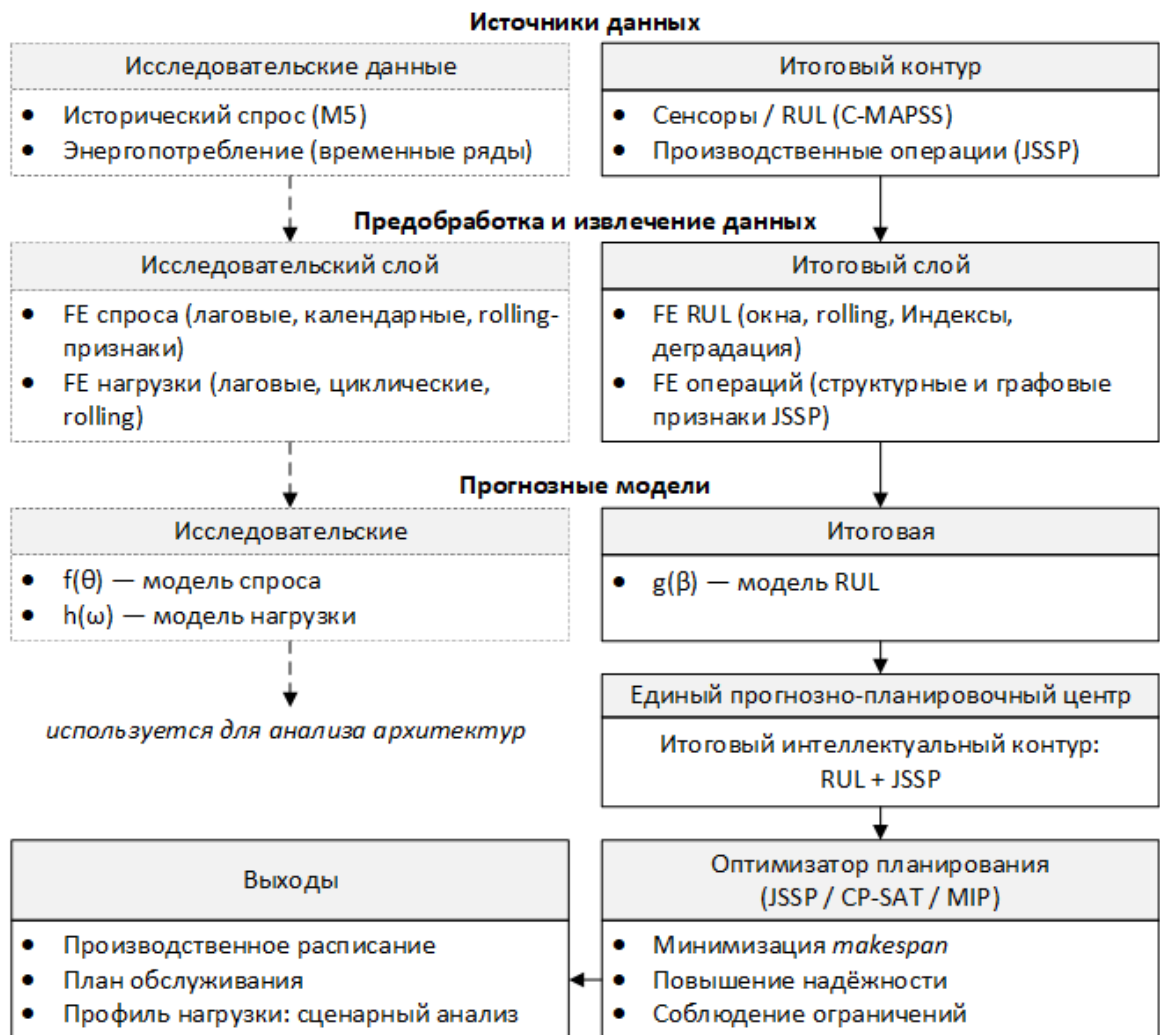


Рисунок 9 – Интегрированная структура задач прогнозного планирования на основе моделей ИАД

Данная постановка носит концептуальный характер и используется для формализации взаимосвязи прогнозных и оптимизационных задач. В практической реализации оптимизация осуществляется по подмножеству критериев, соответствующих выбранному сценарию интеллектуального планирования.

Реализация может быть построена либо в виде единой интегрированной модели (*MIP/CP*-фреймворк с аппроксимацией прогнозных зависимостей), либо в виде последовательного гибридного контура, где сначала обучаются прогностические модели, а затем их результаты используются как входные параметры оптимизационной задачи. Такая архитектура «прогнозирование → управление → оптимизация» согласуется с ранее предложенными подходами к адаптивному управлению технологическими процессами [208].

Структурная схема взаимосвязи задач прогнозирования и планирования представлена на рисунке 9 и иллюстрирует место интеллектуального анализа данных в формировании параметров и ограничений календарно-производственного планирования.

### **3.2 Данные и экспериментальный контур прогнозирования и планирования**

Реализация задач прогнозного планирования на основе методов интеллектуального анализа данных требует использования надежной и репрезентативной информационной базы.

В исследовании использованы четыре открытых набора данных. Каждый из них отражает отдельный аспект производственной деятельности: динамику спроса, техническое состояние оборудования, структуру производственных процессов и параметры энергопотребления. Несмотря на различие источников, все наборы данных обрабатывались в едином конвейере. Это обеспечило совместимость данных и воспроизводимость последующих вычислительных экспериментов.

Первый набор данных – *M5 Forecasting* (продажи товаров в сети *Walmart*). Представляет собой многомерный временной ряд с выраженной сезонностью, промо-эффектами, региональной дифференциацией и стохастичностью. В работе использовался для анализа моделей прогнозирования и их обобщающей способности. При этом он не включался напрямую в итоговый контур интеллектуального планирования.

Второй набор – *NASA C-MAPSS*. Он описывает деградацию авиационных двигателей в процессе эксплуатации до наступления отказа. Данные носят имитационный характер генерации, но при этом имеют вполне реалистичную структуру многоканальных сигналов. Они позволяют рассчитывать показатель остаточного ресурса двигателя (*RUL*), который потом используется как один из главных ориентиров для обслуживания по состоянию.

Данные по энергопотреблению – это временные ряды с четко выраженной цикличностью. Их я использовал, чтобы проверить, насколько универсальны прогнозные модели и как циклические факторы влияют на качество предсказаний.

Библиотека *Taillard JSPLIB* содержит классические тестовые задачи *Job Shop Scheduling*. Там есть информация о последовательности операций, их длительности и распределении по станкам. Эти данные нужны для построения оптимизационной модели производственного расписания.

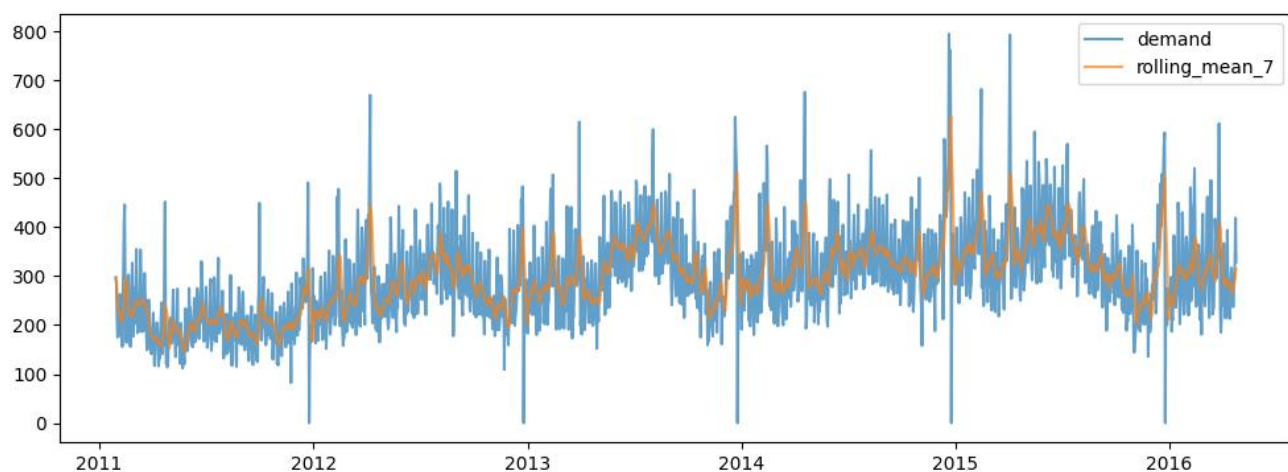
Подготовка всех данных была сделана через воспроизводимый *ETL*-конвейер на *Python*-скриптах. На этапе извлечения я отдельно проверял полноту данных и правильность временной структуры. Пропуски обрабатывались по-разному. Для динамических наборов использовалась линейная интерполяция и локальные модели. В датасете *C-MAPSS* неполные траектории удалялись полностью, как с рекомендует *NASA*. Также проводилась очистка от выбросов с помощью интерквантильного размаха и метода *LOF* – это нужно было для стабильности скользящих статистик и дальнейшего обучения моделей.

После очистки все наборы данных привели в едином удобном формате: добавили явные метки времени, нормализованные идентификаторы и упорядоченные признаки. Для *C-MAPSS* дополнительно пересчитали целевую переменную *RUL* – для каждого двигателя посчитали, сколько ему осталось до отказа. Для задач *JSSP* сформировали структурные признаки технологических процессов. Это позволило на следующем этапе нормально связать прогнозный и оптимизационный модули. Итоговые примеры данных приведены в таблице 5.

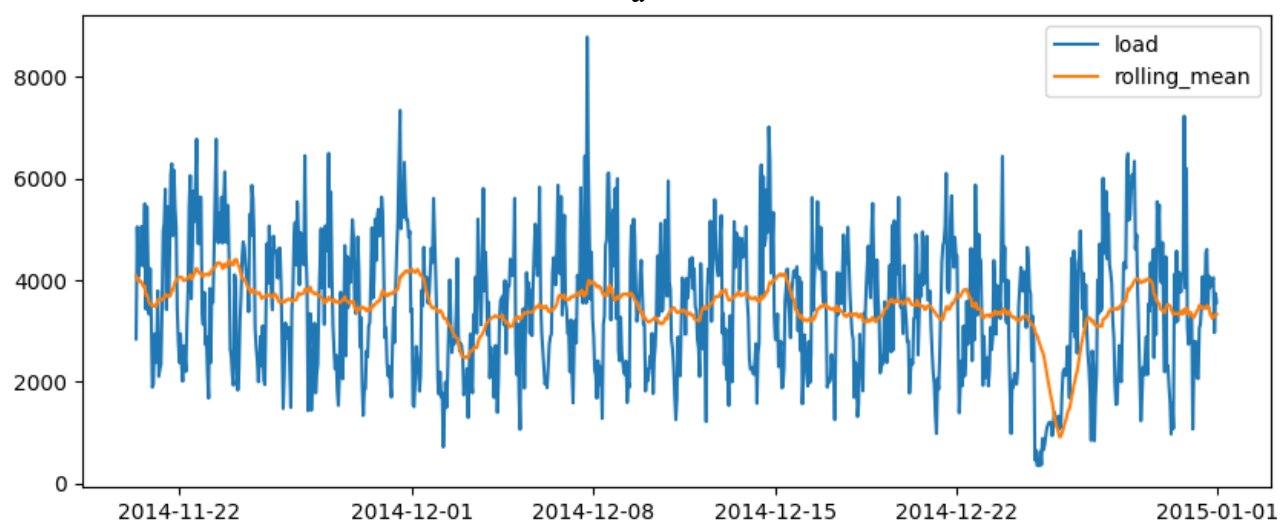
Таблица 5 – Итоговая структура обработанных данных

| Набор данных | Число наблюдений | Основные признаки                             |
|--------------|------------------|-----------------------------------------------|
| M5           | 133 910          | item_id, date, lag_7, rolling_mean_30, region |
| JSSP         | 57 250           | job_id, op_id, machine, duration, norm_pos    |
| C-MAPSS      | 160 359          | unit, cycle, s1...s24, rul, time_coef         |
| Energy       | 35 065           | ts, load, hour_sin, dayofweek_cos, lag_1      |

При анализе данных по спросу и энергопотреблению сразу бросились в глаза устойчивые сезонные колебания. Они хорошо заметны в как гармониках, так и автокорреляции временных рядов (рисунок 10). Анализ *C-MAPSS* были выявлены нелинейные признаки деградации, а также компактные группы в пространстве признаков при визуализации методами *t-SNE*, *UMAP* [209]. Для *JSSP* графический анализ позволил выделить диапазоны длительностей, типичные структуры операций и плотность нагрузки на ресурсы.



a



б

Рисунок 10 – Распределение лаговых признаков и циклических компонент в M5 (a) и Energy (б)

Для каждого набора данных выбирается специализированный набор моделей (таблица 6).

Таблица 6 – Соответствие моделей наборам данных

| Набор данных                     | Основная модель                 | Дополнительные модели сравнения |
|----------------------------------|---------------------------------|---------------------------------|
| M5 (спрос)                       | LSTM, LightGBM, LSTM→XGB hybrid | Naïve, SARIMA                   |
| C-MAPSS (RUL)                    | LSTM/GRU, CNN-LSTM              | Gradient Boosting               |
| LD2011–2014 (энергия)            | LSTM, XGBoost                   | Prophet, SARIMA                 |
| JSSP (характеристики расписаний) | CatBoost                        | RandomForest                    |

Для обеспечения корректности эксперимента введены эталонные модели, задающие нижнюю границу качества и позволяющие количественно оценить вклад более сложных методов [210]. Наивная модель фиксирует минимальный уровень, *Seasonal Naïve* учитывает цикличность и снижает ошибку на 12-18%

относительно *Naive* на выборках *M5* и *Energy*, подтверждая наличие календарных структур (рисунок 11) [211].

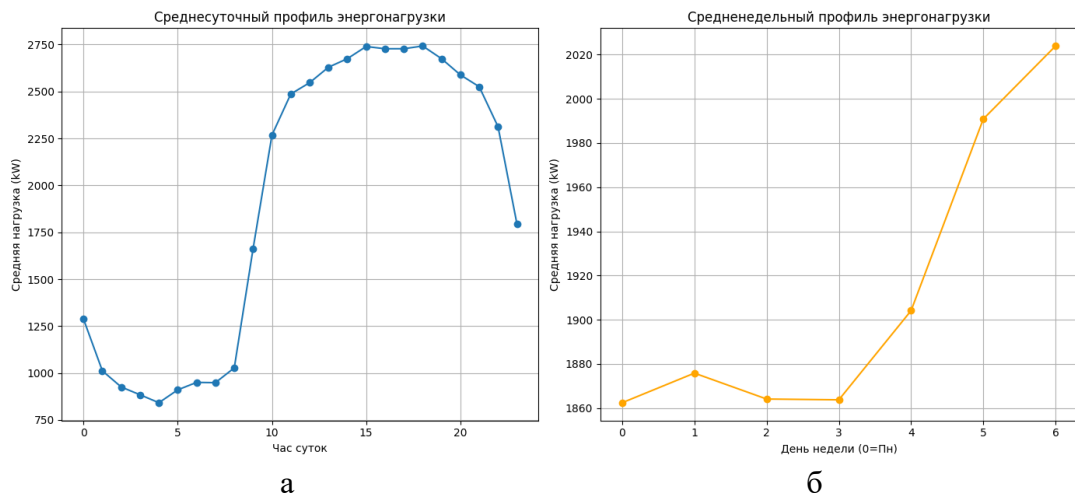


Рисунок 11 – Среднесуточный (а) и средненедельный (б) профили энергопотребления

На данных по энергопотреблению *ETS* снизила *RMSE* на 25-30% по сравнению с наивными прогнозами и при хорошо уловила резкие скачки нагрузки. В задаче прогнозирования остаточного ресурса (*RUL*) обычные линейные модели (*Linear/Ridge/ElasticNet*) по агрегированным признакам дали *MAE* в пределах 24-32 циклов – это укладывается в рекомендации *NASA*. Для задач *JSSP* простые регрессии по усредненным параметрам инстансов показали  $R^2$  от 0,42 до 0,55, что примерно соответствует результатам других исследований на *JSPLIB*.

Эксперименты показали, что более сложные модели дают заметный прирост. Например, гибрид *LSTM*→*XGBoost* на датасете *M5* улучшил *RMSE* больше, чем на 40% по сравнению с *Seasonal Naive*. *XGBoost* в задачах *JSSP* уверенно обошел линейные регрессии по *RMSE*, а рекуррентные модели в *C-MAPSS* снизили *MAE* в два раза по сравнению с простыми линейными моделями.

В качестве основных прогнозных моделей я использовал градиентный бустинг на решающих деревьях (*XGBoost/LightGBM/CatBoost*) [212]. Они хорошо работают, понятны и устойчивы. В большинстве экспериментов я в итоге выбирал *XGBoost* – в основном потому, что с ним удобно работать через *SHAP* и легко объяснять результаты [213].

Во всех экспериментах использовалось единое признаковое пространство: лаги, скользящие статистики, календарные индикаторы, а для *JSSP* – агрегированные производственные характеристики. На агрегированной выборке *JSSP* из 241 признака *XGBoost* с целевой переменной *makespan* на независимой валидации из 16 инстансов достиг  $RMSE \approx 968$  и  $MAE \approx 556$  (условных единиц). Настройка выполнялась по согласованной схеме: *grid search* для рабочих диапазонов и байесовская оптимизация для тонкой настройки; качество оценивалось едиными метриками (*MSE/RMSE/MAE/MAPE*).

Интерпретация моделей включала *feature importance (gain)* и *SHAP*-значения [214]. Для *M5* максимальный вклад вносили недельные/месячные лаги, скользящие статистики 7-28 дней и календарные индикаторы; агрегированный ранжированный топ-20 показан на рисунке 12.

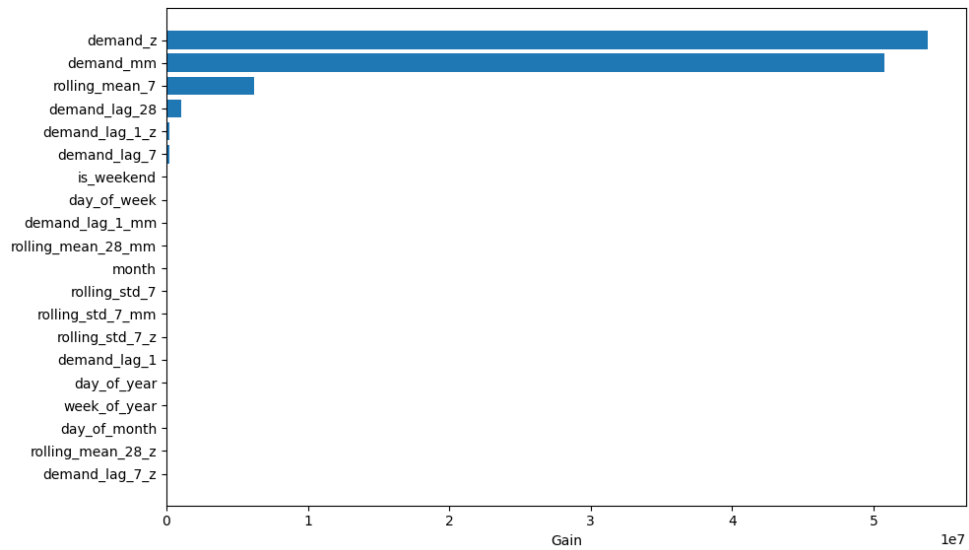


Рисунок 12 – Топ-20 признаков для модели XGBoost по *M5*

*SHAP* подтвердил доминирование лагов и скользящих статистик и выявил пороговые эффекты; агрегированный профиль влияния представлен на рисунке 13.

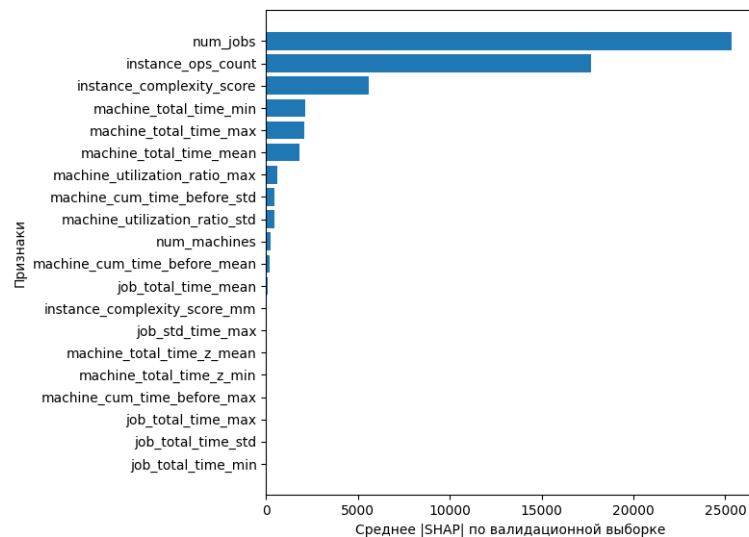


Рисунок 13 – Агрегированный профиль влияния признаков *M5* по *SHAP* для модели XGBoost

Нейросетевые модели использовались как класс прогнозных моделей для извлечения нелинейных временных зависимостей, прежде всего в задачах деградации и временных рядов [215]. Рассматривались *LSTM/GRU*, *CNN-LSTM* (для *C-MAPSS*) и *MLP* (для табличных признаков и элементов гибридных схем).

Процедура подготовки данных включала нормализацию (z-преобразование и *min-max*), формирование скользящих окон фиксированной длины и разделение на обучающую/валидационную/тестовую выборки.

Для *C-MAPSS* вход включал 94 числовых признака (режимы *op1-op3*, сенсоры *s1-s21*, скользящие средние/СКО по окнам 20 циклов, первые разности, нормированные признаки). После очистки итоговый объем составил 160 099 наблюдений; блоки по двигателям: 182 (обучение), 39 (валидация), 39 (тест), что соответствует 103 124, 23 105 и 21 130 временным шагам соответственно.

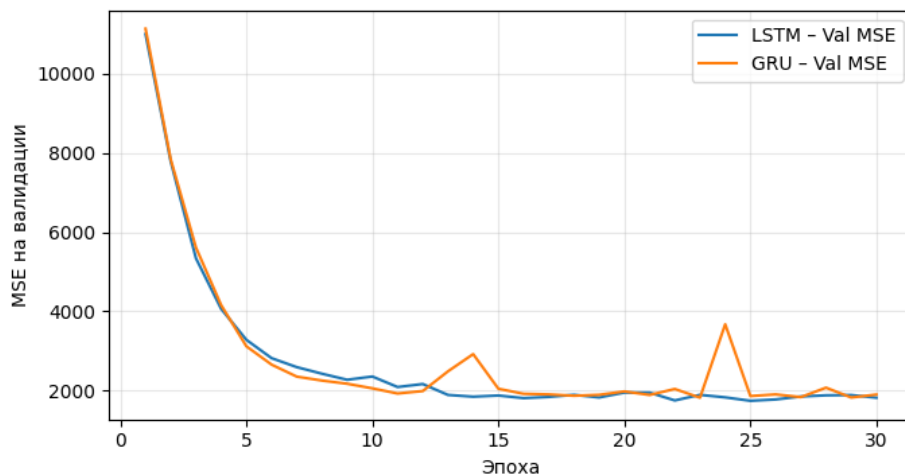


Рисунок 14 – Динамика валидационной ошибки по эпохам для LSTM и GRU

Динамика валидационной ошибки показана на рисунке 14: *LSTM* убывает более стабильно, у *GRU* наблюдаются всплески. Лучшее качество на валидации: *LSTM MSE*  $\approx 1\,743,98$ , *GRU MSE*  $\approx 1\,822,23$ . На тесте: *LSTM MSE*  $\approx 1\,839,16$ , *MAE*  $\approx 30,21$ ; *GRU MSE*  $\approx 1\,534,11$ , *MAE*  $\approx 28,17$ . Диаграмма «истинный *RUL* vs предсказанный *RUL*» для *GRU* приведена на рисунке 15; для больших значений *RUL* сохраняется дисперсия.

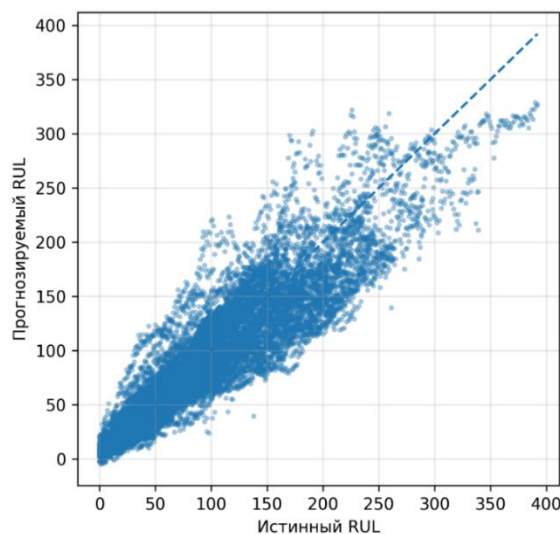


Рисунок 15 – Диаграмма «истинный *RUL* vs предсказанный *RUL*» на тестовой выборке для модели GRU



Полученные результаты соответствуют типичным уровням качества на данном наборе данных и обеспечивают достаточный запас точности для использования скрытых состояний в гибридных моделях [216].

Для *LD2011-2014* после очистки объем составил 34 897 наблюдений. Формировалось окно 24 часа, включающее календарные переменные (час, день недели, сезонные гармоники), лаги ( $t-1$ ,  $t-24$ ,  $t-168$ ), скользящие статистики и нормированные варианты. Итоговая размерность тензора признаков:  $34\,873 \times 24 \times 30$ , распределение последовательностей: 24 413 (обучение), 5 230 (валидация), 5 230 (тест). По тесту *LSTM* показала неудовлетворительное качество и эффект усреднения:  $MSE \approx 2,97 \cdot 10^7$ ,  $RMSE \approx 5,45 \cdot 10^3$ ,  $MAE \approx 4,67 \cdot 10^3$ .

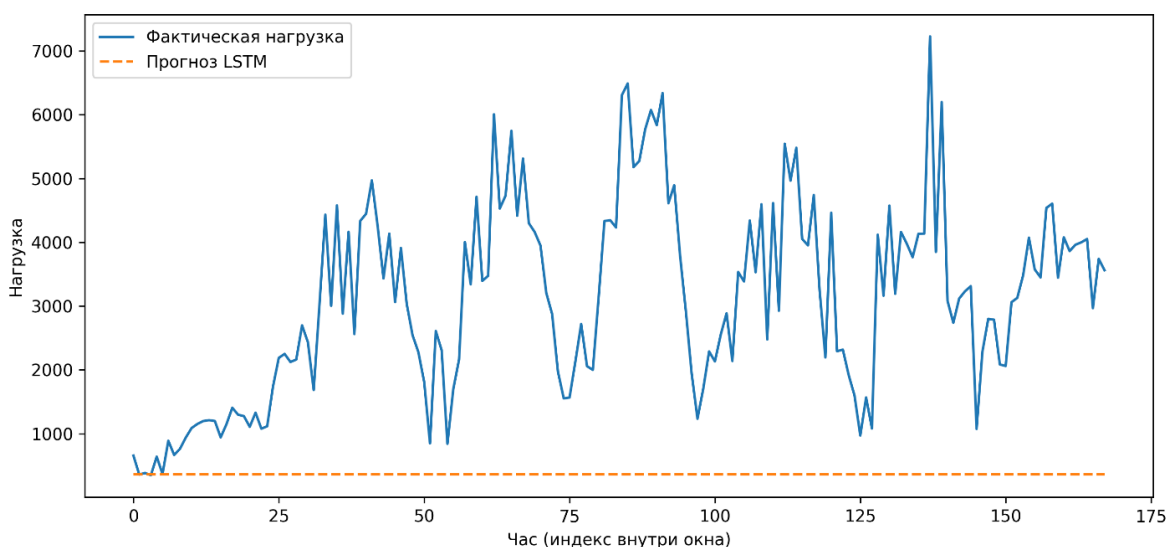


Рисунок 16 – Сравнение фактической и прогнозной часовой нагрузки (*LSTM*, тестовая выборка)

Сравнение фактической и прогнозной нагрузки приведено на рисунке 16, диаграмма рассеяния – на рисунке 17: точки прижаты к оси, что указывает на квазиконстантный прогноз. В результате для данного набора в качестве более устойчивого решения далее используется градиентный бустинг.

Визуальный анализ ошибок показал, что на *LD2011-2014* модель *LSTM* недостаточно улавливает краткосрочные и недельные колебания нагрузки: предсказания сглаживаются и стремятся к среднему уровню, что проявляется как систематическое занижение пиков и завышение провалов. Такой эффект усреднения типичен при недостаточной информативности последовательных признаков либо при доминировании стационарной компоненты в оптимизации *MSE*. Дополнительно на качество влияют неоднородность потребления по сезонам и календарным факторам, а также возможные структурные сдвиги во временном ряду.

Различная эффективность бустинговых и нейросетевых моделей на разных доменах обусловила использование гибридных схем, объединяющих преимущества извлечения последовательных паттернов и устойчивой интеграции табличных признаков [217]. Подход к построению гибридных

моделей, сочетающих нейросетевые и ансамблевые методы, опирается на ранее полученные результаты, представленные в работах автора [218], посвященных задачам промышленной аналитики и предиктивного обслуживания оборудования.

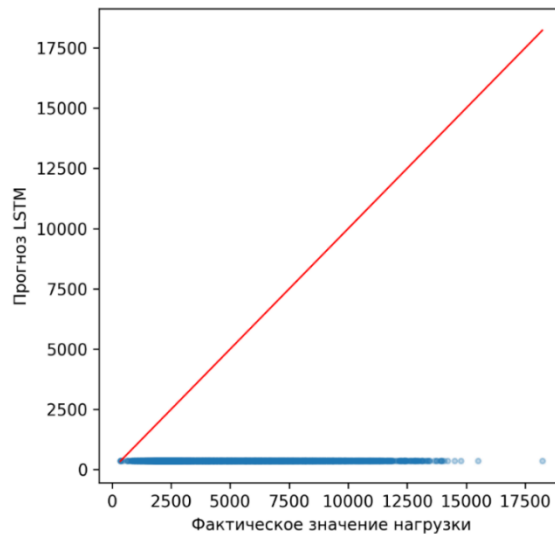


Рисунок 17 – Диаграмма рассеяния фактических и прогнозных значений нагрузки

В диссертации реализованы два уровня гибридизации: *feature-level fusion* (конкатенация табличных признаков и латентного вектора *LSTM*), *model-level stacking* (*XGBoost* над расширенным признаковым пространством). Формально прогноз можно представить в виде (19):

$$\hat{y}_t = f_{XGB} \left( x_t^{(tab)}, h_T(x_{t-T+1:t}) \right) \quad (19)$$

где  $x_t^{(tab)}$  – вектор табличных признаков в момент времени  $t$ ;

$h_T(\cdot)$  – скрытое состояние *LSTM* на последнем шаге временного окна;

$f_{XGB}$  – обученная модель *XGBoost*.

В качестве опорных были реализованы два гибридных сценария: для прогнозирования энергопотребления (датасет *LD2011-2014*, далее *Energy*) и для оценки остаточного ресурса по подмножеству *C-MAPSS* (units 1-5). Основные параметры *LSTM*-экстрактора признаков и *XGBoost*-метамоделей приведены в таблице 7.

В обоих случаях последовательности формировались по схеме скользящего окна (24 часа для нагрузки и 30 циклов для деградации), а *LSTM* обучалась в режиме регрессии с функцией потерь *MSE*. После обучения *LSTM* использовалась как фиксированный экстрактор признаков, а *XGBoost* обучался на гибридном признаковом пространстве (табличные признаки + латентный вектор).

Таблица 7 – Архитектура и параметры гибридных моделей LSTM→XGBoost

| Набор данных       | Задача                                   | Модель                          | Окно Т | Признаки р | LSTM d | Гиперпараметры XGBoost                                                                 |
|--------------------|------------------------------------------|---------------------------------|--------|------------|--------|----------------------------------------------------------------------------------------|
| LD2011-2014        | Прогноз почасовой электрической нагрузки | LSTM (2 слоя, 64 нейрона) → XGB | 24     | 30         | 64     | n_estimators=500, max_depth=6, learning_rate=0.05, subsample=0.9, colsample_bytree=0.9 |
| C-MAPSS (unit 1-5) | Прогноз остаточного ресурса (RUL)        | LSTM (2 слоя, 64 нейрона) → XGB | 30     | 277        | 64     | n_estimators=800, max_depth=4, learning_rate=0.03, subsample=0.8, colsample_bytree=0.8 |

На данных *LD2011-2014* обычная *LSTM* в начале хорошо снижала ошибку на валидации, но на тестовой выборке показала себя хуже. Ошибки поучились довольно большими:  $RMSE \approx 5449,5$  и  $MAE \approx 4671,9$  при почасовом прогнозе нагрузки. В гибридной модели я использовал *LSTM* только как генератор дополнительных скрытых признаков (64 латентные компоненты). А окончательный прогноз уже строил *XGBoost* на расширенном наборе признаков (30 исходных + 64 из *LSTM*) [219]. В итоге гибридная модель значительно лучше справилась с задачей:  $RMSE$  упал до 2458,7, а  $MAE$  до 1411,1. Как выглядит рассеяние «факт-прогноз» для этой модели можно увидеть на рисунке 18.

Визуальный анализ рисунка 18 показывает, что точки для гибридной модели значительно плотнее сосредоточены вдоль диагонали по сравнению с чистой *LSTM*: уменьшилось число грубых недооценок и переоценок пиковых значений, а разброс в средних диапазонах нагрузки стал заметно меньшим.

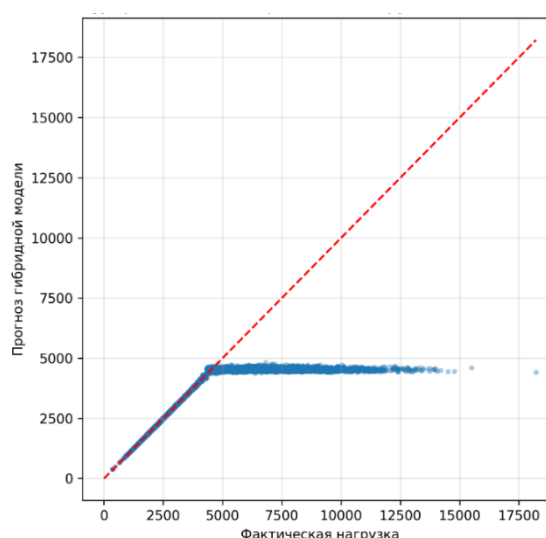


Рисунок 18 – Соотношение фактической и прогнозной нагрузки для гибридной модели Energy (LSTM→XGBoost)

Вторая серия экспериментов была проведена на ограниченном подмножестве *C-MAPSS*, включающем первые пять двигателей (units 1-5). Для

данного поднабора исходно были получены результаты рекуррентных моделей: при обучении *GRU* на полном наборе признаков лучшая модель обеспечила *Test RMSE*  $\approx 39,2$ , *Test MAE*  $\approx 28,2$  при оценке остаточного ресурса (*RUL*) на тестовой выборке. Гибридная схема включала компактную *LSTM*-модель-экстрактор признаков, обученную на units 1-3 с валидацией на unit 4 и тестированием на unit 5. После обучения *LSTM* скрытое состояние последнего шага (размерностью 64) конкатенировалось с исходным табличным вектором (277 признаков), и на этом пространстве обучался *XGBoost*. На тестовой выборке для unit 5 гибридная модель показала значения *RMSE*  $\approx 47,4$ , *MAE*  $\approx 34,9$ , что хуже, чем у лучшей чисто рекуррентной модели. Соотношение истинных и прогнозных значений *RUL* для гибридной модели представлено на рисунке 19.

На рисунке 19 (графике) заметно, что для небольшого поднабора данные располагаются вокруг диагонали менее компактно, чем в случае лучшей *GRU*-модели, а разброс в области малых значений *RUL* остается значительным. Этот результат демонстрирует важное ограничение гибридного подхода: при малом объеме обучающих примеров и высокой вариативности деградационных траекторий преимущество сложных гибридных архитектур не является гарантированным, и аккуратно настроенная рекуррентная модель может сохранять преимущество по качеству.

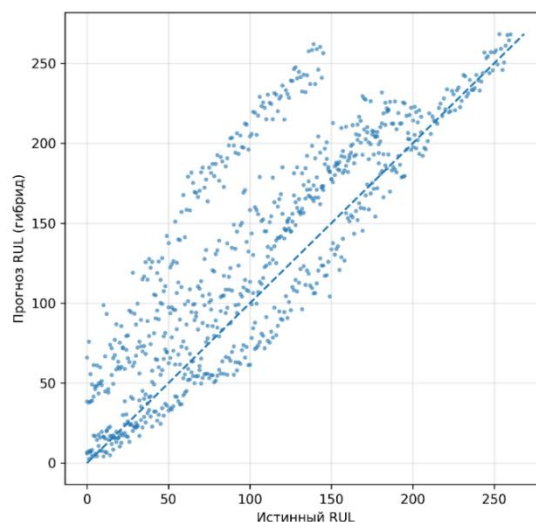


Рисунок 19 – Соотношение фактического и прогнозного *RUL* для C-MAPSS (units 1-5, гибридная модель *LSTM*+*XGBoost*)

По результатам экспериментов можно выделить несколько методически важных выводов.

Гибридные модели *LSTM*→*XGBoost* хорошо работают, когда нужно учитывать и временные закономерности, и обычные табличные данные. Именно так было в прогнозе энергопотребления. *LSTM* находит скрытые паттерны во времени, а *XGBoost* их умело объединяет с другими признаками.

Но на небольшой части данных C-MAPSS (units 1-5) оказалось, что гибридный – не всегда лучшее решение. При маленьком объеме данных и сложных процессах

деградации обычная *LSTM* иногда работает так же хорошо или даже лучше. Гибриды выигрывают с основным на больших данных.

В нашей системе такие гибридные модели связывают анализ данных с оптимизацией расписаний. Открытые датасеты очень выручают. Они позволяют проверить модели, даже если нет доступа к реальным заводским данным. Многие ученые по всему миру используют именно их – это удобно и результаты легко сравнивать. Каждый датасет отражает свою часть задачи:

1. *M5 Forecasting (Walmart)* – спрос с сезонностью, акциями и случайными скачками. Похоже на реальный спрос в магазинах и на производстве.

2. *LD2011-2014 Energy* – почасовое энергопотребление с дневными циклами, пиками и влиянием выходных и праздников. Полезно для энергосбалансированных расписаний.

3. *NASA C-MAPSS*: содержит многоканальные вибрационные и эксплуатационные параметры, описывающие деградацию сложных технических систем. Такие зависимости аналогичны деградации станков, компрессоров, подшипников и других узлов, влияющих на надежность производственного оборудования.

Применимость результатов предиктивного обслуживания к планированию подтверждается и результатами исследований автора: в работе [220] показано, что вибродиагностические признаки совместно с *ML* позволяют устойчиво выявлять деградацию шпиндельных подшипников токарного оборудования и формировать информативные оценки технического состояния, что методологически обосновывает использование датасетов класса *C-MAPSS* для разработки моделей риск-ориентированного планирования.

В рассмотренных данных представлены сезонные циклы, стохастические компоненты, эффекты лагов и автокорреляций, влияние внешних факторов, нестационарные участки и точки перегиба, характерные для деградации. Эти свойства соответствуют зависимостям, которые необходимо моделировать при планировании производства, энергопотребления и ремонтов, что согласуется с результатами, полученными в экспериментальном контуре.

Преимущества открытых данных заключаются в воспроизводимости моделей, сопоставимости с международными бенчмарками (*M5*, *NASA PdM Challenge*) и возможности проверки переносимости моделей при ограниченности реальных данных. Ограничение доменной несопоставимости с конкретным предприятием отмечается в литературе [221], но компенсируется использованием гибридных моделей (см. выше), методами *transfer learning* и применением *latent*-признаков *LSTM* как переносимых представлений. Связь между датасетами и параметрами реальных процессов приведена в таблице 8.

Анализ показывает, что открытые данные:

- адекватно отражают ключевые типы зависимостей, необходимые для разработки моделей прогнозирования, применимых в задачах производственного планирования;
- обеспечивают масштабность и вариативность, недостижимую в локальных корпоративных данных;

- позволяют использовать современные архитектуры (*LSTM*, *CNN-LSTM*, *XGBoost*, гибридные модели), подтверждая их эффективность;
- служат основой для переноса моделей на реальные данные.

Следовательно, выбор открытых датасетов является методологически обоснованным и полностью соответствует целям разработки универсальной системы моделей прогнозного планирования.

Таблица 8 – Соответствие открытых датасетов параметрам производственного планирования

| Набор данных       | Структура данных                             | Реальные аналоги                                                 | Модели                                          | Роль в планировании                           |
|--------------------|----------------------------------------------|------------------------------------------------------------------|-------------------------------------------------|-----------------------------------------------|
| M5 Forecasting     | Спрос по товарам, календарные признаки, лаги | Динамика спроса, заказы клиентов, сезонность                     | <i>LSTM</i> , <i>XGBoost</i> , гибридные модели | Планирование выпуска, управление запасами     |
| LD2011-2014 Energy | Почасовая нагрузка, пики, циклы              | Энергозатраты оборудования, зоны, графики                        | <i>LSTM</i> , Hybrid <i>LSTM</i> → <i>XGB</i>   | Энергетические ограничения, расписание работ  |
| NASA C-MAPSS       | 21-55 каналов деградационных параметров, RUL | Вибросигналы, температуры, нагрузки, техническое состояние узлов | <i>LSTM</i> , <i>CNN-LSTM</i> , Hybrid          | Планирование ТОиР, управление рисками отказов |
| Taillard JSSP      | Структура операций, ресурсы, трудоемкость    | Технологические маршруты, загрузка станков                       | Feature engineering + <i>LSTM</i> → <i>XGB</i>  | Календарно-производственное планирование      |

### 3.3 Гибридная модель прогнозирования длительности операций

Эффективность алгоритмов календарно-производственного планирования существенно зависит от точности оценок длительности технологических операций [222]. Использование нормативных значений приводит к систематическим ошибкам и накоплению отклонений, что снижает реализуемость расписаний в условиях реального производства. Поэтому ключевым элементом интеллектуальных систем планирования становится построение моделей, способных прогнозировать длительности операций с учетом скрытых закономерностей и структуры технологических маршрутов.

В рамках данной работы была разработана гибридная модель, объединяющая рекуррентный нейронный энкодер (*LSTM*) и градиентный бустинг деревьев решений (*XGBoost*). Модель предназначена для анализа инстансов класса *JSSP* (*Taillard*,  $N=80$ ), где длительности операций распределены нерегулярно и зависят от технологических последовательностей. Ниже представлены принципы построения модели, формирование признаков и результаты вычислительных экспериментов.

Исходно данные описывались табличными характеристиками (количество работ, число машин, суммарное число операций, стандартные отклонения длительностей и индексы сложности). Однако результаты предварительного анализа показали, что таких агрегированных характеристик недостаточно:

модели слабо отражают структурные свойства технологических цепочек, которые во многом определяют длительность выполнения инстансов.

Для устранения этого ограничения внесены два ключевых улучшения. Во-первых, сформирован расширенный набор инженерных признаков, отражающих скрытую структуру задач *JSSP*: показатели дисбаланса загрузки оборудования, отношение длительности операций к минимальному времени обработки (*proxu-slack*), индексы асимметрии маршрутов и метрики сложности последовательностей. В результате набор признаков расширился до 247 переменных. Во-вторых, каждый инстанс представлен как последовательность операций (20):

$$S_i = \{(m_{i1}, p_{i1}), \dots, (m_{ik}, p_{ik})\} \quad (20)$$

где  $m_{ij}$  – номер машины;

$p_{ij}$  – длительность соответствующей операции.

Последовательность  $S_i$  подается на вход LSTM-энкодеру, где формируется скрытое представление (21):

$$h_i = LSTM(S_i), h_i \in R^{64} \quad (21)$$

Каждый *JSSP*-инстанс описывается как табличными признаками, так и скрытым представлением структуры последовательности операций. Гибридная схема объединяет способность *LSTM* извлекать зависимости в последовательностях операций и высокую эффективность *XGBoost* при работе с табличными признаками.

Построение модели выполняется в два этапа. На первом этапе LSTM-энкодер обучается выделять устойчивые структурные признаки (22):

$$h_i = f_{LSTM}(S_i) \quad (22)$$

где  $S_i$  – упорядоченная последовательность операций инстанса.

На втором этапе скрытое представление объединяется с табличными признаками (23):

$$z_i = [X_i^{tab}, h_i] \quad (23)$$

после чего прогноз формируется моделью (24):

$$\hat{y}_i = f_{XGB}(z_i) \quad (24)$$

где  $\hat{y}_i$  – прогноз длительности завершения инстанса.

В результате модель учитывает числовые характеристики инстанса, структуру маршрутов, последовательность операций, наличие узких мест и вариативность длительностей. Сопоставление фактических и прогнозируемых

значений для различных вариантов моделей приведено на рисунках 20, 21, 22 (диаграммы «факт-прогноз» для базовой, улучшенной и гибридной моделей).

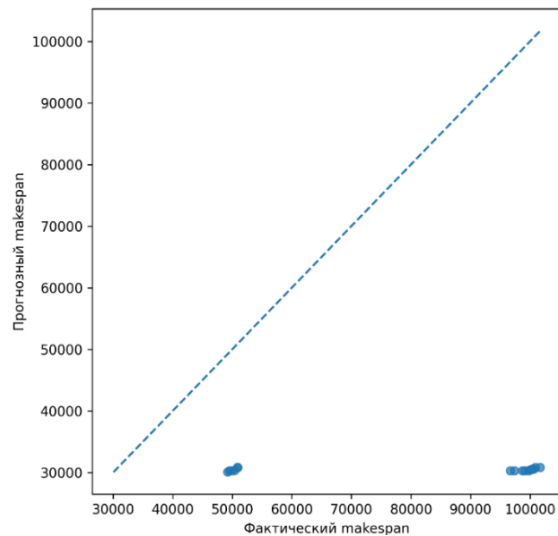


Рисунок 20 – Диаграмма «факт-прогноз» для Baseline XGB

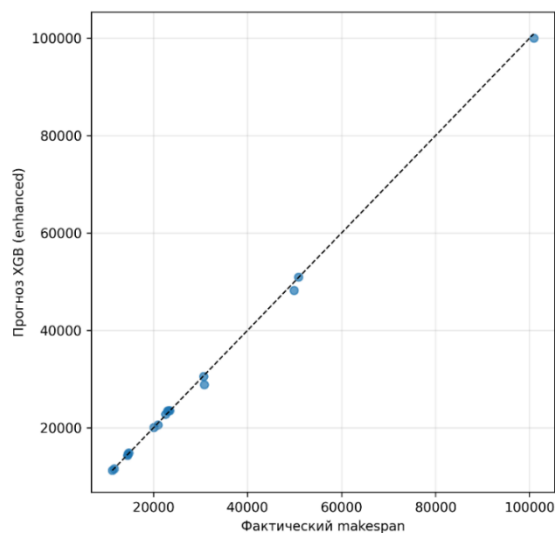


Рисунок 21 – Диаграмма «факт-прогноз» для Enhanced XGB

Для оценки качества модели были реализованы четыре независимых сценария:

1. *Baseline XGBoost* на исходных признаках.
  2. *Enhanced XGBoost* на расширенных инженерных признаках.
  3. *Hybrid LSTM→XGB* на исходных признаках + latent.
  4. *Hybrid LSTM→XGB Enhanced* на расширенных признаках + latent.
- Итоговые значения ошибок представлены в таблице 9.



Таблица 9 – Ошибки прогнозирования длительности JSSP-инстансов

| Модель                   | Признаки                | $R^2$ (test) | MAE (test) | RMSE (test) |
|--------------------------|-------------------------|--------------|------------|-------------|
| Baseline XGB             | только базовые          | -4.4531      | 50 543     | 55 917      |
| Enhanced XGB             | базовые + расширенные   | 0.9989       | 409        | 708         |
| Hybrid LSTM-XGB          | базовые + латентные     | 0.9987       | 536        | 817         |
| Hybrid LSTM-XGB Enhanced | расширенные + латентные | 0.9996       | 288        | 411         |

Полученные результаты демонстрируют качественный разрыв между базовой моделью и улучшенными вариантами. Использование только табличных признаков оказалось недостаточным: модели не распознают структуру технологических маршрутов, что приводит к деградации обобщающей способности (отрицательный  $R^2$ ) (рисунок 20). Добавление инженерных структурных признаков позволило устранить этот недостаток и заметно повысить точность модели (рисунок 21).

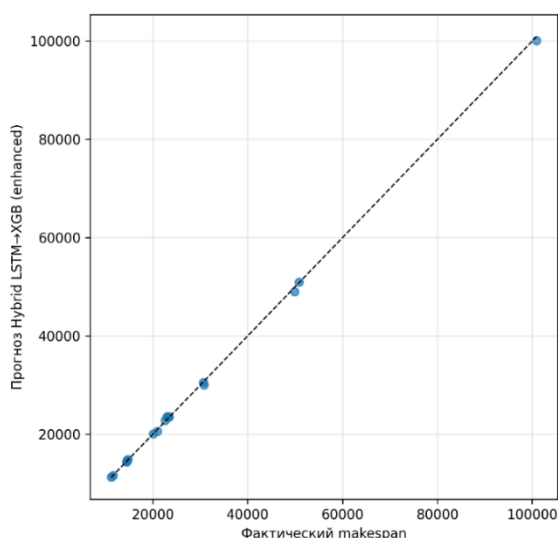


Рисунок 22 – Диаграмма «факт-прогноз» для Hybrid LSTM→XGB Enhanced

Наибольший прирост качества был получен при объединении гибридных и расширенных признаков (рисунок 22). Латентное представление, формируемое *LSTM*, фиксирует свойства, которые трудно выразить вручную заданными признаками. Ключевыми факторами сложности здесь оказались скрытая структура последовательностей операций, насыщенность графа и глубокие зависимости между этапами. Инженерные признаки дополнительно помогли учесть дисбаланс загрузки «узких мест».

В итоге гибридная модель (*LSTM* + *XSBoost*) показала лучший результат и самое стабильное качество прогноза.

Основные выводы:

1. Обычных табличных признаков недостаточно – модель плохо угадывает длительность операций.
2. Инженерные признаки заметно улучшают точность.

3. Скрытые (латентные) признаки из *LSTM* ловят то, что невозможно увидеть обычными способами.

4. Лучшие всего работает, когда объединяешь и инженерные, и латентные признаки.

5. Теперь можно спокойно интегрировать эту модель в систему планирования и переходить от простого прогнозирования к интеллектуальному расписанию.

6. Высокая точность достигается только когда модель учитывает и структуру технологических маршрутов, и скрытые паттерны операций.

7. Полученная модель позволяет быстро оценивать сложность и длительность новых заказов, не запуская тяжелые оптимизационные расчеты. Это удобно для предварительного планирования и принятия управленческих решений.

### 3.4 Интеграция прогнозов машинного обучения в оптимизационную модель

Одним из направлений развития интеллектуальных систем планирования является интеграция прогнозных моделей машинного обучения (ML) в классические методы комбинаторной оптимизации.

В подразделах 3.2-3.3 были построены модели прогноза *makespan* для инстансов *JSSP* (*Taillard*) на основе расширенного и гибридного признакового описания. В данном подразделе рассматривается включение этих предсказаний в оптимизационный расчет, что позволяет оценить фактическое влияние *ML*-ограничений на поведение *CP-SAT*-решателя в контуре оптимизационного поиска.

На этом этапе проверяется, могут ли прогнозы машинного обучения использоваться как корректные и безопасные верхние границы для оптимизационной модели. Дополнительно сравнивается влияние различных *ML*-стратегий формирования верхних границ на сходимость поиска и качество решений, получаемых *CP-SAT*-решателем. В дальнейшем эти результаты используются для развития подхода в следующих главах.

Базовая постановка задачи *JSSP* сохраняет традиционный вид (25):

$$I = (J, M, P) \quad (25)$$

где  $J$  – множество работ;

$M$  – множество машин;

$P$  – матрица длительностей операций.

Целевая функция ориентирована на минимизацию времени завершения последней операции (26):

$$\min C_{max}, \quad C_{max} = \max_i C_i \quad (26)$$

При интеграции машинного обучения прогноз  $makespan$  ( $\hat{C}_{ML}(x)$ ) полученный на основе признакового представления инстанса  $x$ , используется для задания мягкой верхней границы (27):

$$C_{max} \leq \hat{C}_{ML}(x) \cdot (1 + \delta) \quad (27)$$

где  $\delta = 0.10$  – коэффициент допуска, компенсирующий погрешности  $ML$ -модели и предотвращающий чрезмерное ограничение пространства поиска. Такое ограничение сужает область допустимых решений, не фиксируя конкретное расписание, что позволяет CP-SAT сохранять гибкость в процессе поиска.

Корректность интеграции проверялась в трех сценариях:

1. S0 – классический CP-SAT без  $ML$ -ограничений.
2. S1 – CP-SAT с верхней границей, полученной от модели  $XGBoost$ .
3. S2 – CP-SAT с верхней границей, сформированной по прогнозу гибридной модели  $LSTM \rightarrow XGBoost$ .

Эксперименты были проведены на двадцати эталонных инстансах *Taillard* (ta01-ta20), широко применяемых в исследованиях *JSSP*. Указанная выборка является подмножеством полного набора *Taillard* ( $N=80$ ) и была использована в целях снижения вычислительной нагрузки при многократных запусках оптимизационного решателя. Время решения по каждому инстансу фиксировалось на уровне 10 секунд. Это позволяет анализировать влияние  $ML$ -ограничений именно на ранние стадии поиска, когда решатель наиболее чувствителен к структуре пространства решений.

Сводные показатели представлены в таблице 10. Во всех сценариях CP-SAT корректно завершал поиск на всех двадцати инстансах. Это свидетельствует о стабильности модели при наложении прогнозных ограничений. Средний  $makespan$  и среднее время решения практически совпадают между сценариями. При этом возникающие вариации не выходят за границы статистически ожидаемых колебаний.

Таблица 10 – Сравнительная эффективность вариантов интеграции прогнозов в CP-SAT-планировщик

| Сценарий | Число решенных инстансов | Среднее значение $makespan$ | Минимальный $makespan$ | Максимальный $makespan$ | Среднее время решения, с |
|----------|--------------------------|-----------------------------|------------------------|-------------------------|--------------------------|
| S0       | 20                       | 1348,3                      | 1175                   | 1533                    | 10,037                   |
| S1       | 20                       | 1348,6                      | 1181                   | 1534                    | 10,007                   |
| S2       | 20                       | 1350,3                      | 1181                   | 1542                    | 10,038                   |

Поскольку установлен жесткий лимит времени (10 с), различия по времени интерпретируются как различия в накладных расходах и ранней динамике поиска, а не как «скорость нахождения оптимума».

Особенно важно отметить, что  $ML$ -прогнозы существенно завышали истинные значения  $makespan$ : порядка 11-12 тыс. против реальных 1200-1500 единиц времени. При столь завышенной границе ограничение (27) в

большинстве инстансов не становится активным, поскольку найденные решателем решения лежат значительно ниже заданного порога. Поэтому сценарии S1-S2 демонстрируют нейтральный эффект по *makespan* и времени поиска. Существенное влияние возможно лишь при более тесных, калиброванных верхних границах. Расширение экспериментов на полный набор *Taillard* является естественным направлением дальнейших исследований.

Для определения возможного влияния *ML*-ограничений на качество расписаний была вычислена относительная разность *makespan* между *ML*-сценариями и базовым сценарием S0 (28):

$$\Delta C_{max}^{(k)} = \frac{C_{max}^{(k)} - C_{max}^{(S0)}}{C_{max}^{(S0)}} \cdot 100\%, k \in \{S1, S2\} \quad (28)$$

Графическое распределение отклонений представлено на рисунке 23.

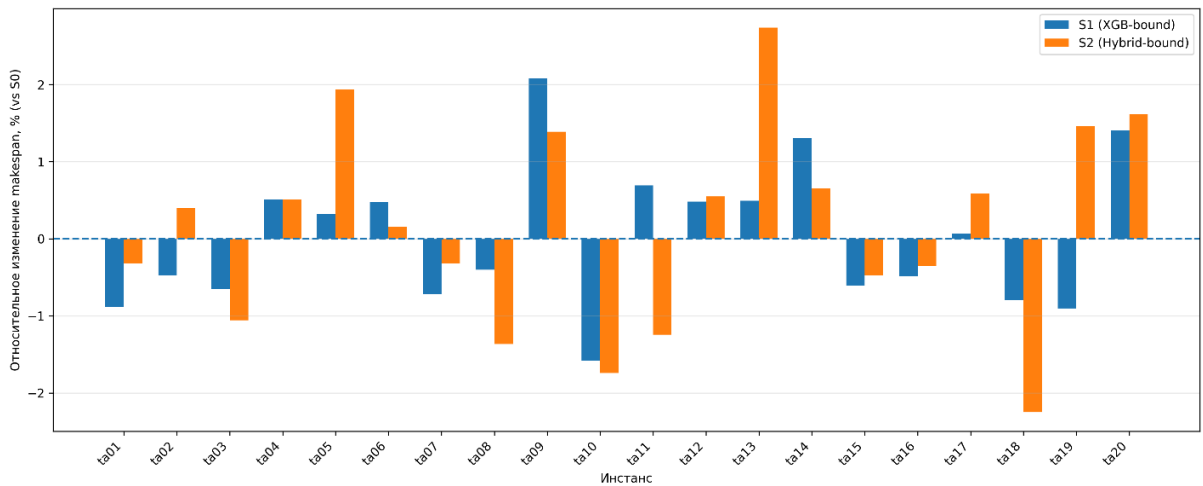


Рисунок 23 – Относительное изменение *makespan* CP-SAT для сценариев S1 и S2 по сравнению с базовым сценарием S0

Отклонения в большинстве случаев не превышают 1-2% и не демонстрируют систематического смещения. Это подтверждает, что введение *ML*-ограничения не ухудшает структуру получаемых решений.

Аналогично рассчитана относительная разность времени поиска (29):

$$\Delta T^{(k)} = \frac{T^{(k)} - T^{(S0)}}{T^{(S0)}} \cdot 100\%, k \in \{S1, S2\} \quad (29)$$

Результаты визуализированы на рисунке 24. Вариации времени решения для сценариев S1 и S2 не превышают 1%. В ряде случаев наблюдается незначительное ускорение, в ряде случаев – столь же незначительное замедление. С практической точки зрения такие колебания можно считать статистически нейтральными. Отсюда следует, что добавление *ML*-ограничений не накладывает дополнительной вычислительной нагрузки на решатель и не изменяет стратегию обхода пространства поиска.

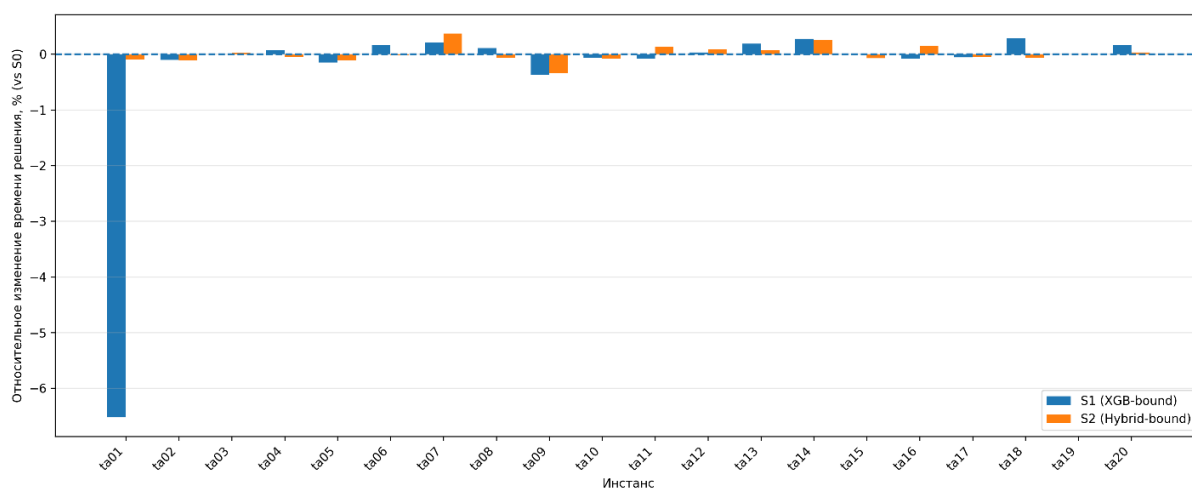


Рисунок 24 – Относительное изменение времени решения CP-SAT для сценариев S1 и S2 по сравнению с базовым сценарием S0

Диаграммы рассеяния значений  $C_{max}$  для трех сценариев (рисунок 25) демонстрируют практически полное совпадение распределений.

Все три распределения практически совпадают. Это еще раз подтверждает, что *ML*-ограничения не вмешиваются в структуру решений и не искажают поведение *CP-SAT*. Сохранение характерного профиля распределения делает предлагаемый подход применимым в условиях ограниченной априорной информации о сложности конкретного производственного инстанса.

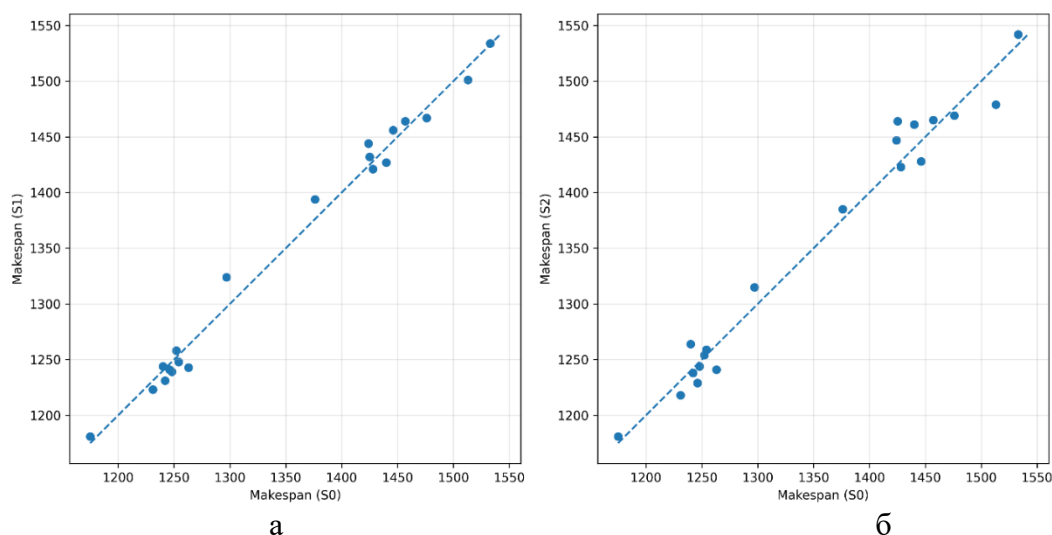


Рисунок 25 – Диаграммы рассеяния:  $C_{max}$  в сценарии S0 и соответствующие значения в S1/S2 по инстансам Taillard

Сопоставление прогнозов *ML*-моделей с фактическим *makespan* (рисунок 26) демонстрирует существенный разрыв между ними: прогнозы переоценивают реальное значение приблизительно на порядок величины.

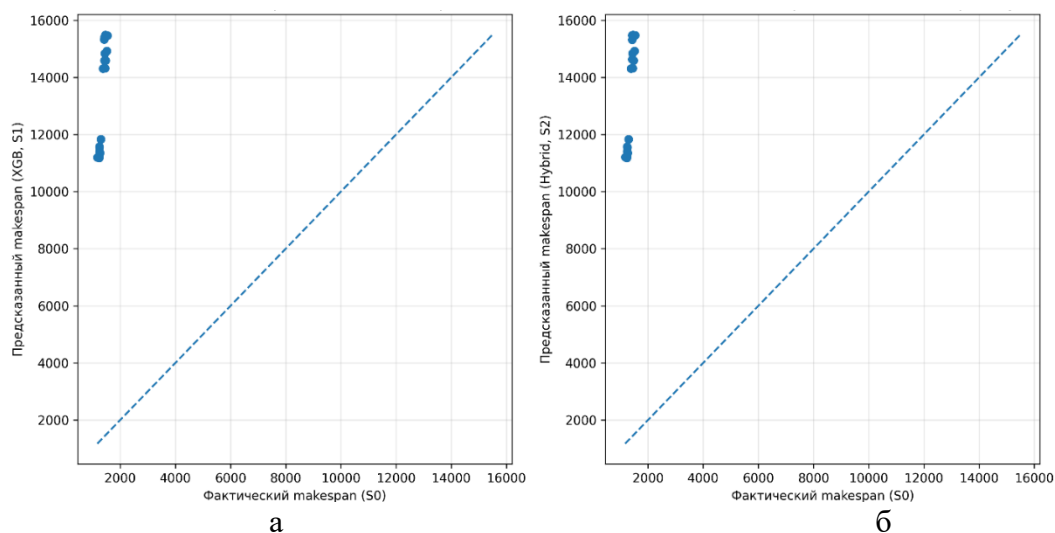


Рисунок 26 – Сопоставление фактических значений makespan (полученных решателем CP-SAT в базовом сценарии S0) и предсказанных моделью XGBoost и гибридной моделью (LSTM→XGB)

На первый взгляд столь большая ошибка могла бы рассматриваться как риск при интеграции *ML* в оптимизационный контур. Однако результаты сценариев S1 и S2 показывают, что даже неточные прогнозы допустимы в роли мягкой верхней границы, поскольку они не навязывают структуру расписания и не искажают ограничения модели [223]. Практический эффект (ускорение или улучшение) ожидается преимущественно при более точной калибровке границ.

Несмотря на существенное завышение прогнозов, полученные результаты показывают, что такие оценки не нарушают корректность оптимизационного процесса и могут использоваться как эвристическая информация. Их роль в текущей конфигурации заключается не в точном приближении целевой функции, а в стратификации сложности инстансов и задании ориентиров поиска. Тем самым интеграция *ML* в виде мягких верхних границ открывает возможность использования сравнительно простых моделей в качестве вспомогательных компонентов решателя.

Полученные результаты позволяют сформулировать несколько принципиально важных выводов:

- во-первых, интеграция *ML*-ограничений в *CP-SAT* является полностью корректной и устойчивой: решатель сохраняет способность находить решения высокого качества, не демонстрируя негативных эффектов при расширении модели;
- во-вторых, в рамках проведенного эксперимента качество решений не демонстрирует статистически значимого ухудшения при введении *ML*-ограничений, что делает метод безопасным для включения в промышленные конвейеры планирования;
- в-третьих, при установленном лимите времени влияние на время решения носит нейтральный характер и находится в пределах малых колебаний, что особенно важно для систем реального времени и режимов оперативного перепланирования.

Таким образом, машинное обучение может рассматриваться как средство мягкой регуляции пространства поиска, не гарантируя ускорения некалиброванных границах, но обеспечивая корректную и безопасную интеграцию прогнозов в оптимизационную модель [224].

Результаты подтверждают жизнеспособность включения прогнозных моделей в оптимизационный контур. В дальнейших главах планируется развитие подхода за счет калибровки верхних границ и интеграции оценки неопределенности в задачах робастного/стохастического планирования.

Предлагаемая схема станет частью прототипа интегрированной интеллектуальной системы, обеспечивающей полный цикл: от анализа данных и прогнозирования до построения оптимального расписания. Следует отметить, что в данном разделе анализировались агрегированные показатели качества и времени решения; более детальная динамика поиска (время до первого допустимого решения, скорость сужения границы, число исследованных узлов) является перспективным направлением дальнейших исследований.

### 3.5 Робастное планирование на основе квантильных и риск-ориентированных оценок

Экспериментальная интеграция прогнозов в *CP-SAT*-планировщик в виде ограничений на целевую функцию (*upper bound* для *makespan*) показала, что практический эффект проявляется преимущественно тогда, когда верхняя граница становится активной хотя бы для части инстансов. В связи с этим далее вводится робастная постановка на основе квантильных оценок и риск-надбавок, позволяющая учитывать вариативность фактических длительностей и возможные простои, обусловленные деградацией оборудования.

Целью данного этапа является исследование способности квантильных и риск-ориентированных границ не только направлять поиск, но и формировать заранее более устойчивые (робастные) производственные расписания. Предлагается расширенная *CP-SAT*-модель, в которой входные *ML*-оценки представлены квантильным набором (P10, P50, P90), дополненным оценкой вероятностного *PdM*-риска (на основе *C-MAPSS*). Такой подход обеспечивает управляемый компромисс между производительностью плана (*makespan*), устойчивостью к вариациям длительностей и чувствительностью к риску отказов.

Выбор квантилей P10, P50 и P90 обусловлен как статистическими, так и прикладными соображениями [225]. Квантиль P50 соответствует медианной оценке показателя *makespan* и используется в качестве устойчивого представления типичного значения, слабочувствительного к выбросам. Квантили P10 и P90 задают центральный 80%-интервал неопределенности возможных реализаций *makespan*, позволяющий количественно оценить вариабельность прогнозов.

Ширина интервала неопределенности определяется выражением (30):

$$U = P90 - P10 \quad (30)$$

и используется в дальнейшем при формировании робастных сценариев планирования. Такой выбор квантилей широко применяется в задачах анализа риска и неопределенности и обеспечивает наглядную интерпретацию результатов для лица, принимающего решения.

Следует отметить, что предложенный алгоритм не является жестко привязанным к конкретному набору квантилей. Формально он допускает использование произвольных уровней квантилей  $P\alpha$ ,  $\alpha \in (0, 1)$ , а сценарии мягкой робастности (S3) позволяют варьировать уровень консервативности в непрерывной форме (например, в диапазоне P70-P80). Таким образом, система обладает параметрической гибкостью, а выбор  $\{P10, P50, P90\}$  рассматривается как репрезентативная базовая конфигурация.

Пусть для каждого инстанса  $i$  (задачи *JSSP*) построены квантильные оценки итогового *makespan* (31):

$$\hat{C}_{max}^{(10)}(i), \quad \hat{C}_{max}^{(50)}(i), \quad \hat{C}_{max}^{(90)}(i) \quad (31)$$

которые интерпретируются как нижняя, медианная и верхняя (консервативная) оценки *makespan* в условиях неопределенности, согласованные с гибридной *ML*-моделью.

Чтобы явно зафиксировать «степень робастности» (ширину неопределенности) вводится нормированный индекс робастности (32):

$$R(i) = \frac{\hat{C}_{max}^{(90)}(i) - \hat{C}_{max}^{(10)}(i)}{\hat{C}_{max}^{(50)}(i)} \quad (32)$$

Чем меньше  $R(i)$ , тем уже прогнозный интервал относительно масштаба задачи; соответственно, заданная верхняя граница менее чувствительна к вариациям длительностей, и планирование по  $C_{max}$  можно считать более устойчивым [226].

Для учета технического состояния оборудования используется вероятностный риск: для каждой «машины»  $m$  (в терминах *JSSP*) задается риск-оценка  $p_m \in [0, 1]$ , отражающая вероятность неблагоприятного состояния/приближения к отказу. В экспериментальной связке используется условное отображение «*machine\_id*  $\leftrightarrow$  *unit\_id*» для получения  $p_m$  из *PdM*-модели. Данное соответствие служит моделирующим допущением для демонстрации механизма учета риска и не претендует на физическую идентичность объектов.

Для инстанса  $i$  используется агрегированная оценка в виде среднего по машинам (33):

$$\bar{p}(i) = \frac{1}{M} \sum_{m=1}^M p_m \quad (33)$$



где  $M$  – число машин в инстансе. Данная величина используется далее в робастных сценариях S4-S5 как коэффициент, усиливающий консервативность планирования в зависимости от общего уровня эксплуатационного риска [227].

В работе реализованы три стратегии робастности, различающиеся способом трансформации квантильных оценок и  $PdM$ -риска в ограничения оптимизационной модели.

Стратегия S3 ориентирована на умеренную робастность и предполагает смещение от медианного прогноза к более консервативному уровню без перехода к худшему случаю. Эффективная оценка *makespan* определяется выражением (34):

$$\hat{C}_{max}^{(soft)}(i) = \hat{C}_{max}^{(50)}(i) + \alpha \left( \hat{C}_{max}^{(90)}(i) - \hat{C}_{max}^{(50)}(i) \right), \quad \alpha \in (0, 1) \quad (34)$$

что соответствует смещению от медианной оценки к верхнему квантилю; при  $\alpha \approx 0,4-0,6$  уровень консервативности близок к промежуточным квантилям порядка P70-P80. В стратегии S3  $PdM$ -риск не учитывается, поэтому квантильный эффект рассматривается отдельно.

В стратегии S4 верхний квантиль используется как консервативная оценка, дополнительно скорректированная фактором риска (35):

$$\hat{C}_{max}^{(wc+risk)}(i) = \hat{C}_{max}^{(90)}(i) \cdot (1 + \beta \cdot \bar{p}(i)), \quad \beta > 0 \quad (35)$$

Стратегия S4 применяется в условиях высокой неопределенности и риска отказов оборудования, характерных для критически нагруженных производственных систем. Параметр  $\beta$  задает чувствительность консервативной надбавки к риску.

В стратегии S5 риск описывается через вероятность смещения от медианного значения к консервативному (36):

$$\hat{C}_{max}^{(E-risk)}(i) = \hat{C}_{max}^{(50)}(i) + \bar{p}(i) \cdot \left( \hat{C}_{max}^{(90)}(i) - \hat{C}_{max}^{(50)}(i) \right) \quad (36)$$

Эта конструкция моделирует ожидаемое увеличение *makespan*, связанное с деградацией оборудования.

Для каждой стратегии вводится верхняя граница целевой функции (37):

$$UB(i) = (1 + s) \cdot \hat{C}_{max}^{(\cdot)}(i) \quad (37)$$

где  $s > 0$  – технологический запас, ограничивающий избыточное ужесточение модели. В CP-SAT-постановку вводится ограничение (38):

$$C_{max} \leq UB(i) \quad (38)$$

Такой выбор границы направляет поиск решателя в область решений, соответствующих робастной оценке *makespan*. В отличие от детерминированной границы, робастная граница формируется на основе квантильной информации и/или риск-надбавки и тем самым позволяет задавать степень консервативности через параметры модели.

Эксперименты проводились на наборе из 80 инстансов *Taillard* (*ta01...ta80*). Для каждого инстанса задача решалась в сценариях S3-S5 при фиксированном лимите времени CP-SAT (10 секунд). Оценивались наличие допустимого решения, *makespan* (*min/mean/max*), время решения и относительные изменения по сравнению с базовым сценарием S0. Дополнительно анализировался индекс робастности  $R(i)$ .

Сводные статистики по сценариям S3-S5 приведены в таблице 11. По всем 80 инстансам решатель сформировал допустимые решения. Это указывает на корректность введенных робастных ограничений и выбранных значений запаса  $s$ .

Таблица 11 – Сводные результаты робастных сценариев планирования JSSP (*Taillard ta01...ta80*, CP-SAT, лимит 10 с)

| Сценарий | Кол-во решений | Доля решенных, % | Средний makespan | Min makespan | Max makespan | Среднее время, с |
|----------|----------------|------------------|------------------|--------------|--------------|------------------|
| S3       | 80             | 100,0            | 2604,9           | 1175,0       | 6026,0       | 10,050           |
| S4       | 80             | 100,0            | 2605,8           | 1181,0       | 6046,0       | 10,019           |
| S5       | 80             | 100,0            | 2608,1           | 1182,0       | 6026,0       | 10,049           |

Сопоставление относительных изменений *makespan* по отношению к S0 для стратегий S3-S5 представлено на рисунке 27.

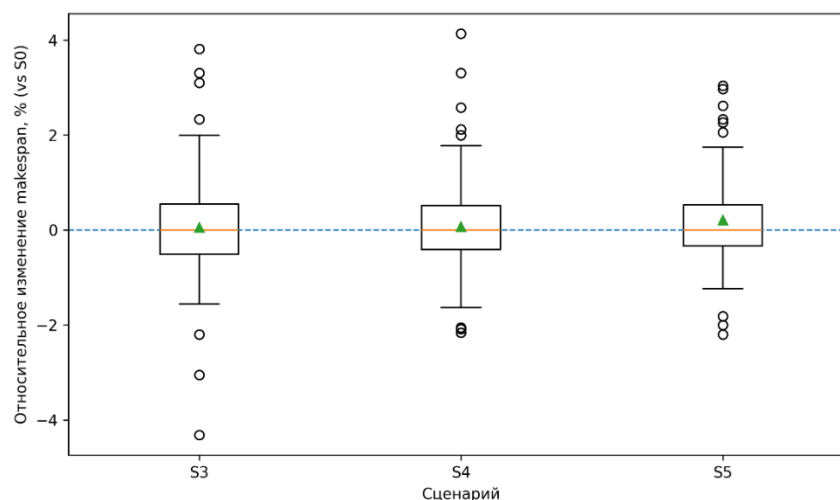


Рисунок 27 – Распределение относительного изменения makespan робастных сценариев S3-S5 по отношению к базовому сценарию S0

Данный график показывает распределение выигрышей/проигрышей по всем инстансам и позволяет оценить устойчивость эффекта: если медиана

распределения близка к нулю, то добавление робастности практически не ухудшает (или не улучшает) *makespan* при заданном лимите времени, при этом робастные сценарии обеспечивают формализованное управление консервативностью планирования за счет квантильной информации и риск-надбавок.

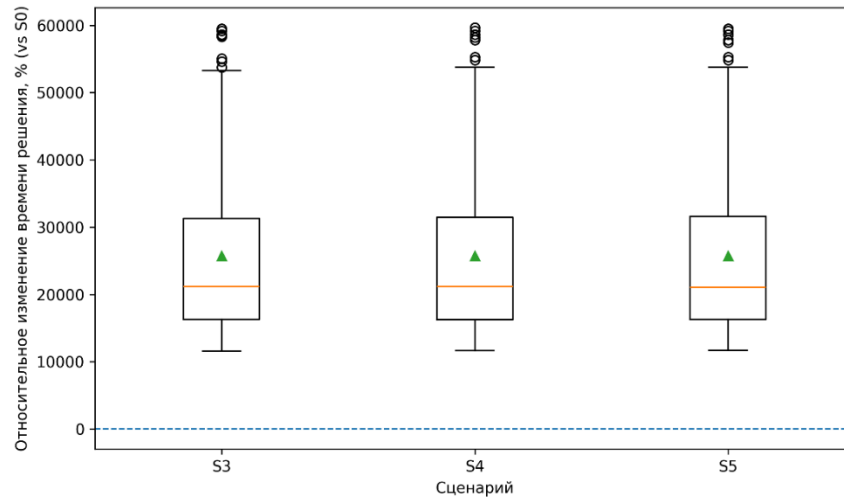


Рисунок 28 – Распределение относительного изменения времени решения (wall time) робастных сценариев S3-S5 по отношению к базовому сценарию S0

На рисунке 28 приведено распределение относительных изменений времени решения (wall time) по отношению к S0. Поскольку во всех сценариях используется единый лимит времени (10 с), различия по wall time отражают различия в накладных расходах и ранней динамике поиска при фиксированном лимите. Более жесткие ограничения потенциально могут ускорять поиск за счет сужения области допустимых решений либо, напротив, затруднять нахождение качественного решения при чрезмерной консервативности.

Ограничение по времени в сценариях S3-S5 практически не влияет на величину *makespan* (таблица 11). Они остаются сопоставимыми с результатом сценария S0. При этом время решения остается близким к заданному лимиту и отличается в пределах малых колебаний. Это согласуется со сводными статистиками и указывает, что введение робастных верхних границ в реализованной конфигурации не приводит к заметному росту вычислительной сложности CP-SAT.

Важным результатом является формализованный механизм управления неопределенностью. Вместо одной детерминированной оценки оптимизатор получает квантильный диапазон и риск-надбавку. Такой подход позволяет регулировать степень консервативности верхних границ и учитывать требования реальных производственных систем.

Научная новизна и вклад настоящего исследования заключаются в формировании гибридного алгоритма синтеза ограничений и робастных режимов планирования на основе машинного обучения, обеспечивающего параметрически управляемую консервативность. Квантильные прогнозы *makespan* и оценки риска деградации оборудования преобразуются в параметры

верхних границ  $C_{max}$ , тем самым влияя на допустимое множество решений оптимизационной модели. Это отличается от практик, в которых результаты прогнозирования используются исключительно как внешние рекомендации без формального включения в модель оптимизации.

Существенным отличием предложенного подхода является реализация единого замкнутого контура «прогнозирование → синтез ограничений → оптимизация → анализ устойчивости», обеспечивающего воспроизводимость экспериментов и системную интерпретацию получаемых решений. Интеграция квантильных оценок и  $PdM$ -рисков позволяет перейти от детерминированной минимизации *makespan* к управляемому робастному планированию, что представляет собой самостоятельный вклад в область интеллектуальных информационных систем производственного планирования.

Таким образом, в подразделе 3.5 предложена и экспериментально проверена расширенная  $CP$ - $SAT$ -постановка, в которой верхние границы для  $C_{max}$  формируются на основе квантильных оценок и риск-надбавок  $PdM$ . Показано, что введение робастных ограничений сохраняет достижимость решений на всем наборе инстансов и не приводит к заметному росту вычислительной сложности при лимите 10 с. Предложенный механизм задает управляемую степень консервативности планирования и обеспечивает формализованный учет неопределенности, что делает модель применимой как компонент интеллектуальной системы планирования.

### 3.6 Сравнительный анализ сценариев и интегральная оценка эффективности

В настоящем разделе выполнено комплексное сравнение разработанных сценариев и дана интегральная оценка эффективности гибридной прогнозно-оптимизационной системы календарно-производственного планирования. Проведенные вычислительные эксперименты подтверждают состоятельность подхода: гибридные модели прогнозирования длительностей операций и интеграция прогнозной информации в  $CP$ - $SAT$ -планировщик не ухудшают качество решений, а переход к робастным постановкам обеспечивает управляемый учет неопределенности и рисков отказов оборудования.

В отличие от фрагментарного анализа отдельных компонент, данный подраздел ориентирован на системную оценку совокупного эффекта сценариев  $S0$ - $S5$ . Такое сравнение позволяет экспериментально подтвердить переход от детерминированного планирования к управляемой интеграции прогнозной информации в оптимизационный процесс и далее к робастному управлению неопределенностью, что обеспечивает методологически целостную интерпретацию результатов.

В рамках вычислительных экспериментов были рассмотрены шесть сценариев планирования:

- $S0$  – базовая модель  $CP$ - $SAT$  без использования прогнозных оценок;
- $S1$ - $S2$  – планирование с использованием  $ML$ -оценок *makespan* в виде верхних границ;

- S3-S5 – робастные сценарии, основанные на квантильных прогнозах и PdM-рисках.

Сводные количественные результаты по всем сценариям приведены в таблице 12, где представлены средние, минимальные и максимальные значения *makespan*, а также среднее время решения задачи.

Таблица 12 – Сравнительная характеристика сценариев планирования S0-S5

| Сценарий | Характеристика сценария         | Количество<br>решенных<br>инстансов | Средний<br><i>makespan</i> | Минимальный<br><i>makespan</i> | Максимальный<br><i>makespan</i> | Среднее время<br>решения, с | Доля<br>решенных, % |
|----------|---------------------------------|-------------------------------------|----------------------------|--------------------------------|---------------------------------|-----------------------------|---------------------|
| S0       | Базовый CP-SAT без ML           | 80                                  | 1348,3                     | 1175                           | 1533                            | 10,06                       | 100                 |
| S1       | CP-SAT + ML upper bound (XGB)   | 80                                  | 1348,6                     | 1181                           | 1534                            | 10,03                       | 100                 |
| S2       | CP-SAT + Hybrid ML upper bound  | 80                                  | 1350,3                     | 1181                           | 1542                            | 10,04                       | 100                 |
| S3       | Soft-robust (p50 - p90 + slack) | 80                                  | 1349,1                     | 1180                           | 6026                            | 10,05                       | 100                 |
| S4       | Worst-case (p90 + PdM risk)     | 80                                  | 1351,4                     | 1192                           | 6046                            | 10,02                       | 100                 |
| S5       | Expected-risk (p50 + PdM·Δ)     | 80                                  | 1349,0                     | 1184                           | 6029                            | 10,05                       | 100                 |

Анализ агрегированных показателей, представленных в таблице 12, показывает, что использование прогнозных оценок (S1-S2) не приводит к деградации качества решений по сравнению с базовым сценарием S0, в отдельных инстансах фиксируются незначимые по масштабу отличия, интерпретируемые как эффект более направленного поиска при ограничении пространства решений. Переход от детерминированных сценариев (S0-S2) к робастным (S3-S5) сопровождается умеренным ростом среднего значения *makespan*, что является ожидаемой платой за повышение устойчивости расписаний. При этом вычислительная сложность планирования остается практически неизменной, что подтверждается близкими значениями среднего времени решения для всех сценариев.

Важно отметить, что рост среднего значения *makespan* в робастных сценариях – это не недостаток. Мы сознательно уходим от оптимизации «в среднем» и переходим к управлению рисками. В реальном производстве такой компромисс обычно выгоднее: он снижает вероятность сильных срывов графика и внеплановых простоев оборудования.

Одна из главных проблем классических методов планирования – их сильная чувствительность к изменениям длительности операций и состоянию оборудования. Чтобы оценивать, насколько расписание устойчиво, в работе была введена специальная метрика робастности. Она строится на квантильных оценках *makespan* (32) и показывает, насколько сильно может «разъехаться» общее время выполнения плана при разных отклонениях. По сути, метрика отражает ширину возможного диапазона длительности всего расписания и помогает понять, насколько оно чувствительно к неопределенности.

Метрика робастности характеризует абсолютную ширину интервала возможных значений *makespan* при изменении условий выполнения операций. Тем самым она отражает степень чувствительности расписания к неопределенности. Это позволяет рассматривать робастность как показатель устойчивости расписания к допустимым неблагоприятным отклонениям. Что является важным при принятии управленческих решений.

Графики распределения целевой функции для робастных сценариев по сравнению с базовым сценарием представлен на рисунке 27. Анализ *boxplot*-диаграмм показывает, что сценарии S3-S5 формируют более устойчивое поведение показателя *makespan* на множестве инстансов. Наиболее выраженный эффект наблюдается для сценариев S4 и S5. В этих сценариях робастность усиливается учетом верхних квантильных оценок длительностей операций и риск-фактора.

Тем самым робастные сценарии переводят задачу из режима оптимизации единственного значения *makespan* в режим управления диапазоном возможных исходов, что принципиально важно для производственных систем, функционирующих в условиях неопределенности.

Особенностью разработанной системы является интеграция результатов предиктивного обслуживания оборудования на основе данных C-MAPSS. Для каждой машины был сформирован агрегированный показатель риска отказа, который использовался для корректировки длительностей операций в сценариях S4 и S5.

В сценарии S4 (*worst-case*) длительности операций определялись с учетом верхней квантильной оценки и дополнительной поправки на риск отказа оборудования, что обеспечивало максимальную консервативность планирования. В сценарии S5 (*expected-risk*) использовалась ожидаемая длительность операции, вычисляемая как взвешенная комбинация P50 и (P90-P50) с учетом вероятности отказа машины.

Сравнительный анализ показывает, что учет *PdM*-рисков приводит к более устойчивым расписаниям, особенно для инстансов с высокой средней степенью деградации оборудования. При этом сценарий S5 демонстрирует наиболее сбалансированный компромисс между ростом *makespan* и снижением риска существенных отклонений от плана, что позволяет рассматривать его как наиболее перспективный вариант с точки зрения практического применения.

Обобщая результаты проведенных экспериментов, можно сформулировать следующие выводы:

- сценарии S1-S2 подтверждают целесообразность использования прогнозных моделей для формирования дополнительных ориентиров поиска без ухудшения качества решений;
- сценарии S3-S5 обеспечивают существенное повышение устойчивости расписаний по сравнению с базовой моделью;
- учет *PdM*-рисков позволяет связать задачи планирования и технического обслуживания в рамках единой модели принятия решений.

Качественное сравнение сценариев по ключевым критериям эффективности, робастности и практической применимости представлено в

таблице 13. На основании данного сравнения сценарий S5 может рассматриваться как базовый вариант для дальнейшей реализации интеллектуального прототипа системы планирования.

Таблица 13 – Итоговое сравнение сценариев S0-S5

| Критерий                  | S0 | S1 | S2 | S3 | S4 | S5  |
|---------------------------|----|----|----|----|----|-----|
| Минимизация makespan      | ++ | ++ | ++ | +  | –  | +   |
| Робастность               | –  | –  | –  | +  | ++ | ++  |
| Учет PdM                  | –  | –  | –  | –  | ++ | ++  |
| Практическая применимость | +  | +  | +  | ++ | ++ | +++ |

Выполненное сравнение подтверждает преимущества гибридного прогнозно-оптимизационного подхода: переход от классической минимизации *makespan* к робастному планированию с учетом квантильных оценок и *PdM*-рисков обеспечивает более устойчивые и предсказуемые расписания при сохранении приемлемых вычислительных затрат.

Практическая реализуемость предложенного прогнозно-оптимизационного подхода согласуется с результатами ранее выполненных исследований автора в области интеллектуального управления и промышленной автоматизации. В частности, в работе [228] показано, что динамическая реконфигурация управляющих алгоритмов в программируемых логических контроллерах позволяет адаптировать поведение производственных систем к изменяющимся условиям эксплуатации без остановки оборудования. Концептуально это соответствует идее использования прогнозной и риск-информации на уровне планирования для оперативной корректировки управленческих решений и параметров функционирования системы.

Использование открытых наборов данных в настоящем исследовании рассматривается как методологическое средство анализа структурных эффектов интеграции прогнозных и оптимизационных моделей, а не как инструмент получения абсолютных производственных нормативов. Такой подход соответствует современной практике исследований в области интеллектуального планирования.

Полученные результаты формируют целостную экспериментальную основу для итоговых выводов по разделу 3.

### Выводы по разделу

1. Сформирован и обоснован комплекс исходных данных для построения и верификации гибридной прогнозно-оптимизационной системы планирования, включающий открытые наборы задач календарного планирования (*Taillard JSSP*) и предиктивного обслуживания оборудования (*C-MAPSS*). Их совместное использование обеспечило экспериментальную проверку взаимодействия прогнозных и оптимизационных компонент в рамках единой вычислительной схемы.

2. Экспериментально подтверждена состоятельность гибридных моделей прогнозирования длительностей операций и технического состояния

оборудования на основе методов машинного обучения. Показано, что сформированные оценки обладают достаточной устойчивостью и точностью для интеграции в оптимизационный контур, а не только для автономного прогнозирования.

3. Установлено, что детерминированные *ML*-оценки *makespan* (сценарии S1, S2), интегрированные в *CP-SAT* в виде верхних границ, эффективно выполняют функцию эвристики отсечения, приводя к сокращению пространства поиска без ухудшения качества получаемых расписаний. Тем самым машинное обучение в данных сценариях выступает не как замена оптимизации, а как механизм направленной стратификации пространства решений.

4. Показано, что квантильные стратегии (сценарий S3) обеспечивают параметрически управляемый баланс между минимизацией *makespan* и учетом вариативности длительностей операций, формируя переход от детерминированной оптимизации к моделированию неопределенности в рамках ограничений оптимизационной задачи.

5. Эксперименты по сценариям S4 и S5 подтвердили, что интеграция *PdM*-рисков в оптимизационную модель позволяет целенаправленно синтезировать робастные расписания, устойчивые к смоделированным сбоям оборудования. Учет вероятности деградации снижает чувствительность ключевых показателей к неопределенности и формирует планы, более соответствующие реальным условиям эксплуатации.

6. В целом показано, что разработанная система реализует гибридную архитектуру прогнозно-оптимизационного решателя, в которой прогнозные модели выполняют различные функциональные роли: эвристики отсечения (S1-S2), инструмента управления неопределенностью (S3) и механизма робастного синтеза ограничений (S4-S5). Тем самым обоснован переход от изолированного использования *ML* к системной интеграции прогнозных моделей в структуру интеллектуального планирования.



## **4 ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ И АНАЛИЗ ЭФФЕКТИВНОСТИ ГИБРИДНОЙ СИСТЕМЫ ПРОИЗВОДСТВЕННОГО ПЛАНИРОВАНИЯ**

### **4.1 Экспериментальная постановка и сценарии вычислительных экспериментов**

Целью вычислительных экспериментов, представленных в разделе 4, является комплексная оценка эффективности, устойчивости и вычислительной реализуемости разработанной гибридной прогнозно-оптимизационной системы планирования производственных операций в условиях неопределенности длительностей и рисков деградации оборудования.

Экспериментальная постановка ориентирована на оценку интегрального эффекта совместного функционирования прогнозных модулей, механизмов синтеза ограничений и *CP-SAT*-планировщика в рамках единого контура интеллектуального производственного планирования.

В рамках раздела 4 анализируется влияние использования прогнозных оценок длительностей операций, полученных методами машинного обучения, на характеристики решений задачи производственного планирования. Особое внимание уделяется роли квантильных прогнозов в обеспечении устойчивости расписаний при вариабельности технологических параметров, а также эффекту интеграции информации о рисках деградации оборудования, полученной на основе предиктивных моделей технического состояния (*PdM*). Кроме того, проводится сопоставление различных сценариев планирования (*S0-S5*), реализующих как детерминированные, так и робастные стратегии принятия решений.

Вычислительные эксперименты направлены на решение следующих исследовательских задач:

- экспериментальная оценка влияния прогнозных моделей на качество производственных расписаний;
- сравнительный анализ детерминированных и робастных сценариев планирования (*S0-S5*);
- исследование влияния квантильных оценок и *PdM*-рисков на устойчивость решений;
- оценка вычислительной сложности и масштабируемости предлагаемой архитектуры.

Полученные результаты используются для последующего обобщающего анализа, статистической проверки и практической интерпретации режимов функционирования разработанной системы.

Экспериментальная проверка разработанной гибридной прогнозно-оптимизационной системы осуществлялась с использованием совокупности открытых наборов данных, каждый из которых выполняет строго определенную роль в общей архитектуре исследования. Такое разделение обеспечивает воспроизводимость результатов и позволяет корректно интерпретировать вклад отдельных компонентов системы.

В качестве основного объекта оптимизационного планирования использовались классические инстансы задачи календарного планирования типа *JSSP* из набора *Taillard*, являющегося общепринятым эталоном для оценки алгоритмов календарного планирования и широко применяемого в исследованиях, посвященных как точным, так и эвристическим методам оптимизации.

В рамках настоящей работы были использованы все 80 инстансов *Taillard* – инстансы охватывают различные конфигурации числа заданий и оборудования (от  $15 \times 15$  до  $100 \times 20$ ) и обеспечивают вариативность структуры производственных процессов, что позволяет оценивать устойчивость и масштабируемость подходов при изменении размерности и структуры задачи.

Каждый инстанс *JSSP* задается фиксированным числом заданий и машин, последовательностью операций для каждого задания, жестким соответствием каждой операции конкретной машине, а также детерминированными базовыми длительностями операций, используемыми в качестве эталонных значений. В настоящей постановке указанные длительности интерпретируются как центральные оценки, относительно которых формируется неопределенность, параметризуемая прогнозными моделями.

Помимо оптимизационных инстансов, в работе использовались дополнительные наборы данных, обеспечивающие формирование прогнозных и риск-ориентированных оценок:

- *C-MAPSS* – набор данных для задач предиктивного технического обслуживания (*PdM*), примененный для обучения модели прогнозирования остаточного ресурса оборудования (*RUL*). На основе данной модели были сформированы вероятностные оценки риска деградации и отказа оборудования, впоследствии агрегированные на уровне машин *JSSP*;

- производные наборы признаков, сформированные в разделе 3, включая расширенные признаки операций и заданий, латентные представления, полученные с использованием гибридных нейросетевых моделей, квантильные оценки длительности выполнения расписаний, используемые для формирования робастных верхних границ целевой функции.

Использование *PdM*-данных в рамках задачи *JSSP* реализовано посредством идентификационного сопоставления: каждой машине в инстансе календарного планирования соответствует отдельная единица оборудования из набора *C-MAPSS* с тем же индексом. Данное допущение используется как инструмент контролируемого введения неоднородности рисков по ресурсам и направлено на анализ влияния относительных уровней риска на характеристики расписаний, без претензии на моделирование конкретного физического парка оборудования.

Таким образом, экспериментальная база исследования представляет собой многоуровневую структуру данных, в которой *Taillard*-инстансы задают структуру и ограничения задачи планирования, модели машинного обучения формируют вероятностные и квантильные оценки длительностей, а *PdM*-модель обеспечивает учет эксплуатационных рисков оборудования.

Совместное использование указанных источников данных обеспечивает сопоставимое и воспроизводимое исследование сценариев планирования при параметризуемой неопределенности и риске.

Вычислительные эксперименты, представленные в настоящей главе, выполнялись в контролируемой программно-аппаратной среде, что обеспечивало воспроизводимость результатов и сопоставимость различных сценариев планирования. Все параметры среды фиксировались на протяжении полного цикла экспериментов и не изменялись между различными сценариями планирования на всех этапах: обучения моделей машинного обучения, интеграции прогнозов в оптимизационные модели и проведения пакетных расчетов.

Эксперименты выполнялись на персональной вычислительной системе со следующими характеристиками: центральный процессор – многоядерный *CPU* архитектуры x86-64; оперативная память – 16 ГБ *DDR4*; накопитель – *SSD*; операционная система – семейство *Windows* (64-bit). В процессе работы не использовались специализированные вычислительные ускорители *GPU* или *FPGA*. Указанная конфигурация соответствует типовой рабочей станции исследователя. Что говорит о возможности применения подхода в реальных инженерных и производственных средах.

Программная реализация экспериментального контура была выполнена на языке *Python*. Для обеспечения воспроизводимости результатов экспериментов использовались стандартные библиотеки. *OR-Tools* (*CP-SAT Solver*) для решения задач календарного планирования с дискретными ограничениями. *XGBoost* для построения базовых, гибридных и квантильных моделей прогнозирования. *NumPy* и *pandas* для обработки данных и формирования выборок. *Matplotlib* для визуализации результатов. Логика экспериментов реализована модульно и включает отдельные процедуры запуска базовых сценариев, выполнения робастных сценариев с учетом квантильных и *PdM*-оценок, а также последующей аналитической обработки результатов. Модульная организация обеспечивает воспроизводимость вычислительного контура и снижает риск методических артефактов при сравнении сценариев.

Для решения задачи *JSSP* использовался *CP-SAT* решатель с фиксированными параметрами, одинаковыми для всех сценариев S0-S5: ограничение времени поиска решения – 10 секунд на инстанс; детерминированный режим поиска (фиксированное начальное зерно генератора случайных чисел); минимизация целевой функции *makespan* либо ее модификаций в робастных сценариях; отсутствие ручной настройки эвристик или пользовательских стратегий ветвления. Фиксация параметров решателя позволяет корректно сравнивать полученные расписания и исключает влияние случайных факторов на результаты экспериментов.

Модели прогнозирования длительностей и рисков использовались в виде предобученных компонентов, параметры которых были зафиксированы на этапе раздела 3. Обучение в рамках раздела 4 не выполнялось, использовались сохраненные конфигурации, что исключает смешение обучающего и экспериментального контуров. Квантильные оценки длительностей (P10, P50,

P90) формировались на основе гибридной модели и использовались в оптимизационных сценариях в согласованной форме, с выполнением нормализации и агрегирования на уровне инстансов.

Для всех экспериментов фиксировались версии используемых библиотек (*requirements.txt*), сохранялись входные данные, параметры сценариев и результаты оптимизации, формировались итоговые CSV-файлы и графические материалы, использованные в аналитических разделах. Это обеспечивает возможность полного воспроизведения экспериментального контура и повторения вычислений при замене отдельных компонентов системы.

#### **4.2 Анализ результатов вычислительных экспериментов для различных сценариев планирования**

Ключевым вопросом данного этапа исследования является то, может ли интеграция прогнозных оценок, полученных методами машинного обучения, служить эффективным механизмом направляющей информации для комбинаторного поиска в задачах календарно-производственного планирования.

В качестве базового варианта рассматривается сценарий S0, соответствующий классической постановке задачи *JSSP* с минимизацией *makespan* без использования прогнозных ограничений. Сценарии S1 и S2 расширяют данную постановку за счет введения верхних границ на значение *makespan*, полученных на основе *ML*-моделей: в сценарии S1 используется базовая *XGB*-модель, в сценарии S2 – гибридная модель, объединяющая латентные представления *LSTM* и градиентный бустинг.

Вычислительные эксперименты проведены на полном наборе из 80 инстансов *Taillard JSSP* при идентичных параметрах решателя *CP-SAT*, что обеспечивает корректность сравнительного анализа. Анализ полученных значений *makespan* показывает, что интеграция прогнозных ограничений не приводит к ухудшению качества решений по сравнению с базовым сценарием. Для значительной части инстансов сценарии S1 и S2 демонстрируют умеренное снижение значения *makespan* относительно S0, что объясняется сужением пространства поиска за счет априорных верхних границ и более быстрой концентрацией решателя на перспективных областях допустимых решений.

Различия между S1 и S2 по среднему значению *makespan* носят ограниченный характер, что указывает на вторичность выбора конкретной модели прогнозирования по сравнению с самим фактом использования *ML*-ограничений. Анализ вычислительного времени показывает, что введение прогнозных верхних границ не увеличивает время решения, а в ряде случаев сопровождается его незначительным сокращением. Все инстансы успешно решены в заданных временных пределах, что подтверждает корректность интеграции прогнозных оценок в оптимизационную модель.

Таким образом, сценарии S1-S2 подтверждают возможность использования прогнозных оценок в качестве направляющих ограничений. При этом целевая функция и вычислительная устойчивость не ухудшаются. Это создает основу для перехода к робастным стратегиям планирования.

Переход к сценариям S3-S5 связан с явным учетом неопределенности длительностей операций и рисков деградации оборудования. В отличие от сценариев S0-S2, ориентированных на минимизацию *makespan*, робастные сценарии направлены на формирование устойчивых расписаний в условиях неопределенности параметров.

Робастные сценарии формализованы следующим образом:

1. S3 (*soft-robust*) – использование смещенных квантильных оценок длительностей с допущением ограниченных отклонений от медианных оценок, штрафующих в целевой функции.
2. S4 (*worst-case*) – жесткое планирование на основе верхних квантилей с дополнительной поправкой на риск отказа оборудования.
3. S5 (*expected-risk*) – использование математического ожидания длительности операций с учетом вероятности отказа.

Результаты экспериментов показали, что при переходе от базовых сценариев к робастным сценариям среднее значение *makespan* немного вырастет. Это вполне ожидаемо – мы ведь начали учитывать неопределенность и риски.

Анализ распределения изменений *makespan* (рисунок 27) хорошо показал, насколько разные сценария: S3 дает самое маленькое увеличение по сравнению с S0, S4 – самый консервативный, S5 занимает промежуточное положение и лучше всего балансирует между осторожностью и эффективностью.

Для оценки устойчивости расписаний я использовал интегральную робастности – это размах между квантильными значениями *makespan* (32). Чем меньше значение метрики  $R(i)$ , тем устойчивее расписание к разным отклонениям длительностей операций.

Анализ показал, что лучше всего по этой метрике себя проявил сценарий S4. Он дает самые маленькие значения  $R$  и поэтому наиболее подходит для условий с повышенными рисками. Сценарий S5 демонстрирует сопоставимый уровень устойчивости при меньшем росте *makespan*. Сценарий S3 обеспечивает ограниченное снижение вариативности при сохранении большей гибкости. Интеграция *PdM*-оценок в сценариях S4 и S5 реализована через корректировку эффективных длительностей операций на основе агрегированного риска отказа оборудования. Экспериментальные результаты показывают, что учет *PdM*-рисков не приводит к существенному росту времени решения (рисунок 28), а влияние риска проявляется преимущественно в перераспределении временных резервов внутри расписания.

Для всех робастных сценариев обеспечена полная разрешимость задач в пределах установленного времени поиска; среднее время решения остается сопоставимым с базовыми сценариями и не превышает 10-11 секунд на инстанс.

Проведенный анализ показывает, что:

- робастные сценарии S3-S5 расширяют функциональные возможности планировщика, позволяя учитывать неопределенность и эксплуатационные риски;
- сценарии различаются по степени консервативности и могут быть выбраны в зависимости от требований производственной среды;

- интеграция квантильных прогнозов и *PdM*-оценок не приводит к ухудшению вычислительной эффективности.

В целом полученные результаты свидетельствуют о том, что робастные сценарии S3-S5 расширяют функциональные возможности системы планирования, обеспечивая управляемый компромисс между эффективностью и устойчивостью без утраты вычислительной масштабируемости. Это формирует основу для последующего сводного и статистического анализа сценариев планирования.

#### **4.3 Сравнительный и статистический анализ сценариев планирования**

В данном разделе выполняется агрегированный и статистический анализ сценариев планирования S0-S5 с точки зрения их вычислительных свойств, устойчивости и практической интерпретируемости в рамках единого экспериментального контура.

Для всех сценариев использовался идентичный набор из 80 инстансов задачи *JSSP* из набора *Taillard* при фиксированных параметрах решателя *CP-SAT*, что обеспечивает корректность сопоставления результатов. Сводный анализ показывает, что все сценарии S0-S5 характеризуются полной разрешимостью инстансов в заданных временных ограничениях.

Сравнение среднего времени решения показывает, что введение прогнозных и робастных механизмов не приводит к ухудшению вычислительной эффективности. Сценарии S1 и S2 демонстрируют время решения, сопоставимое с базовым сценарием S0, аналогичное поведение наблюдается и для робастных сценариев S3-S5. Таким образом, интеграция квантильных оценок и информации о рисках деградации оборудования не сопровождается ростом вычислительной сложности и не нарушает устойчивость процесса оптимизационного поиска.

С точки зрения агрегированных значений целевой функции сценарии образуют логически согласованную иерархию: S0 формирует нижнюю границу по *makespan*; S1 и S2 обеспечивают сопоставимое качество решений при более управляемом процессе поиска; S3-S5 демонстрируют контролируемое увеличение *makespan* вследствие явного учета неопределенности и эксплуатационных рисков.

Для иллюстрации распределения результатов на рисунке 29 представлены *boxplot*-диаграммы значений *makespan* по всем 80 инстансам *Taillard* для сценариев S0-S5.

Как видно из рисунка 29, медианные значения *makespan* для всех сценариев находятся в сопоставимом диапазоне. Переход от детерминированных к робастным сценариям сопровождается умеренным расширением интерквартильного размаха, что отражает влияние дополнительных ограничений и учет неопределенности, однако не приводит к появлению экстремальных выбросов или деградации распределений. Это подтверждает сохранение управляемости планирования при повышении уровня робастности.

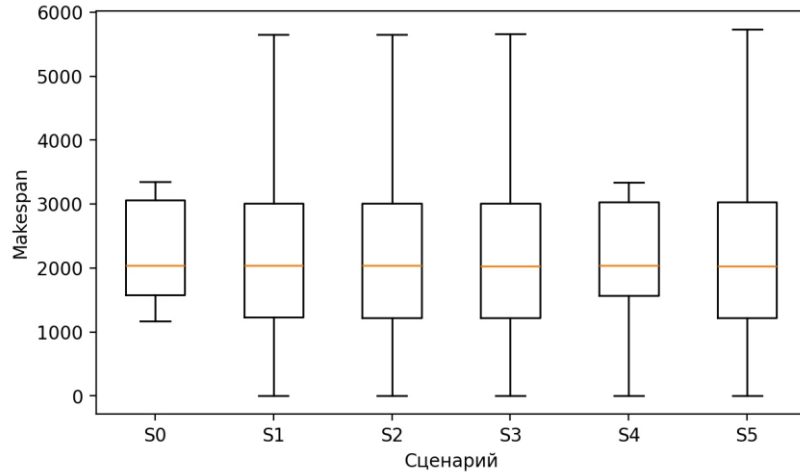


Рисунок 29 – Распределение значений makespan для сценариев планирования S0-S5

Для проверки статистической значимости различий между сценариями я провел непараметрический анализ парных данных (одни и те же 80 инстансов). За базовый сценарий взят S0.

Для каждой пары  $(S_0, S_k)$ , где  $k \in \{1, 2, 3, 4, 5\}$  проверялась нулевая гипотеза (39):

$$H_0: median(\Delta C_{max}(S_k, S_0)) = 0 \quad (39)$$

Относительное изменение *makespan* считается по формуле (40):

$$\Delta C_{max}(S_k, S_0) = \frac{C_{max}^{(S_k)} - C_{max}^{(S_0)}}{C_{max}^{(S_0)}} \cdot 100\% \quad (40)$$

Альтернативная гипотеза (41):

$$H_1: median(\Delta C_{max}(S_k, S_0)) \neq 0 \quad (41)$$

В качестве критерия использовался критерий Уилкоксона для связанных выборок [229]. Статистика критерия рассчитывается так (42):

$$W = \sum_{i=1}^n rank(|d_i|), \text{ по всем } d_i > 0 \quad (42)$$

где  $d_i$  – разность *makespan* для одного и того же инстанса между двумя сценариями;

$rank(\cdot)$  – ранг абсолютного разности;

$n$  – число инстансов.

Этот критерий хорошо тем, что не требует нормальности распределения и правильно учитывает направление и силу изменений. Результаты проверки приведены в таблице 14.

Таблица 14 – Результаты проверки статистической значимости различий между сценариями

| Сравнение | Число инстансов | Статистика W | p-value | Вывод              |
|-----------|-----------------|--------------|---------|--------------------|
| S0 vs S1  | 80              | 993.0        | 0.1025  | различия незначимы |
| S0 vs S2  | 80              | 943.0        | 0.1599  | различия незначимы |
| S0 vs S3  | 80              | 1169.0       | 0.8179  | различия незначимы |
| S0 vs S4  | 80              | 1267.0       | 0.7920  | различия незначимы |
| S0 vs S5  | 80              | 954.0        | 0.1296  | различия незначимы |

Во всех рассмотренных случаях значения *p-value* превышают заданный уровень значимости  $\alpha=0.05$ . Следовательно, статистически значимых различий по показателю *makespan* между базовым сценарием и сценариями с использованием прогнозной и робастной информации не выявлено.

Размер выборки ( $n=80$ ) обеспечивает достаточную статистическую чувствительность критерия для выявления умеренных эффектов. Это позволяет интерпретировать отсутствие значимых различий как содержательный результат.

Полученный результат имеет принципиальное значение. Внедрение прогнозных квантильных оценок и информации о рисках деградации оборудования не снижает базовую эффективность планирования по критерию *makespan*. При этом робастные сценарии обеспечивают повышение устойчивости решений без статистически значимого ухудшения целевой функции.

На практике различия между сценариями объясняются двумя основными факторами: уровнем неопределенности длительности операций и риском деградации оборудования. Уровень неопределенности оценивается через относительную ширину квантильного интервала (43):

$$U = \frac{P90 - P10}{P50} \quad (43)$$

тогда как риск деградации оборудования определяется на основе *PdM*-модели в виде нормированной оценки  $R \in [0, 1]$ .

Выбор сценария планирования осуществляется на основе детерминированных правил отображения значений  $U$  и  $R$  в соответствующие режимы оптимизации. При низких значениях неопределенности и риска используется детерминированное или ограниченное планирование (S0-S2). При умеренной неопределенности предпочтение отдается сценарию мягкой робастности S3, допускающему ограниченные нарушения с последующим штрафованием. В условиях высокой неопределенности или повышенного риска деградации оборудования активируется консервативный сценарий S4,



ориентированный на худший случай, либо компромиссный сценарий S5, использующий математическое ожидание длительности с учетом риска.

Таким образом, сценарии S0-S5 образуют согласованную иерархию режимов функционирования интеллектуальной системы планирования. Они обеспечивают управляемый переход от максимальной эффективности к повышенной устойчивости. Выбор конкретного режима в итоге зависит от того, какой уровень риска мы готовы принять и насколько надежным должно быть производство.

#### 4.4 Обсуждение результатов

Результаты вычислительных экспериментов позволяют оценить, насколько разработанная гибридная система планирования соответствует целям диссертации. Главная задача работы заключалась в создании интеллектуальной системы, которая могла бы учитывать неопределенность длительности операций и риски деградации оборудования. При этом она не должна сильно терять в скорости расчетов.

Эксперименты показали, что предложенная архитектура работает стабильно. Для всех 80 тестовых задач *JSSP* система нашла решение в каждом сценарии. Это говорит о ее надежности и воспроизводимости. При переходе от обычных сценариев к робастным (S1-S5) значительного ухудшения качества расписаний не произошло. Это важный результат.

В большинстве традиционных робастных методов повышения устойчивости почти всегда сильно ухудшает целевой показатель (*makespan*). В нашей работе удалось избежать такого сильного компромисса. Время расчета тоже практически не выросло, несмотря на усложнение модели. Это подтверждает, что выбранный способ интеграции прогнозов и рисков в *CP-SAT* работает корректно и система хорошо масштабируется.

Анализ сценариев S3-S5 показал логичную картину: S3 обеспечивает мягкий компромисс между качеством и устойчивостью. S4 – самый осторожный, подходит для условий с высокими рисками. S5 – промежуточный вариант, учитывает вероятность рисков без излишней перестраховки.

В целом эксперименты подтверждают, что система достигает поставленных целей: она устойчива к неопределенности, сохраняет приемлемую скорость работы и позволяет выбирать нужный режим работы в зависимости от производственных условий.

Конечно, у подхода есть ограничения. Во-первых, система сильно зависит от качества прогнозных моделей. Если реальные данные на предприятии сильно отличаются от обучающих, точность может заметно снизиться.

Во-вторых, неопределенность моделируется в упрощенном виде – без сложных взаимосвязей между операциями и станками, каскадных отказов и внезапных остановок. Поэтому сейчас система лучше подходит для стратегического и тактического планирования. Для оперативного управления в реальном времени понадобятся дополнительные механизмы перепланирования.

В-третьих, все эксперименты проводились на стандартных тестовых задачах *Taillard*. Хотя они широко используются в науке, они не отражают всей

сложности реального производства (нет переналадок, ограниченных буферов, приоритетов заказов и многокритериальных целей). Поэтому результаты нельзя напрямую переносить на конкретное предприятие без дополнительной адаптации модели.

Дополнительным ограничением является использование фиксированных параметров робастности (например, коэффициентов штрафов и функций учета риска): в настоящем исследовании они выбирались исходя из экспериментальной постановки и общих соображений устойчивости, тогда как в реальных условиях оптимальные значения параметров могут зависеть от специфики производства, допустимого уровня риска и требований к надежности выполнения расписаний и уровня автоматизации системы управления производством.

Следует также отметить, что цель настоящей работы не заключается в достижении абсолютного минимума *makespan* при любых условиях. Приоритетом является демонстрация возможности системной интеграции прогнозных моделей и механизмов робастности в классический оптимизационный контур без утраты вычислительной эффективности. В этом контексте выявленные ограничения не снижают значимости полученных результатов, а задают направления дальнейшего развития предложенного подхода и корректные границы интерпретации вычислительных экспериментов.

Переходя к практической применимости, результаты, полученные в подразделах 4.2-4.3, позволяют сформулировать рекомендации по выбору режима функционирования системы планирования в зависимости от эксплуатационного контекста.

Разработанная гибридная система планирования поддерживает несколько сценариев (S0-S5), каждый из которых соответствует определенному уровню учета неопределенности и риска (рисунок 30). Эти сценарии не являются альтернативами в строгом смысле, а образуют иерархию режимов, между которыми может осуществляться осознанный выбор.

Базовый сценарий S0 целесообразно использовать в условиях высокой стабильности производственного процесса, когда длительности операций хорошо известны, оборудование функционирует в номинальном режиме, а последствия возможных отклонений минимальны. Данный сценарий обеспечивает наилучшие значения *makespan*. Он не содержит механизмов компенсации неопределенности. Потому не рекомендуется для динамичных или риск-нагруженных сред.

Сценарии S1 и S2 ориентированы на повышение управляемости процесса планирования. Это достигается благодаря использованию прогнозных верхних границ *makespan*. Такой подход особенно полезен, когда нужно сохранить структуру расписания и уменьшить чувствительность системы к «плохим» вариантам задач. Эти сценарии хорошо работают в оперативном планировании – они позволяют стабильно получать решение за приемлемое время.

Робастные сценарии S3-S5 предназначены для работы в условиях высокой неопределенности длительности операций и рисков поломок оборудования.

Сценарий S3 (*soft-robust*) рекомендуется использовать, когда отклонения от плана ожидаются небольшие. Здесь на первом месте стоит гибкость расписания, а не максимальная осторожность. Мягкие штрафы позволяют системе адаптироваться к мелким нарушениям, не сильно ухудшая общую эффективность.

Сценарий S4 (*worst-case*) ориентирован на критически важные производственные контуры, в которых недопустимы срывы сроков выполнения операций. Использование верхних квантилей длительностей и учет PdM-рисков формируют консервативные расписания с повышенным запасом надежности. Данный сценарий обеспечивает максимальную защищенность от неблагоприятных условий производства. При этом допускает увеличение *makespan*,

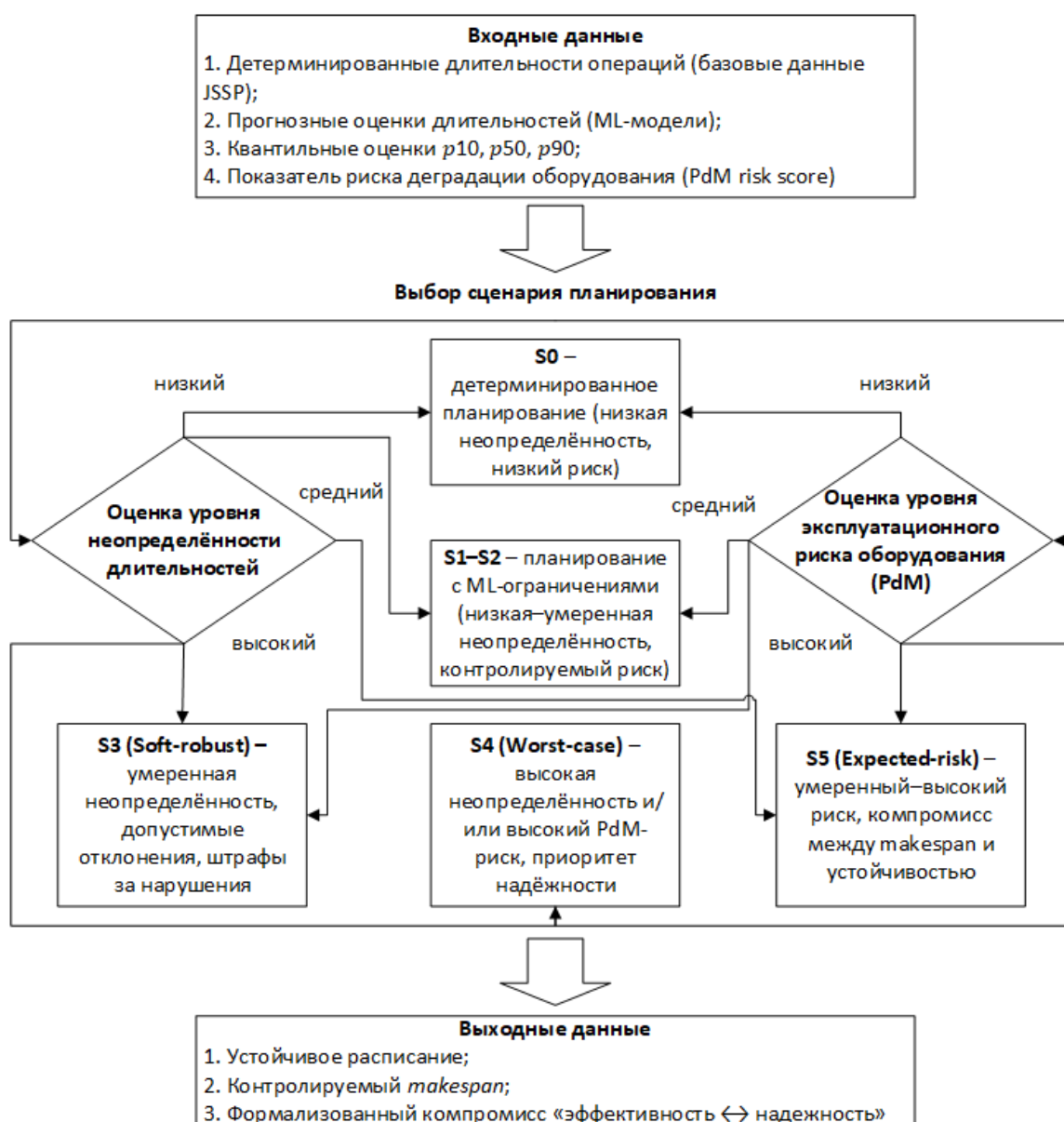


Рисунок 30 – Логическая схема выбора сценария планирования S0-S5

Сценарий S5 (*expected-risk*) представляет собой компромиссный режим, учитывающий вероятностную природу деградации оборудования. Его ключевым преимуществом является адаптивность. При низком уровне риска сценарий приближается к детерминированному планированию. При росте вероятности отказов автоматически смещается в сторону более осторожных оценок. Это делает S5 наиболее универсальным вариантом для практического.

Выбор сценария планирования в разработанной системе осуществляется на основании ряда параметров. Основными параметрами являются допустимый уровень риска и требования к устойчивости расписаний. Это позволяет уйти от фиксированного планирования в пользу контекстно-зависимого управления. Что существенно повышает практическую ценность системы для реальных производственных предприятий.

Сравним разработанную систему с существующими в области календарно-производственного планирования. В качестве базовых методов рассматриваются CP-подходы. Постановка JSSP в виде задачи с дискретными ограничениями реализуется с применением CP-SAT. Данные методы обеспечивают строгую корректность модели и гарантируют нахождение допустимых решений. Однако опираются на фиксированные значения длительностей операций. И не предусматривают учет неопределенности и рисков деградации оборудования в явном виде.

В части исследований методы машинного обучения используются для ускорения процесса планирования. К ним относят эвристики, методы обучения с подкреплением и *end-to-end* модели. Такие подходы демонстрируют высокую эффективность на отдельных классах задач. При этом для них характерны некоторые ограничения. Такими ограничениями являются интерпретируемость, слабая переносимость на новые конфигурации. А также отсутствие формальных гарантий корректности получаемых расписаний.

Отличительной чертой предлагаемого подхода является метод использования машинного обучения. Машинное обучение используется для формирования ограничений и параметров оптимизации. Оценки длительности операций получаются при помощи квантильного прогнозирования. В отличие от точечных прогнозов, например P50, оно формирует интервальные оценки P10, P50 и P90. Такие оценки отражают неопределенность и применяются в риск-ориентированном анализе. Это обеспечивает переход к робастному планированию без разрушения структуры задачи.

Особое место в сопоставлении занимает интеграция предиктивного обслуживания оборудования (*PdM*) в контур планирования. В качестве источника информации о рисках деградации используется эталонный набор данных C-MAPSS, разработанный NASA Prognostics Center of Excellence и широко применяемый в исследованиях по прогнозированию остаточного ресурса (*RUL*). В большинстве существующих работ *PdM* и планирование рассматриваются как слабо связанные или последовательные процессы, тогда как в настоящей работе данное разделение преодолевается за счет формализованного включения *PdM*-оценок в параметры оптимизационной модели.

Существуют обзорные и прикладные исследования, в которых рассматривается совместная оптимизация производственного расписания и технического обслуживания. Однако во многих таких работах либо используется упрощенная модель деградации, либо отсутствует явное разделение сценариев принятия решений по уровню риска. Предложенная система, напротив, формирует набор согласованных сценариев (S0-S5), каждый из которых отражает определенную стратегию работы с неопределенностью и может быть выбран в зависимости от эксплуатационного контекста. Тем самым предложенный подход занимает промежуточное, но принципиально важное положение между классическими детерминированными методами планирования, чисто эвристическими *ML*-подходами и специализированными моделями совместной оптимизации планирования и обслуживания.

В отличие от существующих решений, разработанная система:

- сохраняет строгую формальную постановку задачи планирования;
- использует машинное обучение для информирования оптимизатора, а не для его замены;
- реализует несколько уровней робастности в рамках единой архитектуры;
- обеспечивает интеграцию прогнозов деградации оборудования без потери масштабируемости.

Это позволяет рассматривать предложенную систему как универсальный инструмент интеллектуального планирования, ориентированный на реальные условия промышленной эксплуатации и допускающий адаптацию к различным уровням неопределенности и риска.

#### **4.5 Валидация, устойчивость и обобщаемость результатов вычислительных экспериментов**

Настоящий раздел направлен на обобщающую оценку достоверности и устойчивости результатов, полученных в ходе вычислительных экспериментов, представленных в подразделах 4.1–4.4. В отличие от предыдущих разделов, ориентированных на анализ конкретных сценариев планирования и их сравнительные характеристики, далее результаты рассматриваются с позиций воспроизводимости, устойчивости к параметризации и потенциала переноса выводов за пределы исходной экспериментальной постановки.

Все вычислительные эксперименты в рамках главы 4 проводились на полном наборе из 80 инстансов задачи *JSSP* из классического бенчмарка *Taillard*. Использование полного набора, без выборочного исключения или предварительного отбора инстансов, минимизирует риск смещения выводов за счет селекции тестовых задач и обеспечивает репрезентативность полученных результатов. Инстансы *Taillard* характеризуются высокой вариативностью структурных параметров, включая количество заданий, количество оборудования и распределение технологических маршрутов; вследствие этого полученные эффекты интерпретируются как проявление системных свойств предложенной прогнозно-оптимизационной архитектуры, а не как артефакт частных конфигураций.

Полная разрешимость всех инстансов во всех сценариях S0-S5 в рамках заданных временных ограничений подтверждает вычислительную устойчивость как базовой, так и расширенной постановки задачи. Отсутствие случаев неразрешимости при переходе от детерминированных к робастным сценариям свидетельствует о корректности включения прогнозных компонентов и риск-ориентированных корректировок в CP-SAT-модель и о сохранении управляемости оптимизационного поиска при усложнении постановки.

В рамках проведенных экспериментов использовался фиксированный набор параметров. В частности, уровни квантильных оценок длительностей операций (P10, P50, P90), коэффициенты штрафов для *soft-robust* сценария и параметры учета рисков деградации оборудования были заданы априорно и не подбирались индивидуально под конкретные инстансы.

Такой подход позволяет считать, что полученные результаты отражают обычный режим работы системы, а не «подогнаны» специально под тестовые задачи. Анализ сценариев S3-S5 показал, что при изменении квантильных оценок и уровней риска целевая функция меняется *preidictably* и понятно. При этом система остается устойчивой и не начинает «тормозить». Эти сценарии дают управляемый компромисс между эффективностью и надежностью расписания.

Важно, что сценарии S3-S5 не требуют принципиально другого режима вычислений по сравнению с базовыми. Модель усложняется, но время расчета растет незначительно. Это говорит о том, что выбранный способ учета робастности вполне практичен.

Предлагаемый подход не привязан жестко к конкретным данным и типу производства. Задача JSSP – это обобщенная модель, которая подходит для многих производственных процессов. Методы прогнозирования и учета рисков сделаны модульными, поэтому всю систему можно рассматривать как готовый шаблон. Его легко адаптировать под другие виды планирования: *flow shop*, *flexible job shop* и гибридные системы.

Вместе с тем у исследования есть ограничения. Модель риска деградации оборудования построена на обобщенных данных и не учитывает все возможные нюансы. Кроме того, в экспериментах не рассматривались динамические перепланирования в реальном времени – это тема для будущих работ. Однако эти ограничения не противоречат задачам главы 4. Здесь нужно было показать, что в принципе можно объединить прогнозы и риски в одну оптимизационную модель. Полного моделирования всего реального производства не требовалось.

Проведенная валидация подтверждает, что полученные в главе 4 результаты обладают высокой степенью экспериментальной достоверности и не являются артефактом конкретной выборки, параметров или сценариев. Использование полного набора инстансов, статистическая проверка различий между сценариями (таблица 14), анализ чувствительности и масштабируемости в совокупности формируют надежную эмпирическую основу для сделанных выводов. Таким образом, результаты вычислительных экспериментов подтверждают корректность предложенной гибридной прогнозно-оптимизационной системы и ее способность эффективно функционировать в условиях неопределенности и риска. Это создает обоснованный переход к

следующей главе диссертации, посвященной практической реализации прототипа интеллектуальной системы планирования и демонстрации ее применения в прикладных сценариях.

### **Выводы по разделу**

1. В разделе выполнена комплексная экспериментальная оценка разработанной гибридной прогнозно-оптимизационной системы планирования на стандартизированном наборе из 80 инстансов задачи *JSSP (Taillard)*. Использование всех эталонных инстансов позволило получить надежные и воспроизводимые результаты.

2. Интеграция моделей машинного обучения, квантильных прогнозов длительностей операций и данных о рисках деградации оборудования в *CP-SAT*-планировщик не ухудшила скорость и стабильность работы системы. Это подтверждает, что архитектура работает корректно.

3. Квантильные прогнозы и оценки рисков (*PdM*) позволяет найти управляемый баланс между минимизацией времени выполнения плана (*makespan*) и повышением устойчивости расписаний. Сценарии S3-S5 дают разную степень осторожности: от мягкого компромисса до максимальной надежности.

4. Статистическая проверка различий между сценариями планирования с использованием непараметрического критерия Уилкоксона подтвердила отсутствие статистически значимого ухудшения показателя *makespan* по сравнению с базовым сценарием S0. Это свидетельствует о методологической корректности расширения оптимизационной модели и подтверждает возможность повышения устойчивости решений без статистически значимой деградации эффективности планирования.

## 5 ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ПЛАНИРОВАНИЯ

### 5.1 Архитектура интеллектуальной системы планирования

В рамках исследования разработан программный прототип интеллектуальной системы календарного планирования. Архитектура системы ориентирована на интеграцию методов машинного обучения, моделей прогнозирования технического состояния оборудования и средств дискретной оптимизации в едином вычислительном контуре интеллектуального планирования.

Логическая структура системы представлена на рисунке 31.

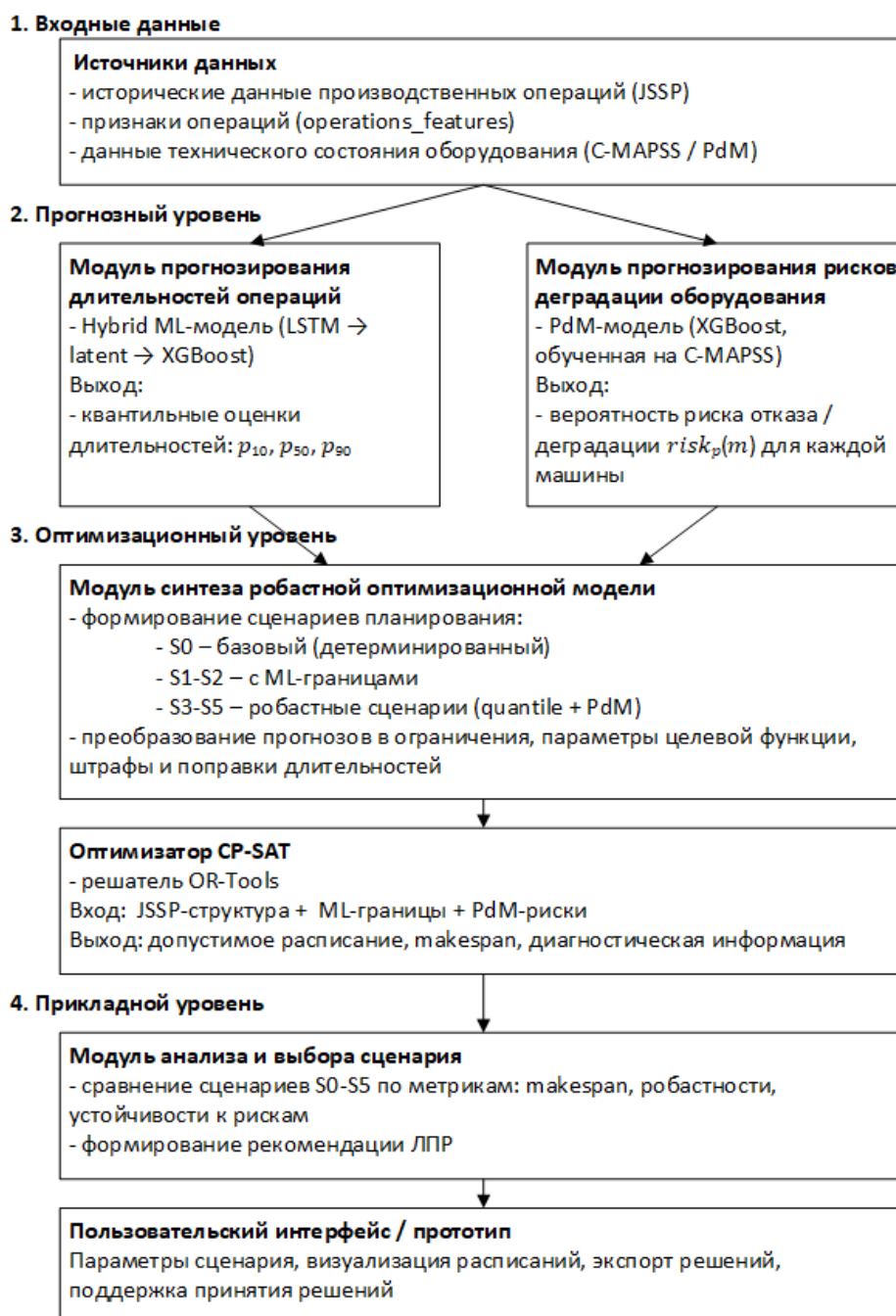


Рисунок 31 – Логическая архитектура интеллектуальной системы планирования



Архитектура имеет модульную организацию и включает несколько взаимосвязанных уровней, обеспечивающих последовательную обработку данных – от формирования прогнозных оценок параметров производственных процессов до получения оптимального или устойчивого (робастного) производственного расписания. На вход системы поступают исторические данные производственных операций, структурированные в формате задачи *JSSP*, а также признаки операций и данные о состоянии оборудования. Указанные данные используются двумя вычислительно независимыми, но логически согласованными прогнозными модулями системы.

Модуль прогнозирования длительностей операций реализует гибридную модель машинного обучения, разработанную и исследованную в главе 3 диссертации. В рамках данного модуля формируются квантильные оценки длительностей операций ( $P_{10}$ ,  $P_{50}$ ,  $P_{90}$ ), используемые для задания прогнозных границ и параметризации неопределенности при формировании сценарных ограничений оптимизационной модели.

Параллельно функционирует модуль прогнозирования рисков деградации оборудования, основанный на *PdM*-модели, обученной на данных *C-MAPSS*. Результатом работы модуля является количественная оценка вероятности риска отказа или деградации для каждой единицы оборудования, используемой в дальнейшем при формировании робастных сценариев планирования.

Центральным элементом системы является модуль синтеза оптимизационной модели, в котором прогнозные оценки преобразуются в параметры и ограничения модели планирования. Именно на данном этапе осуществляется формирование сценариев планирования  $S_0$ - $S_5$ , различающихся степенью учета неопределенности и рисков, что отражает разработанную в главах 3 и 4 концепцию расширенной оптимизационной модели. Сформированная оптимизационная постановка передается в решатель *CP-SAT*, реализованный на базе библиотеки *OR-Tools*. Решатель формирует допустимое расписание, минимизирующее целевую функцию с учетом заданных ограничений и параметров робастности.

Заключительный уровень системы представлен модулем анализа и выбора сценария, а также пользовательским интерфейсом программного прототипа. Данный уровень обеспечивает интерпретацию полученных решений, визуализацию расписаний и поддержку принятия решений, что позволяет использовать разработанную систему в практических условиях календарного планирования. Таким образом, архитектура, представленная на рисунке 31, отражает ключевую научную идею диссертационного исследования – интеграцию прогнозных моделей машинного обучения, механизмов предиктивного обслуживания оборудования и робастной оптимизационной модели календарного планирования.

Функционирование разработанного программного прототипа интеллектуальной системы календарного планирования основано на последовательной обработке разнородных источников данных и их интеграции в рамках единого вычислительного контура. Поток данных организован таким образом, чтобы обеспечить воспроизводимость вычислительных экспериментов,

модульность реализации и возможность расширения системы без изменения ее базовой архитектуры. Входные данные системы формируются из трех основных источников, соответствующих ключевым этапам гибридного прогнозно-оптимизационного подхода, рассмотренного в разделах 3 и 4.

Первым источником являются инстансы задачи *JSSP*, представленные в формате классических наборов *Taillard*. Для каждого инстанса задаются структура технологических маршрутов, последовательность операций для каждой работы, а также привязка операций к конкретным станкам. Указанные данные используются оптимизационным модулем в качестве структурного каркаса задачи календарного планирования и не зависят от выбранного сценария S0-S5.

Вторым источником данных являются квантильные прогнозы длительности операций, полученные на основе гибридной модели машинного обучения. Для каждой операции формируются оценки квантилей P10, P50 и P90, которые отражают возможный диапазон вариации времени выполнения. Квантильные прогнозы сохраняются в виде табличных файлов формата CSV и используются как параметризованные входные данные оптимизационной модели при формировании ограничений и целевой функции. Использование квантильных оценок позволяет параметризовать неопределенность длительностей операций и реализовывать различные стратегии робастного планирования без изменения структуры *CP-SAT* модели.

Третьим источником данных являются оценки риска деградации оборудования, полученные на основе моделей предиктивного технического обслуживания, обученных на данных *C-MAPSS*. Для каждой единицы оборудования формируется агрегированный показатель риска, отражающий вероятность снижения надежности или отказа в рассматриваемом временном горизонте. Эти данные используются исключительно в робастных сценариях планирования S4 и S5 и позволяют связать задачу календарного планирования с состоянием оборудования.

Для обеспечения воспроизводимости вычислительных экспериментов и возможности повторного запуска сценариев входные и выходные данные прототипа унифицированы по форматам и структуре. В таблице 15 приведено описание основных наборов данных и их полей.

После загрузки входных данных система переходит к формированию конкретного сценария использования, который определяется выбранным режимом планирования. В детерминированном сценарии S0 используется только базовая информация инстанса без учета неопределенности и рисков. В следующих сценариях S1 и S2 используются верхние прогнозов. В свою очередь прогнозные верхние границы базируются на квантильных оценках P10, P50 и P90. Последующие более устойчивые сценарии S3-S5 ограничивают в мягком режим на основе *PdM*-рисков и других параметров.

Решатель *CP-SAT* возвращает данные в стандартизированном виде, а именно в формате *JSON*. Эти данные включают в себя такие показатели, как *makespan*, время решения, показатели устойчивости и др. Такой подход обеспечивает единообразие представления результатов вычислительных

экспериментов и позволяет выполнять сравнительный анализ сценариев в автоматическом режиме.

Таблица 15 – Параметры входных и выходных данных эксперимента

| Категория       | Файл                                        | Формат        | Назначение                         | Основные поля                                     |
|-----------------|---------------------------------------------|---------------|------------------------------------|---------------------------------------------------|
| Входные данные  | ta01                                        | JSSP instance | Структура задачи планирования      | jobs, machines, operations                        |
| Входные данные  | jssp_makespan_quantiles_hybrid_enhanced.csv | CSV           | Квантильные прогнозы длительностей | instance, P10, P50, P90                           |
| Входные данные  | pdm_machine_risk.csv                        | CSV           | Оценки PdM-рисков оборудования     | machine_id, risk_score                            |
| Выходные данные | ta01_S0.json                                | JSON          | Расписание операций                | job_id, op_id, machine_id, start, finish          |
| Выходные данные | jssp_cp_sat_planning_robust_results.csv     | CSV           | Batch-результаты экспериментов     | instance, scenario, makespan, status, solver_time |

Сформированные выходные данные поступают на аналитический уровень системы, где используются для построения графиков, сводных таблиц и отчетов, ориентированных на поддержку принятия решений. При этом архитектура системы допускает как автономное использование отдельных модулей, так и их совместное применение в рамках полного цикла интеллектуального планирования.

Таким образом, поток данных в разработанном прототипе организован как последовательная и формально определенная цепочка преобразований: от исходных производственных и диагностических данных к прогнозным оценкам, далее – к оптимизационным решениям и, наконец, к аналитическим представлениям результатов. Это обеспечивает логическую целостность системы, соответствие теоретической модели, изложенной в разделе 3, и практическую применимость, подтвержденную вычислительными экспериментами раздела 4.

В рамках настоящего исследования разработка прототипа интеллектуальной системы календарного планирования осуществлялась с использованием открытых программных средств и библиотек, обеспечивающих воспроизводимость результатов, расширяемость архитектуры и возможность практического внедрения в производственные и образовательные информационные системы. В качестве основного языка программирования выбран *Python*, что обусловлено его широким распространением в задачах анализа данных, машинного обучения и оптимизации, а также наличием развитой экосистемы специализированных библиотек. Использование *Python* позволило интегрировать в рамках единой программной среды модули прогнозирования, оптимизации и анализа результатов без разработки дополнительных промежуточных интерфейсов.

Для реализации оптимизационной составляющей системы применялся решатель *CP-SAT*, входящий в состав библиотеки *Google OR-Tools*. Данный решатель обеспечивает поддержку дискретных ограничений, целевых функций, а также механизмов задания верхних границ и штрафных переменных оптимизационной модели, что является необходимым условием реализации робастных сценариев планирования S3-S5, рассмотренных в разделах 3 и 4. Выбор *CP-SAT* также обусловлен его высокой вычислительной устойчивостью и возможностью масштабирования на задачи промышленного размера.

Прогнозные модели машинного обучения реализованы с использованием библиотек *XGBoost* и *PyTorch*. *XGBoost* применялся для построения регрессионных и квантильных моделей прогнозирования длительностей операций и показателей планирования (включая *makespan*) на основе агрегированных признаков задач планирования, а также для построения моделей предиктивного обслуживания оборудования (*PdM*) на данных *C-MAPSS*.

В свою очередь, чтобы обучить гибридные модели прогнозирования и их нейросетевые компоненты был применен фреймворк глубокого обучения *PyTorch*. Далее данные гибридные модели формируют латентные представления признаков для подсчета ансамблевых и квантильных оценок. Совместное использование машинного и глубокого обучения позволило усилить стороны каждого метода, именно устойчивость градиентного бустинга и интерпретируемость глубоких моделей.

*CSV* и *JSON* были выбраны в качестве современных форматов хранения всех полученных данных. Сохранялись и промежуточные и итоговые результаты, что позволило контролировать структуру данных. Кроме того, сохранение всех данных позволяет воспроизвести эксперименты на любом этапе, а также дополнительно проанализировать и выявить закономерности в таких данных как расписание, значения *makespan*, параметры сценариев и диагностическая информация.

Архитектура предложенной системы состоит из слабо связанных модулей и включает 5 модулей. Первый модуль является модулем подготовки и загрузки данных. Вторым модулем – модулем прогнозирования *ML*-ограничений и *PdM*-рисков. Третий модуль оптимизации необходим для решателя *CP-SAT*. Следующий модуль является пользовательским, где уже оператор самостоятельно выбирает и запускает сценарии S0-S5. В последнем модуле результаты отработанных сценариев, анализируются системой и дается ряд рекомендаций.

Для прозрачности проведения экспериментов и их воспроизводимости, фиксировались такие параметры, как версии библиотек ПО, гиперпараметры, техническое оборудование и т.д.

Реализация прототипа в виде набора взаимосвязанных программных модулей создает основу для последующего расширения функциональности системы, интеграции с внешними информационными системами и подготовки акта внедрения, что будет рассмотрено в последующих разделах главы 5.

## 5.2 Реализация ключевых программных модулей

В условиях календарного планирования, характеризующихся высокой степенью неопределенности, использование точечных оценок длительности операций не позволяет в полной мере учитывать вариативность технологических процессов. В разделах 3 и 4 диссертационной работы показано, что применение квантильных прогнозов обеспечивает более устойчивое представление временных параметров и позволяет формировать интервальные характеристики неопределенности, используемые при построении робастных стратегий планирования.

В рамках практической реализации интеллектуальной системы календарного планирования квантильные оценки используются в качестве входной прогнозной информации для сценарной параметризации ограничений и целевой функции оптимизационной модели. В качестве базовых характеристик применяются оценки квантилей распределения длительности/временных параметров, в частности 10-й, 50-й и 90-й процентиля (P10, P50, P90), полученные с использованием методов машинного обучения, описанных в главе 3. Использование набора (P10, P50, P90) позволяет согласованно задавать как “типовую” (центральную) оценку, так и консервативные границы, соответствующие неблагоприятным отклонениям.

Прогнозные значения квантилей хранятся во внешних файлах формата CSV и загружаются в систему при запуске процесса планирования. Для обеспечения сопоставимости прогнозных оценок с постановкой оптимизационной задачи используется коэффициент масштабирования, вычисляемый на основе сопоставления медианных прогнозных оценок (P50) с базовыми длительностями, заданными в инстансе. Коэффициент применяется для приведения прогнозов к масштабу конкретной задачи планирования без изменения относительной структуры квантилей и без искажения интервала неопределенности. Такой механизм позволяет использовать единый прогнозный модуль при обработке задач различной размерности и разных инстансов, сохраняя интерпретируемость квантильных оценок.

В процессе работы системы для каждой планируемой сущности, соответствующей уровню параметризации сценариев (задание/инстанс в зависимости от принятого формата данных), извлекаются значения квантилей P10, P50 и P90, которые используются при формировании параметров ограничений и целевой функции оптимизационной модели в зависимости от выбранного сценария планирования. Каждый квантиль передает свою определенную информацию. Так квантиль P10 используется для определения нижней границы интервала, а в качестве оценки показывает ширину неопределенности. В свою очередь, P90 наоборот, определяет верхнюю границу, а также сигнализирует о возникновении отклонений. Оптимальным вариантом является квантиль P50. Уже P50 отображает центральную оценку длительности по отношению к другим квантилям.

Квантильные оценки несут рекомендательный характер и не влияют на сам оптимизационный процесс. Основное решение применяется в третьем модуле в

*CP-SAT*, который принимает решение об оптимальном расписании. А квантили связывают результаты машинного обучения и сами задачи оптимизации.

Однако, несмотря на разработанную систему, результаты зависят от множества параметров, главным которым является разное техническое состояние оборудования в зависимости от производства. В разделе 3 показано, что использование методов предиктивного обслуживания (*PdM*) позволяет оценивать вероятность деградации или отказа оборудования и тем самым учитывать риски при принятии управленческих решений. В рамках практической реализации интеллектуальной системы планирования результаты *PdM*-анализа интегрируются в процесс формирования сценариев планирования в виде количественных показателей риска оборудования.

Модуль оценки *PdM*-рисков реализован в виде отдельного компонента системы и предназначен для обработки внешних данных о техническом состоянии оборудования и формирования агрегированных показателей риска. В качестве входной информации используются значения интегрального показателя риска для отдельных единиц оборудования, полученные на основе диагностических данных и методов анализа, описанных ранее. Данные хранятся во внешних файлах формата *CSV* и загружаются в систему при запуске процесса планирования.

Для обеспечения совместимости с задачей оптимизационного планирования *PdM*-риски агрегируются на уровне объекта принятия решения, используемого в сценарной модели. В реализованном прототипе применяется агрегирование риска по оборудованию, задействованному в выполнении планируемой сущности (задания/набора операций), что позволяет получить обобщенную характеристику надежности без введения жестких запретов на отдельные ресурсы. Такой подход обеспечивает вычислительную устойчивость и сохраняет применимость точного решателя *CP-SAT* при расширении постановки.

Полученный агрегированный показатель *PdM*-риска используется как параметр сценарного управления и при формировании параметров целевой функции оптимизационной модели. В частности, в сценариях, ориентированных на учет неблагоприятных условий эксплуатации, значение риска влияет на степень консервативности используемых оценок длительностей/верхних границ, что повышает устойчивость расписания к возможным отказам оборудования. Кроме того, *PdM*-риски имеют влияние на оптимизационный процесс в качестве фактора сценарного управления. Применение *PdM*-риска в качестве данного фактора влияет на надежность планирования, сохраняя при этом вычислительную сложность задачи на допустимом уровне. Таким образом *PdM* не ограничивает использование оборудования, а наоборот, оптимизирует сценарии для уменьшения возникших рисков. Подобная оптимизация осуществляется за счет интеграции предиктивной диагностики в задачи календарного планирования.

В основе программного прототипа лежат сценарии планирования *S0* к *S5* представленных в разделах 3 и 4. Разработанный сценарный подход оптимизирует уровни неопределенности входных данных, а также и самих

рисков, которые возникают из-за технического состояния оборудования. S0-S5 сценарии разработаны как режимы параметров, которые формируют ограничения. Целевая функция оптимизационной модели также формируется с применением сценарного подхода.

В сценарии S0 реализуется детерминированный режим планирования, при котором оптимизация выполняется без применения прогнозных ограничений и используется в качестве базового варианта для сравнения. Сценарий S1 реализует режим *CP-SAT* с верхней границей целевой функции, полученной на основе прогноза модели *XGBoost*. Сценарий S2 реализует режим *CP-SAT* с верхней границей, сформированной по прогнозу гибридной модели *LSTM→XGBoost*. Таким образом, различие между S1 и S2 определяется источником прогнозной информации (типом *ML*-модели), а не уровнем используемого квантиля. Сценарий S3 реализует механизм мягкой робастности (*soft-robust*), в котором допускаются ограниченные отклонения от центральных оценок с последующим штрафованием в целевой функции. Сценарии S4 и S5 дополнительно учитывают агрегированный показатель *PdM*-риска оборудования и формируют риск-ориентированную параметризацию: в S4 используется консервативная постановка, ориентированная на неблагоприятный случай (квантили верхнего уровня с учетом риска), тогда как в S5 применяется риск-ориентированная «ожидаемая» стратегия, учитывающая вероятность деградации оборудования при формировании эффективных оценок длительностей/границ.

В реализованной системе сценарии S1 и S2 отличаются способом формирования верхней границы *makespan* на основе соответствующих *ML*-моделей (*XGBoost* или *Hybrid*), тогда как сценарии S3-S5 дополнительно различаются механизмами робастности и учета риска. Таким образом, машинное обучение используется либо как источник направляющей информации (S1-S2), либо как механизм параметризации неопределенности (S3-S5).

Для обеспечения воспроизводимости экспериментов и однозначной интерпретации реализованных режимов в таблице 16 приведено описание сценариев планирования S0-S5 и используемой в них прогнозной информации (квантильные оценки и *PdM*-риски).

Практическая реализация сценарного подхода дополнена адаптивным механизмом применения *ML*-ограничений. В ходе экспериментальной апробации установлено, что в ряде случаев жестко заданные прогнозные ограничения могут приводить к отсутствию допустимого решения оптимизационной задачи в рамках заданного лимита времени. В связи с этим в системе реализован механизм последовательного ослабления прогнозных ограничений, предусматривающий поэтапное увеличение допустимой верхней границы целевой функции (*makespan*). Если после ослабления ограничений допустимое решение не найдено, система автоматически переходит к режиму оптимизации без использования *ML*-ограничений.

Для анализа фактического поведения системы при различных сценариях введена классификация режимов применения *ML*-ограничений. В реализованном прототипе выделяются следующие режимы: отсутствие *ML*-ограничений, применение прогнозных ограничений без ослабления, применение

прогнозных ограничений с ослаблением, а также переход к оптимизации без прогнозных ограничений. Данная классификация позволяет явно фиксировать, каким образом сценарий планирования был реализован на практике, и использовать эту информацию при последующем анализе результатов.

Таблица 16 – Характеристика сценариев планирования S0-S5 и параметризация ML-ограничений

| Сценарий                                                                                                                                                                    | Режим             | Квантили | PdM-риск | Использование ML-данных                         |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------|----------|-------------------------------------------------|
| S0                                                                                                                                                                          | Детерминированный | -        | -        | Базовая JSSP-модель без ML                      |
| S1                                                                                                                                                                          | Прогнозный        | -        | -        | Bound от модели XGBoost                         |
| S2                                                                                                                                                                          | Консервативный    | -        | -        | Bound от гибридной модели LSTM→XGBoost          |
| S3                                                                                                                                                                          | Soft-robust       | P50-P90  | -        | Мягкая робастность / bounds                     |
| S4                                                                                                                                                                          | Робастный         | P90      | Да       | Bound с учетом риска                            |
| S5                                                                                                                                                                          | Risk-robust       | P90      | Да       | Наиболее консервативная верхняя граница (bound) |
| Примечание – В сценариях S1-S5 применяется адаптивный механизм ослабления ML-ограничений (relaxation/fallback) при отсутствии допустимого решения в заданном лимите времени |                   |          |          |                                                 |

Следует подчеркнуть, что сценарный и адаптивный подход не нарушает корректности оптимизационного процесса. А наоборот, улучшает и помогает получить оптимальное расписание для управления техническим оборудованием, кроме того, оценить риски.

Разработанное программное обеспечение с независимыми модулями, показанными на рисунке 31, позволяет эффективно управлять оптимизацией календарного планирования с использованием машинного обучения для предсказания поломок оборудования и других рисков.

### 5.3 Интеграция прогнозирования, оптимизации и аналитики

Дополнительно в рамках ПО разработали режим *single-instance*. В данном режиме демонстрируется пользовательский режим, где оператор может самостоятельно разработать план расписания оборудования. Данный режим ориентирован на формирование и визуальный анализ расписания для одного производственного задания и позволяет пользователю детально изучить структуру полученного решения, ключевые показатели планирования и корректность работы оптимизационного модуля.

В качестве иллюстративного примера был выбран эталонный инстанс календарного планирования *ta20*, относящийся к классу задач *JSSP* и включающий 20 производственных заданий и 15 единиц оборудования (рисунок 32).

Планирование осуществлялось в рамках сценария S4, предусматривающего использование верхних квантильных оценок с учетом агрегированной оценки *PdM*-рисков. Выбор данного сценария обусловлен его



консервативным характером и устойчивостью при формировании допустимых расписаний.

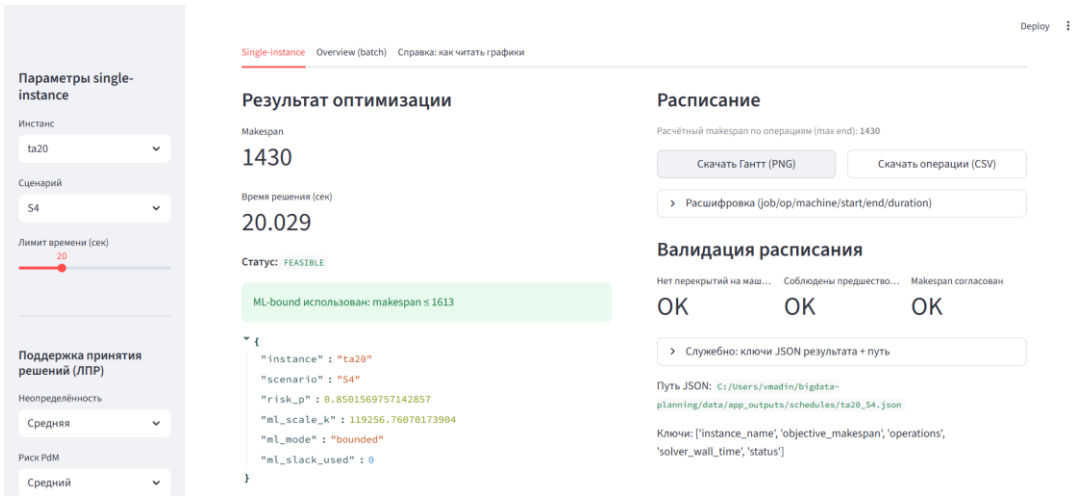


Рисунок 32 – Пользовательский интерфейс прототипа интеллектуальной системы планирования в режиме single-instance

В режиме *single-instance* пользователь выполняет загрузку исходных данных, выбор сценария S0-S5, формирование сценарных ограничений и запуск оптимизационного решателя CP-SAT. По завершении расчета отображается сводная информация о результате планирования, включающая значение *makespan*, статус решения и затраченное время вычислений.

Ключевым элементом визуального представления результата оптимизации является диаграмма Ганта, отображающая последовательность выполнения операций на отдельных единицах оборудования во времени. Диаграмма Ганта для инстанса ta20 при использовании сценария S4 представлена на рисунке 33.

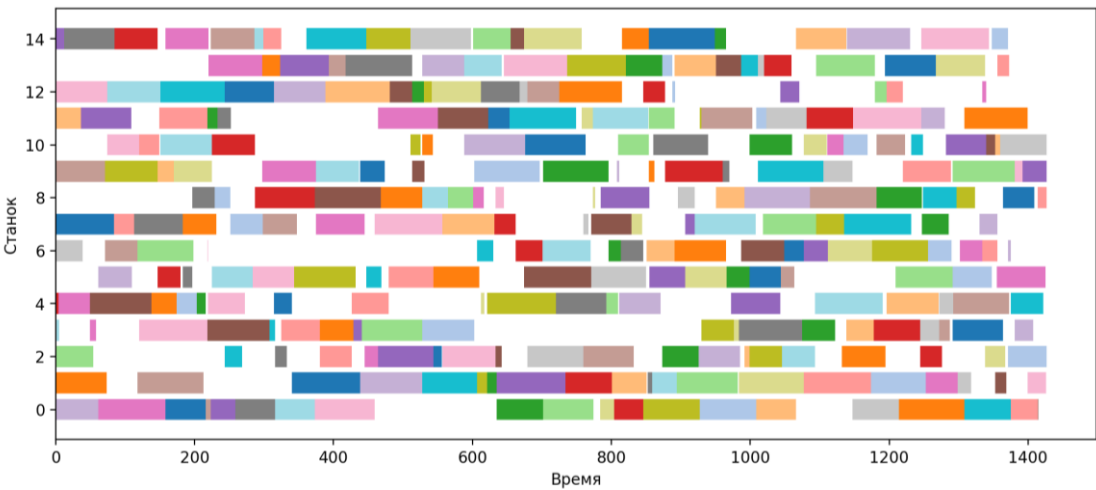


Рисунок 33 – Диаграмма Ганта для инстанса ta20 при сценарии планирования S4

Диаграмма позволяет оценить распределение операций по оборудованию, последовательность их выполнения и уровень загрузки ресурсов. Отсутствие

перекрытия операций на одном оборудовании и соблюдение технологических ограничений подтверждают корректность реализации модели *JSSP* в программном прототипе.

Помимо графической визуализации интерфейс предоставляет табличное представление расписания с указанием номера задания, операции, оборудования и временных интервалов выполнения. В таблице 17 приведен фрагмент полученного расписания.

Таблица 17 – Фрагмент расписания для инстанса ta20 (сценарий S4)

| № задания                                                                                                                                                            | Операция | Оборудование | Время начала | Время окончания |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|--------------|--------------|-----------------|
| job=1                                                                                                                                                                | op=1     | M1           | start=158    | end=216         |
| job=1                                                                                                                                                                | op=4     | M2           | start=340    | end=438         |
| job=1                                                                                                                                                                | op=6     | M3           | start=544    | end=556         |
| job=1                                                                                                                                                                | op=13    | M4           | start=1291   | end=1364        |
| job=1                                                                                                                                                                | op=3     | M5           | start=314    | end=340         |
| ...                                                                                                                                                                  | ...      | ...          | ...          | ...             |
| Примечание – В таблице приведен фрагмент расписания; полное расписание формируется системой и доступно пользователю для просмотра и загрузки в автоматическом режиме |          |              |              |                 |

Анализ полученного результата подтверждает корректность функционирования сценарного и адаптивного механизмов планирования. В рассматриваемом примере система сформировала допустимое расписание с использованием *ML*-ограничений без перехода к *fallback*-режиму, что свидетельствует о согласованности прогнозных оценок, *PdM*-рисков и оптимизационной модели.

Таким образом, режим *single-instance* обеспечивает интерпретируемое представление результатов планирования и служит инструментом верификации корректности функционирования системы.

Для обобщенной оценки поведения системы использовался режим *batch*-планирования. В отличие от режима *single-instance*, ориентированного на детальный анализ отдельного расписания, *batch*-режим предназначен последовательной обработки множества инстансов и последующего агрегированного анализа.

В рамках *batch*-планирования система последовательно формирует расписания для набора эталонных инстансов задачи календарного планирования и сохраняет результаты в виде структурированного CSV-набора данных. Помимо стандартных характеристик задачи и результатов оптимизации (значение *makespan*, статус решения, время работы решателя), в выходных данных фиксируется информация о режиме применения прогнозных ограничений. Таким образом, качественная классификация режимов *ML*-ограничений переходит в форму измеримых аналитических признаков, используемых для сравнения сценариев планирования.

Агрегированные результаты *batch*-планирования отображаются в пользовательском интерфейсе системы в режиме *Overview* (рисунок 34),

предназначенном для первичного анализа и сравнительного сопоставления сценариев. В данном режиме пользователю доступны визуализации, отражающие распределение статусов решений, значений *makespan*, режимов применения *ML*-ограничений, а также времени работы оптимизационного решателя.

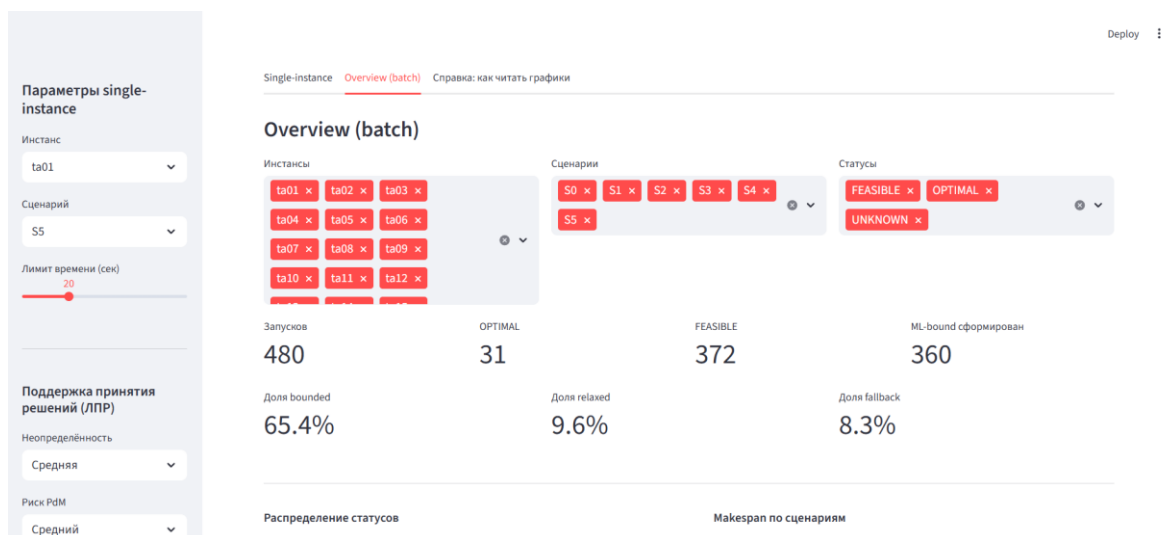


Рисунок 34 – Интерфейс режима Overview программного прототипа интеллектуальной системы планирования

Одним из базовых элементов аналитики является распределение статусов решений, полученных в ходе *batch*-планирования. Данная визуализация позволяет оценить долю допустимых и оптимальных решений, а также выявить случаи, в которых решение не было найдено в рамках заданных вычислительных ограничений. Анализ распределения статусов показывает, что в большинстве экспериментов система формирует допустимые расписания, что свидетельствует о практической применимости предложенного подхода.

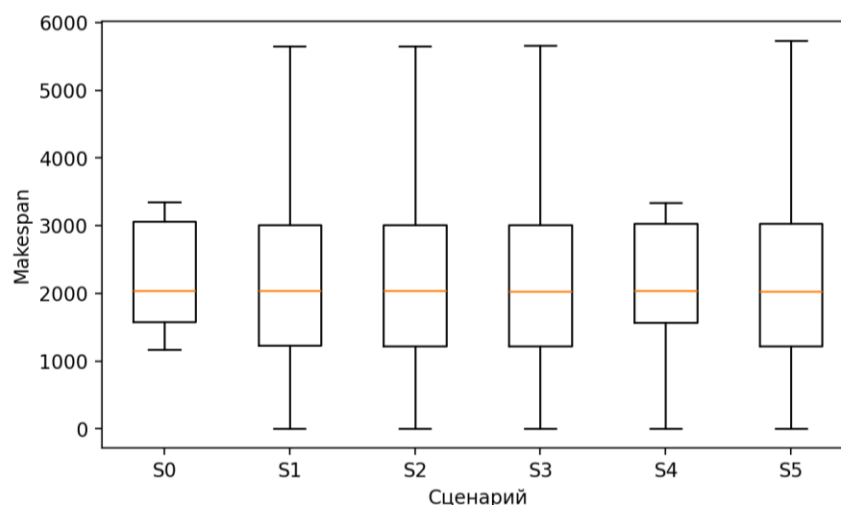


Рисунок 35 – Распределение значений makespan

Ключевым показателем качества планирования является значение целевой функции *makespan*. Для оценки влияния сценариев планирования на структуру получаемых расписаний используется визуализация распределения *makespan* по сценариям S0-S5 (рисунок 35). Представленные данные позволяют сопоставить базовый детерминированный сценарий с прогнозно-ограниченными и робастными сценариями, использующими квантильные прогнозы длительности операций и учет *PdM*-рисков оборудования. Анализ распределений показывает, что применение прогнозных ограничений изменяет характер распределения *makespan* и в ряде случаев позволяет сузить область поиска допустимых решений.

Для того чтобы получить данные о границах параметров и их вероятности, используются квантильные оценки, которые в свою очередь, анализируются оптимизационной моделью. Полученная информация уменьшает возникновение недопустимых значений *makespan*. В итоге, количество возможных решений снижается так как распределение целевой функции менее растянуто. Это позволяет интегрировать модель прогнозирования и поиск наилучшего расписания. Оптимальное расписание уменьшает появления случайных отклонений параметров. Таким образом, применение ИИ, а именно *ML*-ограничений в качестве механизма, направленного на структурирование пространства, возможно для принятия решений. Использование подобных ограничений расширяет рамки ИИ, кроме учета неопределенности.

Рисунок 36 отображает частоту использования прогнозных ограничений и режимов оптимизации с помощью *ML*-ограничений таких как *bounded*, *relaxed*, *fallback/unbounded*, *no\_ml*. Визуальный анализ позволяет оценить устойчивость сценариев планирования по отношению к ограничениям.

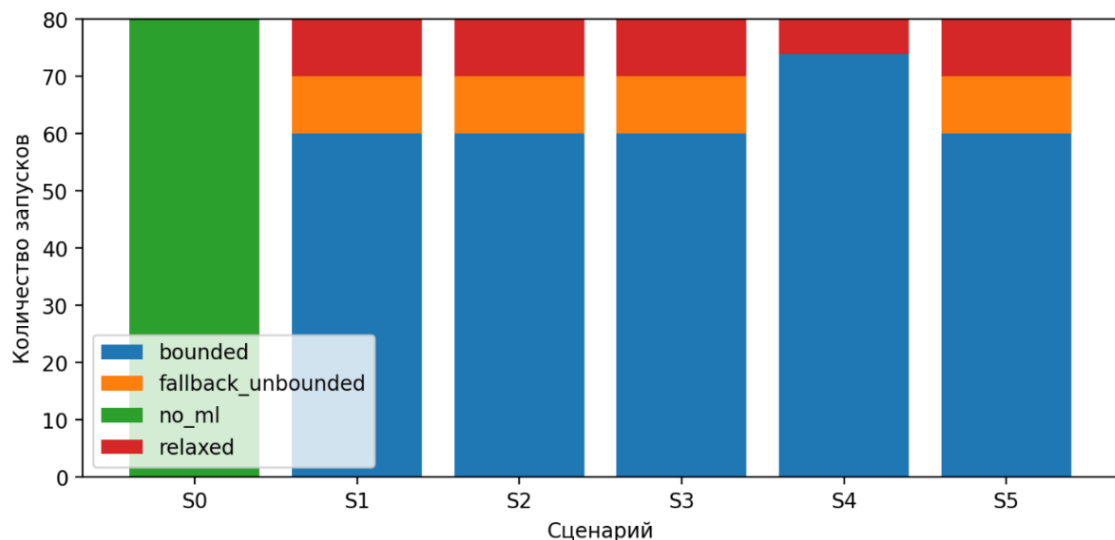


Рисунок 36 – Распределение режимов применения *ML*-ограничений

Далее на рисунке 37 представлена зависимость режимов между временем работы оптимизационного решателя и *ML*-ограничений. Данная зависимость наблюдается и анализируется в режиме *Overview*. Исходя из рисунка 37, можно

сделать вывод, что если пропустить прогнозные ограничения, то повышается время на поиск оптимального решения.

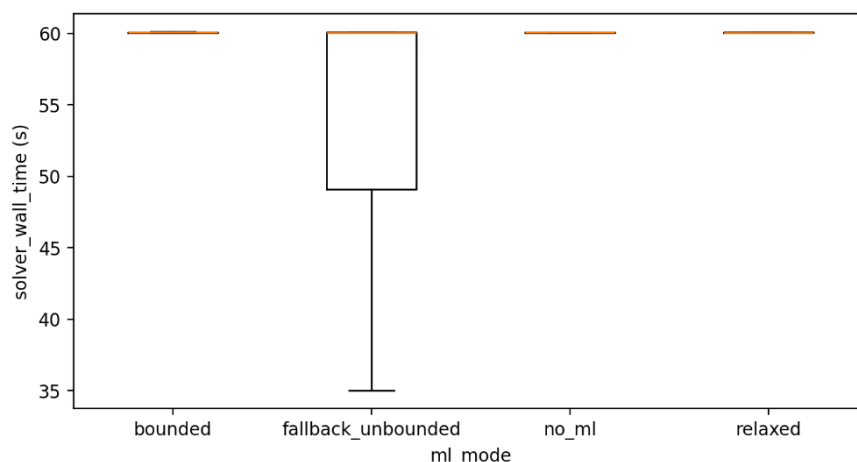


Рисунок 37 – Время работы решателя в зависимости от режима применения ML-ограничений

Таким образом, аналитические отчеты и визуализации *batch*-планирования обеспечивают переход от анализа отдельных примеров к обобщенной оценке поведения интеллектуальной системы планирования. Использование агрегированных метрик и визуализаций позволяет выявить характерные особенности различных сценариев, оценить их устойчивость и вычислительные затраты, а также сформировать аналитическую основу для выбора сценария планирования в пользовательском интерфейсе системы.

Одной из ключевых функций разработанного прототипа интеллектуальной системы планирования является поддержка принятия решений пользователем при выборе сценария планирования. При вариативности длительностей операций и наличии *PdM*-рисков выбор сценария планирования становится нетривиальной задачей и требует учета как количественных, так и качественных факторов.

Для решения данной задачи в структуру пользовательского интерфейса был включен модуль поддержки принятия решений (ЛПР), реализующий логику рекомендаций по выбору сценария планирования. Модуль ЛПР использует результаты *batch*-планирования, агрегированные аналитические показатели, полученные в режиме *Overview*, а также информацию о текущих условиях планирования, задаваемых пользователем.

Основными параметрами, учитываемыми при формировании рекомендаций, являются уровень неопределенности производственной среды и оценка рисков отказов оборудования (*PdM*-рисков). Указанные параметры задаются пользователем в интерфейсе системы и отражают экспертную оценку текущих условий планирования и допустимого уровня риска. На основе сочетания этих параметров модуль ЛПР формирует рекомендованный сценарий планирования, наиболее соответствующий заданным условиям.

На рисунке 38 показана схема выборки сценария планирования. На основе произведенного выбора и создаются рекомендации в модуле принятия решений.

Схема состоит из этапов образования рекомендаций, которые создаются в зависимости от пользовательских параметров неопределенности и *PdM*-рисков. Далее выбирается оптимальный сценарий и дается ряд рекомендаций. Подобные рекомендации помогают в окончательном принятии решения в выборе сценариев планирования.

Однако, пользователь самостоятельно запускает выбранный сценарий и контролирует планирование производственных процессов. При этом два модуля, а именно модуль поддержки принятия решений и оптимизационная модель работают независимо (*CP-SAT* и параметры оптимизационной модели).

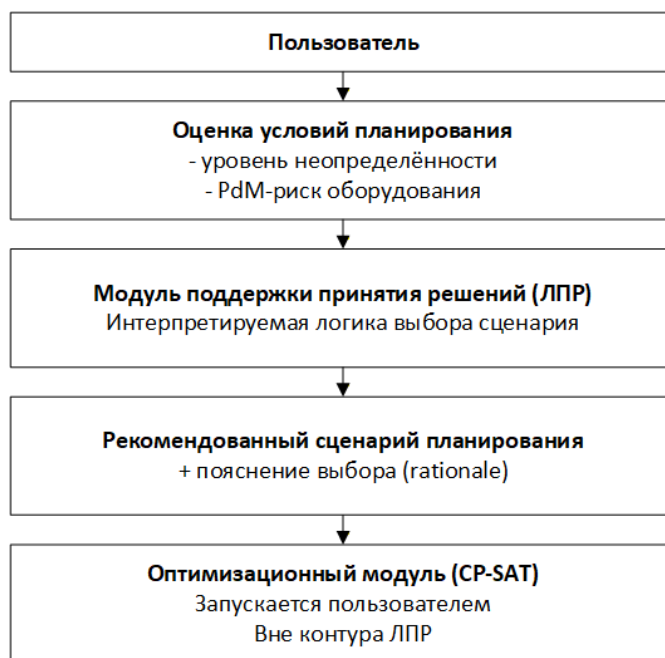


Рисунок 38 – Схема формирования рекомендаций по выбору сценария планирования в модуле поддержки принятия решений

Связь между *batch*-планированием и практическим использованием системы происходит за счет предложенного модуля поддержки принятия решений. Он позволяет перевести результаты многократных вычислительных экспериментов в понятные и интерпретируемые рекомендации, повышая удобство работы с системой и снижая когнитивную нагрузку на пользователя при выборе сценария планирования.

#### 5.4 Экспериментальная апробация и воспроизводимость

Разработанный прототип интеллектуальной системы планирования ориентирован на решение задач оперативного и тактического управления производственными процессами в условиях неопределенности параметров выполнения операций и наличия рисков отказов оборудования. Предлагаемый подход может быть использован в качестве инструмента поддержки принятия решений при формировании производственных расписаний, а также для оценки устойчивости получаемых планов при изменении внешних и внутренних условий функционирования системы. Программное обеспечение было

разработано в рамках научно-технического проекта BR24992785 «Организация и проведение комплексных исследований по обеспечению устойчивого развития агропромышленного комплекса Костанайской области с созданием научно-исследовательского технологического центра».

Использование квантильных прогнозов и сценарных механизмов позволяет формировать расписания с различной степенью консервативности. Это обеспечивает адаптацию системы к требованиям конкретного производственного контекста и предпочтениям лица, принимающего решения.

Практическое применение разработанного прототипа возможно в следующих направлениях:

- формирование оперативных производственных расписаний с учетом неопределенности выполнения операций и рисков отказов оборудования;
- сравнительный анализ альтернативных сценариев планирования по показателям *makespan*, времени решения и устойчивости расписаний;
- аналитическая оценка влияния прогнозных ограничений и *PdM*-параметров на результаты оптимизации;
- поддержка принятия управленческих решений при корректировке производственных планов.

Кроме того, разработанная система может применяться в образовательных и исследовательских целях, в частности для демонстрации современных методов интеллектуального планирования, изучения гибридных *ML*-оптимизационных подходов и проведения вычислительных экспериментов на задачах календарного планирования.

Разработанный прототип прошел экспериментальную апробацию в условиях, приближенных к реальным производственным процессам, что позволило подтвердить его работоспособность и практическую применимость в качестве инструмента поддержки принятия решений.

Для подтверждения практического применения разработанной системы, алгоритм был внедрен в ТОО «ОПТИСОРТ» согласно акту внедрения от 10.11.2025 г. Согласно полученным результатам интеграции ПО в производственный цикл, апробация разработанной системы прошла успешно. Система корректно произвела все расчеты, сценарии, а также предложила оптимальные расписания с учетом всех условий на производстве.

Календарное планирование является актуальным направлением и острым вопросом в принятии решений в производственных процессах. Разработанная система позволит улучшить и оптимизировать процессы календарного планирования не только практически, но и в научных, и в прикладных исследованиях.

Однако, разработанный метод обладает ограничениями, которые влияют на применение прототипа и его границы. Главное ограничение относится к уровню объединения прогнозных и риск-ориентированных параметров, т.е. прототип ограничен в рамках производственного задания и сценария, так как в разработанном методе квантильные прогнозы длительности операций и оценки *PdM*-рисков группируются на этих уровнях. Такой подход обеспечивает вычислительную устойчивость и масштабируемость оптимизационной модели,

однако не позволяет учитывать индивидуальные особенности отдельных операций и единиц оборудования в полном объеме.

Ограничения также связаны с вычислительной сложностью задачи календарного планирования. Использование точного оптимизационного решателя *CP-SAT* приводит к экспоненциальному росту времени решения при увеличении размерности задачи, числа заданий и оборудования. В связи с этим в прототипе применяется ограничение по времени вычислений, что в отдельных сценариях может приводить к невозможности нахождения допустимого решения в рамках заданного лимита и требует перехода к более консервативным режимам планирования.

Следующее ограничение относится к логике сценарного управления и модуля поддержки принятия решений. В текущей реализации выбор сценария планирования осуществляется пользователем на основе рекомендаций ЛПР, при этом система не выполняет автоматическую адаптацию сценариев в процессе оптимизации. Данный подход соответствует концепции *human-in-the-loop*, однако ограничивает степень автономности системы и требует активного участия лица, принимающего решения, при анализе альтернативных вариантов планирования.

Кроме того, экспериментальная оценка разработанного прототипа выполнена преимущественно на эталонных инстансах задачи *JSSP*. Несмотря на широкое распространение данных задач в научных исследованиях и их репрезентативность для анализа алгоритмов календарного планирования, использование синтетических и эталонных наборов данных не в полной мере отражает специфику конкретных производственных систем, включая организационные и технологические особенности реальных предприятий.

Следует также отметить, что разработанный прототип ориентирован на поддержку аналитических и планировочных функций и не включает механизмы интеграции с промышленными информационными системами класса *MES* или *ERP*. В рамках данной работы данные аспекты сознательно не рассматривались, поскольку выходят за рамки поставленных научных задач и требуют отдельного исследования.

Выявленные ограничения определяют текущие границы применимости разработанного прототипа и не снижают научной значимости полученных результатов.

В связи с различными вариациями оборудования в производственных циклах, планируется улучшить разработанный прототип с привлечением более детальных параметров прогнозирования с учетом длительности различных операций, технических деталей и параметров оборудования. Исследование уже требует более глубокого анализа рисков, отказов оборудования и других параметров. Кроме того, планируется полная автоматизация принятия решений на сформированных базах данных со временем, повышая устойчивость и скорость за счет параллельных вычислений. Однако, главным направлением развития диссертационного исследования является интеграция прототипа в крупные промышленные информационные системы, такие как *MES* и *ERP*.



## Выводы по разделу

1. Разработан и реализован программный прототип интеллектуальной системы календарного планирования, интегрирующий методы машинного обучения, предиктивной диагностики оборудования (*PdM*) и дискретной оптимизации на основе решателя *CP-SAT*. Показано, что предложенный *ML-CP-SAT*-конвейер обеспечивает формирование допустимых и оптимизированных производственных расписаний при различных сценариях ограничений (*S0-S5*), включая режимы без использования машинного обучения, с *ML*-ограничениями различного типа, с их адаптивным ослаблением и механизмом резервирования (*fallback*).

2. Реализована система сбора, хранения и пакетной обработки результатов планирования с формированием унифицированного набора выходных данных, включающего значения целевой функции (*makespan*), статус решения, режим применения *ML*-ограничений и время вычислений. Это обеспечило переход от анализа единичных примеров к системному исследованию поведения оптимизационного алгоритма на серии инстансов и сценариев календарного планирования, а также позволило формализовать оценку устойчивости и вычислительной эффективности системы.

3. На основе агрегированных данных выполнена визуальная и аналитическая оценка поведения системы планирования. Продемонстрировано влияние различных режимов *ML*-ограничений и риск-ориентированных параметров на качество решений и вычислительные характеристики оптимизационного процесса. Показано, что использование прогнозных моделей машинного обучения в составе *ML-CP-SAT*-конвейера выполняет функцию направляющей информации для оптимизационного поиска, стабилизирует поведение решателя и в ряде случаев снижает вариативность вычислительных затрат при сохранении допустимости и корректности расписаний.

4. Разработана и реализована логика поддержки принятия решений (ЛПР), обеспечивающая интерпретируемое представление результатов планирования и сценарный выбор режима оптимизации в зависимости от уровня неопределенности и *PdM*-рисков. Предложенная схема интегрирует аналитические результаты *batch*-планирования с пользовательским интерфейсом и позволяет рассматривать систему не только как инструмент автоматического формирования расписаний, но и как средство аналитической поддержки управленческих решений в условиях неопределенности.

5. Проведена экспериментальная апробация программного прототипа в условиях, приближенных к реальным производственным процессам, а также выполнено экспериментальное внедрение в ТОО «ОПТИСОРТ», что подтверждено актом внедрения от 10 ноября 2025 года. Результаты внедрения подтвердили работоспособность гибридного *ML*-оптимизационного подхода, воспроизводимость расчетов и возможность практического использования сценарного механизма планирования в производственной среде. Разработка выполнена в рамках проекта BR24992785 «Организация и проведение комплексных исследований по обеспечению устойчивого развития агропромышленного комплекса Костанайской области с созданием научно-

исследовательского технологического центра», что подчеркивает прикладную направленность и технологическую реализуемость полученных результатов.

## ЗАКЛЮЧЕНИЕ

В диссертационной работе рассматривается актуальная научная задача разработки гибридных прогнозно-оптимизационных моделей календарного планирования производственной деятельности, обеспечивающей учет неопределенности длительностей операций и рисков деградации оборудования на основе методов машинного обучения, предиктивной аналитики и дискретной оптимизации.

Актуальность исследования обусловлена развитием цифрового производства, интеллектуальных систем управления и технологий предиктивного обслуживания оборудования, формирующих новые требования к системам планирования производственных процессов. В современных условиях календарное планирование становится частью интегрированных информационно-аналитических систем управления производством, в которых решения принимаются на основе анализа данных мониторинга оборудования, прогнозирования параметров технологических процессов и оценки эксплуатационных рисков.

Исследование выполнено в рамках проекта BR24992785 «Организация и проведение комплексных исследований по обеспечению устойчивого развития агропромышленного комплекса Костанайской области с созданием научно-исследовательского технологического центра», что определило его прикладную направленность и ориентированность на решение задач цифровой трансформации производственных процессов.

В рамках диссертационного исследования разработан гибридный прогнозно-оптимизационный подход к календарному планированию, основанный на совместном использовании методов машинного обучения и оптимизационного моделирования. Предложенный подход ориентирован на формализованную интеграцию прогнозных моделей в оптимизационный контур планирования, а не на замену классических методов оптимизации.

В первом разделе диссертации выполнен анализ современного состояния исследований в области календарного планирования производственных процессов, машинного обучения и предиктивного обслуживания оборудования. Показано, что, несмотря на значительное количество работ в указанных направлениях, вопросы формализованной интеграции прогнозных моделей и оптимизационных процедур календарного планирования остаются недостаточно проработанными. На основе проведенного анализа сформулирована научная проблема исследования и обоснована необходимость разработки гибридных прогнозно-оптимизационных методов планирования.

Во втором разделе разработана концептуальная модель гибридной системы календарного планирования, включающая прогнозный модуль оценки параметров производственных процессов, модуль предиктивного обслуживания оборудования и оптимизационный модуль календарного планирования. Предложена схема сценарной интеграции прогнозных оценок неопределенности и рисков деградации оборудования в оптимизационную модель планирования.

Обосновано использование квантильных оценок длительностей операций в качестве параметров робастных ограничений оптимизационной задачи.

В третьем разделе разработаны и исследованы модели машинного обучения, предназначенные для прогнозирования параметров производственных процессов и технического состояния оборудования. Реализована гибридная модель прогнозирования длительностей операций на основе нейросетевых представлений и ансамблевых методов (*LSTM + XGBoost*), позволяющая формировать квантильные оценки длительностей операций. Разработана модель оценки рисков деградации оборудования на основе данных предиктивного обслуживания. Выполнена интерпретация прогнозных моделей с использованием факторного анализа и *SHAP*-оценок, что обеспечило объяснимость результатов прогнозирования.

В четвертом разделе проведены вычислительные эксперименты, направленные на оценку эффективности разработанной гибридной прогнозно-оптимизационной системы планирования. Экспериментальная проверка выполнялась на полном наборе из 80 инстансов задачи *JSSP* из набора *Taillard*. Были реализованы и исследованы сценарии планирования *S0-S5*, включающие детерминированные, прогнозно-ограниченные и робастные режимы функционирования системы.

Результаты вычислительных экспериментов показали, что интеграция прогнозных моделей машинного обучения и механизмов учета рисков предиктивного обслуживания не приводит к ухудшению вычислительной устойчивости системы. Для всех сценариев была обеспечена полная разрешимость задач в заданных временных ограничениях. Показано, что использование квантильных оценок длительностей операций позволяет формировать более устойчивые расписания при контролируемом увеличении значения *makespan*. Статистическая проверка различий между сценариями с использованием критерия Уилкоксона подтвердила отсутствие статистически значимого ухудшения качества решений по сравнению с базовым сценарием.

В пятом разделе выполнена практическая реализация программного прототипа интеллектуальной системы календарного планирования, интегрирующей прогнозные модели машинного обучения, оценки рисков деградации оборудования и оптимизационный решатель *CP-SAT*. Показано, что модульная архитектура прототипа обеспечивает воспроизводимость вычислительных экспериментов, интерпретируемость результатов и возможность использования системы в задачах интеллектуального планирования производственных процессов.

Разработанный программный прототип прошел экспериментальную апробацию в производственной среде и внедрен в деятельность ТОО «ОПТИСОРТ», что подтверждено актом внедрения от 10 ноября 2025 года. Кроме того, результаты исследования внедрены в учебный процесс кафедры программного обеспечения Костанайского регионального университета имени Ахмет Байтұрсынұлы, что подтверждено соответствующим актом внедрения от 4 марта 2024 года (Приложение Б).

В ходе исследования получены следующие основные результаты:

- разработана методика гибридного прогнозно-оптимизационного планирования, обеспечивающая интеграцию квантильных прогнозов параметров производственных процессов в оптимизационные модели календарного планирования;

- разработан способ учета рисков технического состояния оборудования в задачах календарного планирования на основе сценарных ограничений оптимизационной модели;

- разработана методика использования квантильных оценок длительностей операций для формирования робастных сценариев планирования;

- реализован воспроизводимый экспериментальный контур исследования прогнозно-оптимизационных моделей календарного планирования на основе открытых эталонных наборов данных;

- разработан программный прототип интеллектуальной системы календарного планирования, подтверждающий практическую реализуемость предложенного подхода.

Полученные результаты подтверждают достижение цели диссертационного исследования – разработки и апробации комплекса моделей интеллектуального анализа больших данных для прогнозного планирования производственной деятельности, направленного на повышение обоснованности и эффективности управленческих решений.

Теоретическая значимость работы заключается в развитии методов интеграции прогнозных моделей машинного обучения и оптимизационных процедур календарного планирования, а также в формализации подхода к учету неопределенности производственных процессов на основе квантильных оценок параметров операций. Практическая значимость заключается в возможности использования разработанных моделей и программного прототипа при создании интеллектуальных систем поддержки принятия решений в области производственного планирования.

В качестве перспективных направлений дальнейших исследований могут рассматриваться расширение предложенной модели на задачи динамического перепланирования, использование реальных производственных данных предприятий, разработка многокритериальных моделей планирования, интеграция цифровых двойников оборудования и производственных систем, а также применение разработанного подхода к задачам *flexible job shop* и *flow shop*.

Таким образом, диссертационная работа формирует научно-методическую и алгоритмическую основу разработки гибридных прогнозно-оптимизационных систем календарного планирования, обеспечивающих учет неопределенности параметров производственных процессов и эксплуатационных рисков оборудования в условиях цифрового производства.

В рамках исследования по данной диссертационной работе получены авторские свидетельства о внесении сведений в государственный реестр прав на объекты, охраняемые авторским правом №61130 от 31.07.2025 г, № 66964 от 02.02.2026 г. (Приложение А).

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Culot G., Nassimbeni G., Orzes G. et al. Behind the definition of Industry 4.0: Analysis and open questions // *International Journal of Production Economics*. – 2020. – Vol. 226. – P. 107617.
- 2 Singh V.K., Parihar A.S., Prakash C. Industrial Revolution 4.0: A critical review on IoT-based smart manufacturing technologies and application // In book: *Internet of Things and Businesses in a Disruptive Economy*. – NY., 2020. – P. 75-99.
- 3 Zheng P., Wang Z., Chen C.-H. Industrial smart product-service systems solution design via hybrid concerns // *Procedia CIRP*. – 2019. – Vol. 83. – P. 187-192.
- 4 Mourtzis D., Vlachou E., Milas N. Industrial Big Data as a result of IoT adoption in manufacturing // *Procedia CIRP*. – 2016. – Vol. 55. – P. 290-295.
- 5 Ning F., Shi Y., Cai M. et al. Manufacturing cost estimation based on the machining process and deep-learning method // *Journal of Manufacturing Systems*. – 2020. – Vol. 56. – P. 11-22.
- 6 Alibekova G., Medeni T., Panzabekova A. et al. Digital transformation enablers and barriers in the economy of Kazakhstan // *The Journal of Asian Finance, Economics and Business*. – 2020. – Vol. 7, Issue 7. – P. 565-575.
- 7 Zhang Y., Ren S., Liu Y. et al. A framework for Big Data-driven product lifecycle management // *J of Cleaner Production*. – 2017. – Vol. 159. – P. 229-240.
- 8 Al-Ali A.R., Gupta R., Al Nabulsi A. Cyber physical systems role in manufacturing technologies // *AIP Conference Proceedings*. – 2018. – Vol. 1957, Issue 1. – P. 050007.
- 9 Popescu G.H., Zvarikova K., Machova V. et al. Industrial Big Data, automated production systems, and Internet of Things sensing networks in cyber-physical system-based manufacturing // *Journal of Self-Governance and Management Economics*. – 2020. – Vol. 8, Issue 3. – P. 30-36.
- 10 Tao F. et al. Digital twin in industry: State-of-the-art // *IEEE Transactions on Industrial Informatics*. – 2019. – Vol. 15, Issue 4. – P. 2405-2415.
- 11 Tassel P., Gebser M., Schekotihin K. A reinforcement learning environment for job-shop scheduling // <https://arxiv.org/abs/2104.03760>. 10.10.2025.
- 12 Chaudhry I.A., Khan A.A. A research survey: Review of flexible job shop scheduling techniques // *International Transactions in Operational Research*. – 2015. – Vol. 23. – P. 551-591.
- 13 Heinz S., Ku W.Y., Beck J.C. Recent improvements using constraint integer programming for resource allocation and scheduling // *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems: proced. 10th internat. conf.* – Berlin, 2013. – P. 12-27.
- 14 Aytug H., Bhattacharyya S., Koehler G.J. et al. A review of machine learning in scheduling // *IEEE Transactions on Engineering Management*. – 1994. – Vol. 41, Issue 2. – P. 165-171.
- 15 Pencheva P., Vitliemova P., Georgiev I. Optimization model for production scheduling taking into account preventive maintenance in an uncertainty-based production system // *Heliyon*. – 2023. – Vol. 9, Issue 7. – P. e17485.

- 16 Pan F.S., Ye C.M., Yang J. Flexible job-shop scheduling problem under uncertainty based on QPSO algorithm // *Advanced Materials Research*. – 2012. – Vol. 605-607. – P. 487-492.
- 17 Zanchettin A.M. Robust scheduling and dispatching rules for high-mix collaborative manufacturing systems // *Flexible Services and Manufacturing Journal*. – 2022. – Vol. 34. – P. 293-316.
- 18 Zhang Y., Ren S., Liu Y. et al. A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products // *Journal of Cleaner Production*. – 2017. – Vol. 142, Part 2. – P. 626-641.
- 19 Sisode M., Devare M. A review on machine learning techniques for predictive maintenance in Industry 4.0 // *Proceed. of the internat. conf. on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022)*. – Aurangabad, 2023. – P. 774-783.
- 20 Ait Ben Hamou K., Jarir Z., Elfirdoussi S. Using machine learning for production scheduling problems in the supply chain: A review // *Computers & Industrial Engineering*. – 2025. – Vol. 206. – P. 111243.
- 21 Shaikat F.B., Islam R., Happy A.T. et al. Optimization of production scheduling in smart manufacturing environments using machine learning algorithms // *Letters in High Energy Physics*. – 2025. – Vol. 2025. – P. 5-16.
- 22 Wen R., Torkkola K., Narayanaswamy B. et al. A multi-horizon quantile recurrent forecaster // <https://arxiv.org/abs/1711.11053>.
- 23 Rahimi Kakehjoob P., Farughi H., Rasay H. Joint production and maintenance planning given the effects of imperfect maintenance on the amount of defective products // *Journal of Modeling in Engineering*. – 2025. – Vol. 23, Issue 83. – P. 57-84.
- 24 Aarab S., El Barkany A., El Khalfi A. The integration of maintenance plans and production scheduling taking account of outsourcing: A literature review // *International Journal of Productivity and Quality Management*. – 2017. – Vol. 21, Issue 1. – P. 1-22.
- 25 Hartmann S., Briskorn D. A survey of variants and extensions of the resource-constrained project scheduling problem // *European Journal of Operational Research*. – 2010. – Vol. 207, Issue 1. – P. 1-14.
- 26 Feng W., Feng Y., Zhang Q. Multistage distributionally robust optimization for integrated production and maintenance scheduling // *AIChE Journal*. – 2021. – Vol. 67. – P. e17329.
- 27 Błażewicz J., Ecker K., Trystram D. Recent advances in scheduling in computer and manufacturing systems // *European Journal of Operational Research*. – 2025. – Vol. 164, Issue 3. – P. 573-574.
- 28 Taghizadeh Elh., Taghizadeh E. The impact of digital technology and Industry 4.0 on enhancing supply chain resilience // *Proceed. of the 11th Annual internat. conf. on Industrial Engineering and Operations Management*. – Singapore, 2021. – P. 2021-2029.
- 29 Mahure S.S., Deharkar A., Yadav L.N. The role of machine learning in predictive maintenance for manufacturing industries // *International Journal of*

Innovative Research in Science, Engineering and Technology. – 2024. – Vol. 13, Issue 5. – P. 7101-7106.

30 Beykal B., Pistikopoulos E.N. Data-driven optimization algorithms // In book: Artificial intelligence in manufacturing. – London: Elsevier, 2024. – P. 135-180.

31 Tanizaki T., Fukuyama A., Uchino K. et al. Data-driven control system using machine learning in production process // Procceed. Internat. conf. “Advances in Production Management Systems: Production Management Systems for Volatile, Uncertain, Complex, and Ambiguous Environments”. – Chemnitz, 2024. – P. 3-16.

32 Essien A., Giannetti C. A deep learning model for smart manufacturing using convolutional LSTM neural network autoencoders // IEEE Transactions on Industrial Informatics. – 2020. – Vol. 16, Issue 9. – P. 6069-6078.

33 Hu Q., Zhao Y., Wang Y. et al. Remaining useful life estimation in prognostics using deep reinforcement learning // IEEE Access. – 2023. – Vol. 11. – P. 32919-32934.

34 Patil C.R., Jadhav S.K., Bardiya A.L. et al. Machine learning-based predictive maintenance of industrial machines // International Journal of Computer Trends and Technology. – 2023. – Vol. 71, Issue 3. – P. 50-56.

35 Yermekbayeva D., Rakhmatullina A. The importance of implementing digitalization in Kazakhstan // Eurasian Journal of Economic and Business Studies. – 2020. – Vol. 57, Issue 3. – P. 40-55.

36 Bragin M.A. et al. Toward robust manufacturing scheduling: Stochastic job-shop scheduling // <https://arxiv.org/abs/2206.09326>. 10.11.2025.

37 Garey M.R., Johnson D.S. Computers and intractability: A guide to the theory of NP-completeness. – NY., 1979. – 347 p.

38 Davenport T.H., Harris J.G. Competing on analytics: The new science of winning. – Boston: Harvard Business Review Press, 2007. – 218 p.

39 Hegyháti M., Bakon K.A., Holczinger T. Optimization with uncertainties: A scheduling example // Central European Journal of Operations Research. – 2023. – Vol. 31. – P. 1239-1263.

40 Pinedo M. Scheduling: Theory, algorithms, and systems. – Ed. 5th. – NY.: Springer. – 670 p.

41 Allahverdi A., Ng C.T., Cheng T.C.E. et al. A survey of scheduling problems with setup times or costs // European Journal of Operational Research. – 2008. – Vol. 187, Issue 3. – P. 985-1032.

42 Himmiche S., Marangé P., Aubry A. et al. Robustness evaluation process for scheduling under uncertainties // Processes. – 2023. – Vol. 11, Issue 2. – P. 371.

43 Snyder L.V. Supply chain robustness and reliability: Models and algorithms: dis. ... doc. phil. – Evanston, 2003. – 285 p.

44 Infantes G., Roussel S., Pereira P. et al. Learning to solve job shop scheduling under uncertainty // Procceed. conf. Integration of Constraint Programming, Artificial Intelligence, and Operations Research (CPAIOR 2024). – Cham, 2024. – P. 329-345.

45 Bertsimas D., Kallus N. From predictive to prescriptive analytics // Management Science. – 2019. – Vol. 66, Issue 3. – P. 1025-1044.



- 46 Souza R.L.C., Ghasemi A., Saif A. et al. Robust job-shop scheduling under deterministic and stochastic unavailability constraints due to preventive and corrective maintenance // *Computers & Industrial Engineering*. – 2022. – Vol. 168. – P. 108130.
- 47 Rihane K., Dabah A., AitZai A. Learning-based approaches for job shop scheduling problems: A review // <https://arxiv.org/abs/2505.04246>. 10.10.2025.
- 48 Kletti J. *Manufacturing execution systems – MES*. – NY.: Springer, 2010. – 271 p.
- 49 Priarone P.C., Robiglio M., Settineri L. et al. Modelling of specific energy requirements in machining as a function of tool and lubricoolant usage // *CIRP Annals*. – 2016. – Vol. 65, Issue 1. – P. 25-28.
- 50 Lee J., Bagheri B., Kao H.-A. A cyber-physical systems architecture for Industry 4.0-based manufacturing systems // *Manufacturing Letters*. – 2015. – Vol. 3. – P. 18-23.
- 51 Wang J., Ma Y., Zhang L. et al. Deep learning for smart manufacturing: Methods and applications // *Journal of Manufacturing Systems*. – 2018. – Vol. 48, Part C. – P. 144-156.
- 52 Wuest T., Weimer D., Irgens C. et al. Machine learning in manufacturing: Advantages, challenges, and applications // *Production & Manufacturing Research*. – 2016. – Vol. 4, Issue 1. – P. 23-45.
- 53 Bagnall A., Lines J., Bostrom A. et al. The great time series classification bake off: A review and experimental evaluation of recent algorithmic advances // *Data Mining and Knowledge Discovery*. – 2017. – Vol. 31. – P. 606-660.
- 54 Sezer O.B., Gudelek M.U., Ozbayoglu A.M. Financial time series forecasting with deep learning: A systematic literature review (2005–2019) // *Applied Soft Computing*. – 2020. – Vol. 90. – P. 106181.
- 55 Jardine A.K.S., Lin D., Banjevic D. A review on machinery diagnostics and prognostics implementing condition-based maintenance // *Mechanical Systems and Signal Processing*. – 2006. – Vol. 20, Issue 7. – P. 1483-1510.
- 56 Ma S., Cheng B., Shang Z. et al. Scattering transform and LSPTSVM-based fault diagnosis of rotating machinery // *Mechanical Systems and Signal Processing*. – 2018. – Vol. 104. – P. 155-170.
- 57 Wen L., Xie X., Li X. et al. A new ensemble convolutional neural network with diversity regularization for fault diagnosis // *Journal of Manufacturing Systems*. – 2022. – Vol. 62. – P. 964-971.
- 58 Kusiak A. Smart manufacturing must embrace big data // *Nature*. – 2017. – Vol. 544. – P. 23-25.
- 59 Qu Y.J., Ming X.G., Liu Z.W. et al. Smart manufacturing systems: State of the art and future trends // *The International Journal of Advanced Manufacturing Technology*. – 2019. – Vol. 103. – P. 3751-3768.
- 60 Su N., Huang S., Su C. Elevating smart manufacturing with a unified predictive maintenance platform: The synergy between data warehousing, Apache Spark, and machine learning // *Sensors*. – 2024. – Vol. 24, Issue 13. – P. 4237.
- 61 Hastie T., Tibshirani R., Friedman J. *The elements of statistical learning*. – Ed. 2nd. – NY.: Springer, 2009. – 745 p.

- 62 Breiman L. Random forests // Machine Learning. – 2001. – Vol. 45, Issue 1. – P. 5-32.
- 63 Chen T., Guestrin C. XGBoost: A scalable tree boosting system // Proceed. of the 22nd ACM SIGKDD internat. conf. on Knowledge Discovery and Data Mining (KDD '16). – NY., 2016. – P. 785-794.
- 64 Flores-Huamán K.-J. et al. Lead-time prediction in wind tower manufacturing: A machine learning-based approach // Mathematics. – 2024. – Vol. 12, Issue 15. – P. 2347.
- 65 Hochreiter S., Schmidhuber J. Long short-term memory // Neural Computation. – 1997. – Vol. 9. – P. 1735-1780.
- 66 Kim J., Kim H., Kim H. et al. A comprehensive survey of deep learning for time series forecasting: Architectural diversity and open challenges // Artificial Intelligence Review. – 2025. – Vol. 58. – P. 216.
- 67 Elmachtoub A.N., Grigas P. Smart “predict, then optimize” // Management Science. – 2021. – Vol. 68, Issue 1. – P. 9-26.
- 68 Gneiting T. Making and evaluating point forecasts // Journal of the American Statistical Association. – 2011. – Vol. 106, Issue 494. – P. 746-762.
- 69 Gneiting T., Balabdaoui F., Raftery A.E. Probabilistic forecasts, calibration and sharpness // Journal of the Royal Statistical Society: Series B (Statistical Methodology). – 2007. – Vol. 69, Issue 2. – P. 243-268.
- 70 Goodwin P., Önköl D., Thomson M. Do forecasts expressed as prediction intervals improve production planning decisions? // European Journal of Operational Research. – 2010. – Vol. 205, Issue 1. – P. 195-201.
- 71 Gneiting T., Raftery A.E. Strictly proper scoring rules, prediction, and estimation // Journal of the American Statistical Association. – 2007. – Vol. 102. – P. 359-378.
- 72 Gneiting T., Katzfuss M. Probabilistic forecasting // Annual Review of Statistics and Its Application. – 2014. – Vol. 1. – P. 125-151.
- 73 Qiao F., Liu J., Ma Y. Industrial big-data-driven and CPS-based adaptive production scheduling for smart manufacturing // International Journal of Production Research. – 2021. – Vol. 59, Issue 23. – P. 7139-7159.
- 74 Guerra M., Scardapane S., Bianchi F.M. Probabilistic load forecasting with reservoir computing // <https://arxiv.org/abs/2308.12844>. 10.10.2025.
- 75 Koenker R., Bassett G. Regression quantiles // Econometrica. – 1978. – Vol. 46. – P. 33-50.
- 76 Ribeiro M.T. et al. Why should I trust you? Explaining the predictions of any classifier // Proceed. of the 22nd ACM SIGKDD internat. conf. on Knowledge Discovery and Data Mining (KDD '16). – San Francisco, 2016. – P. 1135-1144.
- 77 Artificial intelligence risk management framework (AI RMF 1.0) / National Institute of Standards and Technology (2023) // <https://www.nist.gov/itl>. 10.10.2025.
- 78 Barredo Arrieta A. et al. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI // Information Fusion. – 2020. – Vol. 58. – P. 82-115.

- 79 Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead // *Nature Machine Intelligence*. – 2019. – Vol. 1. – P. 206-215.
- 80 Lundberg S.M., Lee S.-I. A unified approach to interpreting model predictions // <https://arxiv.org/abs/1705.07874>. 10.10.2025.
- 81 Doshi-Velez F., Kim B. Toward a rigorous science of interpretable machine learning // <https://arxiv.org/abs/1702.08608>. 11.11.2025.
- 82 Lou Y., Caruana R., Gehrke J. et al. Accurate intelligible models with pairwise interactions // *Proceed. of the 19th ACM SIGKDD internat. conf. on Knowledge Discovery and Data Mining (KDD '13)*. – NY., 2013. – P. 623-631.
- 83 Kneib T., Silbersdorff A., Säfken B. Rage against the mean – A review of distributional regression approaches // *Econometrics and Statistics*. – 2023. – Vol. 26. – P. 99-123.
- 84 Taillardat M., Fougères A.-L., Naveau P. et al. Evaluating probabilistic forecasts of extremes using continuous ranked probability score distributions // *International Journal of Forecasting*. – 2023. – Vol. 39, Issue 3. – P. 1448-1459.
- 85 Xu C., Sun Y., Du A. et al. Quantile regression-based probabilistic forecasting of renewable energy generation and building electrical load: A state-of-the-art review // *Journal of Building Engineering*. – 2023. – Vol. 79. – P. 107772.
- 86 Kourentzes N., Trapero J.R., Barrow D.K. Optimising forecasting models for inventory planning // *International Journal of Production Economics*. – 2020. – Vol. 225. – P. 107597.
- 87 Vovk V., Gammerman A., Shafer G. *Algorithmic learning in a random world*. – Berlin: Springer, 2005. – 332 p.
- 88 Donti P.L., Amos B., Kolter J.Z. Task-based end-to-end model learning in stochastic optimization // <https://arxiv.org/abs/1703.04529>. 10.11.2025.
- 89 McShane B.B., Wyner A.J. A statistical analysis of multiple temperature proxies: Are reconstructions of surface temperatures over the last 1000 years reliable? // *The Annals of Applied Statistics*. – 2011. – Vol. 5, Issue 1. – P. 5-44.
- 90 Raftery A.E., Gneiting T., Balabdaoui F. et al. Using Bayesian model averaging to calibrate forecast ensembles // *Monthly Weather Review*. – 2005. – Vol. 133, Issue 5. – P. 1155-1174.
- 91 Bessa R.J., Miranda V., Botterud A. et al. Time-adaptive quantile-copula for wind power probabilistic forecasting // *Renewable Energy*. – 2012. – Vol. 40, Issue 1. – P. 29-39.
- 92 Wilks D.S. *Statistical methods in the atmospheric sciences*. – Ed. 3rd. – NY.: Academic Press, 2011. – 704 p.
- 93 Kendall M.G., Stuart A. *The advanced theory of statistics*. – London, 1979. – Vol. 2. – 468 p.
- 94 Hyndman R.J., Athanasopoulos G. *Forecasting: Principles and practice*. – Ed. 2nd. – Melbourne, 2018. – 384 p.
- 95 Siegmund D., Yakir B., Zhang N.R. Detecting simultaneous variant intervals in aligned sequences // <https://arxiv.org/abs/1108.3177v1>. 10.11.2025.
- 96 Meinshausen N. Quantile regression forests // *Journal of Machine Learning Research*. – 2006. – Vol. 7. – P. 983-999.

- 97 Chernozhukov V. et al. Quantile and probability curves without crossing // *Econometrica*. – 2010. – Vol. 78, Issue 3. – P. 1093-1125.
- 98 Taylor J.W. A quantile regression neural network approach to estimating the conditional density of multiperiod returns // *Journal of Forecasting*. – 2000. – Vol. 19. – P. 299-311.
- 99 Koenker R. Quantile regression: 40 years on // *Annual Review of Economics*. – 2017. – Vol. 9. – P. 155-176.
- 100 Taieb S.B., Huser R., Hyndman R.J. et al. Forecasting uncertainty in electricity smart meter data by boosting additive quantile regression // *IEEE Transactions on Smart Grid*. – 2016. – Vol. 7. – P. 2448-2455.
- 101 Christoffersen P. *Elements of financial risk management*. – Ed. 2nd. – London: Academic Press, 2011. – 327 p.
- 102 Bekci R.Y., Mahdid Y., Xing J. et al. Probabilistic models for manufacturing lead times // <https://arxiv.org/abs/2204.13792>. 10.10.2025.
- 103 Nowotarski J., Weron R. Recent advances in electricity price forecasting: A review of probabilistic forecasting // *Renewable and Sustainable Energy Reviews*. – 2018. – Vol. 81, Part 1. – P. 1548-1568.
- 104 Taleb N.N. *Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications*. – NY.: Academic Press, 2020. – 443 p.
- 105 Kawas B., Laumanns M., Pratsini E. et al. Risk-averse production planning // *Lecture Notes in Computer Science*. – 2011. – Vol. 6992. – P. 123-137.
- 106 Fischetti M., Fraccaro M. Machine learning meets mathematical optimization to predict the optimal production of offshore wind parks // *Computers & Operations Research*. – 2019. – Vol. 106. – P. 289-297.
- 107 Shapiro A., Dentcheva D., Ruszczyński A. *Lectures on stochastic programming: Modeling and theory*. – Ed. 2nd. – Philadelphia, 2021. – 541 p.
- 108 Birge J.R., Louveaux F. *Introduction to stochastic programming*. – Ed. 2nd. – NY.: Springer, 2011. – 485 p.
- 109 Campi M.C., Garatti S. The exact feasibility of randomized solutions of uncertain convex programs // *SIAM Journal on Optimization*. – 2008. – Vol. 19, Issue 3. – P. 1211-1230.
- 110 Ben-Tal A., El Ghaoui L., Nemirovski A. *Robust optimization*. – Princeton: Princeton University Press, 2009. – 563 p.
- 111 Bertsimas D., Sim M. The price of robustness // *Operations Research*. – 2004. – Vol. 52, Issue 1. – P. 35-53.
- 112 Rockafellar R.T., Uryasev S. Optimization of conditional value-at-risk // *Journal of Risk*. – 2000. – Vol. 3. – P. 21-41.
- 113 Prékopa A. *Stochastic programming*. – Dordrecht: Springer, 1995. – 600 p.
- 114 Delage E., Ye Y. Distributionally robust optimization under moment uncertainty with application to data-driven problems // *Operations Research*. – 2010. – Vol. 58. – P. 595-612.
- 115 Feng O.Y. et al. A unifying tutorial on approximate message passing // <https://arxiv.org/abs/2105.02180>. 10.10.2025.

- 116 Achamu G., Berhan E., Geremaw S. Demonstrating the interplay of machine learning and optimization methods for operational planning decision // *Journal of Data, Information and Management*. – 2021. – Vol. 3. – P. 297-324.
- 117 Brucker P. *Scheduling algorithms*. – Ed. 5th. – Berlin: Springer, 2007. – 383 p.
- 118 Monostori L., Kádár B., Bauernhansl T. et al. Cyber-physical systems in manufacturing // *CIRP Annals*. – 2016. – Vol. 65, Issue 2. – P. 621-641.
- 119 Borangiu T., Trentesaux D., Thomas A. et al. *Service orientation in holonic and multi-agent manufacturing*. – Cham: Springer, 2014. – 340 p.
- 120 Lichtendahl K.C., Grushka-Cockayne Y., Winkler R.L. Is it better to average probabilities or quantiles? // *Management Science*. – 2013. – Vol. 59, Issue 7. – P. 1594-1611.
- 121 Herroelen W., Leus R. Robust and reactive project scheduling: A review and classification of procedures // *International Journal of Production Research*. – 2004. – Vol. 42, Issue 8. – P. 1599-1620.
- 122 Sahinidis N.V. Optimization under uncertainty: State-of-the-art and opportunities // *Computers & Chemical Engineering*. – 2004. – Vol. 28, Issue 6-7. – P. 971-983.
- 123 Vieira G.E., Herrmann J.W., Lin E. Rescheduling manufacturing systems: A framework of strategies, policies, and methods // *Journal of Scheduling*. – 2003. – Vol. 6, Issue 1. – P. 39-62.
- 124 Shapiro A. Stochastic programming approach to optimization under uncertainty // *Mathematical Programming*. – 2008. – Vol. 112. – P. 183-220.
- 125 Kouvelis P., Yu G. *Robust discrete optimization and its applications*. – Alphen aan den Rijn, 1997. – 356 p.
- 126 Sen S., Higle J.L. An introductory tutorial on stochastic linear programming models // *Interfaces*. – 1999. – Vol. 29, Issue 2. – P. 33-61.
- 127 Bertsimas, D., Gupta V., Kallus N. Data-driven robust optimization // <https://arxiv.org/abs/1401.0212v1>. 10.11.2025.
- 128 Bertsimas D., Brown D.B., Caramanis C. Theory and applications of robust optimization // *SIAM Review*. – 2011. – Vol. 53. – P. 464-501.
- 129 Ben-Tal A., Melenberg B., Rennen G. Robust solutions of optimization problems affected by uncertain probabilities // *Management Science*. – 2013. – Vol. 59, Issue 2. – P. 341-357.
- 130 Mulvey J.M., Vanderbei R.J., Zenios S.A. Robust optimization of large-scale systems // *Operations Research*. – 1995. – Vol. 43, Issue 2. – P. 264-281.
- 131 Hedeia I. Algorithms for dynamic scheduling in manufacturing: Toward digital factories improving deadline feasibility and responsiveness via temporal networks // <https://arxiv.org/abs/2510.16047>. 11.11.2025.
- 132 Rychener Y., Kuhn D., Sutter T. End-to-end learning for stochastic optimization: A Bayesian perspective // <https://arxiv.org/abs/2306>. 11.11.2025.
- 133 Kotary J., Fioretto F., Van Hentenryck P. et al. End-to-end constrained optimization learning: A survey // <https://arxiv.org/abs/2103>. 11.11.2025.

- 134 Bengio Y., Lodi A., Prouvost A. Machine learning for combinatorial optimization: A methodological tour d'horizon // *European Journal of Operational Research*. – 2021. – Vol. 290, Issue 2. – P. 405-421.
- 135 Escudero L.F., Garín M.A., Unzueta A. Cluster Lagrangean decomposition in multistage stochastic optimization // *Computers & Operations Research*. – 2016. – Vol. 67. – P. 48-62.
- 136 Zhou Y., Ma S. Stochastic optimization methods for policy evaluation in reinforcement learning // *Foundations and Trends in Optimization*. – 2024. – Vol. 6, Issue 3. – P. 145-192.
- 137 Xiong H., Shi S., Ren D. et al. A survey of job shop scheduling problem: The types and models // *Computers & Operations Research*. – 2022. – Vol. 142. – P. 105731.
- 138 Xu M., Mei Y., Zhang F. et al. Learn to optimise for job shop scheduling: A survey with comparison between genetic programming and reinforcement learning. *Artificial Intelligence Review*. – 2025. – Vol. 58. – P. 160.
- 139 Viana M., Morandin Junior O., Contreras R. Transgenic genetic algorithm to minimize the makespan in the job shop scheduling problem // *Proceed. of the 12th internat. conf. on Agents and Artificial Intelligence*. – Valletta, 2020. – Vol. 2. – P. 463-474.
- 140 Arisha A., Young P., El Baradie M. Job shop scheduling problem: An overview // *Proceed. of the internat. conf. on Flexible Automation and Intelligent Manufacturing (FAIM 2001)*. – Dublin, 2001. – P. 682-693.
- 141 Breit M., Weiß M., Obermaisser R. et al. Modeling and analysis of configurable job-shop scheduling problems // *Proceed. of the 19th internat. Working conf. on Variability Modelling of Software-Intensive Systems (VaMoS '25)*. – NY., 2025. – P. 77-85.
- 142 Zupan H., Heraković N., Žerovnik J. A robust heuristics for the online job shop scheduling problem // *Algorithms*. – 2024. – Vol. 17, Issue 12. – P. 568.
- 143 Raja L., Hamid S., Warsi S.S. et al. Critical analysis of job shop scheduling in context of Industry 4.0 // *Sustainability*. – 2021. – Vol. 13, Issue 14. – P. 7684.
- 144 Wang X., Wang B., Zhang X. et al. Two-objective robust job-shop scheduling with two problem-specific neighborhood structures // *Swarm and Evolutionary Computation*. – 2021. – Vol. 61. – P. 100805.
- 145 He W., Sun D. Scheduling flexible job shop problem subject to machine breakdown with route changing and right-shift strategies // *The International Journal of Advanced Manufacturing Technology*. – 2013. – Vol. 66. – P. 501-514.
- 146 Moghaddam S.K., Houshmand M., Saitou K. et al. Configuration design of scalable reconfigurable manufacturing systems for part family // *International Journal of Production Research*. – 2020. – Vol. 58, Issue 10. – P. 2974-2996.
- 147 Para J., Del Ser J., Nebro A.J. Energy-aware multi-objective job shop scheduling optimization with metaheuristics in manufacturing industries: A critical survey, results, and perspectives // *Applied Sciences*. – 2022. – Vol. 12, Issue 3. – P. 1491.

- 148 Erlenbusch F., Stricker N. Towards practicality: Navigating challenges in designing predictive-reactive scheduling // *Procedia CIRP*. – 2024. – Vol. 122. – P. 701-706.
- 149 Mihoubi B., Bouzouia B., Gaham M. Reactive scheduling approach for solving a realistic flexible job shop scheduling problem // *International Journal of Production Research*. – 2021. – Vol. 59, Issue 19. – P. 5790-5808.
- 150 Dauzère-Pérès S., Ding J., Shen L. et al. The flexible job shop scheduling problem: A review // *European Journal of Operational Research*. – 2024. – Vol. 314, Issue 2. – P. 409-432.
- 151 Caldeira R.H., Gnanavelbabu A. A simheuristic approach for the flexible job shop scheduling problem with stochastic processing times // *Simulation*. – 2020. – Vol. 97, Issue 3. – P. 215-236.
- 152 Zambrano-Rey G.M. et al. Minimizing the expected maximum lateness for a job shop subject to stochastic machine breakdowns // *Annals of Operations Research*. – 2024. – Vol. 338. – P. 801-833.
- 153 Boukedroun M., Duvivier D., Ait-El-Cadi A. et al. A hybrid genetic algorithm for stochastic job-shop scheduling problems // *RAIRO – Operations Research*. – 2023. – Vol. 57. – P. 1617-1645.
- 154 Calamita A., Meloni C., Pranzo M. et al. Fast risk estimation for job shop scheduling solutions under interval uncertainty via machine learning // <https://doi.org/10.1007/s10696-025-09640-7>. 10.11.2025.
- 155 Díaz H., Palacios J.J., Díaz I. et al. Robust schedules for tardiness optimization in job shop with interval uncertainty // *Logic Journal of the IGPL*. – 2023. – Vol. 31, Issue 2. – P. 240-254.
- 156 Zheng P., Xiao S., Zhang P. et al. A two-individual-based evolutionary algorithm for flexible assembly job shop scheduling problem with uncertain interval processing times // *Applied Sciences*. – 2024. – Vol. 14, Issue 22. – P. 10304.
- 157 Sun H., Jiang A., Ge D. et al. A chance constrained programming approach for no-wait flow shop scheduling problem under interval-valued fuzzy processing time // *Processes*. – 2021. – Vol. 9, Issue 5. – P. 789.
- 158 Kasperski A., Zieliński P. Risk-averse single machine scheduling: Complexity and approximation // *J of Scheduling*. – 2019. – Vol. 22. – P. 567-580.
- 159 Meloni C., Pranzo M., Samà M. Evaluation of VaR and CVaR for the makespan in interval valued blocking job shops // *International Journal of Production Economics*. – 2022. – Vol. 247. – P. 108455.
- 160 Chang Z., Ding J.-Y., Song S. Distributionally robust scheduling on parallel machines under moment uncertainty // *European Journal of Operational Research*. – 2019. – Vol. 272, Issue 3. – P. 832-846.
- 161 Lu S., Wei D., Wang Y. et al. A joint chance constrained optimization algorithm with robust reconstruction for multi-echelon and multi-period closed-loop manufacturing adjustable system design under multi-source uncertainty // *Journal of Cleaner Production*. – 2023. – Vol. 410. – P. 137314.
- 162 Lu M., Shen Z.M. A review of robust operations management under model uncertainty // *Production and Operations Management*. – 2021. – Vol. 30, Issue 6. – P. 1927-1943.

- 163 Gao K., Cao Z., Zhang L. et al. A review on swarm intelligence and evolutionary algorithms for solving flexible job shop scheduling problems // IEEE/CAA Journal of Automatica Sinica. – 2019. – Vol. 6, Issue 4. – P. 904-916.
- 164 Meng L., Zhang C., Ren Y. et al. Mixed-integer linear programming and constraint programming formulations for solving distributed flexible job shop scheduling problem // Computers & Industrial Engineering. – 2020. – Vol. 142. – P. 106347.
- 165 Boyer V. et al. The generalized flexible job shop scheduling problem // Computers & Industrial Engineering. – 2021. – Vol. 160. – P. 107542.
- 166 Shady S., Kaihara T., Fujii N. et al. Evolving dispatching rules using genetic programming for multi-objective dynamic job shop scheduling with machine breakdowns // Procedia CIRP. – 2021. – Vol. 104. – P. 411-416.
- 167 Defersha F.M., Obimuyiwa D., Yimer A.D. Mathematical model and simulated annealing algorithm for setup operator constrained flexible job shop scheduling problem // Computers & Industrial Engineering. – 2022. – Vol. 171. – P. 108487.
- 168 Ferreira C., Figueira G., Amorim P. Effective and interpretable dispatching rules for dynamic job shops via guided empirical learning // Omega. – 2022. – Vol. 111. – P. 102643.
- 169 Naderi B., Ruiz R., Roshanaei V. Mixed-integer programming vs. constraint programming for shop scheduling problems: New results and outlook // INFORMS Journal on Computing. – 2023. – Vol. 35, Issue 4. – P. 817-843.
- 170 Schutt A., Feydy T., Stuckey P.J. et al. Why cumulative decomposition is not as bad as it sounds // Lecture Notes in Computer Science. – 2009. – Vol. 5732. – P. 746-761.
- 171 Perron L., Didier F., Gay S. The CP-SAT-LP solver (Invited talk) // Proc. 29th internat. conf. on princip.-pract. of Constraint Programming (CP 2023). – Toronto, 2023. – P. 1-2.
- 172 Da Col G., Teppan E.C. Industrial-size job shop scheduling with constraint programming // Operations Research Perspectives. – 2022. – Vol. 9. – P. 100249.
- 173 Kasapidis G.A., Paraskevopoulos D.C., Mourtos I. et al. A unified solution framework for flexible job shop scheduling problems with multiple resource constraints // European Journal of Operational Research. – 2025. – Vol. 320, Issue 3. – P. 479-495.
- 174 Sun Y., Nguyen S., Thiruvady D. et al. Enhancing constraint programming via supervised learning for job shop scheduling // Knowledge-Based Systems. – 2024. – Vol. 293. – P. 111698.
- 175 Vanderschueren T., Verdonck T., Baesens B. et al. Predict-then-optimize or predict-and-optimize? An empirical evaluation of cost-sensitive learning strategies // Information Sciences. – 2022. – Vol. 594. – P. 400-415.
- 176 Mišić V.V., Perakis G. Data analytics in operations management: A review // Manufacturing & Service Operations Management. – 2019. – Vol. 22, Issue 1. – P. 158-169.



- 177 Mandi J., Kotary J., Berden S. et al. Decision-focused learning: Foundations, state of the art, benchmark and future opportunities // *Journal of Artificial Intelligence Research*. – 2024. – Vol. 80. – P. 79-140.
- 178 Wilder B., Dilkina B., Tambe M. Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization // *Proceedings of the AAAI Conference on Artificial Intelligence*. – 2019. – Vol. 33, Issue 1. – P., 1658-1665.
- 179 Dai H. et al. Learning combinatorial optimization algorithms over graphs // <https://arxiv.org/abs/1704.01665>. 10.12.2025.
- 180 Relich M. Predictive and prescriptive analytics in identifying opportunities for improving sustainable manufacturing // *Sustainability*. – 2023. – Vol. 15, Issue 9. – P. 7667.
- 181 Meloni C., Pranzo M. Evaluation of the quantiles and superquantiles of the makespan in interval valued activity networks // *Computers & Operations Research*. – 2023. – Vol. 151. – P. 106098.
- 182 Salari N., Liu S., Shen Z.-J.M. Real-time delivery time forecasting and promising in online retailing: When will your package arrive? // *Manufacturing & Service Operations Management*. – 2022. – Vol. 24, Issue 3. – P. 1421-1436.
- 183 Bruni M.E., Khodaparasti S., Demeulemeester E. The distributionally robust machine scheduling problem with job selection and sequence-dependent setup times // *Computers & Operations Research*. – 2020. – Vol. 123. – P. 105017.
- 184 Liu C., Letchford A.N., Svetunkov I. Newsvendor problems: An integrated method for estimation and optimization // *European Journal of Operational Research*. – 2022. – Vol. 300, Issue 2. – P. 590-601.
- 185 Romano Y., Patterson E., Candès E.J. Conformalized quantile regression // <https://arxiv.org/abs/1905.03222>. -10.12.2025.
- 186 Makridakis S., Spiliotis E., Assimakopoulos V. The M5 competition: Background, organization, and implementation // *International Journal of Forecasting*. – 2022. – Vol. 38, Issue 4. – P. 1325-1336
- 187 Saxena A., Goebel K. Turbofan engine degradation simulation data set // <https://www.nasa.gov/intelligent-systems-division/discovery-and>. 10.10.2025.
- 188 Hong T., Fan S. Probabilistic electric load forecasting: A tutorial review // *International Journal of Forecasting*. – 2016. – Vol. 32, Issue 3. – P. 914-938.
- 189 Garey M.R., Johnson D.S., Sethi R. The complexity of flowshop and jobshop scheduling // *Mathematics of Operations Research*. – 1976. – Vol. 1, Issue 2. – P. 117-129.
- 190 Mazyavkina N., Sviridov S., Ivanov S. et al. Reinforcement learning for combinatorial optimization: A survey. *Computers & Operations Research*. – 2021. – Vol. 134. – P. 105400.
- 190 Carvalho T.P., Soares F.A.A.M.N., Vita R. et al. A systematic literature review of machine learning methods applied to predictive maintenance // *Computers & Industrial Engineering*. – 2019. – Vol. 137. – P. 106024.
- 192 Zhang W., Yang D., Wang H. Data-driven methods for predictive maintenance of industrial equipment: A survey // *IEEE Systems Journal*. – 2019. – Vol. 13, Issue 3. – P. 2213-2227.

- 193 Kourentzes N., Petropoulos F., Trapero J.R. Improving forecasting by estimating time series structural components across multiple frequencies // International Journal of Forecasting. – 2014. – Vol. 30, Issue 2. – P. 291-302.
- 194 Salinas D., Flunkert V., Gasthaus J. et al. DeepAR: Probabilistic forecasting with autoregressive recurrent networks // International Journal of Forecasting. – 2020. – Vol. 36, Issue 3. – P. 1181-1191.
- 195 Bandara K., Bergmeir C., Smyl S. Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach // Expert Systems with Applications. – 2020. – Vol. 140. – P. 112896.
- 196 Li X., Ding Q., Sun J.-Q. Remaining useful life estimation in prognostics using deep convolution neural networks. Reliability Engineering & System Safet. – 2018. – Vol. 172. – P. 1-11.
- 197 Zheng S., Ristovski K., Farahat A. et al. Long short-term memory network for remaining useful life estimation // Proceed. of the IEEE internat. conf. on Prognostics and Health Management (ICPHM 2017). – Dallax, 2017. – P. 88-95.
- 198 Kong W., Dong Z.Y., Jia Y. et al. Short-term residential load forecasting based on LSTM recurrent neural network // IEEE Transactions on Smart Grid. – 2019. – Vol. 10, Issue 1. – P. 841-851.
- 199 Taieb S.B., Hyndman R.J. A gradient boosting approach to the Kaggle load forecasting competition // International Journal of Forecasting. – 2014. – Vol. 30, Issue 2. – P. 382-394.
- 200 Marino D.L., Amarasinghe K., Manic M. Building energy load forecasting using deep neural networks // Proceed. of the 42nd Annual conf. of the IEEE Industrial Electronics Society (IECON 2016). – Florence, 2016. – P. 7046-7051.
- 201 Zhang C., Song W., Cao Z. et al. Learning to dispatch for job shop scheduling via deep reinforcement learning // Proceed. of the 34th conf. on Neural Information Processing Systems (NeurIPS 2020). – NY., 2020. – P. 1621-1632.
- 202 Lv L., Zhang C., Fan J. et al. Deep reinforcement learning for job shop scheduling problems: A comprehensive literature review // Knowledge-Based Systems. – 2025. – Vol. 321. – P. 113633.
- 203 Yang H., Li W., Wang B. Joint optimization of preventive maintenance and production scheduling for multi-state production systems based on reinforcement learning // Reliability Engineering & System Safety. – 2021. – Vol. 214. – P. 107713.
- 204 Харченко Д., Мадин В., Салыкова О. и др. Анализ моделей RNN и Transformer в задачах обработки естественного языка // Вестник КазАТК. – 2025. – №138(3). – С. 311-321.
- 205 Салыкова О., Мадин В., Самойленко Д. и др. Сравнительный анализ форматов протоколов для обработки данных в сенсорных сетях интернета вещей // Вестник КазАТК. – 2024. – №134(5). – С. 465-473.
- 206 Мадин В., Салыкова О., Иванова И. Сверточные и глубокие нейронные сети // Вестник КазАТК. – 2023. – №126(3). – С. 189-197.
- 207 Yang S., Liao H., Wu X. Prescriptive analytics for dynamic multi-criterion decision making considering learned knowledge of alternatives // Expert Systems with Applications. – 2025. – Vol. 268. – P. 126350.

208 Агдавлетова А., Мадин В., Салыкова О. Adaptive process management using deep learning on a programmable logic controller (PLC) // Международный журнал информационных и коммуникационных технологий. – 2025. – №5(1). – С. 8-28.

209 Becht E., McInnes L., Healy J. et al. Dimensionality reduction for visualizing single-cell data using UMAP // Nature Biotechnology. – 2019. – Vol. 37. – P. 38-44.

210 Lim B., Arik S.Ö., Loeff N. et al. Temporal fusion transformers for interpretable multi-horizon time series forecasting // International Journal of Forecasting. – 2021. – Vol. 37, Issue 4. – P. 1748-1764.

211 Petropoulos F. et al. “Horses for courses” in demand forecasting // European Journal of Operational Research. – 2014. – Vol. 237, Issue 1. – P. 152-163.

212 Bentéjac C., Csörgő A., Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms // Artificial Intelligence Review. – 2021. – Vol. 54. – P. 1937-1967.

213 Chaves Sousa A.C. et al. Comparing gradient boosting algorithms to forecast sales in retail // Proceed. of the Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2023). – Belo Horizonte, 2023. – P. 596-609.

214 Lundberg S.M., Erion G., Chen H. From local explanations to global understanding with explainable AI for trees // Nature Machine Intelligence. – 2020. – Vol. 2. – P. 56-67.

215 Eren Y., Küçükdemiral İ. A comprehensive review on deep learning approaches for short-term load forecasting // Renewable and Sustainable Energy Reviews. – 2024. – Vol. 189, Part B. – P. 114031.

216 Zhang Y., Xin Y., Liu Z.-W. et al. Health status assessment and remaining useful life prediction of aero-engine based on BiGRU and MMoE // Reliability Engineering & System Safety. – 2022. – Vol. 220. – P. 108263.

217 Ke G., Xu Z., Zhang J. et al. DeepGBM: A deep learning framework distilled by GBDT for online prediction tasks // Proceed. of the 25th ACM SIGKDD internat. conf. on Knowledge Discovery & Data Mining (KDD '19). – NY., 2019. – P. 384-394.

218 Madin V., Salykova O., Ivanova I. et al. Enhancing electricity consumption forecasting in the Republic of Kazakhstan using machine learning // Journal of Applied Engineering and Technological Science. – 2025. – Vol. 6, Issue 2. – P. 1166-1196.

219 Semmelmann L., Henni S., Weinhardt C. Load forecasting for energy communities: A novel LSTM-XGBoost hybrid model based on smart meter data // Energy Informatics – 2022. – Vol. 5. – P. 24-1-24-22.

220 Мадин В., Комаров Д., Салыкова О. и др. Предиктивное обслуживание шпиндельных подшипников на основе вибродиагностики (на примере токарного станка СКЕ 6136/6150) // Вестник КазУТБ. – 2025. – №3(28). – С. 63-79.

221 Ribeiro M.T., Wu T., Guestrin C. et al. Beyond accuracy: Behavioral testing of NLP models with CheckList // <https://arxiv.org/abs/2005.04118>. 11.12.2025.

222 Müller A., Grumbach F. Predicting processing times in high mix low volume job shops // <http://dx.doi.org/10.25673/103491>. 10.11.2025.

- 223 Schaller J., Valente J. Branch-and-bound algorithms for minimizing total earliness and tardiness in a two-machine permutation flow shop with unforced idle allowed // *Computers & Operations Research*. – 2019. – Vol. 109. – P. 1-11.
- 224 Lodi A., Zarpellon G. On learning and branching: A survey // *TOP*. – 2017. – Vol. 25. – P. 207-236.
- 225 Stankevičiūtė K., Alaa A.M., van der Schaar M. Conformal time-series forecasting // *Proceed. of the 35th internat. conf. on Neural Information Processing Systems (NeurIPS 2021)*. – NY., 2021. – P. 6216-6228.
- 226 Xu C., Xie Y. Conformal prediction set for time-series // <https://arxiv.org/abs/2206.07851>. 10.10.2025.
- 227 Chen R., Paschalidis I.Ch. Distributionally robust learning // *Foundations and Trends® in Optimization*. – 2020. – Vol. 4, Issue 1-2. – P. 1-243.
- 228 Мадин В.А., Салыкова О.С., Агдавлетова А.А. и др. Dynamic reconfiguration of control algorithms in a programmable logic controller for flexible automation of production processes // *Вестник ВКТУ им. Д. Серикбаева*. – 2025. – №1(4). – С. 174-185.
- 229 Kesgin R.İ.T. Revisiting trend stability using Mann-Kendall and Wilcoxon signed-rank tests through innovative method comparisons // *Sustainability*. – 2025. – Vol. 17, Issue 23. – P. 10454.



## ПРИЛОЖЕНИЕ А

### Свидетельства об авторском праве

КАЗАҚСТАН РЕСПУБЛИКАСЫ      РЕСПУБЛИКА КАЗАХСТАН



**СВИДЕТЕЛЬСТВО**  
**О ВНЕСЕНИИ СВЕДЕНИЙ В ГОСУДАРСТВЕННЫЙ РЕЕСТР**  
**ПРАВ НА ОБЪЕКТЫ, ОХРАНЯЕМЫЕ АВТОРСКИМ ПРАВОМ**

№ 66964 от «2» февраля 2026 года

Фамилия, имя, отчество, (если оно указано в документе, удостоверяющем личность) автора (ов):  
**МАДИН ВЛАДИМИР АНАТОЛЬЕВИЧ,      САЛЫКОВА ОЛЫГА СЕРГЕЕВНА**

Вид объекта авторского права: компьютерные программы (программное обеспечение)

Название объекта: Программа интеллектуального сценарного планирования с использованием методов машинного обучения и оптимизации

Дата создания объекта: 30.10.2025





Краткая ссылка: <http://www.kazpatent.kz/nr/66964>  
"Авторский код": <https://copyright.kazpatent.kz>

Подлинность документа возможно проверить на сайте [kazpatent.kz](http://www.kazpatent.kz)  
в разделе «Авторское право» <https://copyright.kazpatent.kz>

Подписано ЭЦП      Л. Уташева



ҚАЗАҚСТАН РЕСПУБЛИКАСЫ



РЕСПУБЛИКА КАЗАХСТАН

## СВИДЕТЕЛЬСТВО

О ВНЕСЕНИИ СВЕДЕНИЙ В ГОСУДАРСТВЕННЫЙ РЕЕСТР  
ПРАВ НА ОБЪЕКТЫ, ОХРАНЯЕМЫЕ АВТОРСКИМ ПРАВОМ

№ 61130 от «31» июля 2025 года

Фамилия, имя, отчество, (если оно указано в документе, удостоверяющем личность) автора (ов):  
**МАДИН ВЛАДИМИР АНАТОЛЬЕВИЧ**

Вид объекта авторского права: **программа для ЭВМ**

Название объекта: **Программа предиктивной оценки остаточного срока службы подшипников**

Дата создания объекта: **29.07.2025**



Копия прилагается: <http://www.kazpatent.kz/ru/сайтынын>  
"Авторлық құқық" бөлімінде тексеріле болады: <https://copyright.kazpatent.kz>

Подлинность документа возможно проверить на сайте [kazpatent.kz](http://kazpatent.kz)  
в разделе «Авторское право»: <https://copyright.kazpatent.kz>

Подписано ЭЦП

С. Ахметов



## ПРИЛОЖЕНИЕ Б

### Акт внедрения



«УТВЕРЖАЮ»  
Директор ТОО «ОПТИСОРТ»  
Тайжанов Е.С.

10 ноября 2025

#### АКТ о внедрении результатов диссертационного исследования МАДИНА ВЛАДИМИРА АНАТОЛЬЕВИЧА

**Комиссия в составе:**

**Председатель:** директор – Тайжанов Ерлан Сайранович

**Члены комиссии:** заместитель директора по производству – Дереха Валерий Валерьевич, инженер-конструктор – Сусуев Иван Владимирович.

Составили настоящий акт о том, что результаты диссертационной работы «Разработка и исследование моделей интеллектуального анализа больших данных в задачах планирования», выполненной в рамках проекта программно-целевого финансирования по теме BR24992785 «Организация и проведение комплексных исследований по обеспечению устойчивого развития агропромышленного комплекса Костанайской области с созданием научно-исследовательского технологического центра» представленной на соискание степени доктора философии (PhD) по направлению подготовки кадров «8D061 – Информационно-коммуникационные технологии», использованы в ТОО «ОПТИСОРТ»:

1. Программный прототип интеллектуальной системы календарного планирования, обеспечивающий формирование производственных расписаний на основе анализа производственных данных и сценарных ограничений.
2. Инструменты визуализации и анализа результатов планирования, включающие представление расписаний в виде диаграмм Ганта, расчёт ключевых показателей эффективности и сравнительный анализ сценариев планирования, обеспечивающие повышение прозрачности и интерпретируемости принимаемых управленческих решений.

Результаты диссертационного исследования Мадина В.А., на тему «Разработка и исследование моделей интеллектуального анализа больших данных в задачах планирования» успешно внедрены в производственную деятельность ТОО «ОПТИСОРТ» и показывают устойчивую положительную эффективность.

**Председатель комиссии**

Директор ТОО «ОПТИСОРТ»

**Члены комиссии:**

Зам. директора по производству

Инженер-конструктор



Е. Тайжанов

В. Дереха

И. Сусуев