

Федеральное государственное бюджетное учреждение науки
Институт систем энергетики им. Л.А. Мелентьева Сибирского отделения
Российской академии наук

На правах рукописи



Цыбиков Алексей Ринчинович

**МЕТОДИЧЕСКИЙ ПОДХОД К ПОСТРОЕНИЮ
ЦИФРОВЫХ ДВОЙНИКОВ ЭНЕРГЕТИЧЕСКИХ ОБЪЕКТОВ
С ПРИМЕНЕНИЕМ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ
(НА ПРИМЕРЕ СОЛНЕЧНОЙ ЭЛЕКТРОСТАНЦИИ)**

Специальность 1.2.2. Математическое моделирование, численные методы и
комплексы программ

ДИССЕРТАЦИЯ

на соискание ученой степени
кандидата технических наук

Научный руководитель:
доктор технических наук, профессор,
Массель Людмила Васильевна

Иркутск – 2025

Оглавление

Введение	4
Глава 1. Анализ проблемы и постановка задач.....	14
1.1. Солнечная электроэнергетика.....	14
1.2. Цифровые двойники.....	23
1.3. Применение цифровых двойников в энергетике	33
1.4. Методы искусственного интеллекта и машинного обучения.....	38
1.5. Выводы по главе 1	51
Глава 2. Методический подход к построению умных цифровых двойников ВИЭ	54
2.1. Модель умного цифрового двойника	54
2.2. Агрегированная модель солнечной электростанции как основа вычислительного ядра умного цифрового двойника	58
2.3. Методика построения виртуальной среды	64
2.4. Интеграция методов машинного обучения в виртуальную среду цифрового двойника	70
2.5. Численный метод ансамблирования с выбором моделей на основе состояния внешней среды (EnvSwitch)	77
2.6. Методика построения умного цифрового двойника.....	80
2.7. Проектирование цифрового двойника солнечной электростанции	82
2.8. Выводы по главе 2	88
Глава 3. Разработка и применение цифрового двойника солнечной электростанции.....	91
3.1. Архитектура умного цифрового двойника	91
3.2. Реализация компонентов цифровой тени и цифровой модели.....	97

3.3.	Анализ и подготовка данных для виртуальной среды цифрового двойника солнечной электростанции.....	105
3.4.	Проведение вычислительного эксперимента для эмуляции данных (применение предложенного численного метода).....	114
3.5.	Выводы по главе 3	125
	Заключение	127
	Список сокращений и условных обозначений.....	129
	Список литературы	131
	Приложение А. Свидетельства о регистрации программ и баз данных.....	145

Введение

Актуальность работы. Концепция цифровых двойников (ЦД) является составляющей четвертой промышленной революции, связанной с направлением «Интернет вещей». Эта концепция описывает цифровой двойник как виртуальный прототип физического объекта, предполагающий обмен данными с реальным объектом. Цифровые двойники позволяют оценивать ситуацию, оказывать помощь при проектировании энергетического объекта, прогнозировать его поведение, наблюдать за процессами, связанными с объектом, оценивать ситуацию, а также использоваться в системах управления и поддержки принятия решений.

Особое значение имеет рассмотрение цифрового двойника как инструмента сопровождения объекта на различных стадиях жизненного цикла – от проектирования до эксплуатации и модернизации. Согласно ГОСТ Р 57700.37–2021 [1], цифровой двойник разрабатывается и применяется на всех стадиях жизненного цикла изделия, изменяясь на каждой стадии, а его наполнение и функциональность зависят от текущей стадии жизненного цикла. В стандарте также отмечается, что применение цифрового двойника на стадии разработки позволяет обосновывать проектные решения, проводить цифровые испытания, анализировать влияние изменений параметров и определять критические зоны, включая обоснование состава датчиков для последующей эксплуатации.

На этапе проектирования солнечной электростанции (СЭС) цифровой двойник может выполнять функции цифрового прототипа: обеспечивать формирование структуры объекта, выбор состава оборудования, расчёт ожидаемой выработки, анализ влияния климатических факторов и проверку проектных решений в виртуальной среде. Такой подход соответствует современному пониманию цифрового двойника как средства интеграции физических и виртуальных данных, используемых для поддержки проектирования, производства и сервисного сопровождения объекта.

После ввода солнечной электростанции в эксплуатацию функциональность цифрового двойника изменяется: на первый план выходят сбор и обработка телеметрии, формирование цифровой тени, уточнение параметров моделей, прогнозирование выработки, диагностика отклонений и поддержка решений по техническому обслуживанию. В этом случае цифровой двойник становится динамической системой, в которой физический объект и его цифровое представление связаны потоками данных и механизмами обратной связи.

Для фотоэлектрических систем жизненный цикл цифрового двойника особенно важен, поскольку на разных стадиях решаются различные задачи: при проектировании – выбор конфигурации и оценка ожидаемой генерации; при вводе в эксплуатацию – проверка соответствия расчётных и фактических параметров; при эксплуатации – мониторинг, прогнозирование, выявление аномалий и снижение затрат на обслуживание. В современных обзорах по цифровым двойникам фотоэлектрических установок отмечается, что такие системы применяются для повышения эффективности генерации, прогнозирования энергии и уменьшения затрат на эксплуатацию и техническое обслуживание, а также для более обоснованного принятия решений на основе данных реального времени.

Концепция цифровых двойников не нова, но лишь в 2016 году появилась на кривой стратегии развития значимых технологий для Интернета вещей института Gartner¹, где была обозначена как технология на стадии запуска. Развитие началось в 2017 году, когда эту технологию американская консалтинговая компания Gartner включила в цикл новых технологий с высоким потенциалом [2], в 2019 г. она оказалась уже на пике кривой Гартнера.

В 2021 году в рамках Петербургского международного экономического форума был представлен ежегодный доклад «10 прорывных идей в энергетике на следующие 10 лет» (Global Energy Association, 2021). В это исследование [3] была включена статья директора Института систем энергетики им. Л.А. Мелентьева СО РАН В.А. Стенникова, посвященная цифровым двойникам. В статье

¹ Gartner – исследовательская и консалтинговая компания, публикующая аналитические материалы и прогнозы в области информационных технологий и цифровой трансформации

систематизированы их классификационные признаки, проанализированы ключевые преимущества и технологические ограничения при реализации, рассмотрены аспекты информационного обеспечения, методологические подходы к построению, а также практические кейсы внедрения цифровых двойников в различных отраслях промышленности.

В последнее время появился термин «умный цифровой двойник» (Smart Digital Twin), который до сих пор четко не определен, но очевидно, что в его состав должны входить методы искусственного интеллекта. Необходимы разработка методов построения и модели умного цифрового двойника, выбор моделей машинного обучения для использования в ЦД энергетических объектов, в частности, в ЦД солнечной электростанции. Кроме того, недостаточно проработан вопрос использования научного инструментария (результатов исследования систем и объектов энергетики, неявных знаний экспертов) при построении ЦД. Это обуславливает актуальность предложенной работы, в которой предлагается решение поставленных вопросов.

Степень разработанности проблемы. В настоящее время существует ряд работ, посвящённых изучению данной проблемы.

В ИСЭМ СО РАН исследования энергетических систем рассматривались в работах Воропая Н.И., Криворучского Л.Д., Кононова Ю.Д., Мелентьева Л.Н., Макарова А.А., Санеева Б.Г. и др. Исследования цифровых двойников представлены в работах М. Гривза (M. Grieves), К. Шваба (K. Schwab), Боровкова А.И., Ковалева С.П., Прохорова А.Н., Бахтадзе Н.Н. и др. Вопросы применения интеллектуальных технологий в исследованиях энергетики и построении цифровых двойников рассматривались в работах: Массель Л.В., Масселя А.Г., Новицкого Н.Н., Ковалева С.П., Ворожцовой Т.Н., Макагоновой Н.Н., Томина Н.В., Гальперова В.И., Барахтенко Е.А., Соколова Д.В., Домышева А.В., Сидорова Д.Н., Волошина А.А., Бахтадзе Н.Н., Веселова Ф.В., Илюшина П.В., Кулагина В.А. и др.

В то же время в этих работах не рассматривалась задача разработки интегрированного методического подхода к построению цифровых двойников, объединяющего проектирование компьютерных моделей, цифровой тени,

основанной на моделировании данных, и методов искусственного интеллекта (онтологического моделирования и методов машинного обучения) в рамках умного цифрового двойника. В диссертационной работе ставится и решается эта задача.

Цель работы: Разработка методического подхода к построению умных цифровых двойников энергетических объектов с использованием интеллектуальных технологий, в частности, моделей машинного обучения.

Поставлены и решены следующие задачи:

1. Выполнен анализ существующих математических моделей и методов машинного обучения и возможностей их применения при разработке умных цифровых двойников.

2. Построена агрегированная математическая модель СЭС и предложена формализованная модель умного цифрового двойника на основе модификации базовой модели ЦД.

3. Выполнена адаптация численного метода ансамблирования моделей с динамическим выбором на основе состояния внешней среды для повышения точности прогнозирования временных рядов в составе виртуальной среды цифрового двойника солнечной электростанции.

4. Разработаны методики построения умного цифрового двойника с применением онтологического и математического моделирования и моделей машинного обучения, и виртуальной среды для эмулирования данных цифрового двойника солнечной электростанции на этапе научного прототипирования.

5. Разработана агентно – сервисная архитектура цифрового двойника солнечной электростанции и реализован программный комплекс (научный прототип умного цифрового двойника солнечной электростанции), интегрирующий основные базовые компоненты ЦД.

Объектом исследования являются энергетические объекты – солнечные электростанции, включающие фотоэлектрические модули, инверторное оборудование, системы накопления электроэнергии, коммутационное, измерительное и управляющее оборудование, а также средства сбора и передачи данных.

Предметом исследования являются методы построения умных цифровых двойников для проектирования солнечных электростанций.

Методология и методы исследования: для решения поставленных задач использовались методы математического моделирования, методические основы построения интеллектуальных информационных систем в исследованиях энергетики; методы инженерии знаний; методы анализа данных, методы машинного обучения; методы агентно-ориентированного проектирования, программирования и разработки многоагентных систем; методы проектирования и программирования web – приложений с использованием средств языка Python.

Достоверность научных результатов исследования обеспечивается разработанными и корректными математическими моделями компонентов солнечных электростанций и методами машинного обучения, применением реальных и эмулированных данных при исследовании виртуальной среды цифрового двойника солнечной электростанции, сравнительным анализом результатов работы моделей машинного обучения по стандартным метрикам качества.

Научная новизна диссертационной работы заключается в комплексном подходе к построению умного цифрового двойника солнечной электростанции, объединяющем математическое моделирование энергетического объекта, семантическое описание предметной области, интеллектуальные методы анализа данных и виртуальную среду. Получены следующие новые научные результаты:

1. Предложен методический подход к построению ЦД СЭС, отличающийся интеграцией математических и семантических моделей с интеллектуальными компонентами машинного обучения и включающий методику построения умного цифрового двойника энергетического объекта и численный метод ансамблирования моделей с динамическим выбором на основе состояния внешней среды.

2. Построены агрегированная математическая модель СЭС и формализованная модель умного цифрового двойника на основе модификации

базовой модели цифрового двойника, отличающаяся добавлением описаний моделей искусственного интеллекта, базы знаний и виртуальной среды.

3. Выполнена адаптация численного метода ансамблирования моделей машинного обучения с динамическим выбором, отличающаяся использованием онтологий и учетом состояния внешней среды.

4. Предложена методика построения виртуальной среды, отличающаяся возможностью эмуляции данных для отладки цифрового двойника (научного прототипирования).

5. Разработана агентно-сервисная архитектура умного цифрового двойника энергетического объекта и выполнена реализация программного комплекса, интегрирующего вычислительное ядро, цифровую тень и виртуальную среду, и являющегося научным прототипом цифрового двойника солнечной электростанции.

Положения, выносимые на защиту:

1. Методический подход к построению ЦД СЭС, включающий методику построения умного цифрового двойника энергетического объекта, основанную на интеграции математических и семантических моделей с интеллектуальными компонентами машинного обучения, и применение численного метода ансамблирования моделей с динамическим выбором на основе состояния внешней среды (соответствует п. 2 паспорта специальности).

2. Агрегированная математическая модель СЭС и формализованная модель умного цифрового двойника на основе модификации базовой модели цифрового двойника (соответствует п. 2 паспорта специальности).

3. Численный метод ансамблирования моделей машинного обучения с динамическим выбором на основе состояния внешней среды (соответствует п. 2, 9 паспорта специальности).

4. Методика построения виртуальной среды, обеспечивающей эмуляцию данных для отладки цифрового двойника (научного прототипирования) (соответствует п. 9 паспорта специальности).

5. Агентно-сервисная архитектура умного цифрового двойника энергетического объекта, предусматривающая интеграцию вычислительного ядра, цифровой тени и виртуальной среды для решения задач проектирования и прогнозирования работы солнечной электростанции (соответствует п. 3, 9 паспорта специальности).

Соответствие диссертации паспорту научной специальности.

Содержание диссертационной работы соответствует паспорту научной специальности 1.2.2. Математическое моделирование, численные методы и комплексы программ (технические науки) по следующим пунктам:

- П.2. Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий. (положения 1, 2, 3 новизны);
- П.3. Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно – ориентированных программ для проведения вычислительного эксперимента. (положение 5 новизны);
- П.9. Постановка и проведение численных экспериментов, статистический анализ их результатов, в том числе с применением современных компьютерных технологий (технические науки). (положения 3, 4, 5 новизны).

В диссертации присутствуют оригинальные результаты из трёх областей:

1. *Математическое моделирование.* Агрегированная математическая модель СЭС и формализованная модель умного цифрового двойника на основе модификации базовой модели цифрового двойника, отличающаяся добавлением описаний моделей искусственного интеллекта, базы знаний и виртуальной среды.
2. *Численные методы.* Новый оригинальный метод ансамблирования моделей машинного обучения с динамическим выбором, отличающийся использованием онтологий и учетом состояния внешней среды.
3. *Комплексы программ.* Программный комплекс, являющийся научным прототипом ЦД СЭС и интегрирующий цифровую модель, цифровую тень СЭС и модели машинного обучения.

Теоретическая значимость работы заключается в разработке методического подхода к построению умных цифровых двойников энергетических объектов, основанного на интеграции математических моделей, методов машинного обучения и численного метода ансамблирования моделей с динамическим выбором состояния внешней среды в виртуальной среде, и возможностью его тиражирования.

Практическая значимость работы определяется возможностью применения предложенных методов и алгоритмов для разработки ЦД солнечных электростанций (СЭС), а также программной реализацией научного прототипа умного цифрового двойника СЭС для проектирования ЦД, прогнозирования и расчета выходных значений СЭС.

Получены шесть свидетельств о государственной регистрации программ для ЭВМ и баз данных. Результаты диссертационной работы отмечены именной стипендией Губернатора Иркутской области (2023 г.). В 2024 г. автор в составе молодежного коллектива стал лауреатом Иркутского областного конкурса в сфере науки и техники, в номинации «За лучшие научные, научно – технические и инновационные разработки молодых ученых».

Апробация работы. Результаты работы, а также результаты отдельных исследований и разработок автора докладывались и обсуждались на российских и международных конференциях: VIII Международной конференции «ЗНАНИЯ – ОНТОЛОГИИ – ТЕОРИИ» (2021 г., Новосибирск, Россия); International Conference “Intelligent Systems” (INTELS) (2022 г., Москва, Россия); VIII Международной конференции «Физико – техническая информатика» (2022 – 2024 гг., Нижний Новгород, Россия); Международной конференции "Интеллектуальная индустрия" (SmartIndustryCon) (2024); Байкальской Всероссийской конференции с международным участием «Информационные и математические технологии в науке и управлении» (2021-2025 гг., Иркутск, Байкал, Россия); Международном семинаре «Критические инфраструктуры в цифровом мире» (IWCI), (2021 – 2025 гг., Иркутск, Байкал, Россия); Конференции молодых ученых ИСЭМ СО РАН «Системные исследования в энергетике» (2021-2025 гг., Иркутск, Россия);

семинарах отдела Систем искусственного интеллекта в энергетике ИСЭМ СО РАН (2025 г.).

Результаты исследования использовались при выполнении проектов:

1. По гранту РФФИ № 20-07-00994 «Методы и модели поддержки коллективных экспертных решений по стратегическому инновационному развитию энергетической инфраструктуры РФ с применением технологии Big Data», 2020-2022 гг.

2. По гранту РНФ № 23-21-00382, «Разработка и интеграция методов и технологий построения цифровых двойников энергетических объектов (на примере возобновляемых источников энергии)» 2023-2024 гг.

3. По государственному заданию Министерства науки и высшего образования Российской Федерации (тема FWEU-2021-0007 «Методология построения ИТ-инфраструктуры для разработки интеллектуальных систем управления развитием и функционированием систем энергетики», 2021-2025 гг.).

Публикации. По теме исследований опубликованы 15 работ, в том числе 3 – в рецензируемых научных изданиях, рекомендованных ВАК РФ по специальности 1.2.2. Математическое моделирование, численные методы и комплексы программ (К2), 3 публикации, индексируемые в Scopus и/или Web of Science, 3 – в других изданиях, получены 6 свидетельств о государственной регистрации программ для ЭВМ и баз данных.

Объем и структура диссертационной работы. Диссертационная работа состоит из введения, трех глав, заключения, списка сокращений и условных обозначений, списка литературы из 116 наименований, приложения. Общий объем работы составляет 150 страниц.

Личный вклад. Личный вклад автора состоит в разработке методического подхода, включающего методику построения умного цифрового двойника солнечной электростанции и виртуальной среды. Результаты, составляющие новизну и выносимые на защиту, получены лично автором, в частности, предложены модель умного цифрового двойника, модель солнечной электростанции для интеграции в цифровой двойник, метод ансамблирования

моделей машинного обучения. Автором разработаны архитектура и программные компоненты умного цифрового двойника, а также проведен вычислительный эксперимент для определения погодных характеристик для ЦД.

Часть работы выполнена автором совместно с коллегами: постановка общей задачи исследования и обсуждение методического подхода – совместно с научным руководителем Л. В. Массель, а типовая методика и типовая архитектура ЦД ВИЭ разработаны с участием автора в неделимом соавторстве. Основные положения диссертации, составляющие научную новизну и выносимые на защиту, получены автором лично.

Глава 1. Анализ проблемы и постановка задач

В первой главе рассматривается солнечная электроэнергетика как предметная область исследования, концепция цифровых двойников и их применение в энергетике, а также методы искусственного интеллекта и машинного обучения, используемые для анализа данных и прогнозирования характеристик энергетических объектов. Это позволяет перейти от описания особенностей СЭС как объекта моделирования к постановке задачи разработки методического подхода к построению умного цифрового двойника, объединяющего математические модели, цифровую тень, виртуальную среду и интеллектуальные компоненты.

1.1. Солнечная электроэнергетика

Возобновляемая энергия – это энергия, получаемая из природных источников, которые пополняются со скоростью, превышающей скорость ее потребления. Примерами таких постоянно пополняемых источников являются солнечный свет и ветер [4].

В современном мире существует тенденция развития возобновляемой энергетики. Это обусловлено рядом причин [5]:

1. Стремление к повышению уровня энергетической безопасности и снижение зависимости от других стран. В отличие от природных углеводородных ископаемых, которые расположены во владении ограниченного количества стран, ВИЭ распределены равномерно и доступны в любой географической точке. Политика стран, направленная на развитие ВИЭ, понижает зависимость от стран, которые имеют высокий показатель топливно-энергетических ресурсов. Примером зависимости может послужить энергетический кризис в 1973 г., когда страны ОАПЕК снизили объем добычи нефти на 5%, что повлекло повышение цен на 70%.

2. Негативное влияние использования топливно-энергетических ресурсов на климат и экологию. Влияние на экологию «традиционных» топливных

технологий энергетики существенно выше, чем влияние технологий ВИЭ. Все виды ВИЭ характеризуются существенно более низким выделением парниковых газов, в отличие от тепловых электростанций [5].

3. Неисчерпаемость и неограниченность возобновляемых источников энергии. В отличие от традиционных топливно-энергетических ресурсов, таких, как уголь, нефть и газ, которые являются исчерпаемыми и требуют миллионы лет для своего формирования, ВИЭ обладают способностью восполняться в естественных условиях за относительно короткий промежуток времени. Это делает их устойчивой и долгосрочной альтернативой для удовлетворения растущих энергетических потребностей человечества.

4. Экономическая обоснованность ввода в эксплуатацию генерирующих объектов возобновляемых источников энергии. Согласно данным Международного агентства по возобновляемой энергетике (IRENA), в период с 2010 по 2021 год наблюдалось значительное снижение нормированной стоимости единицы энергии (LCOE), производимой с использованием возобновляемых источников. Это снижение обусловлено технологическими усовершенствованиями, оптимизацией цепочек поставок и накопленным опытом производителей оборудования. Одновременно рост цен на углеводородные энергоресурсы дополнительно повышает экономическую привлекательность и конкурентоспособность технологий ВИЭ на глобальном энергетическом рынке.

Одним из видов ВИЭ является солнечная энергетика. Солнечная энергетика – направление альтернативной энергетики, основанное на преобразовании солнечного излучения в энергию. Солнечное излучение преобразуется в электрическую энергию постоянного тока за счет фотоэлементов, которые производятся из кремниевых полупроводниковых фотодиодов. Отсюда вытекают недостатки:

- Высокая стоимость производства фотоэлементов.
- Применение инверторов для преобразования электроэнергии постоянного тока в электроэнергию переменного тока.

- Сильная зависимость от погодных условий и времени суток, и, как следствие: необходимость установки аккумуляторных батарей и дополнительных генераторов в качестве резервных источников.

Эти недостатки приводят к тому, что стоимость электроэнергии, которая выработана при помощи СЭС, превышает стоимость энергии, вырабатываемой «традиционными» источниками энергии.

Одним из ключевых и часто недооцениваемых преимуществ солнечных электростанций (СЭС) является их автономность. В России это качество приобретает особую значимость для энергоснабжения технологически изолированных и труднодоступных территорий. К технологически изолированным и труднодоступным территориям (ТИТТ) относятся районы, не подключенные к единой энергосистеме России, где электроснабжение обеспечивается локальными автономными источниками, а тарифы регулируются региональными властями на основе фактической себестоимости генерации. Основная часть ТИТТ России расположена в районах Крайнего Севера и приравненных к нему территориях – некоторых районах Мурманской, Архангельской, Магаданской, Амурской и Сахалинской областей, Республик Саха (Якутия) и Карелия, Чукотского, Ханты–Мансийского и Ямало–Ненецкого автономных округов [5]. По данным на 2021 год, 29 из 43 энергообъектов в населённых пунктах ТИТТ уже используют солнечные электростанции или гибридные солнечно-дизельные электростанции (СДЭС). Более подробная информация о действующих объектах ВИЭ в ТИТТ по состоянию на конец 2021 года представлена в Таблице 1.1, источники данных в таблице – публикация [5].

Таблица 1.1 – Примеры СЭС и СДЭС, действующих в технологически изолированных и труднодоступных территориях России на конец 2021 года

№	Населённый пункт	Субъект РФ	Тип генерирующего объекта	Установленная мощность, кВт	Год ввода в эксплуатацию
1	Токко	Республика Саха (Якутия)	СЭС	2	2018
2	Юрэн		СЭС	5	2018
3	Столбы		СЭС	10	2015

Продолжение Таблицы 1.1

№	Населённый пункт	Субъект РФ	Тип генерирующего объекта	Установленная мощность, кВт	Год ввода в эксплуатацию
4	Джарглах	Республика Саха (Якутия)	СЭС	15	2014
5	Куберганя		СЭС	20	2014
6	Иннях		СЭС	20	2016
7	Куду – Кюель		СЭС	20	2013
8	Улуу		СЭС	20	2013
9	Тойон – Ары		СЭС	20	2014
10	Дулгалах		СЭС	20	2013
11	Ючюгей		СДЭС	30	2012
12	Верхняя Амга		СЭС	36	2016
13	Юнкор		СЭС	40	2015
14	Бетенкес		СЭС	40	2015
15	Кыстатыам		СЭС	40	2018
16	Эйик		СЭС	40	2014
17	Орто – Балаган		СЭС	50	2017
18	Себян – Кюель		СЭС	50	2017
19	Батамай		СЭС	60	2012
20	Дельгей		СЭС	80	2016
21	Батагай		СДЭС	1000	2015
22	Табалах		СЭС	350	2021
23	Долиновка	Камчатский край	СДЭС	20	2019
24	Ермаки	Иркутская область	СДЭС	42,9	2020
27	Кызыл – Хая	Республика Тыва	СДЭС	150	2019
28	Мугур – Аксы	Забайкальский край	СДЭС	400	2019
29	Менза		СДЭС	120	2017

Таблица 1.1 отображает, что СЭС и СДЭС применяются в технологически изолированных и труднодоступных территориях России и являются инструментом автономного энергоснабжения удаленных населенных пунктов. Использование СЭС и СДЭС в северных, дальневосточных и других изолированных территорий обусловлено удалённостью потребителей, высокой стоимости доставки топлива и ограниченной связности с централизованной энергосистемой. Таблица показывает, что СЭС и СДЭС уже применяются в реальных автономных энергосистемах, а значит, для таких объектов актуальны задачи проектирования, оценки ожидаемой выработки, прогнозирования, диагностики отклонений и поддержки принятия решений. В Иркутской области рассматриваемая проблематика является актуальной. Это подтверждает вводом в эксплуатацию несколько СДЭС в удалённых населённых пунктах, в том числе в Нерхе, Ермаках, Верхней Гутаре и Алыгджере².

Современные фотоэлектрические системы используют преобразование солнечной энергии в электрическую благодаря различным видам фотоэлементов. Эти элементы изготовлены из различных типов полупроводниковых материалов с различной эффективностью. Национальная лаборатория по изучению возобновляемой энергетики (англ. National Renewable Energy Laboratory (NREL)) приводит эффективность таких популярных технологий производства как [6]:

- Монокристаллические кремниевые панели обладают КПД до 27%.
- Поликристаллические панели имеют КПД до 23%.
- Тонкопленочные технологии имеют КПД до 23%.

В проведенном исследовании показана зависимость рекордной подтверждённой эффективности различных технологий солнечных элементов от времени. Из него следует, что развитие фотоэлектрических технологий

² СЭС «Нерха» // ООО «Байкальская энергетическая компания». URL: <https://www.ogirk.ru/2018/01/09/v-tofalarii-nachala-rabotu-solnechnaya-elektrostantsiya-2/>; В селе Ермаки Казачинско-Ленского района запущена комбинированная солнечно-дизельная электростанция // Официальный портал Иркутской области. URL: <https://irkobl.ru/sites/gkh/news/1060746/>; В Тофаларии запустили солнечно-дизельную электростанцию. URL: <https://www.ogirk.ru/2025/10/19/v-tofalarii-zapustili-solnechno-dizelnuju-jelektrostantsiju/>. Дата обращения: 10.04.2025.

сопровождается постепенным повышением эффективности преобразования энергии, однако значения эффективности существенно зависят от типа используемой технологии. Высокие значения демонстрируют отдельные перспективные технологии, однако для практического применения в составе солнечных электростанций наибольший интерес представляют промышленно распространённые кремниевые фотоэлементы, для которых также наблюдается рост эффективности. Помимо рассматриваемых технологий, на графике показаны и другие подходы, и материалы для построения фотоэлектрических панелей, которые пока не получили массового промышленного применения.

Одной из важных проблем являются эксплуатационные ограничения. Для кристаллического кремния, включая монокристаллические и поликристаллические модули, характерны высокая технологическая зрелость и высокий КПД, однако для них существенны деградационные эффекты, связанные со световым воздействием, температурой, влажностью и потенциально-индуцированной деградацией, изучению по которой проводились в Национальной лаборатории возобновляемой энергии (National Renewable Energy Laboratory, NREL) [7]. В обзоре [8] помимо перечисленных эффектов также дополнительно учитываются загрязнения и точечный нагрев солнечных батарей.

При характеристике солнечной электростанции как объекта моделирования необходимо учитывать, что выходные параметры СЭС определяются не только типом фотоэлектрических модулей, но и условиями их размещения. К таким условиям относятся площадь фотоэлектрической поверхности, ориентация панели относительно сторон света, угол наклона, угол падения солнечных лучей, уровень солнечной радиации, температура окружающей среды, ветровое воздействие, загрязнение поверхности, снежный покров и возможное частичное затенение. Солнечные батареи могут использоваться в неподвижном исполнении, когда угол наклона и ориентация панели задаются при проектировании и остаются постоянными, а также в подвижном исполнении, когда положение панели изменяется во времени с целью увеличения поступающей солнечной радиации. В рамках настоящей работы основное внимание уделяется агрегированному

математическому описанию СЭС, достаточному для построения цифрового двойника и проведения вычислительных экспериментов. Поэтому геометрические особенности солнечных батарей учитываются через площадь панели, угол наклона и ориентацию поверхности, а не через детальное моделирование формы отдельных фотоэлектрических элементов.

В Иркутской области влияние резко континентального климата на солнечные батареи имеет большое значение из-за факторов:

- Влияния температуры (перепад и низкие зимние значения).
- Ветровые и механические нагрузки.
- Длительный снежный покров.

Большая протяженность территории с юга на север влияет на неравномерное поступление годовой суммарной радиации. В северных районах она составляет 80 ккал в см^2 , а в южных – 110 ккал в см^2 , то есть распределяется по поверхности Среднесибирского плоскогорья в зависимости от географической широты [9].

Математическое моделирование солнечных электростанций и фотоэлектрических преобразователей может выполняться на различных уровнях детализации. Наиболее детальный уровень связан с описанием электрических характеристик отдельных фотоэлектрических элементов и модулей. В рамках такого подхода применяются однодиодные и двухдиодные модели фотоэлектрического преобразователя, модели вольт-амперных характеристик, а также модели, использующие W-функцию Ламберта для аналитического или численного описания зависимости тока и напряжения. Эти модели позволяют исследовать электрические характеристики фотоэлектрических преобразователей, влияние внутренних сопротивлений, фототока, обратного тока насыщения, рекомбинационных процессов и режимов работы отдельного модуля [10, 11].

Другой уровень моделирования связан с учетом геометрических условий освещения фотоэлектрических модулей. В таких моделях рассматриваются площадь освещенной поверхности, угол падения солнечного излучения, ориентация панели, возможное затенение и неоднородность распределения солнечной радиации по поверхности модуля. Такие подходы особенно важны при

исследовании фотоэлектрических систем сложной конфигурации, подвижных конструкций, систем слежения за Солнцем, а также солнечных батарей космических аппаратов, для которых требуется учитывать изменение взаимного положения источника излучения, объекта и затеняющих элементов конструкции [12].

Отдельное направление составляют модели контроля и диагностики фотоэлектрических модулей в условиях неоднородного освещения. В таких моделях анализируются локальные затемнения, неравномерность освещенности, работа отдельных элементов в области обратного смещения, локальный перегрев и деградационные процессы. Они позволяют выявлять дефекты, оценивать техническое состояние солнечных модулей и исследовать причины снижения эффективности генерации.

На следующем уровне рассматриваются модели фотоэлектрических панелей, массивов и электротехнических комплексов с фотоэлектрическими модулями. Такие модели позволяют исследовать работу солнечной электростанции с учетом климатических условий, схемы соединения модулей, преобразовательного оборудования, нагрузки и режимов эксплуатации. В русскоязычных исследованиях рассматриваются подходы к моделированию фотоэлектрических панелей, как элемента комплекса распределенной генерации, а также моделирование электротехнических комплексов с фотоэлектрическими модулями в конкретных климатических условиях [13, 14]

Широкое распространение получили специализированные модели и программные средства для оценки производительности фотоэлектрических систем. Модель Sandia PV Array Performance Model учитывает электрические, тепловые и оптические характеристики фотоэлектрических модулей и рассчитана на использование часовых данных солнечного ресурса и метеорологических данных. Она применяется для проектирования и выбора параметров фотоэлектрических массивов, оптимизации режимов работы, а также сопоставления измеренной и ожидаемой производительности системы [15] Модель PVWatts, разработанная National Renewable Energy Laboratory, ориентирована на оценку выработки сетевой

фотоэлектрической системы по ограниченному набору исходных данных и объединяет ряд подмоделей для получения итоговой оценки производительности [16].

Таким образом, существующие математические модели солнечных электростанций различаются по назначению, уровню детализации и требованиям к исходным данным. Детальные электрофизические модели целесообразны при исследовании процессов в фотоэлектрических элементах и модулях, тогда как модели уровня массива, станции или энергетического комплекса применяются для оценки выработки, анализа режимов работы, проектирования и эксплуатационного сопровождения. Для задач цифрового двойника особенно важно обеспечить компромисс между физической обоснованностью модели, доступностью исходных данных, вычислительной сложностью и возможностью последующего уточнения модели при появлении дополнительных данных.

Солнечная энергетика обладает значительным потенциалом для энергоснабжения удалённых и технологически изолированных территорий, однако её эффективность существенно зависит от климатических условий, технических характеристик оборудования, качества исходных данных и точности расчёта режимов работы. Для солнечных электростанций особенно важны задачи предварительного проектирования, оценки ожидаемой выработки, анализа влияния внешней среды и последующего сопровождения объекта в эксплуатации. Эти задачи требуют применения инструментов, способных объединять математические модели, данные о внешней среде и эксплуатации, а также прогнозные алгоритмы. В качестве такого инструмента далее рассматривается технология цифровых двойников, позволяющая связать физический объект с его цифровым представлением и использовать его как на этапе проектирования, так и на этапе эксплуатации.

1.2. Цифровые двойники

Впервые концепция «цифровой двойник» описана в 2002 году Майклом Гривзом [17]. Согласно ей, ЦД состоит из трех основных частей: 1) физические компоненты в реальном пространстве, 2) виртуальные компоненты в виртуальном пространстве, 3) соединение данных и информации, которые связывают виртуальные и реальные компоненты вместе. Работа системы организована следующим образом: встроенные в физический объект датчики непрерывно собирают и передают актуальные данные цифровому двойнику в реальном времени. Цифровая модель постоянно обновляется, учитывая текущие изменения состояния объекта и накапливая исторические данные. Такой подход обеспечивает высокую точность моделирования текущего поведения объекта и достоверность прогнозных расчетов.

На основе этой актуализированной модели появляется возможность оптимизировать эксплуатационные режимы объекта. Поступающие данные позволяют цифровому двойнику не только отражать текущее состояние, но и постепенно совершенствовать точность модели, снижая риски и повышая эффективность управления.

Цифровые двойники делят на три типа [18]:

- Цифровой двойник-прототип (Digital Twin Prototype, DTP) – содержит описание физического объекта и информацию, которая необходима для описания и создания физического объекта.
- Цифровой двойник-экземпляр (Digital Twin Instance, DTI) – содержит описание характеристик и эксплуатационные данные объекта, с которым ЦД связан на протяжении всего периода.
- Агрегированный цифровой двойник (Digital Twin Aggregate, DTA) – вычислительная система, которая имеет цифровые двойники – экземпляры и способна посылать запросы с целью мониторинга состояния объекта.

Цифровая тень – это набор данных, который получен от датчиков, и модель, которая позволяет прогнозировать свойства объекта в определенных пределах [19].

В статье [20] описывается отличие цифровой тени от цифрового двойника, состоящее в том, что цифровой двойник включает описание физических процессов на основе математических моделей. Формирование ЦД на основе только цифровой тени делает цифровой двойник «неполным». Такой цифровой двойник не способен адекватно работать в ситуациях, в которых объект не эксплуатировался.

Первое появление цифрового двойника было отмечено в программе «Apollo» NASA [21]. Он определен как: “Цифровой двойник – это интегрированное мультифизическое, многомасштабное, вероятностное моделирование транспортного средства или системы, которое использует наилучшие доступные физические модели, обновления датчиков, историю парка и т.д., чтобы отразить жизнь своего летающего двойника. Цифровой двойник является ультра – реалистичным и может учитывать одну или несколько важных и взаимосвязанных систем транспортного средства, включая силовую установку/накопитель энергии, авионику, жизнеобеспечение, конструкцию транспортного средства, терморегулирование/TPS и т.д. В дополнение к высокоточным физическим моделям цифровой двойник интегрирует данные датчиков из бортовой интегрированной системы управления состоянием автомобиля (IVHM), историю технического обслуживания и все доступные исторические данные о парке, полученные с помощью интеллектуального анализа данных и текста. Системы, установленные на борту «цифрового двойника», также способны предотвращать повреждения или деградацию, рекомендуя изменения в профиле миссии для увеличения срока службы и вероятности успеха миссии.”

В последующих исследованиях концепция цифрового двойника получила развитие как методологическая основа интеграции математического моделирования, потоков данных и интеллектуальной аналитики. В обзорных работах [22–26] цифровой двойник рассматривается как динамическая система, обеспечивающая согласованное функционирование физического объекта и его цифрового представления на протяжении всего жизненного цикла. Подчеркивается, что переход от статических цифровых моделей к цифровым

двойникам связан с необходимостью учета нестационарности процессов, многоуровневости описаний и интеграции методов искусственного интеллекта.

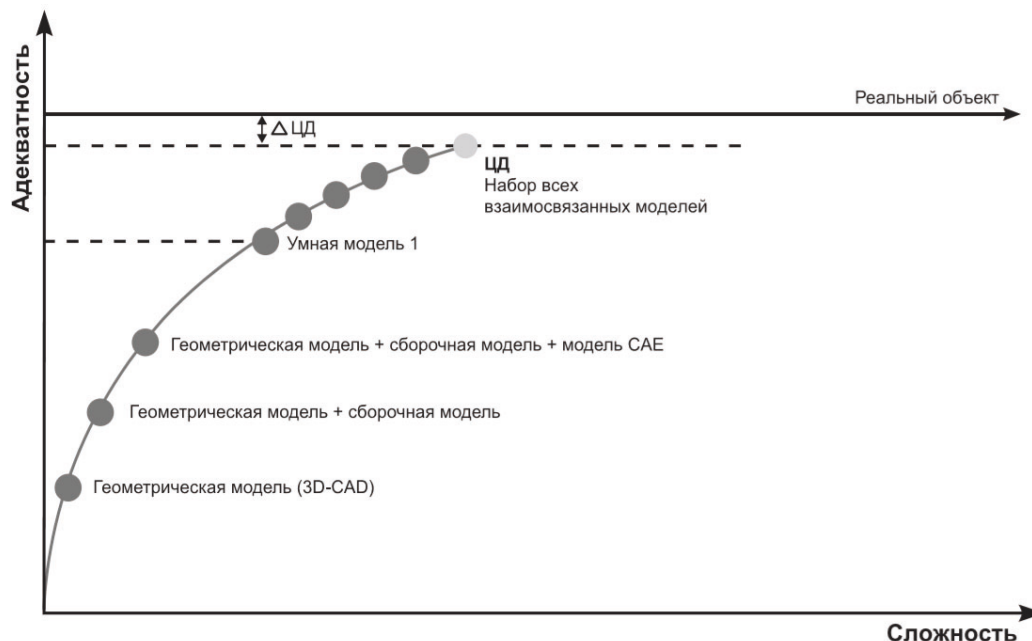


Рисунок 1.1 – Этапы формирования цифровой модели, лежащие в основе ЦД

Технология «цифровой двойник» стала одной из десяти основных тенденций по мнению Gartner в 2017 году [2]. В прогнозе описывается, что цифровые двойники представлены в виде динамической программной модели физического объекта или системы. Используя физические данные работы компонента объекта и реакции на окружающую среду, ЦД можно использовать для анализа и моделирования условий реального мира, реакции на изменения и повышения качества работы. ЦД может функционировать в качестве «посредника» между специалистами и устройствами мониторинга и управления.

В России термин «Цифровой двойник» получил развитие в работах профессора Боровкова А.И. [27–29]. На Рисунке 1.1 представлены этапы развития модели цифрового двойника, которая проходит несколько ключевых этапов. Первоначально создается геометрическая модель объекта, которая затем дополняется сборочной моделью. Следующим этапом становится интеграция

модели инженерного анализа, что в конечном итоге приводит к формированию "умных моделей" – математических моделей с высокой степенью соответствия реальному объекту. Согласно определению А.И. Боровкова, "умные" модели представляют собой комплексное объединение всех знаний, применяемых на этапах проектирования, производства и эксплуатации изделия или технической системы. Помимо обязательных знаний, закладываемых при создании модели, такие модели требуют наличия у разработчиков специальных компетенций. Этими компетенциями являются: кастомизация (способность оперативно откликаться на запросы заказчика), системный инжиниринг (способность видеть всю систему и ее компоненты), многоуровневая матрица целевых показателей и ресурсных ограничений (содержит целевые показатели и требования к продукту), валидация «умных» моделей (необходимость проведения натурных испытаний), «цифровая сертификация» (бизнес – процесс, направленный на прохождение с первого раза всего комплекса натурных, сертификационных, рейтинговых и прочих испытаний) [30].

Профессор А.И. Боровков предлагает классификацию цифровых двойников, выделяя два основных типа: цифровой двойник продукта (DT-1) и цифровой двойник производства (DT-2), каждый из которых включает в себя "умную" модель. Объединение этих двух типов в рамках единой цифровой платформы, осуществляемый через проведение десятков тысяч виртуальных испытаний в процессе специально организованной "цифровой сертификации", приводит к созданию "умного" цифрового двойника первого уровня (SDT-1).

В процессе эксплуатации физического объекта система датчиков непрерывно собирает данные, формируя так называемую цифровую тень. Анализ этих данных позволяет не только точно описывать текущее состояние объекта, но и строить прогнозные модели, например, для оценки остаточного ресурса оборудования. На этом этапе "умный" цифровой двойник на основе своей модели генерирует "умную" цифровую тень (Smart digital shadow, SDS), которая обогащается оперативными данными о работе конкретного объекта через технологии промышленного Интернета вещей и системы диагностики.

Поступившая информация на SDT-1, продолжает его «обучать», что повышает уровень адекватности. Высокий уровень адекватности позволяет моделировать процессы вне эксплуатационных рамок. Большое количество проведенных попыток моделирования позволяет иметь четкое представление о критических зонах, что упрощает расположение датчиков и формирование временного промежутка для отправки данных. Вместо получения большого объема данных, SDT-1 получает необходимые данные в процессе жизненного цикла. В этом случае повышается скорость обработки данных и происходит трансформация в цифровой двойник второго уровня (SDT-2) (Рис. 1.2).

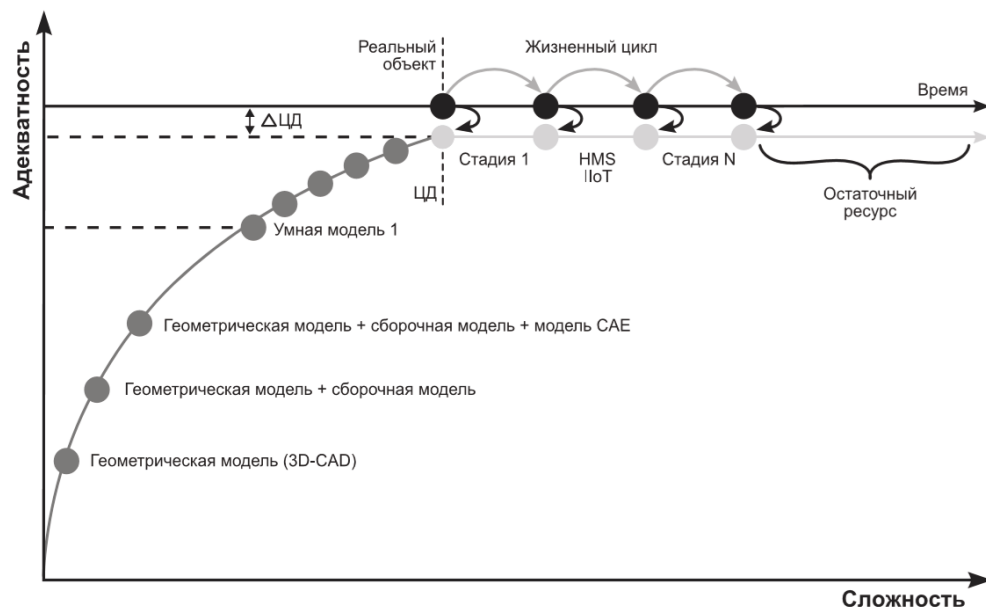


Рисунок 1.2 – Семейство физических и математических моделей

Цифровая модель – это программная реализация математической модели, которая представляет собой формализованное описание объекта, системы или процесса с помощью математических уравнений, алгоритмов и логических правил [27].

Проектирование цифрового двойника, с формированием цифровой модели, является не единственным способом создания ЦД. «Альтернативным» методом является создание аналитической модели, которая основана на данных,

полученных от объекта. Для формирования аналитической модели необходимы датчики, IoT – устройства (устройства Интернета вещей) и системы для сбора и хранения Больших данных.

Отсюда вытекает проблема в потенциальной обработке большого количества совершенно ненужных и «зашумленных» данных, которые усложняют процесс обработки, как с точки зрения затратности (вычислительные ресурсы, хранение), так и с точки зрения точности вычисления (такие данные сложно экстраполировать). Кроме этого, установка и обслуживание датчиков тоже требуют затрат. При использовании такой технологии сложно на ранних этапах определить «нужные» и «ненужные» датчики. В Таблице 1.2 отображены достоинства и сложности этих процессов, которые являются основой цифрового двойника, таблица составлена на основании публикации [31]

Таблица 1.2 – Достоинства и требования/сложности для реализации аналитических и математических моделей физико – механических процессов

	Достоинства	Требования / сложности
Основа – математическое моделирование физических процессов	• Модели строятся на физических законах, определяющих поведение объекта в широких границах • На основе моделирования физических процессов удастся выявить глубинные причинно – следственные связи и важные закономерности в поведении конкретной системы •	• Требуются глубокие знания в области физических процессов, описывающих поведение изучаемого объекта • Даже при больших вычислительных ресурсах требуется существенное время на расчет, что затрудняет использование модели в схемах, где необходима обратная связь в режиме реального времени

Продолжение Таблицы 1.2

	Достоинства	Требования / сложности
Основа – математическое моделирование физических процессов	• Неопределенность контролируется входными данными и точностью моделирования • Модель позволяет делать прогнозы в широких пределах, покрываемых моделью	• До начала моделирования должны быть известны все зависимости между входными и выходными параметрами
Основа – аналитические подходы/машинное обучение	• Модель может улучшаться со временем (обучение с подкреплением) • Выявление взаимосвязей между данными • Модель создается только из данных, не требуя от исследователя наличия глубоких знаний предметной области • Обобщенный гибкий подход – обработка неоднородных потоков данных	• Для создания модели необходимо наличие качественных обучающих данных • Основа – это корреляции, а не причинно – следственные связи •

Аналитиками компании ANSYS приводятся понятия Analytics based DT и Simulation based DT. Эти два определения описывают ситуации выше: Цифровой двойник, основанный на симуляции физических процессов (Simulation based DT), сводится к созданию математической модели и решению ее в частных случаях; Цифровой двойник, основанный на анализе данных (Analytics based DT), сводится к получению данных через технологию Интернета Вещей, последующую обработку с помощью аналитической модели (в том числе методов машинного

обучения) и прогнозирование на основе полученной модели. Интеграция этих двух методов создает гибридный ЦД, который работает по схеме: математическое моделирование + датчики → сбор данных с датчиков → IoT платформа → аналитика [19].

Помимо определений, представленных выше, в Российской Федерации предложен первый в мире стандарт для цифровых двойников: в 2021 г. принят ГОСТ Р 57700.37–2021. “Компьютерные модели и моделирование. Цифровые двойники изделий”. Несмотря на то, что он ориентирован преимущественно на изделия машиностроения, основными его положениями можно руководствоваться при построении ЦД в энергетике. Приведем основные определения из раздела «Общие положения»:

- **Компьютерная модель (электронная модель):** Модель, выполненная в компьютерной (вычислительной) среде и представляющая собой совокупность данных и программного кода, необходимого для работы с данными. *Примечание.* В основе компьютерной модели лежит математическая модель, реализованная в виде программного кода, и данные, определяющие конкретный объект моделирования. Для применения компьютерной модели в процессе моделирования необходимо использовать программное обеспечение компьютерного моделирования и вычислительной техники.

- **Цифровая модель изделия:** Система математических и компьютерных моделей, а также электронных документов изделия, описывающая структуру, функциональность и поведение вновь разрабатываемого или эксплуатируемого изделия на различных стадиях жизненного цикла, для которой на основании результатов цифровых и (или) иных испытаний по ГОСТ 16504 выполнена оценка соответствия предъявляемым к изделию требованиям.

- **Цифровой двойник изделия; ЦД:** Система, состоящая из цифровой модели изделия и двусторонних информационных связей с изделием (при наличии изделия) и (или) его составными частями.

Важным аспектом развития технологии цифровых двойников является их рассмотрение в контексте жизненного цикла технического объекта. Цифровой

двойник не является неизменной моделью: его состав, структура данных и функциональные возможности зависят от стадии жизненного цикла объекта и решаемых на этой стадии задач. На этапе проектирования он может использоваться для анализа требований, проверки проектных решений и проведения вычислительных экспериментов; на этапе реализации – для интеграции моделей, данных и программных компонентов; на этапе эксплуатации – для мониторинга состояния объекта, прогнозирования характеристик и поддержки принятия решений. Обобщённая схема применения цифрового двойника на стадиях жизненного цикла технического объекта представлена на Рисунке 1.3 [32].

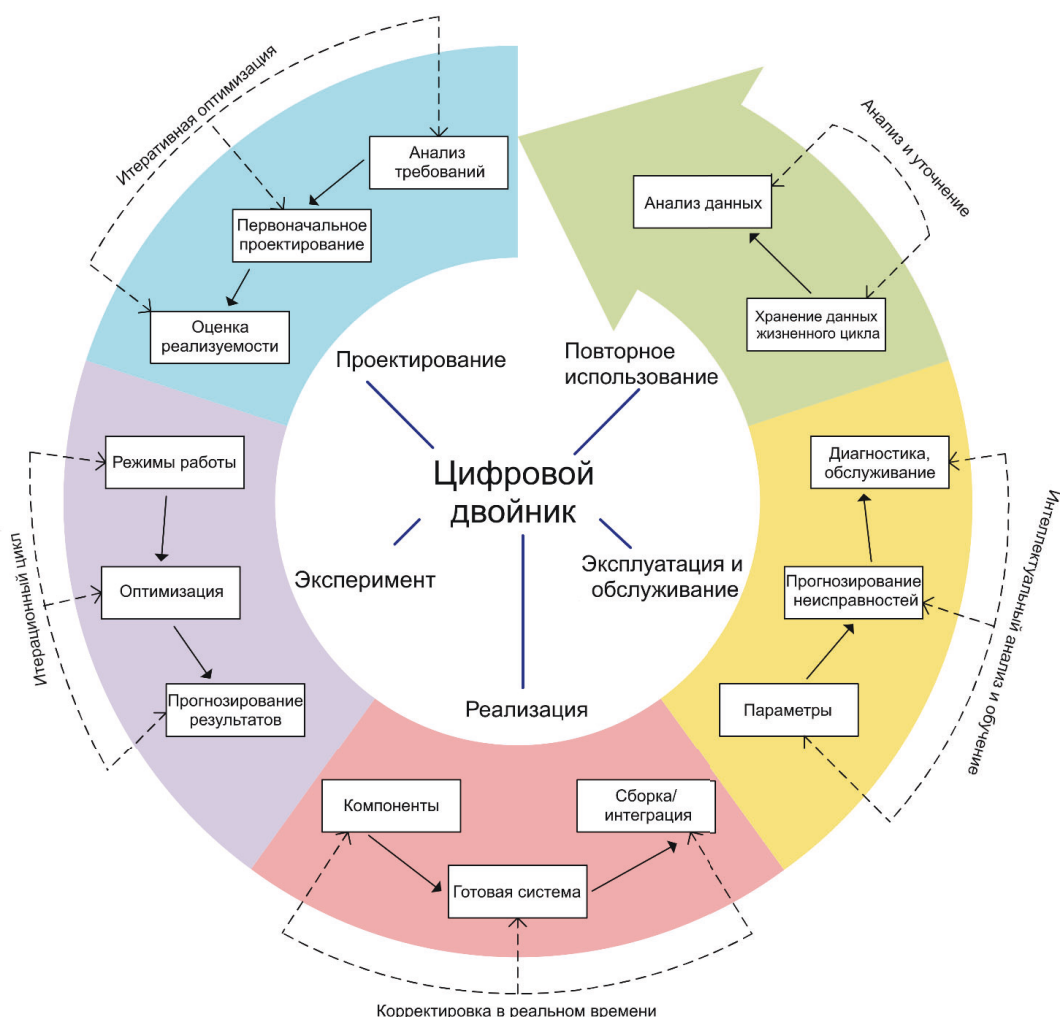


Рисунок 1.3 – Схема применения цифрового двойника на стадиях жизненного цикла энергетического объекта

В работе [33] рассматривается метод синтеза цифровых двойников производственных процессов на основе цифровых идентификационных моделей. Идентификационные модели представляют собой построенные и корректируемые модели объектов/процессов в реальном времени на основе данных функционирования. Сами алгоритмы идентификации выполняют корректирование в каждый момент времени при формировании новой цифровой модели. В построении модели, помимо фактических данных функционирования, интегрируются формализованные знания, полученные на момент формирования модели. Хотя модели могут создаваться различными способами, но в каждой из них предусмотрены получение и интеграция данных из различных источников, анализ полученных данных, формирование и структурирование знаний с последующей формализацией.

Цифровые двойники представляют собой инструмент для моделирования и управления сложными системами, включая энергетические объекты. Их применение в солнечной энергетике, как будет показано далее, позволяет решать задачи проектирования, мониторинга, прогнозирования и повышения эффективности.

Рассмотренные определения и классификации показывают, что цифровой двойник не сводится только к компьютерной модели или набору данных, полученных от датчиков. Его ключевая особенность состоит в объединении цифровой модели, информационных связей с объектом, вычислительных процедур и механизмов обновления состояния модели. Поэтому при переходе к энергетическим объектам цифровой двойник необходимо рассматривать не только как средство визуализации или мониторинга, но и как комплексный инструмент расчёта, прогнозирования и анализа режимов функционирования. В следующем разделе рассматриваются существующие направления применения цифровых двойников в энергетике, что позволяет определить место разрабатываемого подхода среди известных решений.

1.3. Применение цифровых двойников в энергетике

Технология цифровых двойников в энергетике представляет собой одно из перспективных и динамично развивающихся направлений, как в России, так и во всем мире. Компанией Siemens разработан цифровой двойник устройства Siprotec-5 (Siprotec DigitalTwin). Siprotec-5 является цифровым устройством релейной защиты. Этот цифровой двойник содержит все интерфейсы, функции и алгоритмы устройства. При помощи него возможно проводить испытания, как перед эксплуатацией, так и во время наладки оборудования, что сокращает время до ввода системы в эксплуатацию. Данные можно смоделировать и протестировать отдельно от системы.

Другая крупная компания General Electric разработала цифровой двойник, который послужил основой цифровой электростанции Digital Power Plant. Кроме использования цифровых моделей, позволяющих проводить тестирование, оценки каждого объекта на протяжении жизненного цикла и нахождение аномалий, в системе используются датчики для сбора данных с объекта и аналитика больших данных, которая способствует снижению внеплановых простоев, повышению качества технического обслуживания и баланса производства. Цифровые модели, находящиеся в «Цифровой электростанции», способны эмулировать поведение объекта при варьировании эксплуатационных параметров.

Компания GE Renewable Energy успешно внедрила технологию цифровых двойников для управления ветроэнергетическими установками. На примере ветропарков, построенных в 2015 году в Северной Америке, компания разработала цифровые копии физических объектов с использованием собственной программной платформы Predix. Эта система обеспечивает операторам комплексный мониторинг и управление, позволяя собирать телеметрические данные, визуализировать рабочие параметры и проводить углубленный анализ производительности каждого отдельного агрегата. По данным на 2019 год, GE Renewable Energy развернула три специализированных испытательных центра в

различных географических регионах и оснастила технологией цифровых двойников более 15 000 ветрогенераторов собственного производства [34].

В энергетике цифровые двойники рассматриваются, как инструмент системных исследований, ориентированный на анализ режимов функционирования, прогнозирование и поддержку принятия решений. В ряде работ [35–41] показано, что применение цифровых двойников позволяет объединить физико – математические модели энергетических объектов с аналитическими моделями, построенными на данных, что особенно актуально для объектов возобновляемой энергетики. Отмечается, что для ВИЭ ключевыми задачами являются прогнозирование генерации и учет неопределенности внешней среды.

В Российской Федерации принят проект «Цифровая энергетика». Этот проект направлен на преобразование энергетической инфраструктуры Российской Федерации посредством внедрения цифровых технологий и платформенных решений для повышения ее эффективности и безопасности [42]. Технология ЦД является частью высокотехнологической инфраструктуры.

В 2021 году АО «ВНИИАЭС» начала разработку цифрового двойника атомных электростанций малой мощности (АСММ) с реакторной установкой РИТМ – 200Н и «Шельф – М». Этот цифровой двойник содержит цифровые модели, моделирующие физические процессы в различных эксплуатационных режимах, базы данных и сервисное программное обеспечение. По заявлениям разработчиков, цифровой двойник выполняет определенные функции на различных стадиях жизненного цикла АСММ [43, 44].

В 2023 году Центр компетенций НТИ на базе НИУ «МЭИ» представил программно – аппаратный комплекс «Цифровой двойник энергосистемы». Этот программный аппаратный комплекс может применяться для решения определённых задач [45]:

1. Выполнение расчетов электромагнитных и электромеханических переходных процессов.
2. Симуляция поведения энергосистем в режиме реального времени.

3. Разработка алгоритмов защиты и автоматики энергосистем.
4. Разработка проектных решений по планированию развития энергосистем и микроэнергосистем с ВИЭ и СНЭЭ.
5. Разработка проектных решений по цифровым системам защиты и автоматики.
6. Испытания реальных устройств и систем защиты и автоматики для аттестации, сертификации и определения уровня кибербезопасности.
7. Онлайн – управление режимами работы энергосистем.
8. Адаптивная настройка в режиме онлайн.
9. Научные исследования, обучение студентов и курсы повышения квалификации специалистов.

На данный момент в сфере разработки и применения цифровых двойников в энергетике в России имеются трудности. Основными причинами являются низкий уровень цифровизации объектов и процессов на производственных предприятиях, а также отсутствие в РФ поставщиков базовых программных компонентов, необходимых для построения ЦД. Многие энергетические объекты, такие, как электростанции, сети распределения и подстанции, до сих пор используют устаревшие технологии, что затрудняет сбор данных и их интеграцию в цифровые модели.

Кроме того, правовой аспект внедрения цифровых двойников в энергетике требует дополнительной проработки. Энергетические объекты относятся к категории критически важных инфраструктур, что предъявляет повышенные требования к точности, безопасности и надежности используемых технологий. Отсутствие четких нормативных стандартов и регламентов для создания и использования цифровых двойников замедляет их внедрение.

В работе С.П. Ковалева [46] рассматриваются вопросы проектирования информационного обеспечения цифровых двойников энергетических систем, включая структуру потоков данных, требования к архитектуре и принципы интеграции вычислительных моделей с базами данных. Данные положения имеют

методологическое значение при разработке цифровых двойников энергетических объектов.

Подход к использованию цифровых двойников в задачах проектирования интегрированных энергетических систем получил развитие в работах [47, 48], где показано, что цифровой двойник может выступать не только как средство мониторинга, но и как инструмент автоматизированного расчёта и анализа режимов функционирования системы на этапе проектирования.

В статье “Enhanced Solar Photovoltaic System Management and Integration: The Digital Twin Concept” [49] рассматривается применение цифровых двойников для управления и интеграции солнечных электростанций. В этом обзоре внимание уделяется такому аспекту, как применение на различных этапах жизненного цикла: Разработка и производство солнечных элементов, проектирование и установка фотоэлектрических систем, прогнозирование выработки энергии, а также мониторинг и техническое обслуживание. Выделяются аспекты, которые требуют проработки:

- Интеграция мультифизических и структурных моделей и сложность их создания, которая описывает все процессы, происходящие в системе.
- Проблемы передачи данных в удаленных локациях, которые требуют наличие надежных сетей.
- Точность данных датчиков, которые могут повлиять на прогнозирование.
- Точность моделей машинного обучения и искусственного интеллекта, которые используются для прогнозирования характеристик.
- Уязвимость цифровых систем к кибератакам. Чем больше энергетических установок будет переведено в цифровую форму, тем больше будет создано уязвимых киберузлов. Однако архитектура цифровых двойников может быть спроектирована таким образом, чтобы включать возможности мониторинга кибербезопасности.

- Экономические аспекты, это в значительной степени связано с первоначальным капиталом для интеграции цифрового двойника и дополнительными эксплуатационными расходами на хранение облачных данных и вычислительными услугами.

- Отсутствие единых стандартов для цифровых двойников. Стандартизация способствует функциональной совместимости и взаимоподключаемости отдельных элементов. Кроме того, из-за различных уровней аппроксимации, применяемых при моделировании физического двойника, в настоящее время трудно установить точность элементов из разных цифровых двойников.

В работе [50] рассматривается применение технологии цифрового двойника и методов машинного обучения для решения задач мониторинга и оптимизации работы солнечных электростанций. В ней авторы предлагают систему разделить на подсистемы: фотоэлектрические панели, преобразователи и сеть и строить для каждой отдельную модель цифрового двойника. В качестве моделей машинного обучения для цифрового двойника рассматривались ансамблевый метод, многослойная нейронная сеть и градиентный бустинг на основе деревьев решений. В этой работе [50] используются только аналитические методы для построения цифрового двойника солнечной электростанции и рассматривается задача функционирования и управления объектом.

В статье [51] технология цифрового двойника использовалась для решения проблемы частичного затенения, которая снижает эффективность фотоэлектрических панелей, вызывая неравномерное распределение тока, локальные перегревы и уменьшение срока службы. В ней описывается прогнозирование с использованием исторических данных и пиранометра (прибора для определения плотности потока суммарной солнечной радиации, поступающей из всей верхней полусферы на плоскую поверхность), и на основе прогноза составлен режим работы панелей, что повысило мощность на 30 – 45% (в зависимости сценария затенения) и упростило работы контроллеров заряда.

Цифровые двойники уже находят применение в энергетике, однако их внедрение сталкивается с рядом проблем. Анализ существующих примеров применения цифровых двойников в энергетике показывает, что их эффективность во многом определяется не только качеством математических моделей, но и способностью обрабатывать разнородные данные, выявлять скрытые зависимости и адаптироваться к изменяющимся условиям функционирования объекта. Для солнечных электростанций это особенно существенно, поскольку генерация зависит от нестационарных факторов внешней среды: солнечной радиации, температуры, облачности, загрязнения поверхности модулей, сезонности и суточной изменчивости. Следовательно, построение умного цифрового двойника требует включения методов искусственного интеллекта и машинного обучения, обеспечивающих прогнозирование, классификацию состояний, выявление аномалий и повышение точности расчётных моделей.

1.4. Методы искусственного интеллекта и машинного обучения

Основоположниками направления «искусственный интеллект» считаются: А. Тьюринг, Д. Маккарти, М. Мински, Д. А. Поспелов и др. Общепринятыми определениями искусственного интеллекта являются следующие [52]:

- Научное направление, в рамках которого ставятся и решаются задачи аппаратного или программного моделирования тех видов человеческой деятельности, которые традиционно считаются интеллектуальными.
- Свойство интеллектуальных систем выполнять творческие функции, которые традиционно считаются прерогативой человека.

Машинное обучение является разделом искусственного интеллекта, в котором изучаются методы построения алгоритмов, способных обучаться. Обучение состоит из нахождения и использования необходимых признаков и построении модели для решения поставленной конкретной задачи [53].

В машинном обучении принято выделять два основных подхода: обучение по прецедентам (индуктивное обучение) и дедуктивное обучение.

Обучение по прецедентам (индуктивное обучение) основано на анализе эмпирических данных с целью выявления скрытых закономерностей, зависимостей и паттернов. Этот подход предполагает, что модель обучается на наборе примеров (прецедентов), где каждому объекту соответствует набор признаков и, в случае обучения с учителем, целевая переменная (метка). Основная задача индуктивного обучения – обобщение на основе данных, что позволяет модели делать прогнозы для новых, ранее не встречавшихся объектов. К этому виду относятся задачи классификации, регрессии, кластеризации и другие.

Дедуктивное обучение опирается на формализованные знания экспертов, которые структурируются в виде правил, логических утверждений или онтологий и объединяются в базу знаний. В отличие от индуктивного подхода, дедуктивное обучение не требует большого объема данных, так как использует уже известные закономерности и логические выводы. Этот метод широко применяется в экспертных системах, где решения принимаются на основе заранее определенных правил и логики.

Кроме того, существуют гибридные подходы, которые сочетают в себе элементы индуктивного и дедуктивного обучения, например, нейросетевые модели с интеграцией экспертных правил или системы, использующие машинное обучение для дополнения и уточнения баз знаний.

В машинном обучении задачи, основанные на обучении по прецедентам, традиционно классифицируются на несколько ключевых типов [54]:

- Обучение с учителем (Supervised Learning): алгоритмы обучаются на размеченных данных, где каждому объекту соответствует целевая переменная (метка).
- Обучение без учителя (Unsupervised Learning): анализ данных без явных меток, направленный на выявление скрытых паттернов (кластеризация, снижение размерности).
- Частичное обучение (Semi-supervised Learning): комбинация размеченных и неразмеченных данных для улучшения качества моделей.

- Трансдуктивное обучение (Transductive Learning): предсказание меток для конкретного набора объектов без построения общей модели.
- Обучение с подкреплением (Reinforcement Learning): обучение через взаимодействие со средой и максимизацию награды.
- Динамическое обучение (Online Learning): последовательное обновление модели по мере поступления новых данных.
- Активное обучение (Active Learning): обучаемый имеет возможность самостоятельно назначать следующий прецедент, который станет известен.
- Метаобучение (Meta-learning): прецедентами являются ранее решённые задачи обучения.

В рамках диссертационной работы основное внимание уделяется задачам классификации, регрессии и прогнозирования, которые относятся к категории обучения с учителем. Эти задачи предполагают построение моделей, способных предсказывать целевую переменную (дискретную – для классификации, непрерывную – для регрессии) на основе исторических данных. Прогнозирование, как частный случай, часто связано с временными рядами и требует учета временных зависимостей.

Одной из ключевых задач, решаемых в рамках работы, является разработка методов генерации синтетических и аугментированных данных. Для реализации данной задачи предлагается создание специализированной виртуальной среды, предназначенного для моделирования и генерации синтетических данных. Применение синтетических данных позволяет существенно снизить финансовые и временные затраты, связанные с их получением, поскольку процесс их генерации зачастую является менее ресурсоемким по сравнению с процедурой сбора и обработки данных, полученных в результате наблюдения за реальными событиями.

Кроме того, временные затраты на синтез данных могут быть сокращены до нескольких дней, в то время как сбор и обработка реальных данных могут занимать значительно больше времени. Важным аспектом является также проблема ограниченной доступности данных, обусловленная их редкостью или

конфиденциальностью. В таких случаях синтетические данные представляют собой эффективное решение, позволяющее преодолеть указанные ограничения и обеспечить необходимый объем информации для проведения исследований и анализа без ущерба для качества и достоверности результатов.

Аугментированные данные. В отличие от синтетических данных, которые генерируются искусственно, аугментация сохраняет базовую структуру и статистические закономерности исходных данных, но вносит контролируемые изменения. Для табличных данных методы аугментации включают добавление шума в числовые признаки (в пределах допустимой дисперсии), генерацию новых записей через интерполяцию или SMOTE (техника для балансировки классов), перестановку категориальных значений без нарушения их распределения, а также создание производных признаков на основе комбинаций исходных переменных. Такие подходы особенно актуальны для задач машинного обучения, таких, как прогнозирование, классификация или регрессия, где ограниченный объем данных может привести к переобучению моделей. Аугментация табличных данных позволяет повысить устойчивость алгоритмов к вариациям входных параметров, улучшить их обобщающую способность и эффективно работать с несбалансированными наборами данных.

В контексте задач, связанных с анализом и прогнозированием в области солнечной энергетики, выбор методов машинного обучения обусловлен необходимостью обработки сложных временных зависимостей и нелинейных взаимосвязей в данных. Рекуррентные нейронные сети (RNN), включая их модификации – LSTM и GRU, демонстрируют высокую эффективность при работе с последовательностями данных, что подтверждается их успешным применением в задачах прогнозирования выработки энергии. Их ключевое преимущество заключается в способности учитывать долгосрочные временные зависимости, что критически важно для моделирования процессов, зависящих от изменчивых факторов, таких, как погодные условия или суточные колебания солнечной активности.

Рекуррентная нейронная сеть (РНС, англ. Recurrent neural network, RNN) – это семейство нейронных сетей для обработки последовательных данных. Их ключевая особенность заключается в способности учитывать временные зависимости и контекст, что делает их идеальными для задач, где данные имеют временную или последовательную структуру.

В традиционных feedforward-нейронных сетях (например, многослойных перцептронах) информация передаётся только в одном направлении – от входного слоя к выходному. Такие сети не учитывают контекст или предыдущие состояния, что делает их непригодными для обработки последовательностей (Рис. 1.4) [28].

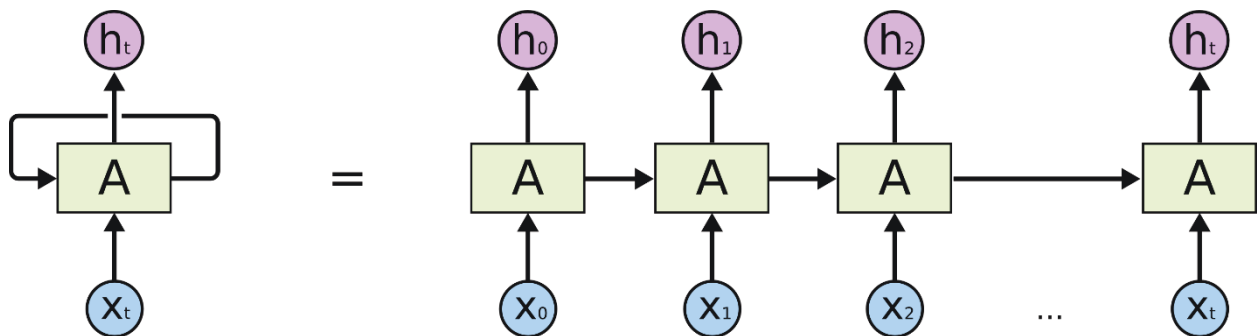


Рисунок 1.4 – Схема рекуррентной нейронной сети в развернутом виде

Основная идея RNN заключается в использовании информации, распространяемой через предыдущие временные шаги, для обработки текущих входных данных. На Рисунке 1.5 [55] RNN изображена, как фрагмент нейронной сети (A), который принимает входные данные (x_t) и выдает результат (h_t), учитывая как текущий ввод, так и контекст, накопленный на предыдущих шагах. Таким образом, RNN можно представить, как последовательность копий одной и той же сети, где каждая копия передает контекстную информацию следующей.

Математически основная идея RNN выражается через использование скрытого состояния (hidden state), которое обновляется на каждом шаге последовательности. Формально это можно описать следующим уравнением

$$h_t = f(h_{t-1} - x_t), \quad (1.9)$$

где h_t – скрытое состояние на шаге t , x_t – выходные данные на шаге t , f – нелинейная функция активации (например, гиперболический тангенс ($\tanh$) или ReLU.).

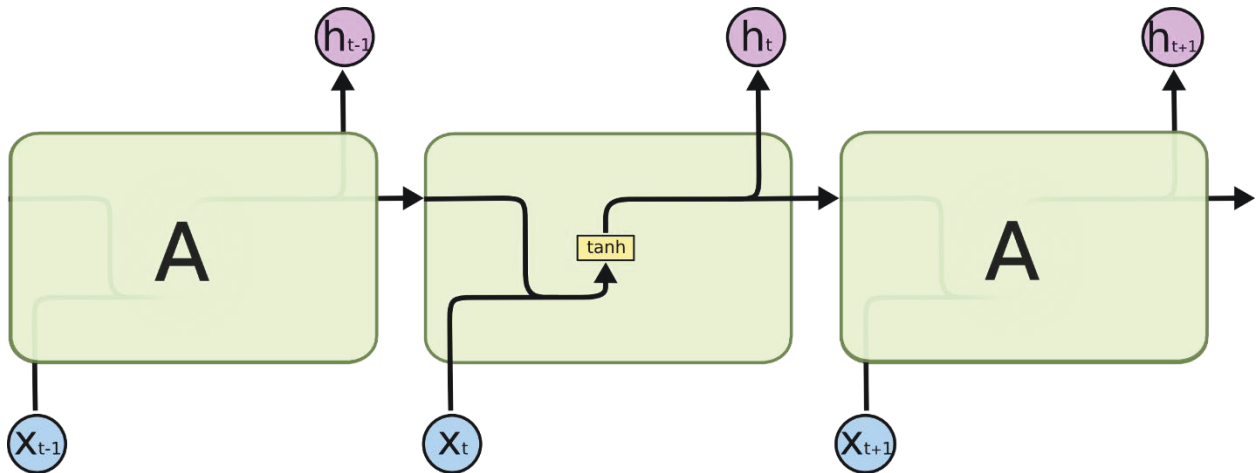


Рисунок 1.5 – Простая RNN сеть содержит один слой

Скрытое состояние h_t служит для передачи информации между временными шагами, что позволяет сети учитывать временные зависимости в данных. Это свойство делает RNN особенно эффективными для задач, связанных с последовательностями, такими, как прогнозирование временных рядов, обработка естественного языка и другие.

Проблема долговременных зависимостей – это одна из ключевых трудностей, с которыми сталкиваются рекуррентные нейронные сети (RNN) при работе с последовательностями данных. Эта проблема возникает, когда сеть должна учитывать информацию, которая была получена на ранних этапах последовательности, чтобы правильно интерпретировать или предсказать данные на более поздних этапах. Однако, из-за особенностей архитектуры и процесса обучения, RNN часто неспособны эффективно сохранять и использовать информацию, относящуюся к долгосрочным зависимостям.

Основные причины проблемы:

- Исчезающие и взрывающиеся градиенты:

В процессе обучения RNN используется метод обратного распространения ошибки через время (Backpropagation Through Time, BPTT). При этом градиенты, которые передаются через множество временных шагов, могут экспоненциально уменьшаться (исчезать) или увеличиваться (взрываться). Это связано с тем, что градиенты на каждом шаге умножаются на матрицу весов, и, если собственные значения этой матрицы меньше 1, градиенты стремятся к нулю, а если больше 1 – к бесконечности.

Исчезающие градиенты приводят к тому, что сеть перестает обновлять веса, связанные с долгосрочными зависимостями, и фокусируется только на краткосрочных связях. Взрывающиеся градиенты, напротив, делают процесс обучения нестабильным.

- Ограниченная память стандартных RNN:

Стандартные RNN имеют простую архитектуру, где скрытое состояние на каждом шаге обновляется на основе текущего входа и предыдущего скрытого состояния. Однако такая архитектура не предназначена для длительного хранения информации. Скрытое состояние "забывает" информацию, полученную на ранних этапах последовательности, особенно если последовательность длинная.

Сеть долгосрочной кратковременной памяти (Long Short Term Memory, LSTM) – разновидность архитектуры рекуррентной нейронной сети, которая способна к обучению долговременной зависимости. Эта архитектура способна решить проблемы исчезающего и взрывающегося градиентов и ограниченной памяти. Предложена была в 1997 году З. Хохрайтером и Ю. Шмидхубером [56].

Архитектурное отличие между рекуррентными нейронными сетями (RNN) (Рис. 1.6) и сетями с долгой краткосрочной памятью (LSTM) (Рис. 1.7) заключается в принципиально иной организации ячеек памяти [55]. В то время как традиционные RNN обладают относительно простой структурой с единственным состоянием, передаваемым между временными шагами, LSTM-сети оснащены сложной системой управляющих механизмов – так называемые "ворота" (с англ. gates).

Эти специализированные ворота (включающие входные, выходные и ворота забывания) выполняют точный контроль над информационными потоками внутри ячейки. Они позволяют LSTM-сетям осуществлять селективное добавление новой информации, целенаправленное удаление ненужных данных и надежное сохранение критически важных состояний на протяженных временных промежутках.

Благодаря такой архитектуре LSTM успешно преодолевает фундаментальные ограничения классических RNN, в частности проблему исчезающих градиентов, которая затрудняет обучение на длинных последовательностях. Кроме того, механизм ворот обеспечивает эффективное запоминание долгосрочных временных зависимостей, что особенно важно для задач обработки последовательных данных с протяженными временными взаимосвязями. Используются 3 типа ворот:

- Входные ворота (input gate): определяют, какая новая информация будет добавлена в состояние ячейки.
- Забывающие ворота (forget gate): решают, какую информацию следует удалить из состояния ячейки.
- Выходные ворота (output gate): контролируют, какая информация будет передана на следующий временной шаг.

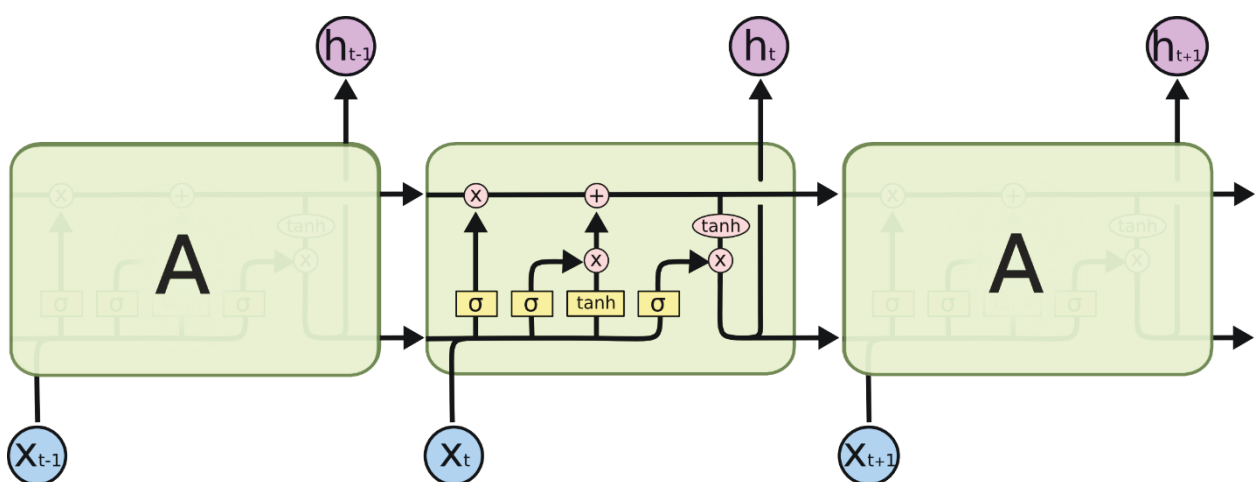


Рисунок 1.6 – LSTM сеть содержит 4 пересекающихся слоя

Эти ворота позволяют LSTM избирательно сохранять или удалять информацию, что делает её более устойчивой к проблемам долгосрочных зависимостей. Работа LSTM описывается следующими уравнениями:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f), \quad (1.10)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i), \quad (1.11)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c), \quad (1.12)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \quad (1.13)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o), \quad (1.14)$$

$$h_t = o_t \odot \tanh(c_t), \quad (1.15)$$

где f_t , i_t , o_t – забывающие, входные и выходные ворота соответственно; x_t – входной вектор на шаге t ; h_t – скрытое состояние на шаге t ; h_{t-1} – скрытое состояние на предыдущем шаге; c_t – состояние ячейки на шаге t ; $\tilde{c}_t$ – кандидат на обновление состояния ячейки; σ – сигмоидная функция активации, которая действует поэлементно; $\tanh$ – функция гиперболического тангенса, которая действует поэлементно; $W_f, W_i, W_c, W_o \in \mathbb{R}^{n \times m}$ – матрицы весов для входного вектора; $U_f, U_i, U_c, U_o \in \mathbb{R}^{n \times m}$ – матрицы весов для скрытого вектора; $b_f, b_i, b_c, b_o \in \mathbb{R}^n$ – матрицы смещения весов для входного вектора; $\odot$ – поэлементное умножение.

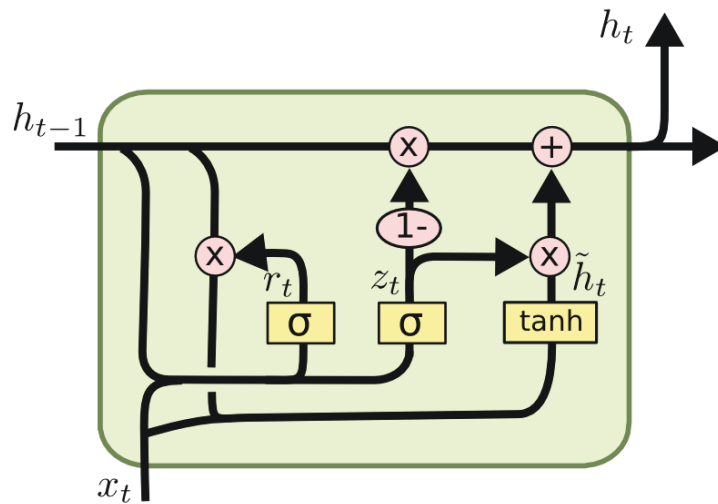


Рисунок 1.7 – GRU сеть

Управляемый рекуррентный блок (Gated Recurrent Units, GRU) представляет собой модернизированную архитектуру рекуррентных нейронных сетей, разработанную для решения проблем "исчезающего градиента" при обработке временных рядов. В отличие от классических LSTM, GRU упрощает вычислительную модель за счет объединения механизмов забывания и обновления информации в единые гейты, сохраняя при этом способность фиксировать долгосрочные временные зависимости [57].

Работа GRU сети описывается следующими уравнениями:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z), \quad (1.16)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r), \quad (1.17)$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h), \quad (1.18)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t, \quad (1.19)$$

где z_t , r_t – ворота обновления, сброса (значения от 0 до 1); $\tilde{h}_t$ – промежуточное скрытое состояние; h_t – новое скрытое состояние на шаге t ; σ – сигмоидная функция активации, которая действует поэлементно; $\tanh$ – функция гиперболического тангенса которая действует поэлементно; $W_z, W_r, W_h \in \mathbb{R}^{n \times m}$ – матрицы весов для входного вектора; $U_z, U_r, U_h \in \mathbb{R}^{n \times m}$ – матрицы весов для скрытого вектора; $b_z, b_r, b_h \in \mathbb{R}^n$ – матрицы смещения весов для входного вектора; $\odot$ – поэлементное умножение. Подробно рекуррентные нейронные сети, LSTM и GRU и принципы работы описаны в [56, 57] и в статье [55] Understanding LSTM Networks.

Градиентный бустинг, представленный такими реализациями, как XGBoost, LightGBM и CatBoost, обеспечивает высокую точность прогнозирования при работе с табличными данными, включая метеорологические параметры и эксплуатационные характеристики оборудования. Этот метод обладает устойчивостью к переобучению благодаря механизмам регуляризации, а также позволяет интерпретировать результаты, что важно для анализа влияния отдельных факторов на эффективность энергосистем.

Градиентный бустинг (Gradient Boosting) – это метод машинного обучения, основанный на идее последовательного построения ансамбля моделей, где каждая

последующая модель корректирует ошибки предыдущих. Этот метод минимизирует заданную функцию потерь путем итеративного добавления слабых моделей (обычно деревьев решений), которые аппроксимируют антиградиент функции потерь. Основная идея градиентного бустинга была формализована Джеромом Фридманом в работе [58], где он предложил использовать градиентный спуск в пространстве функций.

Дана выборка $\{(x_i, y_i)\}_{i=1, \dots, n}$, где x_i – признаки, а y_i – целевая переменная. Цель градиентного бустинга – построить модель F^* , которая минимизирует заданную функцию потерь $L(y, f(x))$:

$$F^* = \arg \min_{f \in \mathcal{F}} \sum_{i=1}^n L(y_i, f(x_i)), \quad (1.20)$$

где f – произвольная предсказывающая модель из множества допустимых моделей $\mathcal{F}$.

Градиентный бустинг использует идею градиентного спуска в пространстве функций. На каждой итерации модель обновляется в направлении антиградиента функции потерь, что позволяет минимизировать ошибку. Формально это можно описать следующим образом:

1. Начальное приближение $F_0(x)$ задается, как константа, минимизирующая функцию потерь:

$$F_0(x) = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, \gamma). \quad (1.21)$$

Для квадратичной функции потерь $L(y, \gamma) = (y - \gamma)^2$ – это среднее значение целевой переменной:

$$F_0(x) = \frac{1}{n} \sum_{i=1}^n y_i. \quad (1.22)$$

2. На каждой итерации $t = 1, \dots, M$, где M – общее количество наблюдений:

Вычисляются псевдо – остатки (антиградиенты):

$$r_{it} = -\left[\frac{\partial L(y_i F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{t-1}(x)}. \quad (1.23)$$

Обучается новая базовая модель $h_t(x)$ на псевдо – остатках $\{(x_i y_i)\}_{i=1, \dots, n}$

Находится оптимальный коэффициент ρ_t , минимизирующий функцию потерь на обучающей выборке $\{(x_i, y_i)\}_{i=1}^n$ для новой базовой модели $h_t(x)$:

$$\rho_t = \arg \min_{\rho} \sum_{i=1}^n L(y_i, F_{t-1}(x_i) + \rho h_t(x_i)). \quad (1.24)$$

Обновляется текущее приближение:

$$F_t(x) = F_{t-1}(x) + p_t h_t(x). \quad (1.25)$$

Итоговая модель:

$$F(x) = F_0(x) + \sum_{t=1}^M p_t h_t(x). \quad (1.26)$$

Градиентный бустинг является одним из самых мощных методов машинного обучения для задач регрессии и классификации на структурированных данных. Его ключевое преимущество – высокая точность, достигаемая за счет последовательного улучшения предсказаний. Однако для достижения наилучших результатов требуется тщательная настройка гиперпараметров и учет особенностей данных.

В задачах прогнозирования солнечной радиации и выработки электроэнергии методы машинного обучения получили широкое распространение. В систематических обзорах [59–64] показано, что рекуррентные нейронные сети, в частности архитектуры LSTM и GRU, обеспечивают более высокую точность прогнозирования по сравнению с традиционными регрессионными моделями за счет учета временных зависимостей. Вместе с тем, для задач с ограниченными объемами данных эффективными оказываются методы ансамблевого обучения и алгоритмы градиентного бустинга, обеспечивающие устойчивость и интерпретируемость результатов.

Отдельное место среди методов машинного обучения занимают ансамблевые методы. Их основная идея заключается в том, что итоговое решение формируется

не одной моделью, а совокупностью нескольких базовых моделей, результаты работы которых объединяются по определенному правилу. Такой подход позволяет повысить устойчивость прогнозирования и уменьшить влияние ограничений, присущих отдельным моделям, поскольку различные алгоритмы по-разному реагируют на шум, выбросы, неполноту данных и нелинейный характер исследуемых зависимостей. Для задач регрессии, к которым относится прогнозирование солнечной генерации, ансамблирование особенно важно, поскольку оно позволяет уменьшить разброс ошибки и повысить обобщающую способность модели.

К числу наиболее распространенных ансамблевых подходов относятся бэггинг (bagging), бустинг (boosting) и стекинг (stacking). В методах бэггинга несколько моделей обучаются на различных подвыборках исходных данных, после чего их прогнозы усредняются. Это позволяет снизить дисперсию и повысить устойчивость результата. В бустинге базовые модели строятся последовательно, при этом каждая следующая модель ориентирована на компенсацию ошибок предыдущих, что способствует повышению точности на сложных участках зависимости. В стекинге объединение выполняется на более высоком уровне: прогнозы нескольких базовых моделей поступают на вход метамодели, которая обучается выбирать или комбинировать их выходы. Таким образом, ансамблевые методы позволяют использовать сильные стороны разнородных алгоритмов в рамках единой прогнозной схемы [65–67].

В задачах прогнозирования внешней среды и выработки электроэнергии ансамблевые методы представляют особый интерес, поскольку исследуемые временные ряды обладают выраженной нестационарностью и зависят от большого числа внешних факторов, включая солнечную радиацию, температуру воздуха, облачность, сезонность и локальные погодные режимы. В обзорных исследованиях показано, что ансамблевые и гибридные модели в ряде случаев обеспечивают более высокую точность по сравнению с одиночными моделями, особенно при комбинировании разнородных предикторов, использовании предварительной обработки данных и учете режимов функционирования внешней среды [68]. Это

делает ансамблирование перспективным направлением для построения моделей прогнозирования в составе цифрового двойника солнечной электростанции.

Таким образом, методы искусственного интеллекта и машинного обучения дополняют математическое моделирование, расширяя функциональные возможности цифрового двойника. Для задач солнечной энергетики их применение позволяет учитывать нелинейные зависимости между метеорологическими параметрами и выработкой электроэнергии, работать с временными рядами, формировать прогнозные модели и адаптировать расчётные алгоритмы к изменяющимся условиям внешней среды. Это обосновывает необходимость разработки методического подхода, в котором математические модели, цифровая тень, виртуальная среда, онтологическое описание и модели машинного обучения рассматриваются как взаимосвязанные компоненты умного цифрового двойника солнечной электростанции

1.5. Выводы по главе 1

В первой главе выполнен анализ предметной области. Рассмотрены преимущества возобновляемых источников энергии (ВИЭ), включая энергетическую безопасность, экологичность и неисчерпаемость, а также специфика солнечной энергетики с её автономностью, что особенно актуально для труднодоступных регионов России. При этом отмечены технологические ограничения, такие, как зависимость от погодных условий, частичное затенение панелей и постепенная деградация фотоэлектрических модулей, снижающие эффективность солнечных электростанций. Выполнен обзор математических моделей отдельных компонентов СЭС и разных уровней детализации при моделировании.

Цифровые двойники представлены, как виртуальные динамические модели, интегрирующие данные с физических объектов для прогнозирования и оптимизации их работы. Подчеркивается различие между подходами на основе математического моделирования (Simulation-based DT) и анализа данных

(Analytics-based DT). Подчеркнута важность развития гибридных моделей, сочетающих оба подхода, для повышения точности прогнозирования в условиях изменяющейся среды. Отмечены существующие проблемы внедрения ЦД, включая необходимость значительных вычислительных ресурсов и отсутствие единых стандартов для энергетических объектов

Методы машинного обучения, такие, как рекуррентные нейронные сети (LSTM) и градиентный бустинг, рассмотрены в контексте их использования для прогнозирования и анализа данных. Особое внимание уделено проблемам обработки нестационарных временных рядов, характеризующихся сезонностью, трендами и зависимостью от внешних факторов, а также способам их решения, включая генерацию синтетических и аугментированных данных для повышения качества моделей. В то же время отмечены проблемы, связанные с необходимостью больших объемов качественных данных для обучения и сложностью интерпретации результатов работы сложных моделей.

Проведённый анализ научных работ показал, что технология цифровых двойников активно развивается и рассматривается как эффективный инструмент моделирования, мониторинга и прогнозирования функционирования энергетических объектов. Вместе с тем существующие исследования в области построения цифровых двойников энергетических систем в большинстве случаев ориентированы либо на использование математических моделей физических процессов, либо на применение методов анализа данных и машинного обучения.

Анализ существующих работ также показал недостаточную разработанность методических подходов к построению умных цифровых двойников, включающих интеллектуальные компоненты и обеспечивающих адаптивное прогнозирование характеристик энергетических объектов.

Таким образом, выявлена научная проблема, заключающаяся в необходимости разработки методического подхода к построению умных цифровых двойников энергетических объектов, обеспечивающего интеграцию математического моделирования, методов машинного обучения и виртуальной

среды для анализа и прогнозирования характеристик объектов возобновляемой энергетики.

На основании этого формулируется постановка задачи:

1. Разработать методический подход к построению умных цифровых двойников, включающий:

1.1. Агрегированную модель солнечной электростанции и формализованную модель умного цифрового двойника.

1.2. Методику построения умного ЦД СЭС.

1.3. Численный метод адаптивного выбора моделей машинного обучения.

1.4. Методику построения виртуальной среды, обеспечивающий эмуляцию данных.

2. Выполнить реализацию программного комплекса (научного прототипа умного цифрового двойника) на основе агентно-сервисной архитектуры, интегрирующего вычислительное ядро (реализация модели ЦД СЭС и численного метода), цифровую тень и виртуальную среду.

Глава 2. Методический подход к построению умных цифровых двойников ВИЭ

Во второй главе сформулированы основные теоретические и методические положения построения умного цифрового двойника солнечной электростанции. Рассмотрены модель умного цифрового двойника, включающая математические, цифровые и интеллектуальные компоненты, описывается моделирование солнечной электростанции в составе цифрового двойника, предложена методика построения виртуальной среды и интеграции моделей машинного обучения. Отдельное внимание уделено предложенному численному методу ансамблирования моделей с выбором на основе состояния внешней среды, предназначенному для повышения точности прогнозирования временных рядов. Результаты опубликованы в работах автора [69–71]. Эти результаты являются составляющими методического подхода к построению умного цифрового двойника СЭС, который далее используется при разработке архитектуры и программной реализации научного прототипа.

2.1. Модель умного цифрового двойника

Для формирования умного цифрового двойника требуется определить его компоненты и взаимосвязь между ними. Четкого определения термина «Умного цифрового двойника» (Smart Digital Twin) пока нет, автор предлагает следующую трактовку.

Умный цифровой двойник – это интеллектуальная система, которая использует не только явные знания, а инженерия знаний, извлекая неявный опыт эксперта и формализуя его в онтологии и базы знаний, и на этой основе самостоятельно диагностирует, прогнозирует и принимает решения. Умный цифровой двойник, как развитие традиционного ЦД, способен не только воспроизводить поведение, но и анализировать и предсказывать работу энергетического объекта, используя интеллектуальные компоненты семантические

(онтологические) модели, экспертные системы, базу знаний, модели машинного обучения), реальные или эмулированные данные из виртуальной среды.

Для управления энергетическими объектами достаточно использовать традиционные ЦД (будем далее использовать сокращение от англоязычного термина Digital Twin – DT). При управлении энергосистемами могут потребоваться элементы интеллектуальных технологий, а при стратегическом управлении (управлении развитием ТЭК) потребуется интеллектуальный или умный DT (Smart Digital Twin – SDT). В отличие от понятия SDT в [72], где SDT понимается как виртуальная модель, которая может воспроизводить проблему планирования, чтобы имитировать ее на основе методологии глубокого обучения с подкреплением, автор предлагает расширить данный термин, и, помимо использования методов машинного обучения, подключить методы работы с экспертными знаниями. В качестве реализации работы с экспертными знаниями предлагается использовать семантические модели, которые формализуют представления знания о предметной области, описывающих сущности, их свойства и отношения между ними.

Методы искусственного интеллекта и перечисленные программные компоненты, необходимые для построения DT, будут интегрированы в SDT. В коллективе, в котором работает автор, предложена следующая модель DT [73]. Формализованная модель цифрового двойника имеет следующий вид:

$$DT = (IM, MM, DB, VT, SE, RC), \quad (2.1)$$

где IM – информационные модели; MM – математические модели; DB – базы данных; VT – средства визуализации, SE – необходимые сервисные компоненты; RC – средства связи между компонентами в режиме реального времени.

Автором диссертации предложена модификация этой модели [70], описывающая умный ЦД:

$$SDT = (IM, MM, DB, VT, SE, RC, SD, KB, VS, AIM), \quad (2.2)$$

где SD – схемы и диаграммы, KB – базы знаний, VS – виртуальная среда, AIM – модели искусственного интеллекта.

Для описания интеграции интеллектуальных технологий в умном цифровом двойнике предлагается использовать фрактальный подход к структурированию знаний.

Основой подхода является построение фрактальной стратифицированной модели (ФС – модели) информационного пространства [74–76], которую графически можно представить в виде совокупности вложенных сферических оболочек (слоев), каждый из которых объединяет множество однотипных информационных объектов, и, в свою очередь, может быть расслоен. Поскольку рассматриваемая предметная область – один из фрагментов информационного пространства, его можно представить в виде вырезки (например, конуса), в котором выделены рассматриваемые слои. Умный ЦД может быть представлен слоями конуса, где выделяются: слой онтологических моделей, слой математических моделей, слой информационных моделей и слой интеллектуальных технологий. Последний, в свою очередь может быть расслоен как:

$$AIM = (SM, ES, ML), \quad (2.3)$$

где SM – семантические модели, ES – экспертные системы, ML – модели машинного обучения.

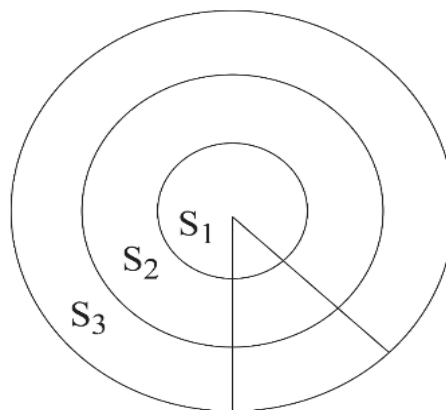


Рисунок 2.1 – Фрактальная стратифицированная модель (ФС-модель)

Для интеграции интеллектуальных технологий в умном ЦД выделены три слоя (Рис. 2.1): S_1 – слой методов машинного обучения, S_2 – слой экспертных

систем и S_3 – слой семантических моделей. Для выделения необходимой части (сектора) для умного цифрового двойника (SDT) представляется конус и описывается отображение слоев (Рис. 2.2), где D_{ES} – SDT с экспертными системами, D_{SM} – SDT с семантическими моделями и D_{ML} – SDT с моделями машинного обучения.

Согласно концепции ФС – модели, можно построить отображения (Рис. 2.2):

- $F_{SM}^{ML}: D_{ML} \rightarrow D_{SM}$ – отображение из слоя моделей машинного обучения SDT в слой семантических моделей SDT ;
- $F_{ES}^{ML}: D_{ML} \rightarrow D_{ES}$ – отображение из слоя моделей машинного обучения SDT в слой экспертных систем SDT ;
- $F_{ES}^{SM}: D_{SM} \rightarrow D_{ES}$ – отображение из слоя семантических моделей SDT в слой экспертных систем SDT .

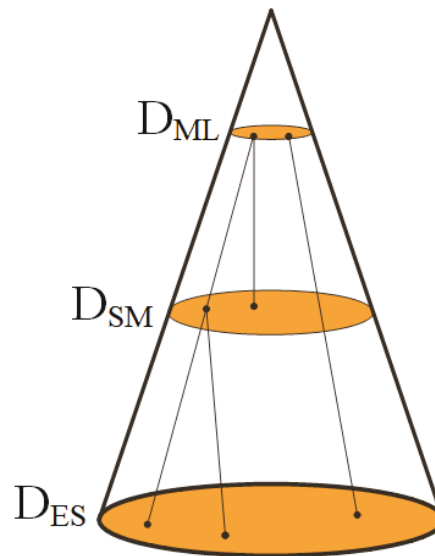


Рисунок 2.2 – ФС-модель интеллектуальных технологий для умного цифрового двойника

В контексте разработки умного цифрового двойника солнечной электростанции, интеграция баз знаний представляет собой компонент, обеспечивающий системе способность к интеллектуальному анализу и принятию решений. Базы знаний выступают в качестве структурированного хранилища, содержащего не только фактические данные об объекте, но и экспертные правила,

семантические связи, а также формализованные знания о предметной области. Их ключевая роль заключается в обеспечении цифрового двойника возможностью интерпретировать данные, выявлять закономерности и генерировать выводы, что значительно расширяет функциональность системы по сравнению с традиционными подходами. Реализацией семантических моделей предлагается использовать онтологии. Применительно к солнечной электростанции семантические модели позволяют формально описать состав объекта, параметры внешней среды и связи между элементами системы.

Отображения слоев поддерживаются специальными, в т.ч. авторскими, инструментальными средствами.

2.2. Агрегированная модель солнечной электростанции как основа вычислительного ядра умного цифрового двойника

При построении умного цифрового двойника солнечная электростанция рассматривается как динамическая система, поведение которой определяется внешними воздействиями и параметрами модели. В этом случае задача моделирования состоит в построении зависимости между входными параметрами среды и выходными характеристиками объекта, необходимой для проведения вычислительных экспериментов, эмуляции данных и решения задач прогнозирования.

В рамках настоящего исследования предлагается построить агрегированную математическую модель солнечной электростанции. Предлагаемый подход позволяет агрегировать существующие модели компонентов ЦД СЭС, подключать новые модели компонентов СЭС и обеспечивать возможность создания разных конфигураций математической модели СЭС для реализации вычислительного ядра (цифровой модели) умного ЦД.

Под агрегированной моделью понимается описание, в котором солнечная электростанция рассматривается не на уровне детального электрофизического моделирования каждого фотоэлектрического элемента, а как система,

связывающая параметры внешней среды, характеристики фотоэлектрических модулей и эксплуатационные коэффициенты с выходными параметрами генерации. Такой уровень детализации является достаточным для задач построения цифрового двойника СЭС, проведения сценарного моделирования, эмуляции данных и последующей интеграции методов машинного обучения.

Агрегированная модель позволяет учитывать основные факторы, влияющие на функционирование солнечной электростанции: солнечную радиацию, угол падения солнечных лучей, угол наклона и ориентацию панели, температуру окружающей среды, скорость ветра, деградацию фотоэлектрических модулей, коэффициенты потерь и загрязнение поверхности. При этом более детальные модели фотоэлектрических преобразователей, модели с использованием W-функции Ламберта, модели неоднородного освещения и модели диагностики состояния модулей рассматриваются как альтернативные уровни математического описания, применимые при детальном исследовании электрических характеристик отдельных модулей и их технического состояния.

В настоящей работе указанные детальные модели не используются как основной расчетный аппарат, поскольку целью исследования является разработка методического подхода к интеграции математических моделей, виртуальной среды и моделей машинного обучения и построение цифрового двойника солнечной электростанции с использованием этого методического подхода. Поэтому в качестве основы цифровой модели СЭС применяется агрегированное описание, отражающее значимые зависимости между внешними параметрами среды и выходными характеристиками станции.

В общем виде данная постановка определяет задачу моделирования солнечной электростанции, как основы вычислительного ядра умного цифрового двойника. Функционирование солнечной электростанции может быть представлено в виде соотношения:

$$Y(t) = F(X(t), \Theta), \quad (2.4)$$

где $X(t)$ – вектор входных параметров среды, $Y(t)$ – вектор выходных характеристик системы (мощность, напряжение, ток, КПД), Θ – параметры модели СЭС: площадь

панели, коэффициенты потерь, температурные коэффициенты, параметры фотоэлектрического преобразователя и др.

Вектор входных параметров среды задается в виде:

$$X(t) = \{G(t), T(t), V(t), p(t)\}, \quad (2.5)$$

где $G(t)$ – солнечная радиация, $T(t)$ – температура воздуха, $V(t)$ – скорость ветра, $p(t)$ – атмосферное давление.

Функция $F(X(t), \Theta)$, определяющая связь между входными и выходными параметрами, может быть конкретизирована математическими моделями функционирования солнечной электростанции. В качестве такого описания могут использоваться формулы (2.6) – (2.14), отражающие основные зависимости, характеризующие процесс преобразования солнечной энергии в электрическую энергию.

Для расчета мощности, вырабатываемой солнечной панелью, используется формула [77, 78]:

$$P_{pv}(t) = G_h(t) \cdot \eta \cdot A, \quad (2.6)$$

где $P_{pv}(t)$ – выходная мощность, Вт; η – КПД панели в стандартных условиях, отн.ед; A – площадь панели, м²; $G_h(t)$ – солнечная радиация, поступающая на поверхность, Вт/м².

В книге Solar Engineering of Thermal Processes [79] описан расчет полной радиации ($G_T(t)$) на солнечную панель

$$G_T(t) = G_b(t) \cdot R_b(t) + G_d(t) \cdot \left(\frac{1 + \cos \beta}{2} \right) + G(t) \cdot \rho \cdot \left(\frac{1 - \cos \beta}{2} \right), \quad (2.7)$$

где $G_b(t)$ – прямая радиация, Вт/м²; $G_d(t)$ – диффузная радиация, Вт/м²; $G(t) = G_b(t) + G_d(t)$ – полная радиация на горизонтальную поверхность, Вт/м²; $R_b(t) = \frac{\cos \theta(t)}{\cos \theta_z(t)}$ – геометрический фактор; $\theta(t)$ – угол падения солнечных лучей на наклонную поверхность; $\theta_z(t)$ – зенитный угол Солнца; β – угол наклона панели; ρ – альбедо (коэффициент отражения поверхности).

Угол падения $\theta(t)$ определяется из соотношения:

$$\theta = \arccos(\sin \delta \cdot \sin \phi \cdot \cos \beta - \quad (2.8)$$

$$\begin{aligned}
& - \sin \delta \cdot \cos \phi \cdot \sin \beta \cdot \cos \gamma + \\
& + \cos \delta \cdot \cos \phi \cdot \cos \beta \cdot \cos \omega + \\
& + \cos \delta \cdot \sin \phi \cdot \sin \beta \cdot \cos \gamma \cdot \cos \omega + \\
& + \cos \delta \cdot \sin \beta \cdot \sin \gamma \cdot \sin \omega),
\end{aligned}$$

где θ – угол падения солнечных лучей на поверхность; δ – склонение Солнца; ϕ – широта местности; β – угол наклона поверхности; γ – азимутальный угол поверхности (0° – юг, 90° – запад); ω – часовой угол Солнца.

В работе [80] выполнен аналитический обзор скорости деградации солнечных панелей. В этом обзоре приведены факторы, влияющие на деградацию солнечных панелей: температура, ультрафиолетовое – излучение и потенциально-индуцированная деградация. В нем выводятся средние показатели по годовой скорости деградации (r):

- Кремниевые модули: $0.5 - 0.8\%$ в год.
- Тонкоплёночные: $0.7 - 1.2\%$ в год.

Показатели деградации влияют на мощность работы солнечных панелей; это отображается в формуле:

$$P(\tau) = P_0(1 - r \cdot \tau), \tau \in [0, \tau_{max}], \quad \tau_{max} < \frac{1}{r}, \quad (2.9)$$

где $P(\tau)$ – мощность через τ лет, Вт; P_0 – начальная мощность, Вт; r – годовая скорость деградации ед. в год.

В статье [81] проводится моделирование солнечной электростанции с учетом изменения параметров окружающей среды, где рассматриваются внешние факторы, влияющие на выработку электроэнергии. Эти факторы отображены в формулах:

$$P_{pv}(t) = G_T(t) \cdot \eta(t) \cdot A \cdot k_L, \quad (2.10)$$

где k_L – коэффициент, учитывающий потери в диодах и из-за загрязнения, принимаемый равным $0,85 - 0,95$ отн. ед;

Эффективность пересчитывается в зависимости от уровня солнечной радиации, температуры наружного воздуха и скорости ветра:

$$\eta(t) = \bar{\eta} \cdot [1 - \beta(T_{pv}(t) - 25)], \quad (2.11)$$

где $\bar{\eta}$ – паспортные значения КПД фотоэлектрического преобразователя, отн. ед; β – температурный коэффициент для кремниевых фотоэлектрических преобразователей (принимается 0,004 – 0,006), отн. ед./°C; T_{pv} – рабочая температура фотоэлектрического преобразователя, °C.

Рабочая температура фотоэлектрического преобразователя определяется по следующей формуле:

$$T_{pv}(t) = T_a(t) + \frac{G_T(t)}{k_0 + k_1 * V_w(t)}, \quad (2.12)$$

где k_0, k_1 – корреляционные коэффициенты Коehl, равные 30,02 и 6,28 [82]; $V_w(t)$ – скорость ветра на поверхности земли, м/с; T_a – температура воздуха, °C.

Далее рассчитывается фототок:

$$I_{PH}(t) = \frac{G_T(t)}{1000} [I_{SC} + k_I(T_{pv}(t) - 25)], \quad (2.13)$$

где I_{SC} – ток короткого замыкания, А; k_I – температурный коэффициент 0,0038 отн.ед./°C.

Далее рассчитывается напряжение:

$$U_{pv}(t) = P_{pv}(t) / I_{PH}(t). \quad (2.14)$$

Таким образом, математические формулы (2.6) – (2.14) позволяют конкретизировать общую зависимость между внешними параметрами среды и выходными характеристиками солнечной электростанции. Они задают основу для последующего построения цифровой модели и использования ее в составе вычислительного ядра умного цифрового двойника. При этом в рамках настоящего исследования основное внимание уделяется не только прямому расчету характеристик по физическим зависимостям, но и задаче прогнозирования выходных величин на основе данных.

В общем случае в качестве целевой выходной переменной может рассматриваться одна из характеристик функционирования солнечной электростанции. Для задач оценки режима генерации такой переменной является мощность генерации СЭС

$$y^*(t) = P(t). \quad (2.15)$$

Для задач прогнозирования в составе виртуальной среды в качестве целевой переменной может также рассматриваться солнечная радиация как один из определяющих факторов генерации. Тогда задача моделирования состоит в построении модели

$$\hat{y}(t) = f(X(t), \Theta), \quad (2.16)$$

обеспечивающей минимизацию функционала ошибки на заданном множестве наблюдений

$$f^* \in \arg \min_{f \in \mathcal{F}} \mathcal{L}(\mathcal{D}, f), \quad (2.17)$$

где $\mathcal{F}$ – множество допустимых моделей; $\mathcal{D}$ – множество наблюдений; $\mathcal{L}$ – функционал ошибки.

В качестве функционала ошибки могут использоваться среднеквадратичная ошибка, средняя абсолютная ошибка и другие метрики, определяемые постановкой вычислительного эксперимента. Такая постановка задачи служит основой для интеграции математических моделей, средств виртуальной среды и методов машинного обучения в составе умного цифрового двойника. Она позволяет перейти от общей модели цифрового двойника к построению вычислительных и интеллектуальных компонентов, обеспечивающих анализ, прогнозирование и эмуляцию данных.

При этом математическая модель (2.4) – (2.17) формирует взаимосвязанную систему зависимостей, описывающих преобразование входных параметров внешней среды $X(t)$ в выходные характеристики $Y(t)$ солнечной электростанции. Входной вектор включает основные факторы внешней среды, такие как солнечная радиация, температура воздуха и скорость ветра, а выходной вектор содержит ключевые эксплуатационные характеристики, включая мощность генерации, напряжение, ток и коэффициент полезного действия. Формулы (2.6)–(2.14) конкретизируют функцию $F(X(t), \Theta)$ и отражают основные физические процессы функционирования СЭС: расчет солнечной радиации с учетом геометрических факторов, влияние температуры на эффективность фотоэлектрического

преобразования, учет потерь и деградации, а также формирование электрических характеристик системы.

В свою очередь, соотношения (2.15) – (2.17) задают постановку задачи прогнозирования, позволяющую аппроксимировать зависимость между входными и выходными параметрами на основе данных. Это обеспечивает возможность применения методов машинного обучения в составе виртуальной среды цифрового двойника. Таким образом, предложенная модель объединяет физико-математическое описание процессов генерации электроэнергии и методы анализа данных, что позволяет использовать ее как основу для построения интеллектуальных компонентов умного цифрового двойника.

В рамках настоящего исследования делается акцент на прогнозировании выходных характеристик (генерируемой мощности) на основе данных виртуальной среды. Решение этой задачи осуществляется с использованием методов машинного обучения, рассматриваемых далее. Дальнейшее построение умного цифрового двойника связано с интеграцией математических и интеллектуальных технологий, обеспечивающих реализацию указанной постановки задачи.

2.3. Методика построения виртуальной среды

Взаимодействие цифрового двойника с реальным объектом предполагает отладку ЦД до его ввода в эксплуатацию (Рис. 2.3 и 2.4).

Автор использует следующие определения из ГОСТа 57721 – 2017 [83]:

Виртуальный эксперимент. Эксперимент, основанный на технологиях мультимедиа, эмуляции, виртуализации и виртуальной реальности и др., способный полностью или частично заменить аналогичный традиционный натурный эксперимент.

Виртуальная среда. Форма имитационного моделирования, использующая средства визуализации для формирования наглядных копий моделируемых систем.

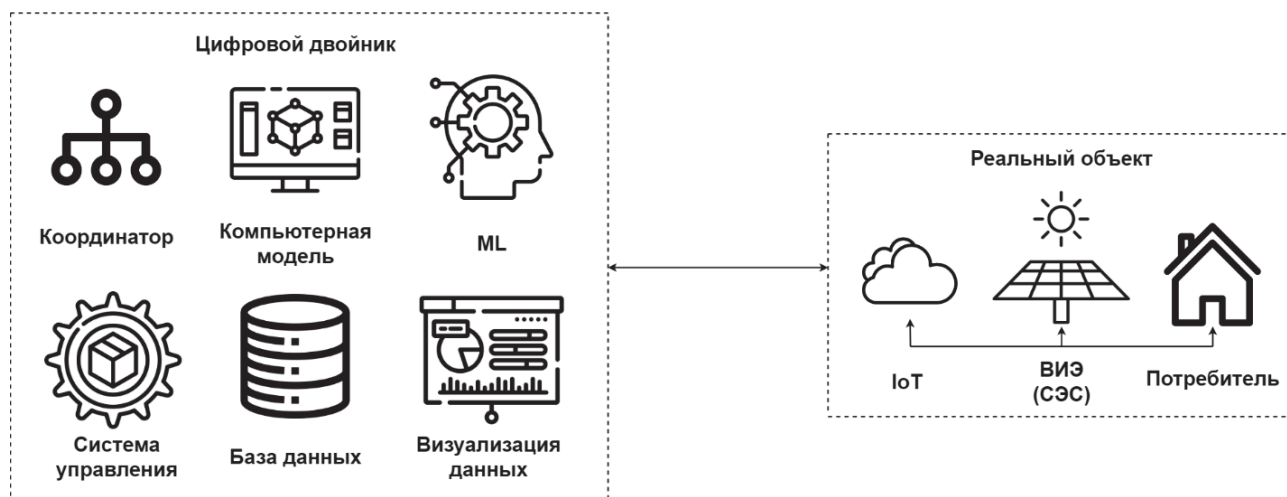


Рисунок 2.3 – Взаимодействие цифрового двойника
с реальным объектом

Виртуализация. Совокупность методов и средств представления числовой информации или физического явления в виде, удобном для зрительного наблюдения и анализа, а также результат такого представления.

Эмуляция. Комплекс программных, аппаратных средств или их сочетание, предназначенное для копирования функции одной вычислительной системы на другой, отличной от первой, вычислительной системе таким образом, чтобы эмулированное поведение как можно ближе соответствовало поведению оригинальной системы.

Примечание. Целью эмуляции является максимально точное воспроизведение поведения в отличие от разных форм моделирования, в которых имитируется поведение некоторой абстрактной модели. Например, моделирование физического процесса или явления не является эмуляцией.

Виртуальная реальность. Высокоразвитая форма виртуальной среды, обладающая высокой степенью достоверности визуализации, имитирующая как воздействие на изучаемый объект, так и реакции на это воздействие.

Автором предложено конкретизировать эти определения для цифровых двойников. **Виртуальная среда** – искусственная среда, созданная с использованием эмуляции и виртуализации, предназначенная для моделирования объектов и процессов реального мира. Реализация виртуальной среды происходит

с помощью программного обеспечения, которое управляет визуализацией и интерактивностью.

В условиях отсутствия нормативных документов, определяющих взаимодействие ЦД и реальных объектов, виртуальная среда используется для отладки научных прототипов цифровых двойников. Цифровой двойник может работать с виртуальной средой, получая симулированные физические данные из нее. Симуляция поведения позволяет проводить сценарное моделирование и анализ данных.

Выделены основные этапы методики построения виртуальной среды для умного цифрового двойника:

1. Построение онтологии для определения характеристик объекта.
2. Определение внешних характеристик объекта с помощью онтологий.
3. Выбор моделей машинного обучения с использованием предложенного алгоритма.
4. Интеграция моделей машинного обучения в виртуальную среду.
5. Использование виртуальной среды для эмуляции данных при отладке цифрового двойника.

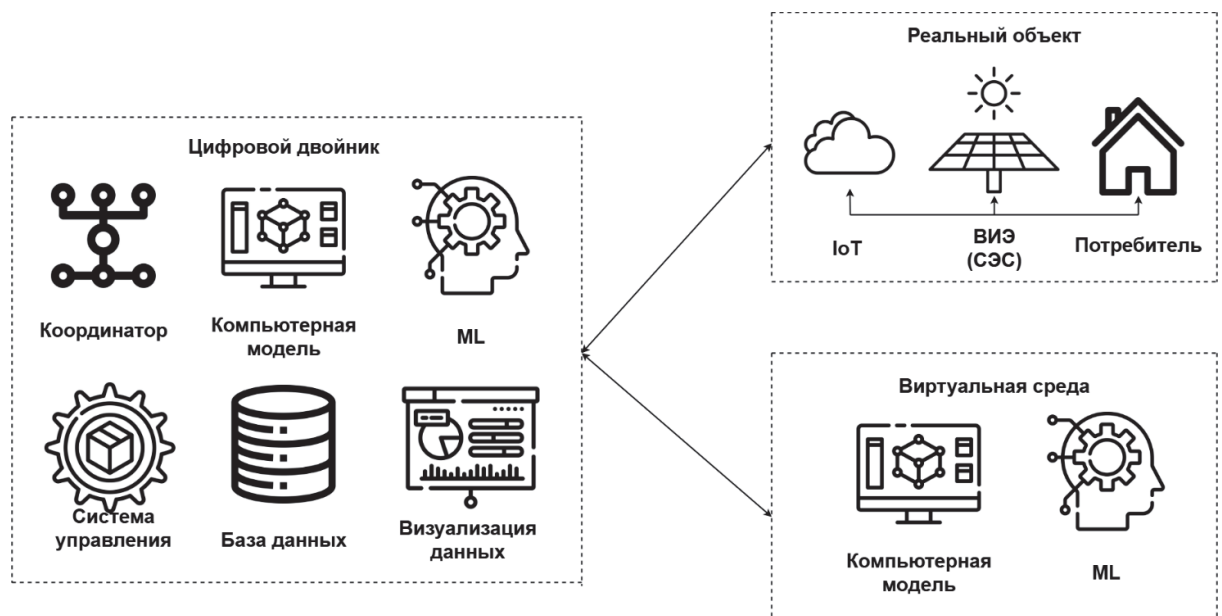


Рисунок 2.4 – Взаимодействие цифрового двойника с реальным объектом и виртуальной средой

На Рис. 2.4 показано взаимодействие цифрового двойника с реальным объектом (вверху справа) и виртуальной средой.

Для работы цифрового двойника необходимо наличие входных и выходных параметров, а также информации о взаимосвязи между всеми параметрами. Для формирования необходимых параметров для создания виртуальной среды предлагается использовать онтологический инжиниринг. Онтологический инжиниринг позволяет структурировать и систематизировать данные, обеспечивая четкое понимание и представление взаимосвязей между различными аспектами модели.

После определения структуры и целей виртуальной среды следующим шагом является определение параметров среды, необходимых для функционирования объекта. Эти характеристики делятся на внешние и внутренние, каждый из которых играет важную роль в моделировании и анализе поведения объекта.

Внутренние параметры представляют собой совокупность объектов и процессов, которые выполняют действия внутри системы. Они описывают состояние и поведение самого объекта, а также зависят от его внутренней структуры и характеристик. Благодаря тому, что природа внутренних параметров обычно связана с физическими законами и механизмами, их можно моделировать с использованием математических методов.

Внешние параметры представляют собой условия окружающей среды, в которых находится и функционирует объект. Они описывают факторы, которые влияют на поведение объекта, но не зависят от его внутренних процессов. В отличие от внутренних параметров, внешние параметры часто носят случайный характер, что делает их моделирование более сложным.

Для цифрового двойника солнечной электростанции требуется определить внешние параметры, которые будут эмулироваться, а также внутренние параметры и связь между ними. Для этого предлагается использовать онтологический инжиниринг, с помощью которого можно определить, какие данные нужно собрать и какие связи между ними важны.

Для визуального проектирования онтологий на примере фотоэлектрической системы был использован четырехступенчатый алгоритм, [73] описанный ниже:

Шаг 1. Цели, стратегия и идентификация границ:

- Определение цели онтологии и потребностей в приобретении знаний в предметной области.

В результате этого этапа была поставлена цель: построить систему онтологий солнечной электростанции и интегрировать её в уже имеющуюся онтологию ТЭК, а также приобрести знания в области солнечной электроэнергетики для построения цифрового двойника;

- Определить область действия или «границы» ограничения онтологии и тип онтологии прежде, чем перейти к шагу 2:

Разрабатываемая онтология относится к типу эвристических и будет использована для построения цифрового двойника.

Шаг 2. Разработка глоссария или определения мета – концепции: необходимо собрать всю информацию, относящуюся к описываемой области. Основная цель этого шага – отбор и вербализация всех основных объектов и понятий в области. Результат шага представлен ниже:

- Система генерации;
- Возобновляемые источники энергии;
- Фотоэлектрические системы (ФС);
- Оборудование ФС;
- Освещенность.

Шаг 3. Распределение по категориям, включая классификацию и спецификацию: Следующим этапом является определение всех основных объектов и концептов предметной области (основных уровней абстракции). Выявление структурированной иерархии является одной из основных целей на данном этапе. Иерархия, построенная с использованием результатов предыдущего шага, представлена ниже:

- электроэнергетические системы;

- системы генерации;
 - возобновляемые источники энергии;
 - фотоэлектрические системы (ФС);
 - оборудование ФС;
 - освещенность.

Шаг 4. Уточнение: последний шаг посвящен обновлению визуальной структуры путем исключения любых излишеств, синонимии и противоречий.

На этом шаге были переработаны концепты и связи между ними, что привело к улучшению понятности и читаемости системы онтологий.

В результате выполнения четырехступенчатого алгоритма автором была построена система онтологий солнечной электростанции, представленная на Рисунке 2.5 [69, 71].

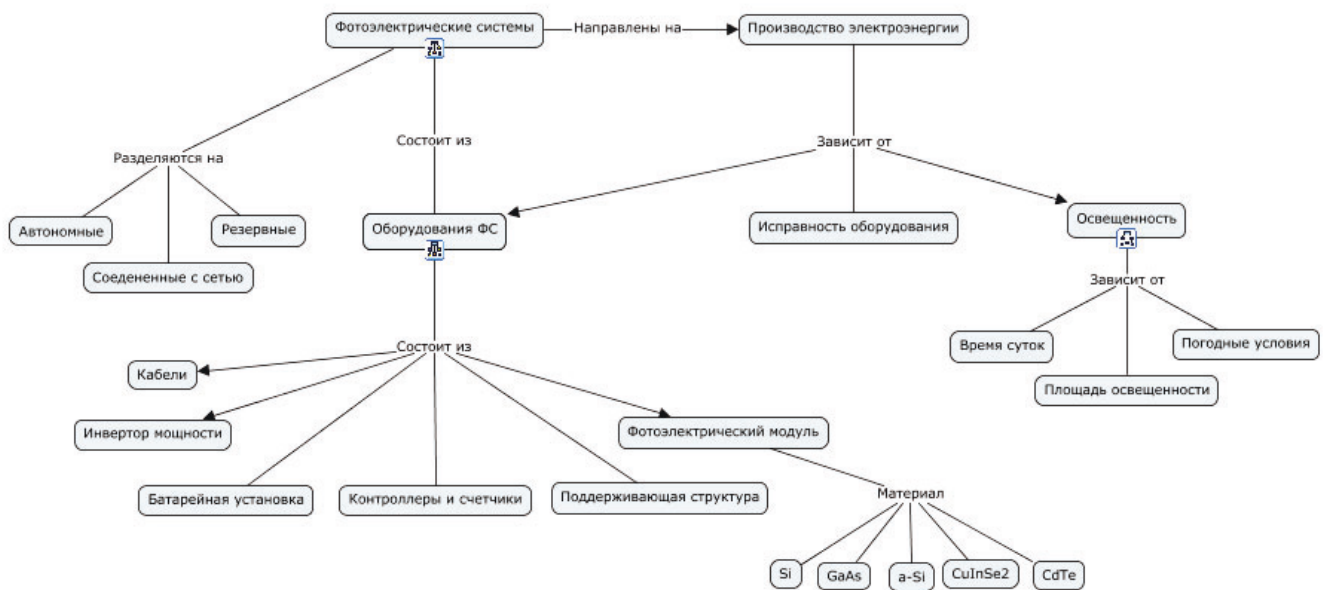


Рисунок 2.5 – Онтология солнечной электростанции

Далее, для последующего использования цифровым двойником, характеристики разделяются на внешние и внутренние. Внешними характеристиками в этом случае являются: потребление электроэнергии,

освещенность, давление, температура и т.д. Внутренними являются характеристики системы (модули, материалы и т.д.).

В современных исследованиях при разработке интеллектуальных моделей широко применяется генерация синтетических и аугментированных данных, что позволяет снизить требования к объему реальных наблюдений и повысить устойчивость моделей машинного обучения [84–86]. Для временных рядов используются методы статистической аугментации и генеративные модели, обеспечивающие сохранение основных закономерностей исходных данных. Применение таких подходов является особенно актуальным на этапе научного прототипирования цифровых двойников.

2.4. Интеграция методов машинного обучения в виртуальную среду цифрового двойника

Для формирования виртуальной среды цифрового двойника предлагается эмулировать данные внешних характеристики объекта цифрового двойника. Для эмуляции данных внешних характеристик в контексте работы цифрового двойника предлагается использовать методы машинного обучения. Эти методы способны эффективно обрабатывать и анализировать большие объемы данных, что позволяет создавать модели поведения системы.

Однако для применения машинного обучения требуется наличие качественных и актуальных данных, которые будут служить основой для функционирования виртуальной среды. Эти данные могут включать историческую информацию, характеристики входных и выходных параметров, а также данные о взаимодействиях между различными элементами системы.

Важно отметить, что эффективность машинного обучения во многом зависит от качества и разнообразия используемых данных. Поэтому необходимо обеспечить предварительную обработку, очистку и агрегирование данных, чтобы они были готовы к обучению моделей. Подходящее качество данных способствует созданию точных моделей, что, в свою очередь, улучшает общую

производительность цифрового двойника и его способность к адаптации в условиях динамических изменений.

В качестве основы для обоснования выбора метода машинного обучения использована методология CRISP-DM [87]. На ее основе разработан модифицированный алгоритм, ориентированный на задачи выбора и интеграции модели машинного обучения в виртуальную среду цифрового двойника солнечной электростанции (Рис. 2.6, 2.7).



Рисунок 2.6 – Основные этапы методологии CRISP-DM

Согласно CRISP-DM, аналитический проект состоит из шести основных этапов, выполняемых последовательно:

- Бизнес – анализ (Business understanding).
- Анализ данных (Data understanding).
- Подготовка данных (Data preparation).
- Моделирование (Modeling).
- Оценка результата (Evaluation).

- Внедрение (Deployment).

Если рассматривать эту методологию в рамках задач проектирования цифрового двойника для солнечной электростанции, то этапы модифицированного алгоритма (Рис. 2.7) выглядят так:

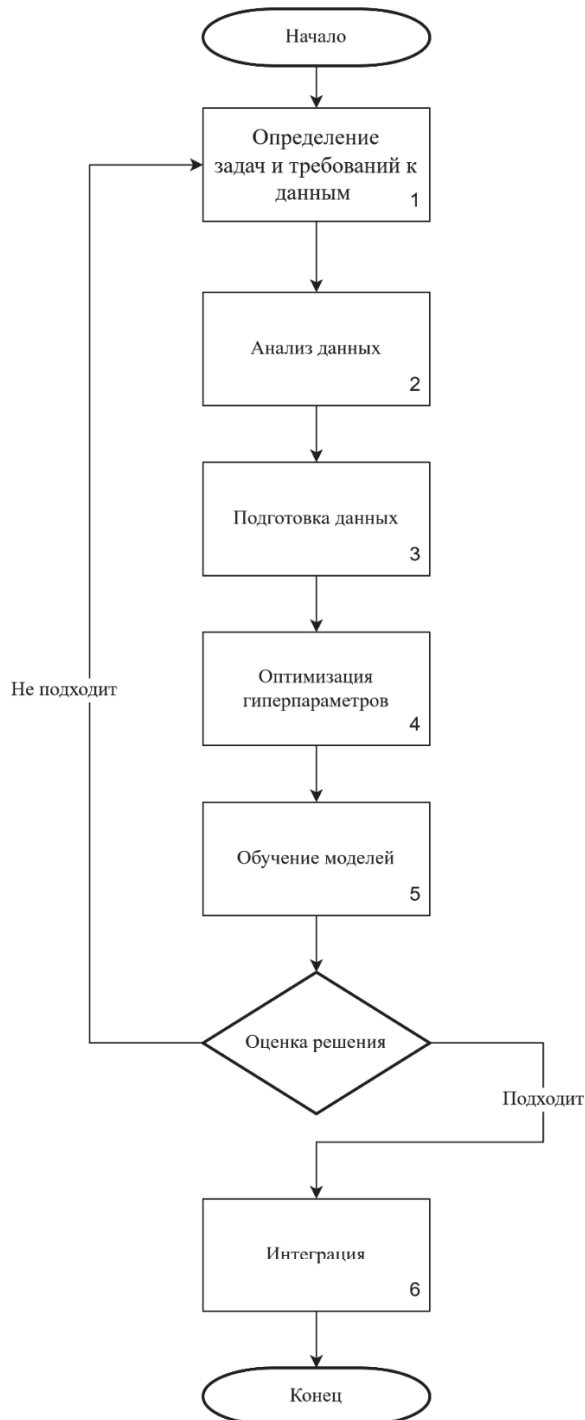


Рисунок 2.7 – Блок-схема модифицированного алгоритма модели машинного обучения для интеграции в виртуальную среду

Модификация базовой методологии CRISP-DM применительно к задачам построения виртуальной среды цифрового двойника солнечной электростанции заключается в выделении дополнительных элементов, обеспечивающих связь процедуры машинного обучения с архитектурой цифрового двойника. В отличие от типовой последовательности этапов анализа данных и построения модели, предложенный алгоритм включает: формализацию задач цифрового двойника и требований к данным, разведочный анализ данных для оценки их пригодности к моделированию, подготовку признакового пространства с учётом параметров внешней среды, оптимизацию гиперпараметров моделей, введение ограничений, направленных на снижение риска переобучения, сравнительную оценку моделей по совокупности метрик качества, интеграцию выбранной модели в виртуальную среду цифрового двойника. Таким образом, модернизированный алгоритм ориентирован не только на построение модели машинного обучения, но и на её последующее использование в составе цифрового двойника как вычислительного и прогнозного компонента.

Рассмотрим этапы алгоритма подробнее:

1. Определение задач и требований к данным – определение задач, которые решаются методом машинного обучения в виртуальной среде; формулирование требований к оценкам качества моделей МО, а также их приоритет; на задачи и требования могут влиять наличие данных для этого объекта, а также их количество.

Основными метриками оценки являются:

- Среднеквадратичная ошибка (RMSE). Измеряется средняя разница между прогнозируемыми и фактическими значениями. Основным преимуществом этой метрики является интерпретируемость в тех же значениях, что и прогнозируемых. Штраф при наличии большой единичной ошибки может сильно повлиять на оценку [88].

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (2.18)$$

- Средне-абсолютная ошибка (MAE). Измеряется средне-абсолютная разница между прогнозируемыми и фактическими значениями. Преимуществом по сравнению с предыдущей метрикой является также интерпретируемость, но, в отличие от RMSE, является менее чувствительной к выбросам. Она имеет приоритет, если в наборе данных присутствуют множественные выбросы, которые не сильно влияют на модель. Следует отметить, что функция абсолютной ошибки является недифференцируемой в точке нулевой ошибки. Однако в задачах машинного обучения это не является существенным ограничением, поскольку используются методы субградиентной оптимизации, позволяющие корректно выполнять обучение моделей. [41]

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (2.19)$$

- Средняя абсолютная процентная ошибка (MAPE). Измеряется средняя абсолютная разница между прогнозируемыми и фактическими значениями, выраженная в процентах от фактических значений. Преимущество MAPE заключается в её интерпретируемости в процентах, что делает её удобной для сравнения ошибок между разными наборами данных. Однако MAPE может быть нестабильной, если фактические значения близки к нулю, так как это приводит к делению на ноль или очень большим значениям ошибки [89]

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}. \quad (2.20)$$

- Коэффициент детерминации (R^2). Измеряется доля дисперсии целевой переменной, которая объясняется моделью. R^2 принимает значения от 0 до 1, где 1 означает, что модель идеально объясняет все изменения в данных, а 0 – что модель не объясняет ничего. Преимущество R^2 заключается в его универсальности для

сравнения моделей на разных наборах данных. Однако R^2 не учитывает масштаб данных и, может быть, введена в «заблуждение», если модель переобучена [90]

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (2.21)$$

Таблица 2.1 – Сравнение метрик оценки моделей

Метрика	Преимущества	Недостатки
RMSE	Чувствительна к большим ошибкам, интерпретируема в единицах целевой переменной.	Сильно штрафует выбросы.
MAE	Устойчива к выбросам, проста для интерпретации.	Не учитывает квадратичные ошибки, менее чувствительна к резким отклонениям.
MAPE	Выражается в процентах, удобна для сравнения моделей.	Не работает при нулевых значениях y_i , искажает ошибки для малых величин.
R^2	Показывает долю объяснённой дисперсии, нормирована на $[0, 1]$.	Может быть завышен при переобучении, не чувствительна к смещениям.

2. Анализ данных – анализ источников данных (например, данные датчиков, исторические записи, внешние базы данных), оценка качества данных (например, наличие пропусков, аномалий, шума); исследование временных зависимостей и сезонности в данных. Для выполнения оценки данных предлагается использовать разведочный анализ данных. Разведочный анализ данных (Exploratory data analysis, EDA) – это процесс исследования и анализа данных с целью выявления закономерностей, паттернов, аномалий и взаимосвязей между переменными. EDA включает в себя использование различных методов и инструментов для описательной статистики, визуализации данных и построения графиков, а также применения статистических тестов и моделей для проверки гипотез и извлечения информации из набора данных.

3. Подготовка данных – очистка, преобразование, агрегирование, выставление признаков, нормализация данных и т.д. На основе выполнения анализа данных выполняются эти шаги при необходимости.

4. Оптимизация гиперпараметров – для обучения сети ставятся задачи оптимизации гиперпараметров и введении дополнительных штрафов и ограничений для избежания переобучения модели [91].

Постановка задачи оптимизации для рекуррентных нейронных сетей:

Минимизация целевой функции $f(\lambda)$ на пространстве гиперпараметров Λ :

$$\lambda^* = \underset{\lambda \in \Lambda}{\operatorname{argmin}} f(\lambda), \quad (2.22)$$

где λ – вектор гиперпараметров; Λ – пространство гиперпараметров; $f(\lambda)$ – целевая функция.

Предлагаемые методы решения:

- Поиск по сетке (Grid Search) – систематическое переборное исследование определенного пространства параметров с целью нахождения наилучшего набора значений, максимизирующих или минимизирующих заданную функцию [92]:

$$\lambda^* = \underset{\lambda \in \mathcal{G}}{\operatorname{argmin}} f(\lambda), \quad (2.23)$$

где $\mathcal{G} \subset \Lambda$ – дискретная сетка значений.

- Случайный поиск (Random Search) – метод оптимизации, который выбирает случайные комбинации параметров из заданного пространства параметров и оценивает их, чтобы найти оптимальное решение [92]:

$$\lambda = \underset{\lambda \sim p(\lambda)}{\operatorname{argmin}} f(\lambda), \quad (2.24)$$

где $p(\lambda)$ – распределение на Λ .

- Байесовская оптимизация – метод оптимизации на основе вероятностной модели, который особенно эффективен для оптимизации сложных, дорогих и шумных целевых функций. В отличие от традиционных методов оптимизации, которые полагаются на информацию о градиенте, байесовская оптимизация создает суррогатную модель целевой функции, обычно используя

гауссовские процессы, для принятия обоснованных решений о том, где делать следующую выборку. Функция приобретения для следующего λ [93]:

$$\lambda_{t+1} = \underset{\lambda \in \Lambda}{\operatorname{argmax}} EI(\lambda), \quad (2.25)$$

где $EI(\lambda)$ – функция ожидаемого улучшения.

Целевая функция для нахождения гиперпараметров LSTM, при используемой метрики RMSE:

$$f(\lambda) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i(\lambda))^2}, \quad (2.26)$$

где $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, λ_1 – число слоев; λ_2 – число нейронов в слое; λ_3 – коэффициент dropout; λ_4 – размер пакета (batch size); $\hat{y}_i(\lambda)$ – предсказание с гиперпараметрами λ .

5. Обучение модели – этап, на котором выполняется обучение выбранных моделей, которое включает введение дополнительных штрафов и ограничений для избежания переобучения моделей.

6. Оценка решения – этап, где выполняется оценка итоговых моделей на основе метрик RMSE, MAE и R2.

7. Интеграция – этап, где выполняется внедрение выбранной модели машинного обучения в виртуальную среду для последующего использования в реальных условиях.

2.5. Численный метод ансамблирования с выбором моделей на основе состояния внешней среды (EnvSwitch)

Анализ литературных источников показывает, что использование одной модели прогнозирования в условиях нестационарных временных рядов не обеспечивает устойчивого качества расчетов. В связи с этим в ряде исследований предлагаются методы ансамблирования и динамического выбора моделей, позволяющие адаптировать процесс прогнозирования к текущему состоянию среды [50, 94–96]. Однако применение подобных подходов в составе цифровых

двойников энергетических объектов, ориентированных на интеграцию математических и интеллектуальных моделей, остается недостаточно проработанным.

При проектировании умного цифрового двойника солнечной электростанции (СЭС) возникает задача прогнозирования выработки электроэнергии в условиях высокой нестационарности внешней среды – изменчивости солнечной радиации, температуры, облачности и других метеорологических факторов. Как показано в разделе 2.3, применение единой модели машинного обучения (например, LSTM, GRU или XGBoost) приводит к снижению точности прогноза в переходных режимах, поскольку такие модели обучаются на усреднённых закономерностях и плохо адаптируются к локальным погодным паттернам.

Для решения этой проблемы автором разработан численный метод ансамблирования моделей с выбором на основе состояния внешней среды (EnvSwitch, Environment-based Switching). Метод реализует принцип адаптивного прогнозирования за счёт комбинирования разнородных моделей машинного обучения в зависимости от текущего режима функционирования внешней среды.

Формальная постановка задачи заключается в следующем. Пусть на входе имеется вектор признаков:

$$x_t = (G_t, T_t, v_t, \dots)^T \in \mathbb{R}^d, \quad (2.27)$$

где G_t – суммарная солнечная радиация (Вт/м²), T_t – температура воздуха (°C), v_t – скорость ветра (м/с) и другие внешние характеристики в момент времени t .

Требуется предсказать целевую переменную – выработку мощность генерации СЭС:

$$y_t = P_{\text{выр}}(t). \quad (2.28)$$

Метод EnvSwitch предполагает использование ансамбля $\mathcal{M} = \{M_1, M_2, \dots, M_K\}$ разнородных моделей машинного обучения:

- рекуррентные нейронные сети (LSTM, GRU), которые для захвата долгосрочных временных зависимостей;

- градиентный бустинг (XGBoost), которые для высокой точности на малых и средних выборках;
- на основе статистических методов.

Каждая модель, входящая в ансамбль $\mathcal{M} = \{M_1, \dots, M_K\}$, проходит индивидуальную процедуру оптимизации гиперпараметров в рамках этапа подготовки (см. раздел 2.3). Для рекуррентных сетей (LSTM, GRU) оптимизируются: количество слоёв, число нейронов, dropout-коэффициенты, размер батча. Для XGBoost оптимизируется: глубина деревьев, скорость обучения, число деревьев и т.д. Оптимизация выполняется отдельно для каждого режима среды $s \in S$ или на полном датасете, в зависимости от объёма данных. Каждая модель M_k специализируется на определённом режиме среды $s \in S$, где S – конечное множество режимов, таких как:

- «ясное утро»,
- «резкое затенение»,
- «зимняя ночь»,
- «облачная стабильная погода» и др.

Режимы определяются двумя способами:

1. С использованием онтологий – на основе экспертных правил, сформулированных в рамках онтологии СЭС (см. Рис. 2.5). Например: *Если $G_t < 10 \text{ Вт/м}^2$ и $T_t < 0^\circ\text{C}$, то $s = \text{«зимняя ночь»}$.*

2. Автоматически – через кластеризацию (k-means) нормализованных признаков внешней среды.

На этапе подготовки для каждого режима s оценивается качество каждой модели M_k по метрике RMSE на соответствующей подвыборке данных:

$$RMSE_{sk} = \sqrt{\frac{1}{|I_s|} \sum_{i \in I_s} (y_i - M_k(x_i))^2}, \quad (2.29)$$

где I_s – индексы наблюдений, принадлежащих режиму s , $|I_s|$ – их количество.

На их основе строится матрица весов $W = [w_{sk}]$, где вес обратно пропорционален ошибке:

$$w_{sk} = \frac{1}{RMSE_{sk} + \varepsilon}, \quad (2.30)$$

где, ε – малая константа (например, 10^{-6}) для избежания деления на ноль.

Такой подход обеспечивает плавный переход между моделями, повышает общую устойчивость прогноза и позволяет эффективно работать как в стационарных, так и в резко меняющихся погодных условиях.

2.6. Методика построения умного цифрового двойника

Автор принимал активное участие в разработке типовой методики построения цифровых двойников, которая представлена в отчете отдела по госзаданию ИСЭМ СО РАН [97].

Предлагаемая методика основывается на том, что при разработке интеллектуальных систем управления развитием и функционированием систем энергетики на основе цифровых двойников и цифровых образов целесообразно выделять две стадии проекта:

1. Разработка научных прототипов цифровых двойников с использованием научного инструментария исследований в энергетике с отладкой информационного взаимодействия в рамках ИТ-инфраструктуры системных исследований в энергетике, с использованием виртуальной среды для эмуляции данных для отладки ЦД.

2. Апробация прототипов ЦД и отладка информационных взаимодействий с реальными информационными потоками для решения конкретных задач; разработка технологии перехода от научно-исследовательских прототипов ЦД к их применению на практике (созданию прототипов интеллектуальных систем управления развитием и функционированием систем энергетики).

Типовая методика построения ЦД и умных ЦД с использованием научного инструментария разработана, как обобщение и детализация ранее предложенных методик [73, 98].

Для целей диссертационной работы автором была выполнена конкретизация типовой методики для разработки ЦД СЭС. При этом из типовой методики исключены этапы, связанные с проектированием хранилища данных и знаний, а также с интеграцией умных цифровых двойников электроэнергетических систем на уровне ТЭК, поскольку они выходят за рамки задач настоящего исследования. Конкретизированная методика включает этапы анализа применимости методов машинного обучения, разработки архитектуры умного ЦД СЭС, интеграции средств искусственного интеллекта и компонентов цифрового двойника.

Конкретизированная методика включает следующие этапы:

1. Онтологический инжиниринг предметной области (соответствующей энергосистемы) и построение онтологической модели.

- 1.1. Определение основных концептов предметной области и отношений между ними.

- 1.2. Разработка метаонтологии предметной области, определяющей систему онтологий, на основе фрактального подхода к структурированию знаний.

- 1.3. Конкретизация необходимых математических и информационных моделей, определяемых онтологической моделью.

2. Разработка цифровой модели физического (энергетического) объекта или системы.

- 2.1. Анализ существующих математических моделей и реализующих их компьютерных программ (программных комплексов).

- 2.2. Модификация, при необходимости, математических моделей и реализация цифровой модели физического объекта на основе модифицированной математической модели.

3. Реализация цифровой тени физического объекта.

- 3.1. Определение исходных данных или потоков данных (состав, источники получения, возможность получения оперативных данных, базы данных и др.) и их взаимодействия с математическими моделями.

- 3.2. Построение модели данных на основе проведенного онтологического инжиниринга.

3.3. Определение способов эмуляции данных для отладки ЦД, разработка виртуальной среды для эмуляции данных (на первой стадии проекта)

3.4. Подключение реальных потоков данных с использованием Интернета вещей (на второй стадии проекта)

4. Разработка цифровых двойников энергетических объектов.

4.1. Разработка архитектуры ЦД (конкретизация типовой архитектуры).

4.2. Разработка, с использованием прошедших реинжиниринг цифровых программ, web-приложений и web – сервисов для реализации цифровых двойников на основе агентно-сервисного подхода.

4.3. Реализация прототипов цифровых двойников энергетических объектов – интеграция цифровой модели и цифровой тени в составе ЦД.

4.4. Отладка ЦД на основе данных, полученных из виртуальной среды.

5. Разработка умных цифровых двойников.

5.1. Анализ возможностей интеграции в умный ЦД интеллектуальных технологий и средств искусственного интеллекта (семантического моделирования).

5.2. Анализ возможности и необходимости применения методов машинного обучения при разработке ЦД.

5.3. Разработка архитектуры умного ЦД и интеграция средств ИИ в умный ЦД.

2.7. Проектирование цифрового двойника солнечной электростанции

При проектировании применяется методология функционального моделирования IDEF0, позволяющая исследовать структуру, параметры и характеристики систем [99].

Для создания функциональной модели используется технология IDEF0, отображающая функции системы, в которых связываются материальные объекты и потоки информации, используя потоки информации, а также их структуру.



Рисунок 2.8 – Проектирование цифрового двойника солнечной электростанции

На Рисунке 2.8 отображен процесс проектирования цифрового двойника солнечной электростанции. На вход в процесс подаются (стрелки слева):

- Погодные данные: содержат информацию об солнечной радиации (прямой, рассеянной и суммарной), температуре и других климатических факторах, влияющий на работу СЭС.
- Данные по потреблению электроэнергии: содержат информацию о нагрузке об электропотреблении.
- Технические характеристики: содержат параметры солнечных панелей, инверторов, систем накопления и другого оборудования.

Регулирование процессов происходит с помощью механизмов управления (стрелки вверх):

- Регуляторные требования и законы: представляют из себя законодательные нормы, стандарты и регламенты, которые необходимо учитывать при проектировании.
- Бизнес-цели: определяют экономические показатели, такие, как минимизация затрат, максимизации эффективности и т.д.

- Технические требования к проектированию: учет спецификации оборудования, местности и ограничения.

Механизмами исполнения (стрелки вниз) являются:

- Конструктор сервисов ЦД СЭС: конструирует инструменты для моделирования, визуализации, прогнозирования и эмуляции данных.
- Эксперт: выполняет анализ данных и формирует требования к ЦД СЭС.
- Разработчик: выполняет реализацию прототипов.

Результатами этого процесса являются (стрелки справа):

- Рекомендации по проектированию ЦД СЭС: включают предложения по выбору методов реализации компонентов ЦД СЭС, которые формируются на основе данных и прогнозов.
- Прогноз по выработке электроэнергии: составляется на основе выбранной модели и реализованной виртуальной среды.

На Рисунке 2.9 Процесс проектирования ЦД СЭС декомпозируется на 4 процесса, а именно: «Сбор данных», «Обработка данных», «Эмуляция и прогнозирование данных» и «Формирование состава сервисов для проектирования ЦД СЭС». На всех процессах механизмы исполнения: «Эксперт», «Разработчик» и отдельные компоненты «Конструктора сервисов ЦД СЭС»

Процесс «Сбор данных» представляет собой получение исторических данных с использованием датчиков, систем мониторинга, а также исторических источников. Выходом этого процесса являются наборы данных (датасеты) для последующего процесса «Обработка данных» (Рис.2.10). Механизмом исполнения является виртуальная среда.

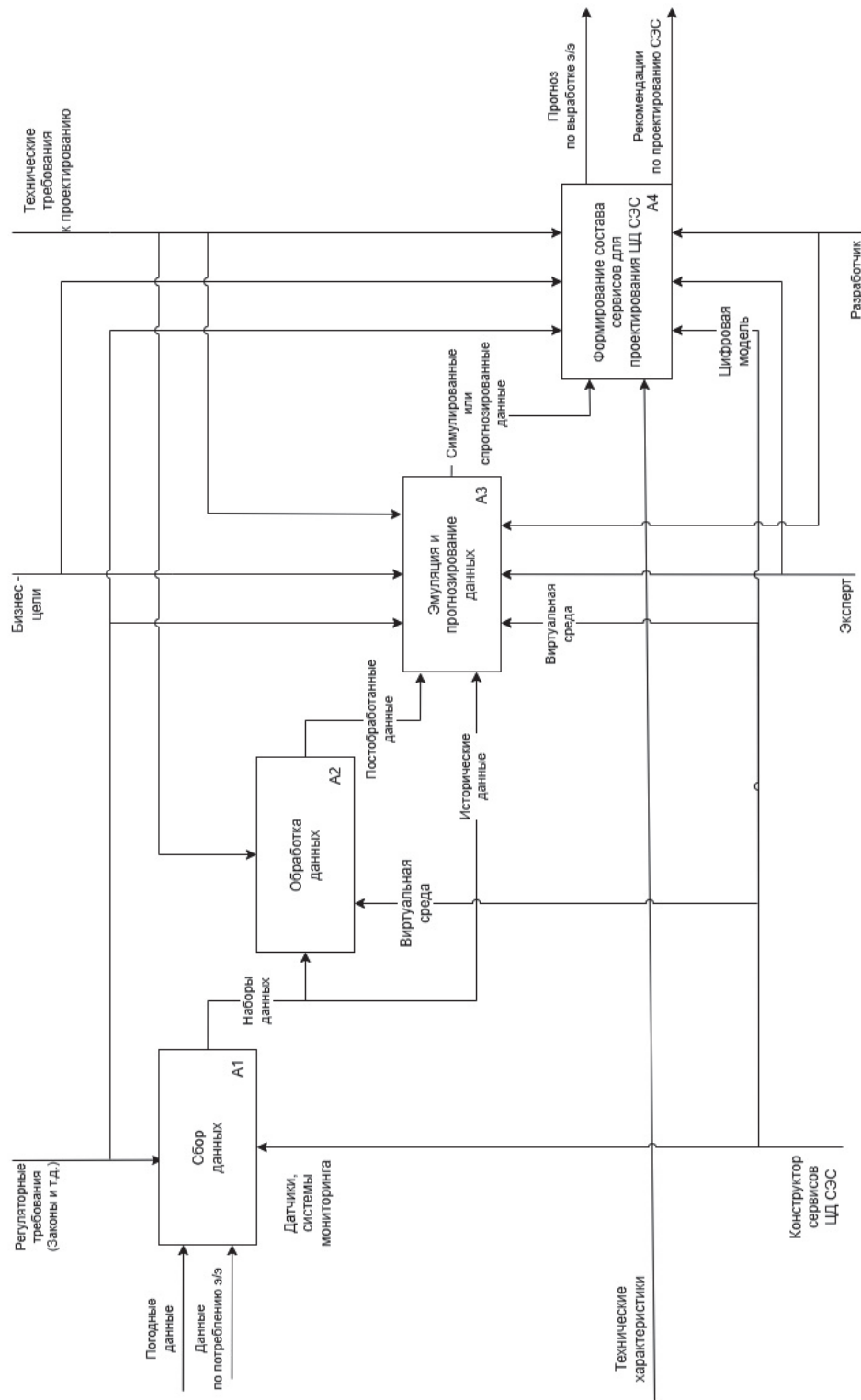


Рисунок 2.9 – Декомпозиция³ процесса «Проектирование ЦД СЭС»

³ Под декомпозицией в нотации IDEF0 понимается разделение моделируемой функции или процесс на функции – компоненты или процессы [99]

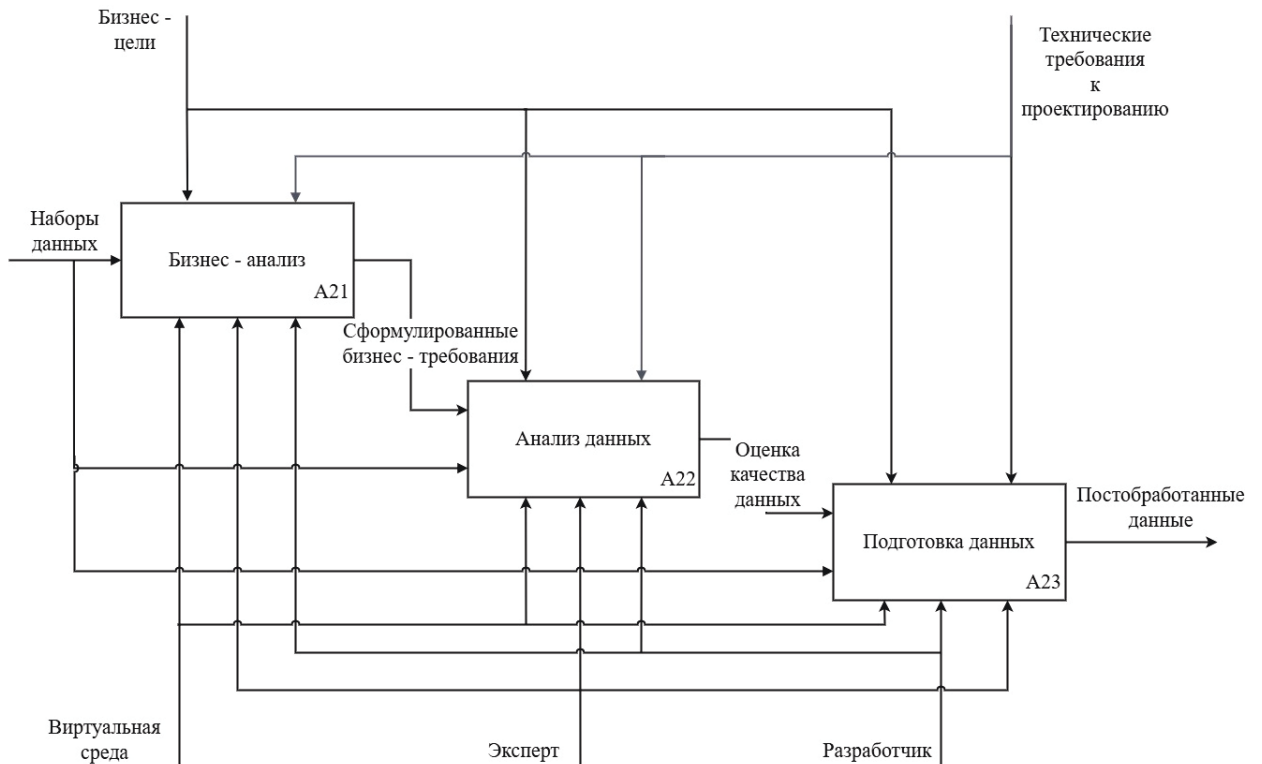


Рисунок 2.10 – Декомпозиция процесса «Обработка данных»

На Рисунке 2.10 процесс «Обработка данных» декомпозируется на 3 процесса: «Бизнес – анализ», «Анализ данных» и «Подготовка данных». Обработка данных в этом процессе является частью алгоритма, описанного в разделе 2.3. После процесса «Бизнес-анализ» процессу «Анализ данных» отправляются сформулированные бизнес-требования и наборы данных к остальным процессам. «Анализ данных» – на основе бизнес-требований выдается оценка качества данных. Процесс «Подготовка данных» включает очистку, преобразование, агрегирование, выставление признаков, нормализацию данных и т.д. на основе выставленных требований и оценки. Выходом этого процесса являются постобработанные данные, которые отправляются в процесс «Эмуляции и прогнозирования данных». Декомпозиция процесса «Эмуляция и прогнозирование данных» представлена на Рисунке 2.11 и включает процессы: «Моделирование», «Оценка решения» и «Внедрение».

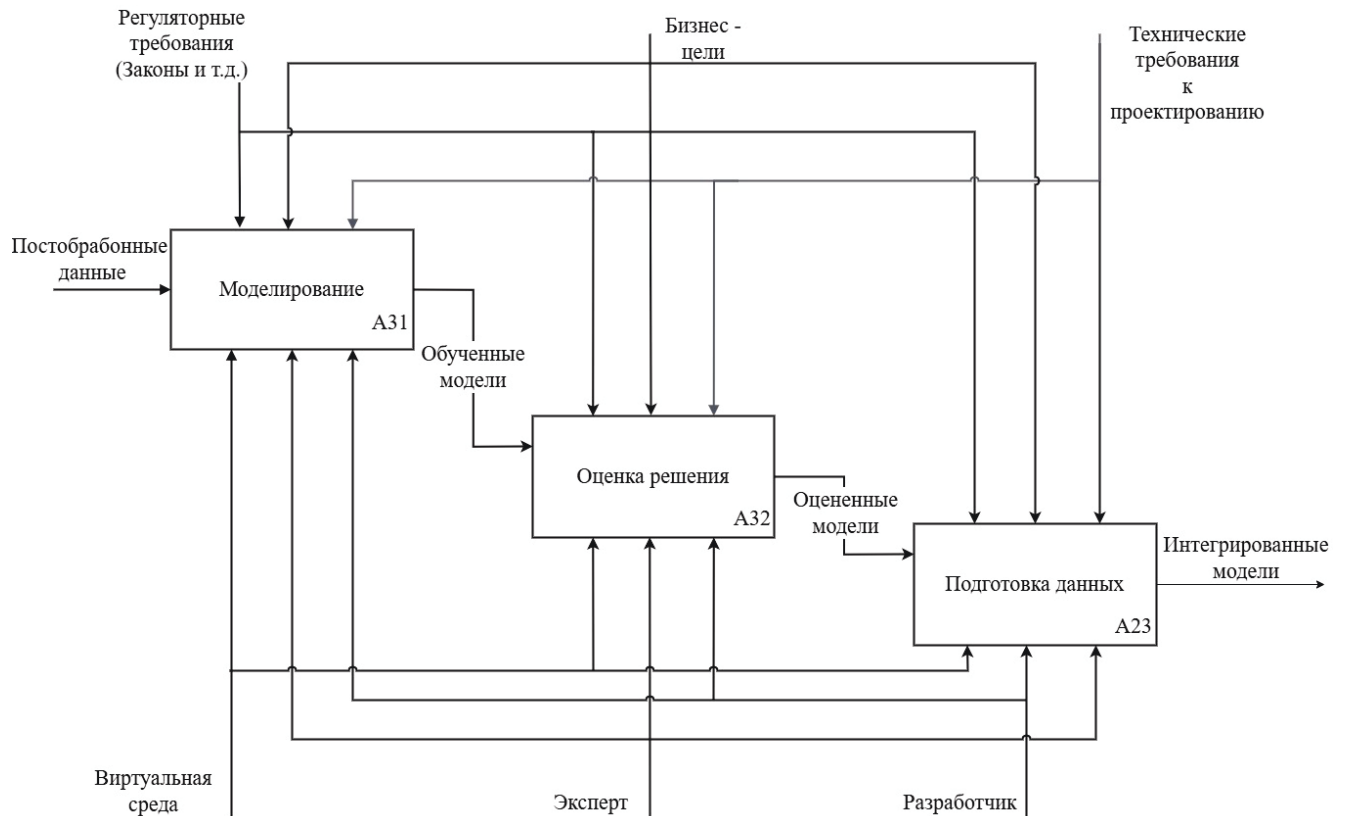


Рисунок 2.11 – Декомпозиция процесса «Эмуляция и прогнозирования данных»

Эти процессы являются следующей, описанной в разделе 2.3. Процесс «Моделирование» получает постобработанные данные, на которых происходит процесс обучения моделей машинного обучения с корректировкой гиперпараметров, архитектуры модели и т.д. В процессе «Оценка решения» оцениваются модели на основе метрик. Оцененные модели передаются в процесс «Внедрение», где происходит подготовка к интеграции и выводу моделей.

На Рисунке 2.12 процесс «Формирование состава сервисов для ЦД СЭС» декомпозируется на 3 подпроцесса: «Анализ технических характеристик», «Сборка сервисов ЦД СЭС» и «Рекомендации по проектированию».

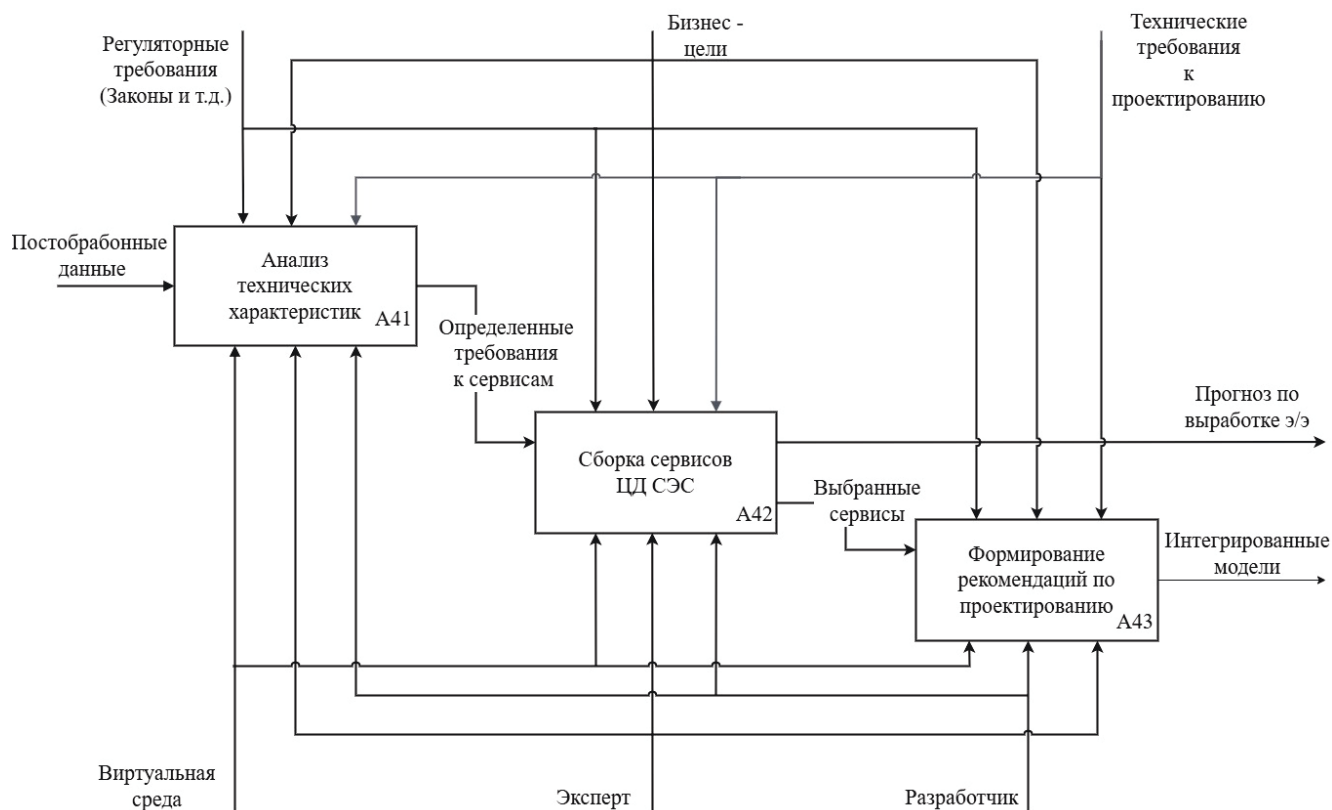


Рисунок 2.12 – Декомпозиция процесса «Формирование состава сервисов для ЦД СЭС»

Интегрированные модели передаются процессу «Анализ технических характеристик», где для процесса «Сборка сервисов ЦД СЭС» формируются технические требования к сервисам, таким, как: производительность, совместимость, аппаратные требования и т.д. После процесса «Сборка сервисов ЦД СЭС» выбранные и модифицированные сервисы передаются процессу «Формирование рекомендаций по проектированию». На основе построенных сервисов возможно сформировать прогнозы по выработке электроэнергии, где основными участниками являются виртуальная среда и компьютерная модель.

2.8. Выводы по главе 2

Во второй главе диссертации автором предложена модель умного цифрового двойника, которая является модификацией модели ЦД и включает базы знаний, виртуальную среду, экспертные системы, семантические модели и методы

машинного обучения. Конкретизировано понятие виртуальной среды в рамках цифровых двойников. Использован онтологический инжиниринг для определения внешних параметров моделирования для виртуальной среды.

Предложена агрегированная модель солнечной электростанции, как основа вычислительного ядра умного цифрового двойника. Солнечная электростанция рассмотрена как динамическая система, функционирование которой определяется вектором входных параметров среды и параметрами модели. Определены входные воздействия, выходные характеристики и целевые переменные моделирования. Показано, что задача моделирования состоит в построении модели, обеспечивающей отображение входных параметров в целевую переменную при минимизации функционала ошибки. Такая постановка служит основой для дальнейшей интеграции математических моделей, виртуальной среды и методов машинного обучения.

В рамках настоящей работы подробное моделирование внутренних характеристик отдельного фотоэлектрического преобразователя не является основной целью исследования. Задача состоит в построении цифрового двойника солнечной электростанции и виртуальной среды, обеспечивающей расчёт и прогнозирование характеристик СЭС на основе доступных параметров внешней среды. Поэтому для включения в состав цифрового двойника используется агрегированный уровень математического описания СЭС, связывающий параметры солнечной радиации, температуры, оборудования и потерь с выходными характеристиками станции. Такой уровень детализации является достаточным для задач проектирования, сценарного моделирования и интеграции с моделями машинного обучения, тогда как однодиодные и двухдиодные модели целесообразны при детальном исследовании электрических характеристик отдельного фотоэлектрического модуля.

Интеграция виртуальной среды с цифровым двойником позволяет проводить сценарное моделирование и тестирование до внедрения в реальных условиях. Это снижает риски и затраты, связанные с проектированием солнечных электростанций.

На основе методологии CRISP-DM разработан модифицированный алгоритм интеграции методов машинного обучения в виртуальную среду цифрового двойника. Алгоритм реализует полный цикл обработки данных, включая этап разведочного анализа для первичной подготовки данных.

Проведённый анализ методов прогнозирования и моделирования процессов выработки электроэнергии показал, что применение единственной модели расчёта не обеспечивает качественного прогнозов во всех режимах функционирования солнечной электростанции. Существенное влияние сезонных и метеорологических факторов, а также нестационарный характер временных рядов обуславливают необходимость адаптивного подхода к выбору модели прогнозирования.

В связи с этим во второй главе предложен численный метод, основанный на оценке текущего состояния внешней среды и характеристик входных данных и реализованный в виде алгоритма адаптивного переключения моделей прогнозирования EnvSwitch, целью которого является повышение точности прогнозирования за счёт динамического выбора модели расчёта.

Применение методологии IDEF0 позволило структурировать процесс проектирования, выделив ключевые входные данные (погодные условия, потребление энергии, технические характеристики), механизмы управления (регуляторные требования, бизнес – цели) и результаты (рекомендации по проектированию, прогноз выработки). Декомпозиция процессов на этапы сбора, обработки, эмуляции данных и выбора сервисов отображает его структуру и подход к интеграции цифрового двойника СЭС.

Предложенная автором конкретизированная для ЦД СЭС методика построения умных цифровых двойников включает все этапы, от онтологического анализа предметной области до интеграции моделей ИИ в виртуальную среду, что обеспечивает комплексное проектирование умного цифрового двойника.

Глава 3. Разработка и применение цифрового двойника солнечной электростанции

В третьей главе рассматривается практическая реализация предложенного методического подхода на примере цифрового двойника солнечной электростанции. Описывается агентно-сервисная архитектура умного цифрового двойника, реализация компонентов цифровой тени и цифровой модели, а также подготовка данных для виртуальной среды. Приводится вычислительный эксперимент, направленный на эмуляцию данных и проверку применения предложенного численного метода ансамблирования моделей машинного обучения. Полученные результаты позволяют оценить работоспособность разработанных моделей, алгоритмов и программных компонентов в составе научного прототипа умного цифрового двойника солнечной электростанции. Результаты отражены в публикациях автора [100–105].

3.1. Архитектура умного цифрового двойника

Для построения архитектуры цифрового двойника предлагается использовать агентно-сервисный подход. Этот подход является авторской интеграцией агентного подхода к проектированию и программированию и сервис-ориентированной архитектуры.

Агент – это автономная сущность, способная воспринимать окружающую среду, интерпретировать данные и воздействовать на неё, вызывая изменения. Агенты могут функционировать как независимо, так и в кооперации с пользователями или другими агентами. В вычислительном контексте агенты представляют собой программные единицы, поддерживающие локальные состояния, параллельные вычисления и способные обмениваться данными с другими агентами для достижения общих целей [106].

Ключевые характеристики агентов [107]:

- Среда: пространство, в котором агент функционирует и с которым взаимодействует.
- Восприятие: способность агента получать данные о среде через сенсоры.
- Интерпретация: анализ полученных данных для принятия решений.
- Действие: воздействие на среду через эффекторы для достижения целей

Когда несколько агентов взаимодействуют в одной среде, они образуют мультиагентную систему. В таких системах агенты действуют параллельно, а контроль и обработка данных логически распределены между ними. Для эффективного решения задач агенты должны уметь:

- Общаться: обмениваться информацией и данными.
- Координироваться: согласовывать действия для достижения общих целей.
- Вести переговоры: разрешать конфликты и находить компромиссы.

Преимущества многоагентных систем:

- Распределенные вычисления: задачи выполняются параллельно, что повышает эффективность вычислений.
- Масштабируемость: система может легко расширяться за счет добавления новых агентов.
- Автономность: агенты способны функционировать независимо, что повышает надежность системы.

Понятие сервисов и сервисной архитектуры взяты из области разработки программного обеспечения. Сервисно – ориентированная архитектура (SOA) – это подход к проектированию программного обеспечения, который направлен на создание функциональных, масштабируемых программных систем из отдельных компонентов, называемых сервисами. Каждый сервис представляет собой автономный компонент, который выполняет определенную бизнес – функцию и предоставляет доступ к своим возможностям через стандартизированные

интерфейсы. В SOA особое внимание уделяется модульности, повторному использованию и совместимости – когда предприятия разбивают сложные приложения на более мелкие и управляемые блоки, в результате повышается гибкость и масштабируемость. SOA широко используется в современных распределенных системах для повышения гибкости, масштабируемости и удобства поддержки [108].

Компания Amazon формулирует четыре фундаментальных принципа сервис-ориентированной архитектуры (SOA), которые обеспечивают эффективность и масштабируемость распределённых систем [109]:

- **Обеспечение совместимости.** Каждый сервис сопровождается документацией, описывающей его функциональность и условия использования. Это позволяет клиентским системам взаимодействовать с сервисом независимо от используемой платформы или языка программирования, обеспечивая кроссплатформенную интеграцию.
- **Слабая взаимозависимость.** Сервисы проектируются как слабосвязанные компоненты с минимальной зависимостью от внешних ресурсов (моделей данных, других систем). Они не сохраняют состояние между вызовами (stateless), что обеспечивает изолированность изменений – модификация одного сервиса не требует корректировки зависимых компонентов
- **Абстрагирование.** Принцип «чёрного ящика» гарантирует, что потребители сервисов взаимодействуют только через строго определённые интерфейсы, без доступа к внутренней логике. Вся необходимая информация предоставляется через сервисные контракты и документацию.
- **Степень детализации.** Каждый сервис инкапсулирует одну законченную бизнес-функцию, сохраняя баланс между атомарностью и полезностью. Для сложных операций сервисы комбинируются, что обеспечивает гибкость без потери модульности.

На Рисунке 3.1 иллюстрируется онтологией общая задача проектирования цифрового двойника фотоэлектрической системы [100, 101]. На основе этой

структуры и модели цифрового двойника (согласно разделу «Методика построения умного цифрового двойника») предлагается типовая агентно-сервисная архитектура цифрового двойника (Рис. 3.2) [102, 103]. Агенты в этом случае будут выступать, как самостоятельные программы, которые объединяются в программный комплекс. Выделены три группы агентов – сервисов: Сервер, Клиент, Оборудование.

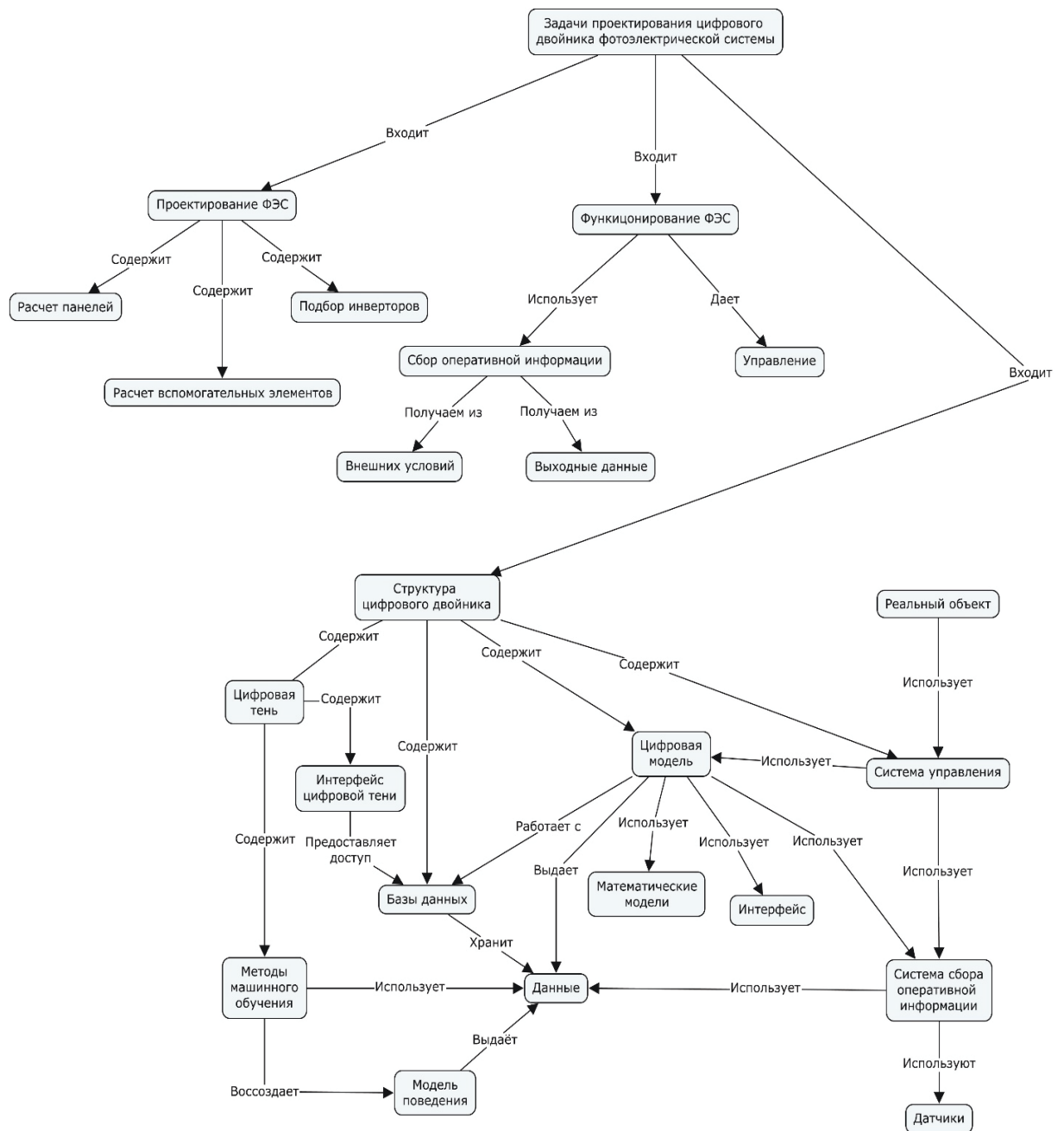


Рисунок 3.1 – Онтология задач и структуры цифрового двойника

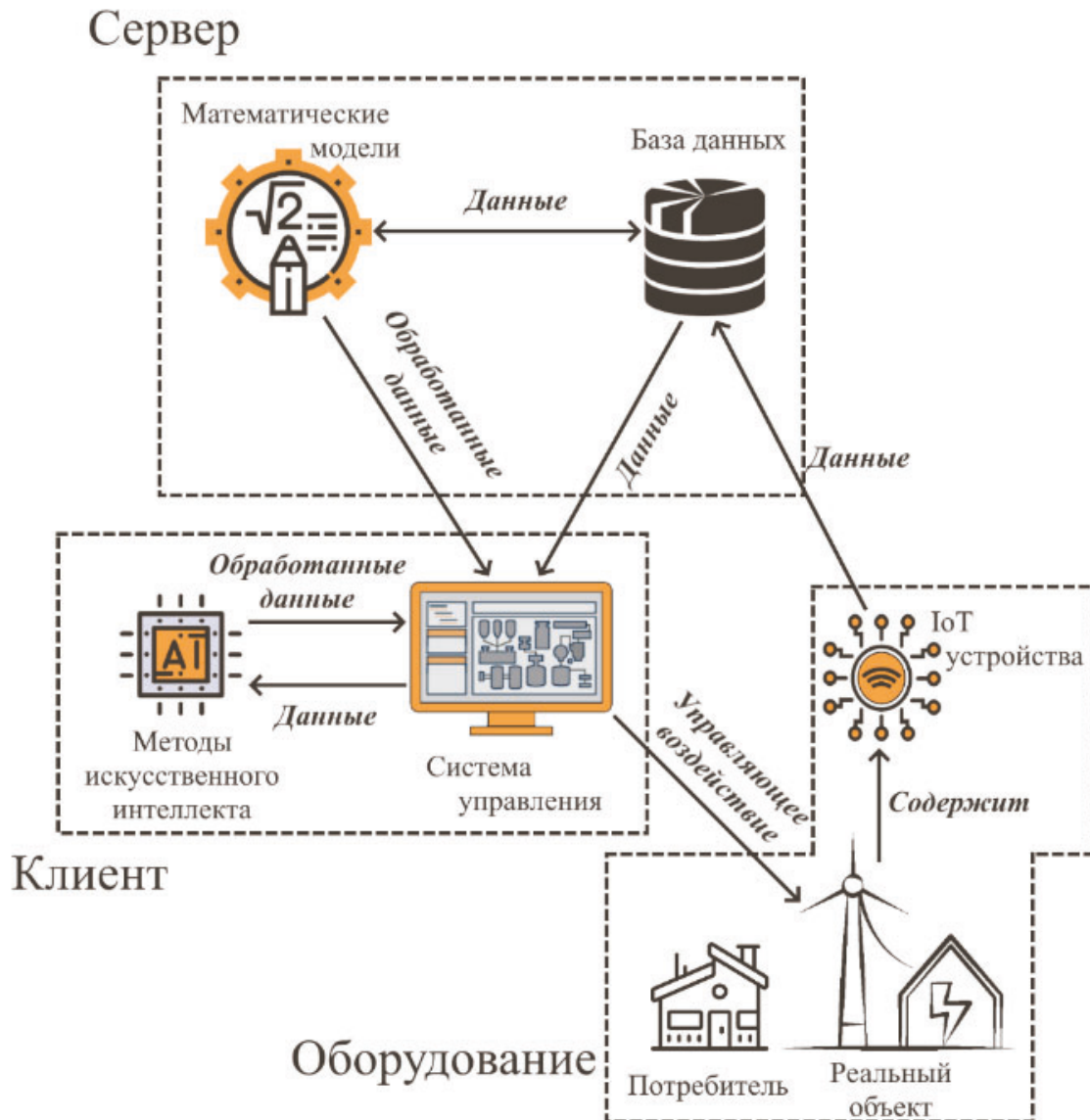


Рисунок 3.2 – Типовая агентно-сервисная архитектура цифрового двойника

В группу «Сервер» входят агенты: сервисы математических моделей (на основе математических моделей реализуются «решатели» в виде web-сервисов) и сервисы управления данными (обеспечивается web-доступ к базам данных).

Агент-решатель является цифровой моделью объекта, которая использует входные данные для формирования новых обработанных данных для цифрового двойника. Агент управления данными содержит обращения к СУБД и БД объекта и системы.

Группа «Клиент» включает агентов: «координатор» и «предсказатель». Для взаимодействия групп агентов друг с другом и передачи промежуточных

результатов требуется наличие Агента-координатора, который входит в систему управления. Его главная задача заключается в формировании сценариев для остальных агентов и связи между агентами. Клиент включает также методы искусственного интеллекта, которые реализованы в составе Агента-предсказателя. Агент-предсказатель реализует различные методы искусственного интеллекта, в том числе методы машинного обучения, которые могут обрабатывать данные и выполнять задачи для формирования сценариев (например, классификацию режимов работы, предсказание погодных характеристик и т.д.).

Группа «Оборудование» реализуется агентами – сервисами сбора данных и управления объектом. Агент управления объектом после получения сценария из Агента-координатора формирует управляющие воздействия на объекты. Агент сбора данных напрямую связан с агентом управления объектом, с реальным объектом и пользователем. Этот агент позволяет при помощи устройств Интернета вещей передавать данные в СУБД. Такой подход позволяет рассмотреть задачу комплексно без конкретизации технологий для реализации программного комплекса Цифрового двойника.

Автором диссертации выполнена конкретизация агентно-сервисной архитектуры цифрового двойника солнечной электростанции, представленная на Рисунке 3.3. Конкретизация заключается в выделении состава агентов, определения их функций и связей между ними.

Два основных агента на Рисунке 3.3: “Агент-координатор” и “Агент сценария”. Задача Агента-координатора: на основе предоставленного сценария взаимодействия агентов скоординировать работу остальных агентов.

Агент информационного обмена (на основе агентов-сервисов управления данными) реализует различные типы эмуляторов Scada–систем для эмуляции данных и организации информационного обмена между объектами реального мира (датчиками) и остальными агентами.

Агент математических моделей (решатель) осуществляет функцию расчетов необходимых характеристик при функционировании объектов.

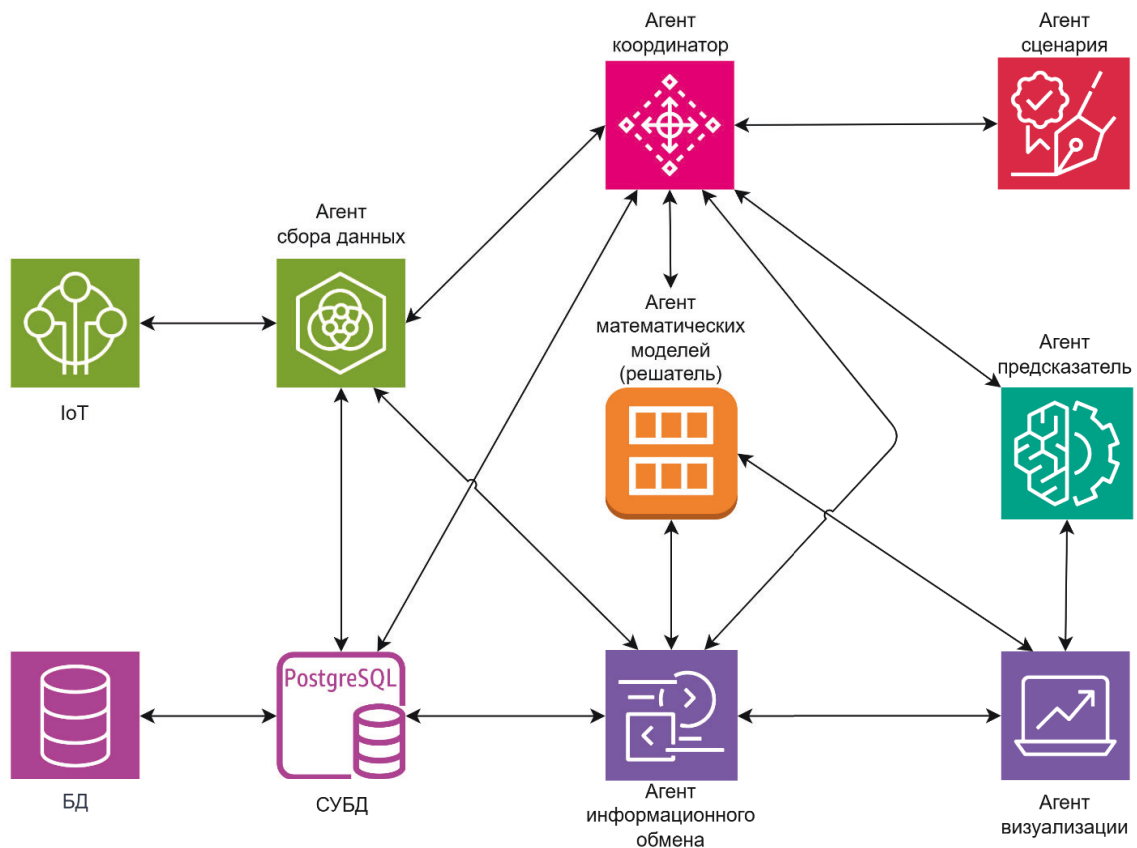


Рисунок 3.3 – Конкретизированная агентно-сервисная архитектура цифрового двойника солнечной электростанции

Агент-предсказатель решает задачи прогнозирования различных характеристик при функционировании объектов; данные, полученные при прогнозировании, используются при проектировании объектов и обосновании их экономической эффективности. Этот тип агентов может использоваться для разного набора задач, которые должны решать цифровые двойники, в том числе, задач предиктивной аналитики.

Агент визуализации выполняет формирование пользовательских интерфейсов и включает инструменты для визуальной аналитики.

3.2. Реализация компонентов цифровой тени и цифровой модели

Разработанное приложение построено на основе клиент–серверной архитектуры. Серверная часть реализована с использованием фреймворка Django и

отвечает за хранение данных, логику и выполнение расчётов. Клиентская часть представляет собой web-интерфейс, обеспечивающий интерактивную работу пользователя с цифровым двойником через браузер. Разработанное решение ориентировано на исследование эксплуатационных характеристик солнечных электростанций и поддерживает проведение расчетов в различных конфигурациях оборудования и климатических условиях.

Цифровой двойник реализован, как совокупность взаимосвязанных компонентов: информационной модели данных, вычислительного ядра [110, 111], блока интеграции метеорологических данных и пользовательского интерфейса. Такой подход обеспечивает разделение ответственности между хранением данных, расчетами и визуализацией, что повышает расширяемость и воспроизводимость результатов. Расширяемость обеспечивается модульной структурой программного комплекса, допускающей подключение дополнительных функциональных компонентов и поддержку других видов ВИЭ [112, 113].

Пользовательский интерфейс web-приложения предназначен для формирования цифровой тени солнечной электростанции и задания исходных параметров моделирования. На Рисунке 3.4 представлен интерфейс расширенного расчета, позволяющий задать основные параметры системы, включая тип и количество солнечных панелей, конфигурацию инвертора и аккумуляторных батарей, а также географические координаты объекта.

Для учета влияния окружающей среды предусмотрен ввод или автоматическое получение погодных и климатических параметров, включая солнечную радиацию, температуру воздуха и скорость ветра. Данные могут поступать как от пользователя, так и из внешних метеорологических источников, что позволяет моделировать работу станции в реальных условиях эксплуатации.

Основные параметры системы

Солнечная панель *

Canadian Solar Canac ▾

Выберите модель панели

Количество панелей *

2

Инвертор

Fronius Fronius Primc ▾

Опционально

Аккумулятор

BYD BYD Battery-Box ▾

Для автономных систем

Количество аккумуляторов

1

Угол наклона панелей

35° (Оптимальный) ▾

Оптимальный угол равен широте местности

География и погодные условия

Широта *

51,5119

Пример: 55.7558 (Москва)

Долгота *

104,1324

Пример: 37.6173 (Москва)

Инсоляция *

850 Вт/м²

Высокая (июнь)

Получить данные по местоположению

Температура воздуха

12 °C

Стандартные условия: 25°C

Скорость ветра

2,0 м/с

Влияет на охлаждение панелей

Коэффициент затенения

0,97

1,0 = нет затенения, 0,8 = 20% потеря

Дополнительные параметры

Температурный коэффициент

0,004 %/°C

Обычно 0,004-0,006 для кремниевых панелей

Годовая деградация

0,005 %/год

Обычно 0,5% в год

☒ Учитывать деградацию панелей

☒ Учитывать сезонные вариации

Рассчитать с использованием формул

Результаты расчета

Мгновенная мощность

921,34 Вт

при текущих условиях

Годовая выработка электроэнергии

763,68 кВт·ч

Панели:

2 × 555,0Вт

Пиковые часы:

4,5 ч/день

Эффективность:

19,4%

Темп. панели:

31,96°C

Автономность:

3,6 дней

Емкость АКБ:

10,24 кВт·ч

Среднесуточно:

2,58 кВт·ч

Среднемесячно:

63,64 кВт·ч

Экологический эффект

Экономия CO₂: 382,0 кг/год

Эквивалентно посадке 5 деревьев

Коэффициенты расчета:

Инсоляция × 0,85

КПД системы × 19,400

Детальная аналитика

Месячная выработка энергии

Январь	Февраль	Март	Апрель	Май	Июнь	Июль	Август	Сентябрь	Октябрь	Ноябрь	Декабрь
10	20	40	80	100	110	115	100	60	30	20	10

Прогноз выработки на 5 лет

Год	1	2	3	4	5
Производство (кВт·ч/год)	750	755	760	765	770

Год	Мощность	Выработка	Деградация
1	1,1 кВт	759,86 кВт·ч	0,5%
2	1,1 кВт	756,04 кВт·ч	1,0%
3	1,09 кВт	752,22 кВт·ч	1,5%
4	1,09 кВт	748,41 кВт·ч	2,0%
5	1,08 кВт	744,59 кВт·ч	2,5%

Использованные формулы

Формула 1.1:
P_{pv(t)} = G_n(t) × η × A
Базовая мощность панели

Формула 1.2:
P(t) = P₀ × (1 - γ × t)
Деградация мощности

Формула 1.4:
η(t) = η × [1 - β(T_{pv(t)} - 25)]
Температурная эффективность

Формула 1.5:
T_{pv(t)} = T_a(t) + G_r(t) / (k₀ + k₁ × V_m(t))
Температура панели

Рисунок 3.4 – Интерфейс расчетов выработки электроэнергии солнечной электростанции



☰

Редактирование станции "DEMO: Байкальск 15 кВт проектный вариант"

📄 Основная информация

Название станции

DEMO: Байкальск 15

Местоположение

Байкальск, Иркутск

Дата установки

дд.мм.гггг

Примечания

Средняя проектная конфигурация для сравнения

Активна ☒

☰ Выбранные компоненты
 7

Jinko Solar PV
модуль mono half-cell 550 Вт

✕

Количество:

-

28

+

Ориентация:

Юг

Мощность: 15400.0 Вт
 Стоимость: 406 000,00 руб.

Huawei
Сетевой инвертор 15 кВт

✕

Количество:

-

1

+

Ориентация:

Юг

Мощность: 15000.0 Вт
 Стоимость: 165 000,00 руб.

Pylontech
LiFePO4 аккумулятор 5.12 кВт·ч

✕

Количество:

-

4

+

Ориентация:

Юг

Мощность: 0.0 Вт
 Стоимость: 500 000,00 руб.

🔍 Доступные компоненты

Фильтр по категории

Все категории

Поиск компонента

Поиск по названию...

Аккумулятор

BYD
BYD Battery-Box

✕

180000,00 руб.

LG Energy Solution
LG RESU

✕

4500,00 руб.

Pylontech
LiFePO4 аккумулятор 5.12 кВт·ч

✕

125000,00 руб.

Добавлен

Рисунок 3.5 – Интерфейс конфигурации состава оборудования солнечной электростанции

После формирования компонентов солнечной электростанции и задания параметров среды выполняется расчет энергетических характеристик солнечной электростанции. Результаты моделирования включают мгновенную мощность, суточную, месячную и годовую выработку электроэнергии, а также прогноз изменения характеристик системы с учетом деградации оборудования.

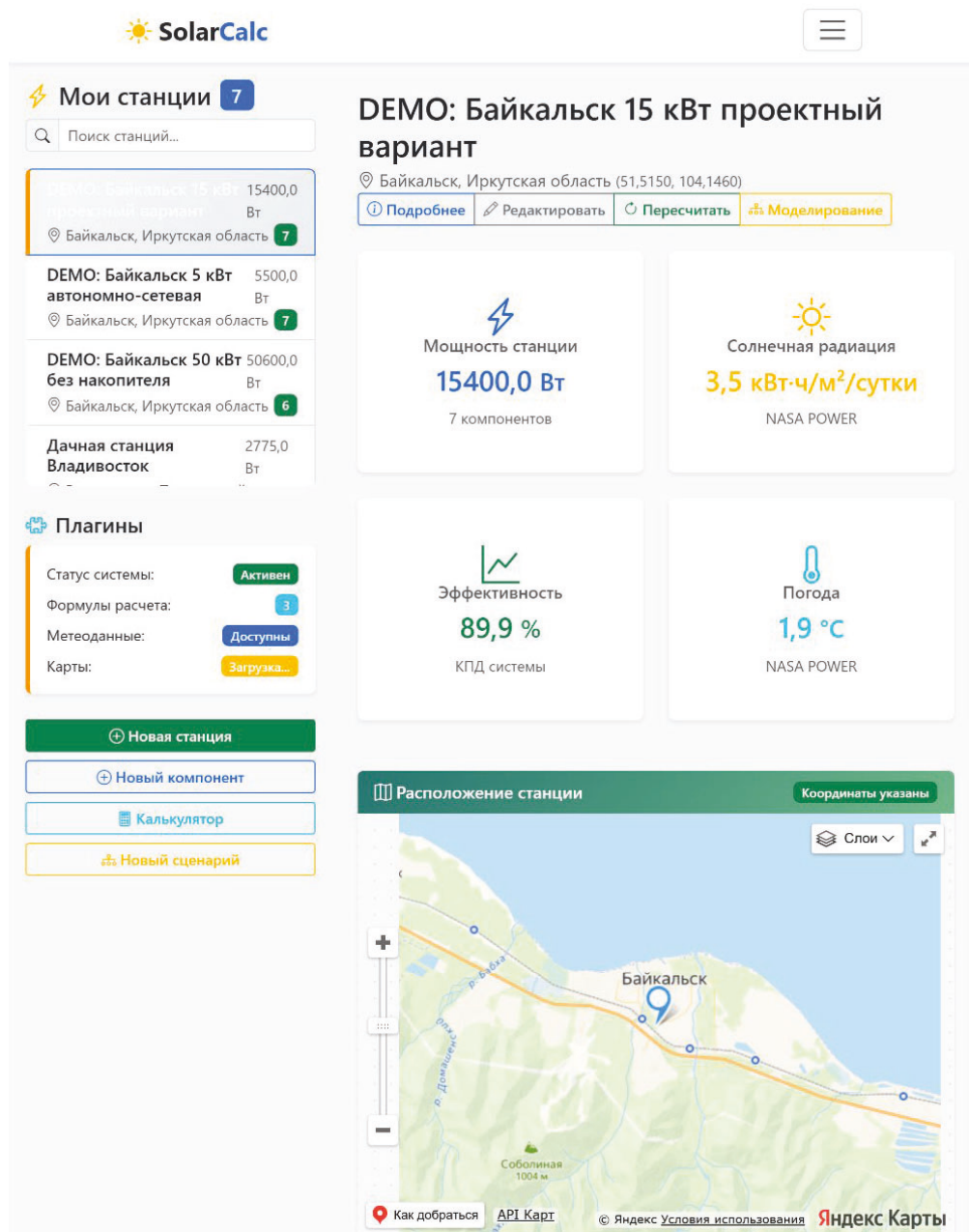


Рисунок 3.6 – Визуализация результатов моделирования цифрового двойника

На Рисунке 3.6 представлен пример визуализации результатов моделирования. Интерфейс отображает ключевые показатели эффективности системы, включая мгновенную мощность, годовую выработку и показатели энергетической эффективности. Дополнительно представлены графики месячной выработки электроэнергии и прогноз изменения выработки на горизонте нескольких лет.

Визуализация временных рядов позволяет наглядно оценить влияние сезонных факторов на работу солнечной электростанции. Существенная неравномерность месячной выработки обусловлена географической широтой объекта и изменением продолжительности светового дня в течение года. Таким образом, пользователь получает возможность анализа как текущего состояния системы, так и долгосрочных сценариев ее эксплуатации.

Для формализации структуры цифровой тени и обеспечения согласованности данных была разработана информационная модель цифрового двойника солнечной электростанции, представленная на Рисунке 3.7. Модель отражает основные сущности предметной области и их взаимосвязи, необходимые для хранения конфигурации станции, исходных данных и результатов моделирования [114].

Центральной сущностью модели является солнечная электростанция (*Solar_Station*), связанная с конкретной географической локацией (*Location*), определяющей координаты размещения объекта. Такое разделение позволяет учитывать пространственные факторы и использовать метеорологические данные, привязанные к заданным координатам.

Состав оборудования электростанции описывается через сущности *Component* и *Station_Component*. Сущность *Component* хранит параметры типовых элементов системы, включая номинальную мощность, коэффициент полезного действия и экономические характеристики. Таблица *Station_Component* реализует связь между станцией и компонентами и содержит информацию о количестве элементов, ориентации и угле наклона солнечных панелей. Данный подход обеспечивает гибкость конфигурации и возможность моделирования различных вариантов компоновки оборудования.

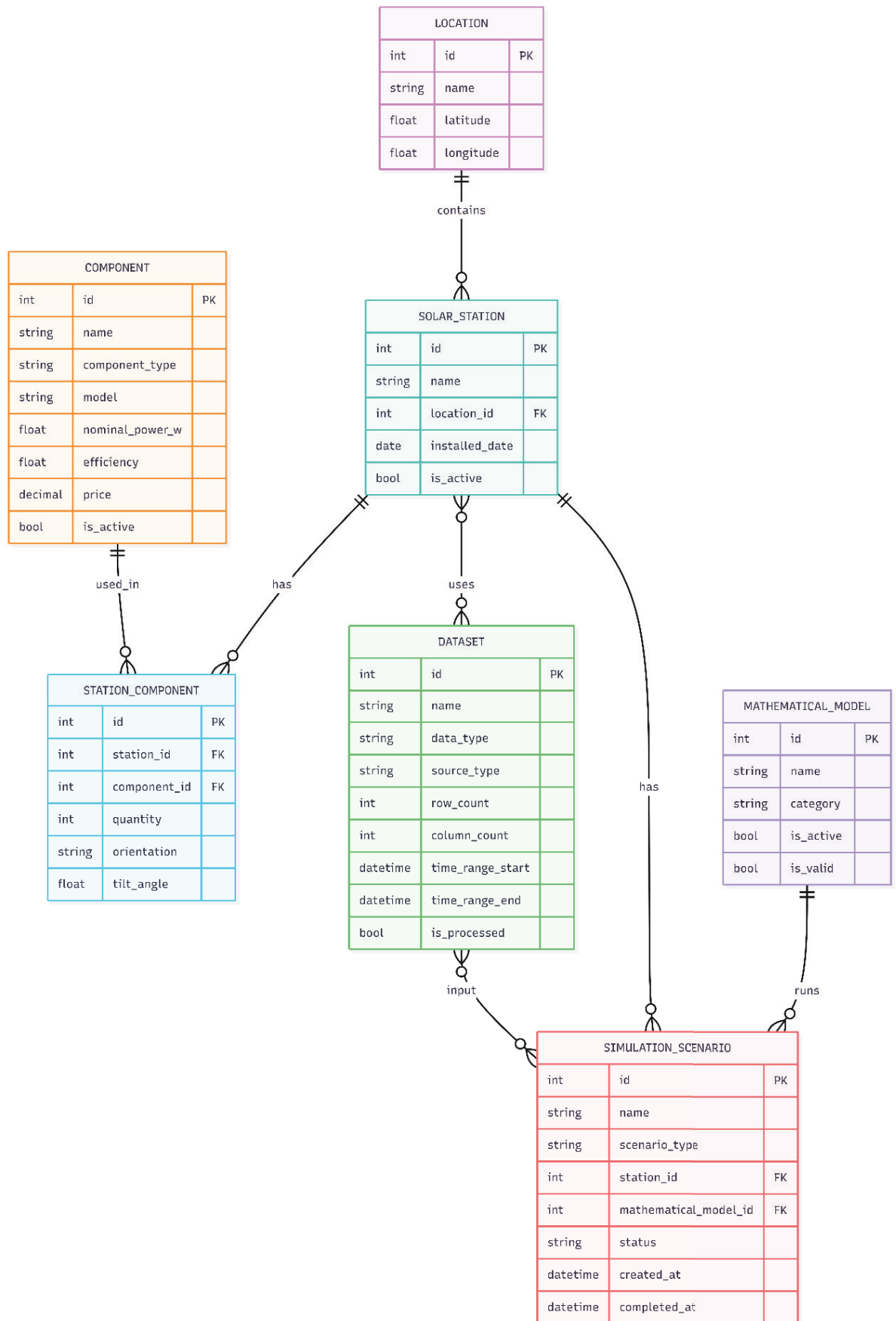


Рисунок 3.7 – Модель данных цифрового двойника солнечной электростанции

Для хранения исходных данных и временных рядов используется сущность *Dataset*, предназначенная для представления метеорологических и экспериментальных данных. Атрибуты временного диапазона и характеристик набора данных позволяют использовать эти данные в различных расчетных сценариях и повторно применять их при проведении вычислительных экспериментов.

Компьютерные модели представлены сущностью *Mathematical_Model*, описывающей используемые математические зависимости и категории расчетов. Запуск моделирования осуществляется в рамках сущности *Simulation_Scenario*, связывающей конкретную солнечную электростанцию, выбранную математическую модель и входные наборы данных. Такой подход обеспечивает воспроизводимость расчетов и возможность анализа различных сценариев эксплуатации.

Разработанный научный прототип ядра цифрового двойника солнечной электростанции обеспечивает целостное представление объекта за счет объединения цифровой тени, данных о внешней среде и компьютерных моделей. Использование web-приложения в качестве платформы реализации позволяет выполнять интерактивный анализ энергетических характеристик системы и проводить вычислительные эксперименты без привязки к специализированному программному обеспечению.

Предложенная архитектура и информационная модель создают основу для расширения функциональности вычислительного ядра цифрового двойника, включая подключение дополнительных моделей расчета и источников данных. В рамках настоящей работы акцент сделан на реализации научного прототипа базового ядра цифрового двойника, достаточного для анализа работы солнечных электростанций и оценки сценариев их эксплуатации в различных климатических условиях.

3.3. Анализ и подготовка данных для виртуальной среды цифрового двойника солнечной электростанции

В качестве примера далее рассматривается задача прогнозирования солнечной радиации, решаемая в составе виртуальной среды цифрового двойника солнечной электростанции.

Для построения модели прогноза погодных характеристик необходимо описать используемые данные (датасет). Данные представляют собой погодные характеристики, которые необходимы для расчета выходных характеристик солнечной электростанции. Конкретный датасет содержит данные о погодных характеристиках в г. Байкальск в размере 88889 записей, сделанных с интервалом в 1 час. Этот датасет содержит столбцы: скорость ветра, температура воздуха, атмосферное давление, прямая солнечная радиация, рассеянная солнечная радиация и суммарная солнечная радиация. Датасет содержит только количественные переменные.

Данные Показано 50 из 88889 строк								Действия
#	N integer	P float	T float	V integer	SumRad float	DirectRad float	ScatterRad float	
1	1	729,4	16,2	2	0,0	0,0	0,0	 
2	2	729,4	16,2	2	0,0	0,0	0,0	 
3	3	729,4	16,2	2	0,0	0,0	0,0	 
4	4	729,4	16,2	2	0,0	0,0	0,0	 
5	5	729,4	16,2	2	0,0	0,0	0,0	 
6	6	729,4	16,2	2	0,0	0,0	0,0	 
7	7	729,4	16,2	2	0,0	0,0	0,0	 
8	8	729,4	16,2	2	0,0	0,0	0,0	 
9	9	729,4	16,2	2	0,0	0,0	0,0	 
10	10	729,4	16,2	2	0,0	0,0	0,0	 
11	11	729,4	16,2	2	64,21	0,43	62,78	 
12	12	729,4	16,2	2	127,21	8,97	117,23	 
13	13	729,4	16,2	2	222,37	45,49	175,87	 
14	14	729,4	16,2	2	345,17	116,91	227,26	 

Страница 1 из 1778 1 2 3 4 » » Всего записей: 88889

Рисунок 3.8 – Часть набора данных погодных характеристик

Последующая реализация анализа и прогноза выполнена при помощи языка Python с использованием встроенных и сторонних библиотек. В Таблице 3.1 приведены сокращения, которые будут использоваться далее. На Рисунке 3.9 изображен график суммарной солнечной радиации за весь период.

Таблица 3.1 – Сравнительная таблица метрик

Сокращение	Полное название
N	Номер часа
V	Скорость ветра, м/с
T	Температура воздуха, °C
P	Атмосферное давление, Мм.рт.ст
DirectRad	Прямая солнечная радиация, Вт/м ²
ScatterRad	Рассеянная солнечная радиация, Вт/м ²
SumRad	Суммарная солнечная радиация, Вт/м ²

В работе используется разведочный анализ данных как подход к предварительному исследованию данных. Этот подход направлен на выявление структуры, закономерностей, выбросов, пропусков и взаимосвязей между данными.

В Таблице 3.2 отображены характеристики набора данных, которые содержат:

- count: количество непропущенных значений в столбце,
- mean: среднее арифметическое значение столбца,
- std: стандартное отклонение (мера разброса) значений в столбце,
- min: минимальное значение в столбце,
- 25%: значение, ниже которого 25% значений в столбце,
- 50%: медиана (значение, разделяющее верхнюю и нижнюю половины разброса значений) столбца,
- 75%: значение, ниже которого 75% значений в столбце,

- **max:** максимальное значение в столбце.



Рисунок 3.9 – График суммарной солнечной радиации за весь период

Таблица 3.2 – Характеристики набора данных

	V	T	P	DirectRad	ScatterRad	SumRad
count	88889	88889	88889	88889	88889	88889
mean	1,29	0,01	721,79	60,59	79,26	140,33
std	1,01	12,43	7,06	130,17	91,64	204,94
min	0,0	-33,5	703,00	0,0	0,0	0,0
25%	1,0	-9,7	716,4	0,0	0,0	0,0
50%	1,0	1,1	721,1	0,0	7,38	7,41
75%	2,0	10,5	726,6	54,51	167,10	227,84
max	45,0	28,0	746,6	763,51	365,3	979,41

На основе приведенных характеристик из Таблицы 3.2 можно сделать выводы о том, что данные не содержат пропущенных значений (NaN). Признаки, характеризующие солнечную радиацию (DirectRad, ScatterRad, SumRad), имеют

высокое стандартное отклонение в отношении остальных характеристик, приведенных выше. Это подтверждают медианные значения, они имеют значения 0,0 для DirectRad; 7,38 для ScatterRad и 7,41 для SumRad, то есть как минимум половина значений не превышает приведенных чисел выше, когда максимальные значения 763,51; 365,3 и 979,41 соответственно. Наглядно это можно продемонстрировать на Рисунке 3.10 на примере SumRad. Приведенный рисунок наглядно отражает частоту появления значений внутри диапазонов 0 – 100, 100 – 200 и т.д.

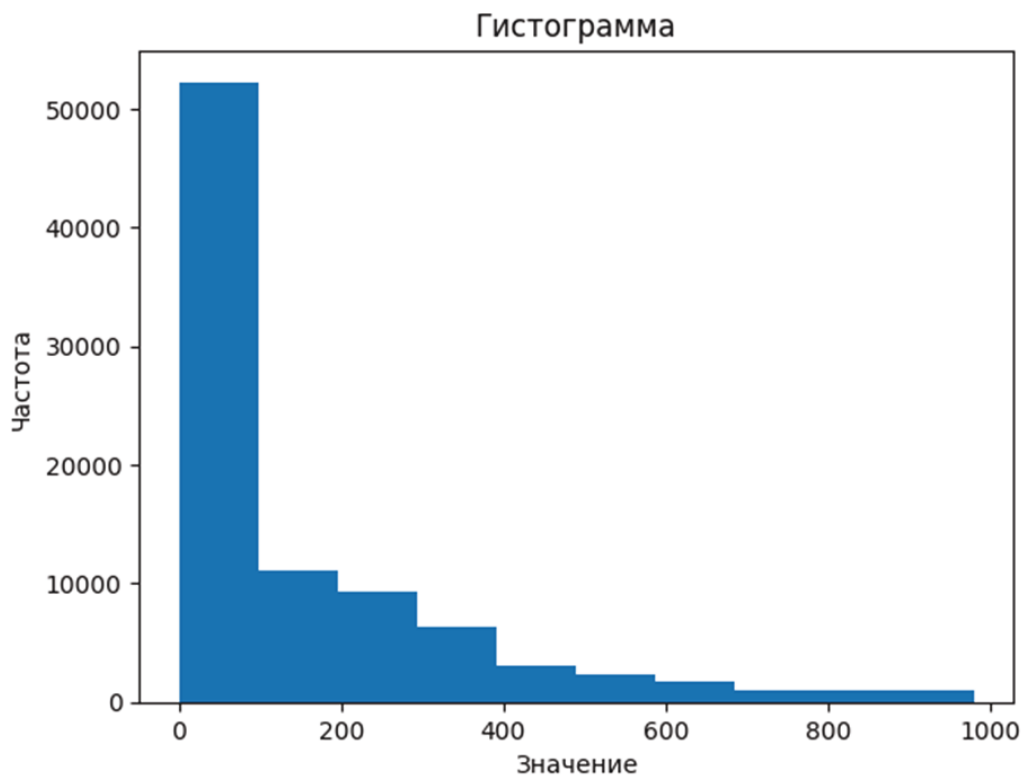


Рисунок 3.10 – Гистограмма частоты появления значений для столбца SumRad

Более 50000 значений лежат в диапазоне 0 – 100, это объясняется тем, что в Байкальске в ночной период значения солнечной радиации близки или равны нулю. Это утверждение наглядно можно увидеть на Рисунке 3.11, где отображен график солнечной радиации за 7 дней.

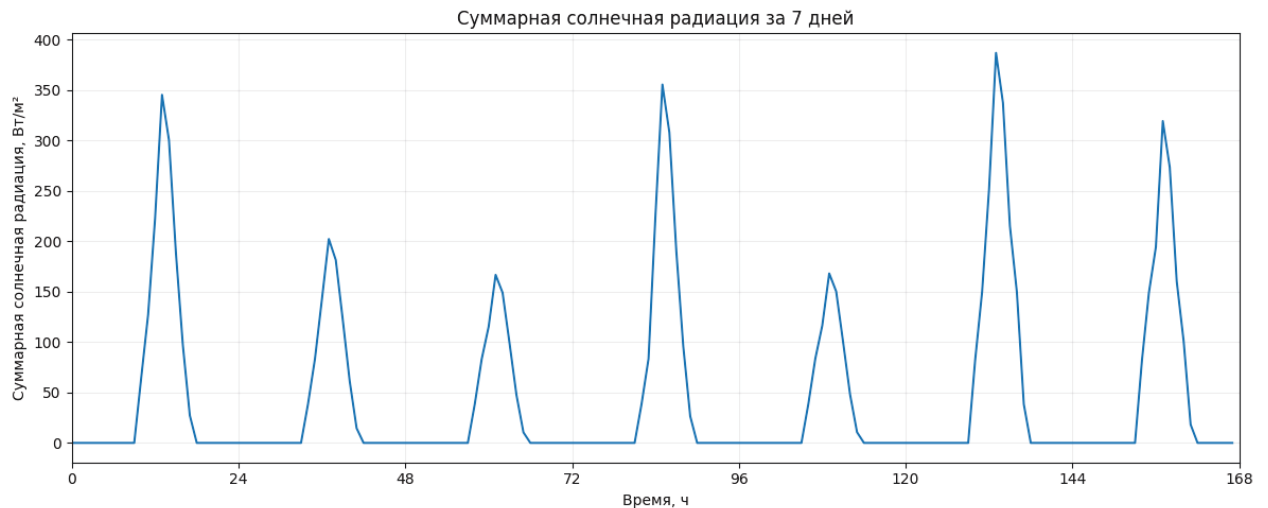


Рисунок 3.11 – График солнечной радиации за 7 дней

На Рисунке 3.12 изображен график средней суммарной солнечной радиации. Повышение солнечной радиации наблюдается с 5 часов утра, достигает пика около 14 часов дня и в дальнейшем идет на спад до 22 часов вечера, где в дальнейшем становится близкой к нулю.



Рисунок 3.12 – Среднесуточная солнечная радиация

Среднемесячная солнечная радиация, полученная за 11 лет (Рис. 3.13), отражает колебания суммарной солнечной радиации. Это объясняется фактором

светового дня, где летний период имеет больший световой день, чем в зимний в г. Байкальск. Амплитуда в этом графике относительно стабильна, что говорит об отсутствии изменений в сезонных условиях за 11 лет.

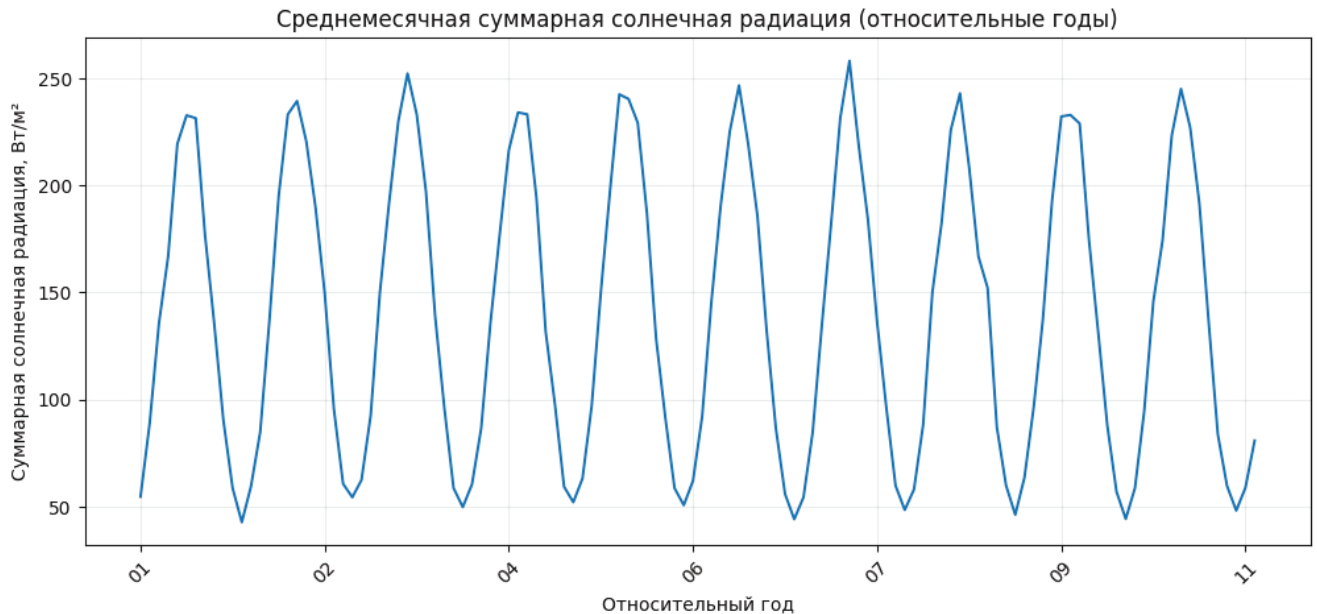


Рисунок 3.13 – Среднемесячная солнечная радиация (относительно года)

На Рисунке 3.14 проводится аддитивная декомпозиция временного ряда, и она же описывается соотношением [115, 116]:

$$y_t = T_t + S_t + \varepsilon_t, \quad (3.1)$$

где y_t – наблюдаемое значение суммарной солнечной радиации в момент времени t ; T_t – трендовая компонента; S_t – сезонная компонента; ε_t – остаточная компонента.

Трендовая компонента определяется как центрированное скользящее среднее:

$$\hat{T}_t = \frac{1}{m} \sum_{k=-(m-1)/2}^{(m-1)/2} y_{t+k}, \quad (3.2)$$

где m – период сезонности. В этом случае $m = 365$, что соответствует годовому циклу и k принимает значения в диапазоне $[-182; 182]$.

После выявления тренда строится детрендрованный ряд:

$$d_t = y_t - \hat{T}_t, \quad (3.3)$$

на основе которого определяется компонента как среднее значение d_t для одинаковых фаз сезонного периода (365 дней):

$$\hat{S}_j = \frac{1}{N_j} \sum_q d_{j+qm}, \quad j = 1, \dots, m, \quad (3.4)$$

где m – период сезонности; N_j – число наблюдений, соответствующих j -й фазе периода; q – индекс, задающий номер сезона, d_{j+qm} – значения детрендрованного ряда, которые попадают в одну и ту же фазу годового цикла.

После оценки сезонных коэффициентов $\hat{S}_j, j = 1, \dots, m$, сезонная компонента для всего временного ряда формируется путем периодического повторения найденного сезонного шаблона. Для каждого момента времени t определяется соответствующая фаза сезона $j = ((t - 1) \bmod m) + 1$, после чего полагается $\hat{S}_t = \hat{S}_j$.

И затем остаточная компонента вычисляется как:

$$\hat{\varepsilon}_t = y_t - \hat{T}_t - \hat{S}_t. \quad (3.5)$$

Визуализация декомпозиции временного ряда разделяет исходные данные на три компоненты: трендовую, сезонную и остаточную:

- Трендовая компонента отражает долгосрочную динамику солнечной радиации. В представленных данных наблюдается относительно стабильный тренд без выраженного роста или снижения на протяжении анализируемого периода. Наблюдается стабильный тренд без выраженного роста/снижения ($\pm 5\%$ от среднего значения $\sim 140 \text{ Вт/м}^2$).
- Сезонная компонента выявляет устойчивые повторяющиеся колебания с четко выраженными годовым и суточным циклом.
- Остаточная компонента представляет собой случайные колебания (случайные погодные колебания, облачность, аномалии, шум и т.д.), не объясненные трендом и сезонностью.

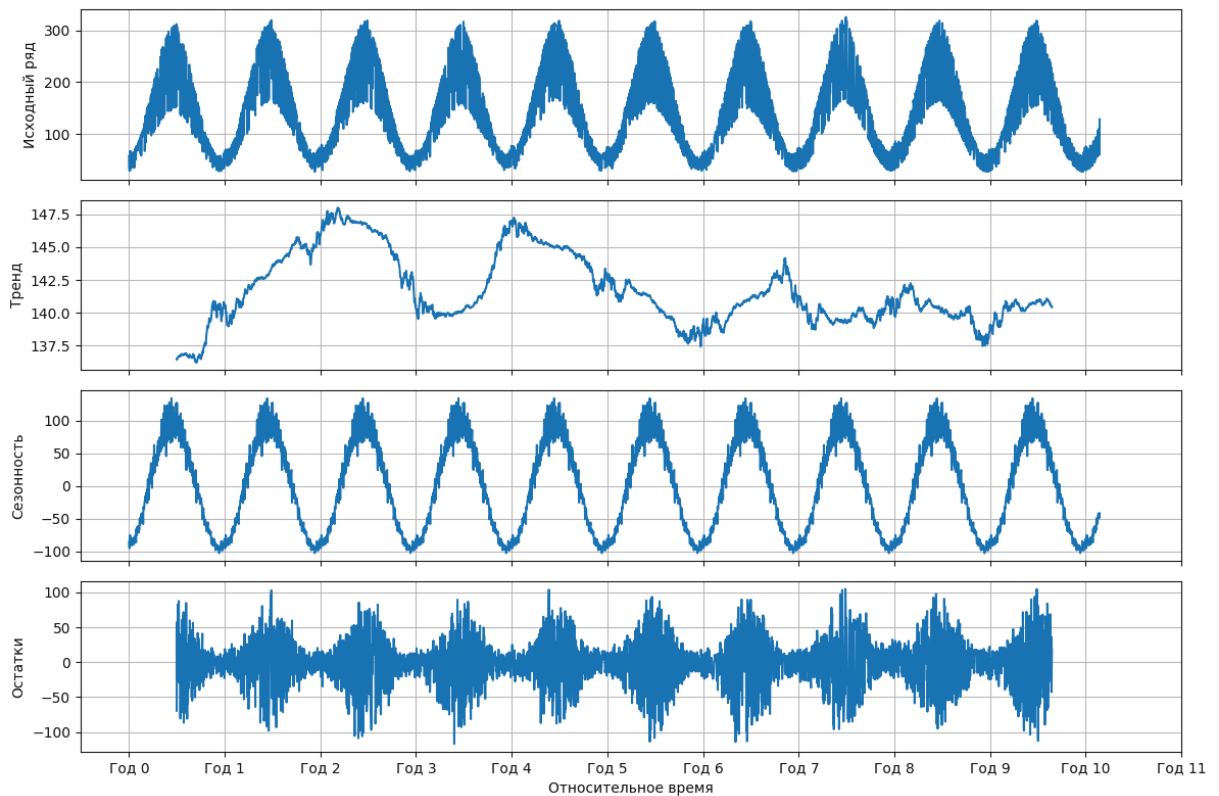


Рисунок 3.14 – Декомпозиция временного ряда

После проведенной декомпозиции основная часть остатков на распределения остатков значений сосредоточена вокруг нуля и само распределение визуально близко к нормальному, что говорит о низком смещении и редких отклонений графике (Рис. 3.15).



Рисунок 3.15 – Распределение остатков

Сравнительный анализ по относительным годам выявляет устойчивость межгодовых закономерностей. Боксовые диаграммы демонстрируют сопоставимые диапазоны значений для аналогичных сезонов разных лет, что подтверждает репрезентативность данных для моделирования. Отдельные выбросы в летних периодах могут отражать периоды аномальной солнечной активности.

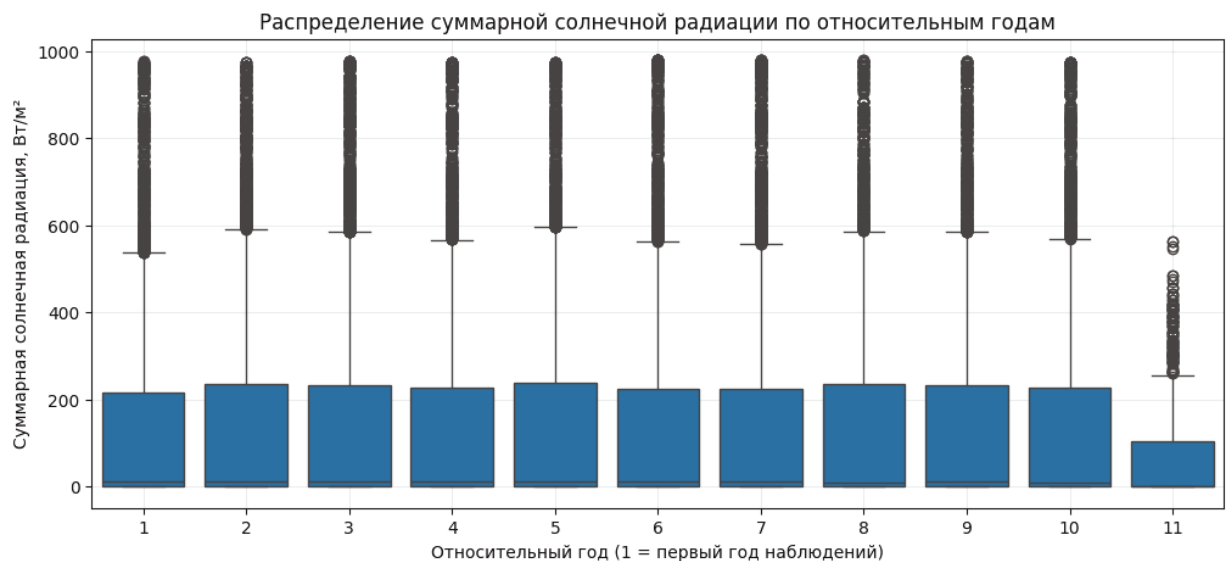


Рисунок 3.16 – График по распределению солнечной радиации

На следующем этапе анализа можно выделить использование тепловых карт (heatmap). Тепловые карты позволяют визуализировать матрицу корреляции наглядно с использованием не только числовых показателей, но и цветовой шкалы. На Рисунке 3.17 представлена тепловая карта датасета. Наибольшие значения коэффициента корреляции Пирсона наблюдаются между признаками прямой, рассеянной и суммарной солнечной радиации. Это объясняется природой рассматриваемых величин: прямая и рассеянная по сути является суммарной солнечной радиацией. Среди остальных умеренную отрицательную корреляцию показывают признаки температуры воздуха и атмосферного давления, а остальные показывают слабую корреляцию.

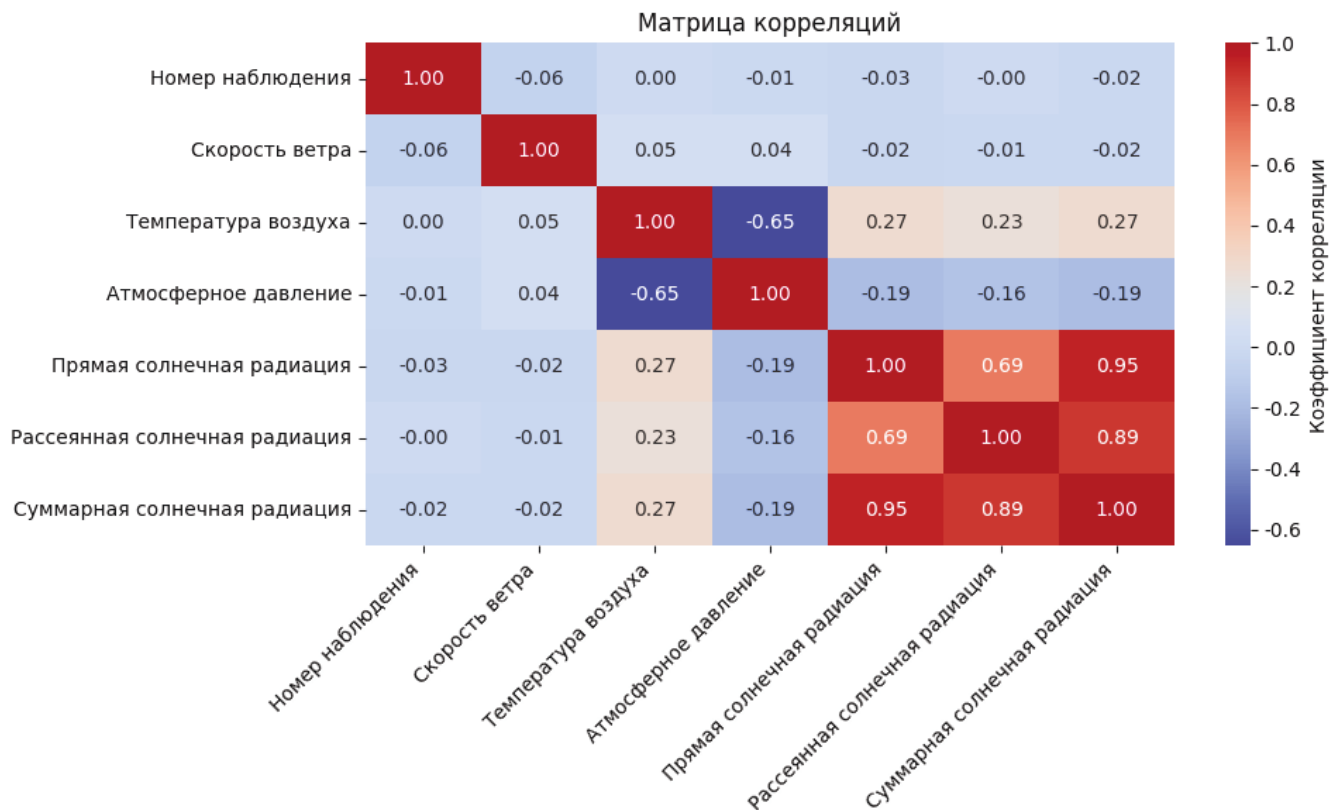


Рисунок 3.17 – Тепловая карта датасета

Предварительная обработка данных позволяет оценить набор данных. Сделанные выводы после каждого этапа обработки:

- датасет не содержит пустые значения и не требуется переобработка данных,
- датасет циклический и имеются сезонность: суточная и годовая,
- столбцы по солнечной радиации имеют большую частоту появления 0,
- тепловая карта показала, что столбцы в большинстве случаев слабо коррелируют между собой, кроме однородных данных о солнечной радиации.

3.4. Проведение вычислительного эксперимента для эмуляции данных (применение предложенного численного метода)

Для проведения эксперимента использовались данные по суммарной солнечной радиации в количестве 88889 записей. Для увеличения

производительности модели использовалась нормализация данных по минимаксу, а также датасет разделен на 2 выборки: 71111 – обучающая выборка и 17778 – тестовая выборка. На Рисунке 3.18 изображены графики предсказания прогнозных значений и тестовых значений. Синим и красным цветом изображены обучающие и тестовые исходные данные, а зеленым и фиолетовым тренировочные и тестовые прогнозируемые значения [104].

На Рисунке 3.18 представлены результаты работы LSTM-сети, состоящей из трех слоев по 128 нейронов. Модель обучалась в течение 12 эпох, общее время обучения составило 17700 секунд. Для предотвращения переобучения использовался метод ранней остановки (EarlyStopping) с мониторингом точности на протяжении 3 эпох. График демонстрирует сравнение фактических и прогнозируемых значений суммарной солнечной радиации, где синим цветом обозначены обучающие данные, красным – тестовые, зеленым и фиолетовым – соответствующие прогнозы модели.

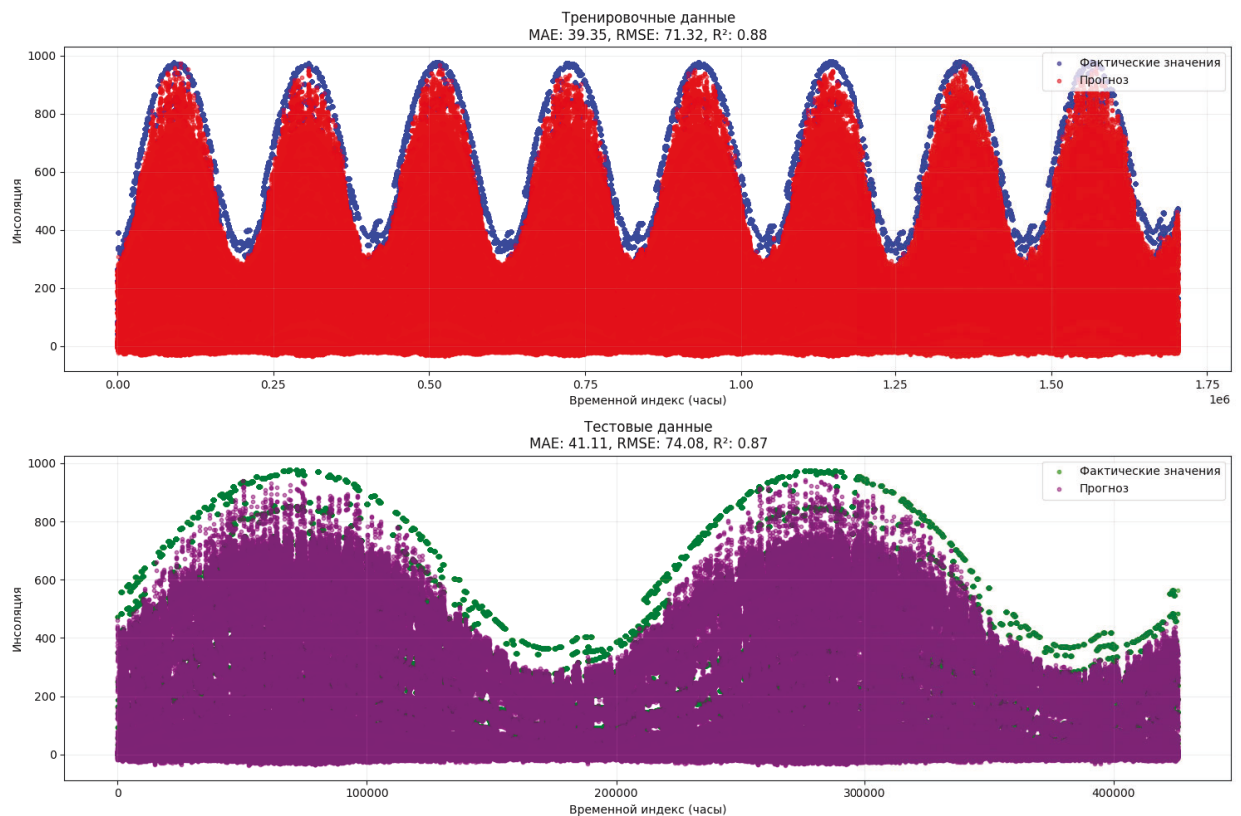


Рисунок 3.18 – Результат работы LSTM-сети

Результаты работы GRU-сети показаны на Рисунке 3.19. Архитектура сети включала три слоя по 150 нейронов, обучение проводилось в течение 9 эпох за 300 секунд. Как и в случае с LSTM, применялся метод ранней остановки. Сравнение времени обучения показывает, что GRU-сеть работает примерно в 60 раз быстрее при сопоставимой точности прогнозирования.

На Рисунке 3.20 представлены результаты работы модели XGBoost, которая продемонстрировала наилучшие показатели точности среди всех тестируемых подходов. Модель состояла из 200 деревьев с максимальной глубиной 5 и также использовала механизм ранней остановки.

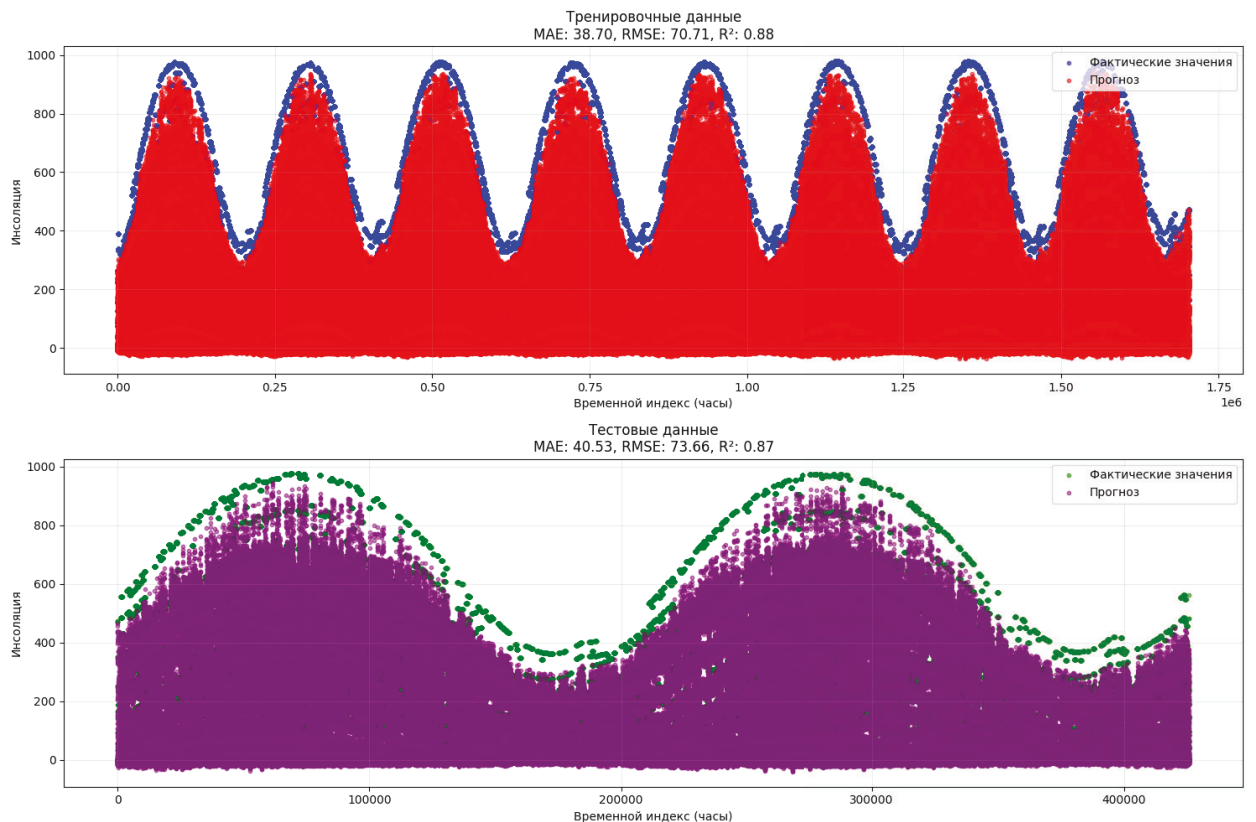


Рисунок 3.19 – Результат работы GRU-сети

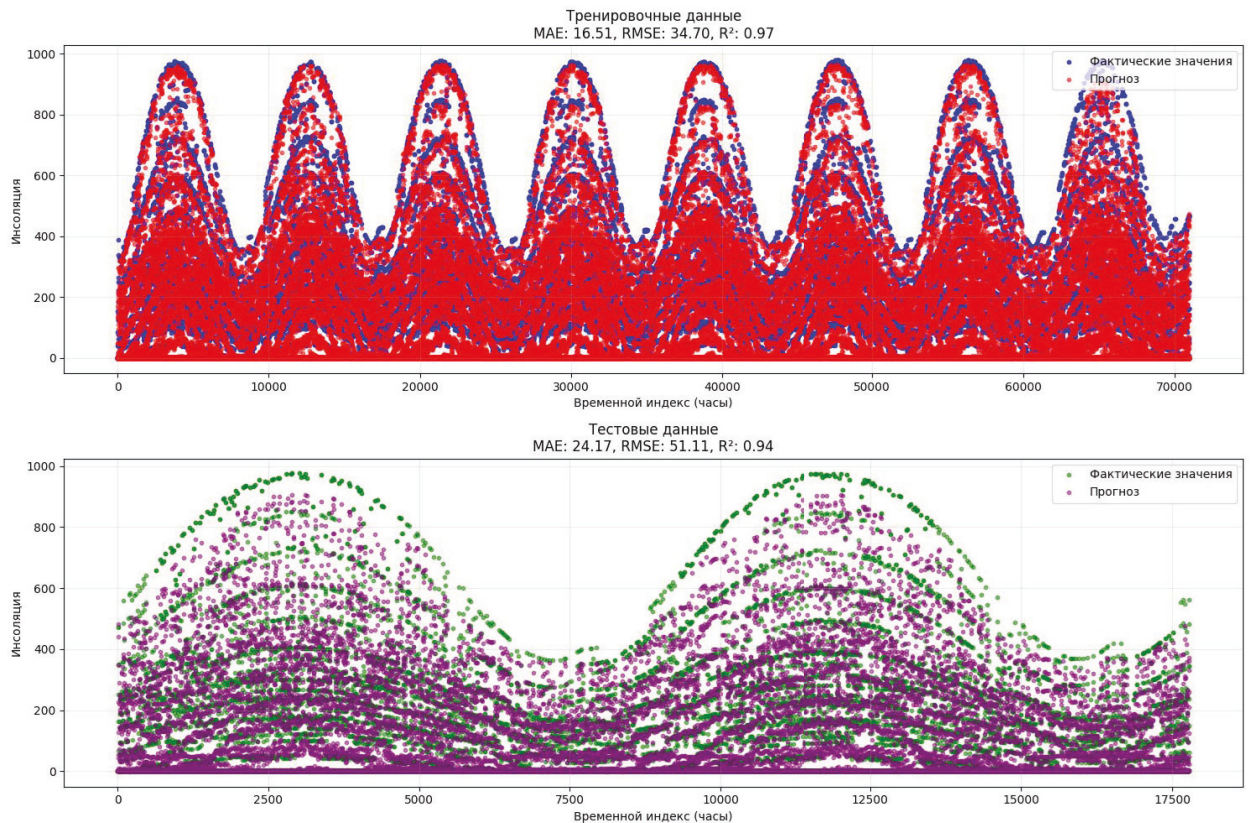


Рисунок 3.20 – Результат работы XGBoost-сети

Для работы алгоритма EnvSwitch строится онтология состояния внешней среды, представленная на Рисунке 3.21. Она формализует основные признаки, влияющие на выбор модели прогнозирования: температурное состояние, состояние солнечной радиации, состояние изменчивости солнечной радиации и состояние достоверности данных. Использование онтологии позволяет перейти от численных значений входных параметров к интерпретируемым режимам среды, на основе которых выполняется выбор модели в составе алгоритма EnvSwitch.

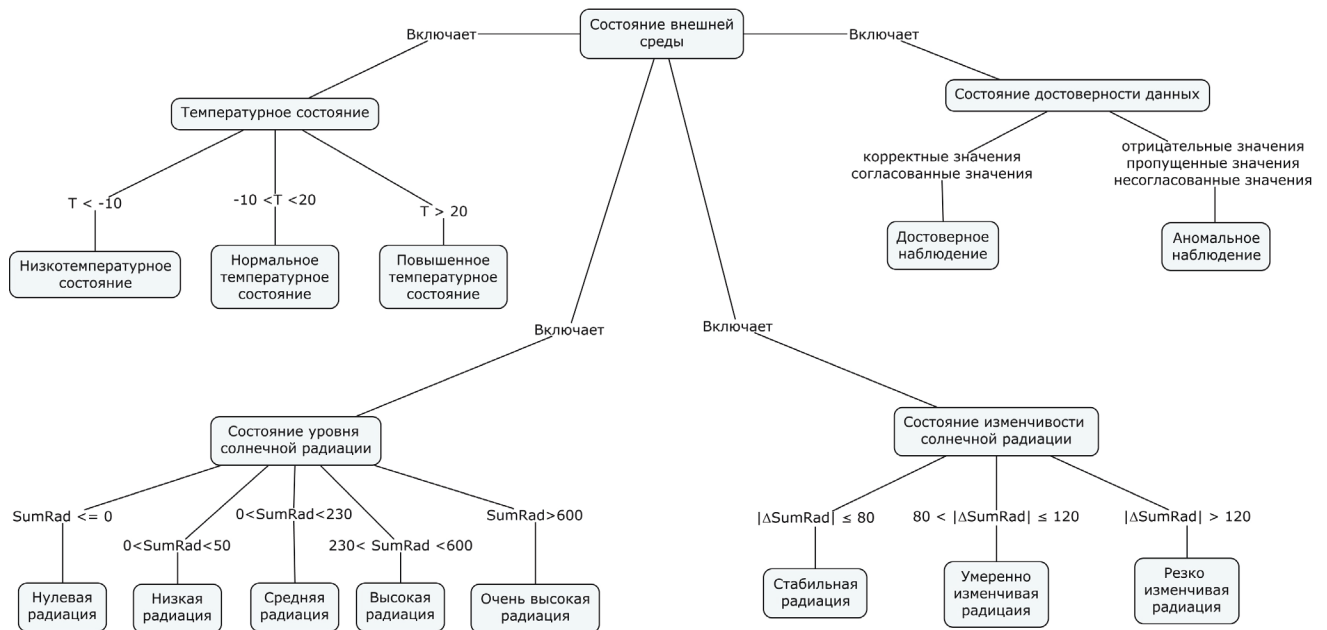


Рисунок 3.21 – Онтология состояния внешней среды

На основе данной онтологии определяются режимы состояния внешней среды, которые используются в алгоритме EnvSwitch для динамического выбора модели прогнозирования. Под режимом состояния внешней среды понимается совокупность признаков, характеризующих текущие условия формирования солнечной радиации: уровень радиации, характер изменения временного ряда, температурное состояние и достоверность исходных данных. Такой подход позволяет перейти от отдельных численных значений входных параметров к интерпретируемому описанию ситуации, в которой выполняется прогнозирование.

Таблица 3.3 – Режимы состояния среды

Режим	Интерпретация	Пояснение
S_0	отсутствие солнечной радиации	Суммарная радиация отсутствует или близка к нулю. Для СЭС это соответствует отсутствию солнечной генерации или ночному участку ряда.
S_1	низкая положительная радиация	Наблюдается малая положительная солнечная радиация. Такой режим может соответствовать началу или завершению периода генерации, сильной облачности или низкому зимнему уровню радиации.

Продолжение Таблицы 3.3

Режим	Интерпретация	Пояснение
<i>S2</i>	стабильная высокая радиация с прямой составляющей	Радиация достаточно высокая, в ее структуре существенна доля прямой составляющей, а временной ряд изменяется плавно.
<i>S3</i>	стабильная рассеянная радиация	Суммарная радиация положительна, но основную часть составляет рассеянная радиация. Это может соответствовать облачным или дымчатым условиям без резких скачков освещенности.
<i>S4</i>	резкая изменчивость радиации	Наблюдается резкое изменение суммарной солнечной радиации или высокая локальная изменчивость временного ряда. Такой режим соответствует изменению облачности, кратковременному затенению или переходным процессам во внешней среде.
<i>S5</i>	низкотемпературный низкоэнергетический режим	Наблюдаются низкая температура и невысокая суммарная радиация. Такой режим отражает сезонно неблагоприятные условия функционирования СЭС.
<i>S6</i>	аномальные или неполные данные	Наблюдение содержит физически некорректные или неполные значения и должно быть направлено в блок предобработки данных.
<i>S7</i>	смешанный режим	Состояние среды не может быть однозначно отнесено к одному из выделенных режимов. Для таких наблюдений используется базовая модель или усредненная комбинация моделей ансамбля.

На Рисунке 3.22 показаны результаты работы алгоритма EnvSwitch на полном наборе данных. В отличие от отдельных моделей, в данном случае прогноз формируется за счёт динамического выбора одной из предобученных моделей в зависимости от текущих характеристик временного окна. В периоды повышенной изменчивости временного сигнала алгоритм преимущественно использует рекуррентные модели (LSTM, GRU), тогда как в условиях относительно

стабильной динамики применяется более простая модель на основе градиентного бустинга (XGBoost), обеспечивающая устойчивость прогноза.

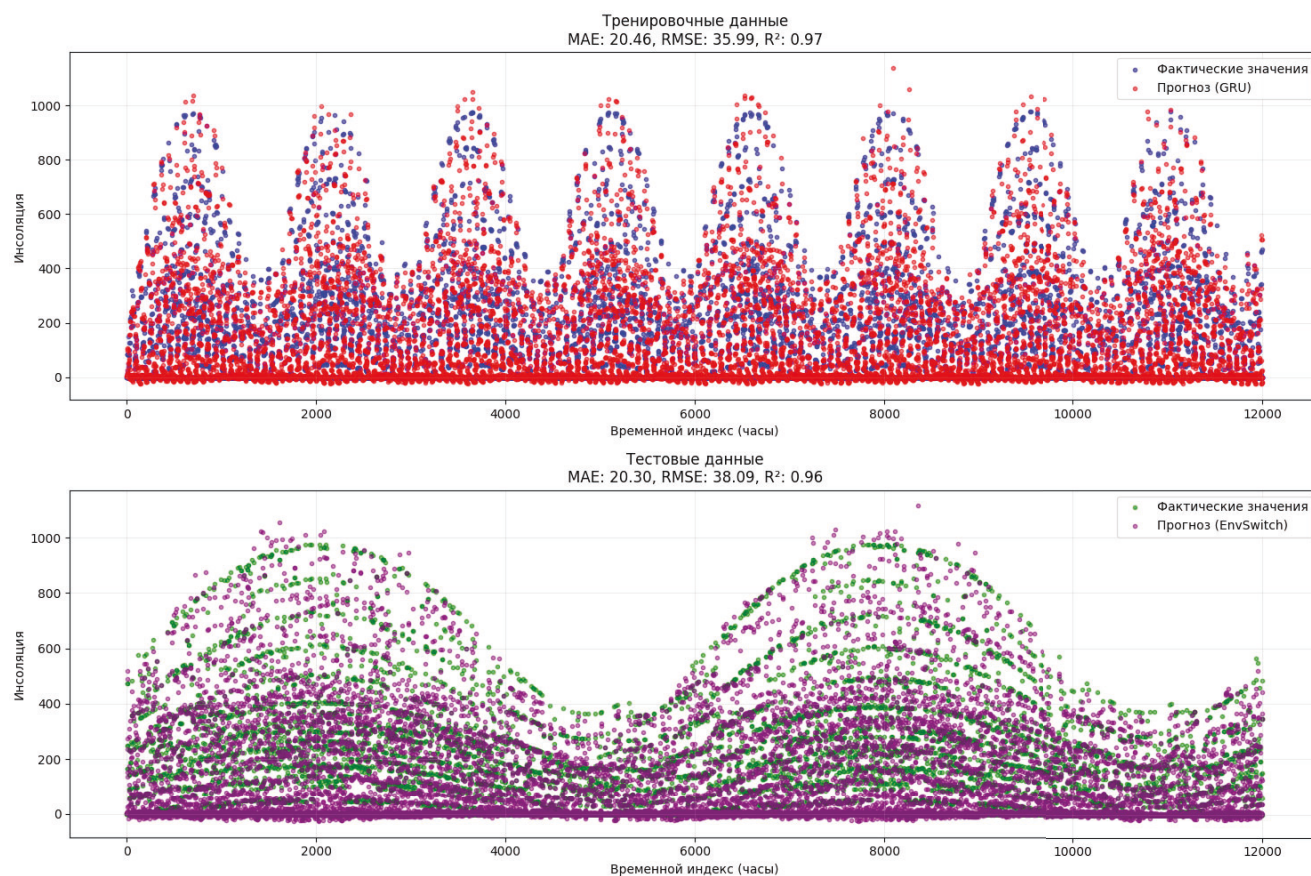


Рисунок 3.22 – Результат работы EnvSwitch

В Таблице 3.4 представлены результат работы сетей, представленных в виде метрик для оценки (RMSE, GRU, XGBoost, EnvSwitch).

Таблица 3.4 – Метрики для оценки методов для датасета из 88889 значений

	LSTM			GRU			XGBoost			EnvSwitch		
	RMSE	MAE	R²	RMSE	MAE	R²	RMSE	MAE	R²	RMSE	MAE	R²
Train	71,32	39,35	0.88	70,71	38,7	0.88	34,7	16,51	0,97	35,99	20,46	0,97
Test	74,08	41.11	0.87	58,7	37.04	0.91	51,11	24,17	0,94	38,09	20,30	0,96

Следующий эксперимент на датасете солнечной радиации в количестве 22000 записей, для работы с которой нормализовались данные и делились на 17500 тренировочных и 4500 тестовых данных.

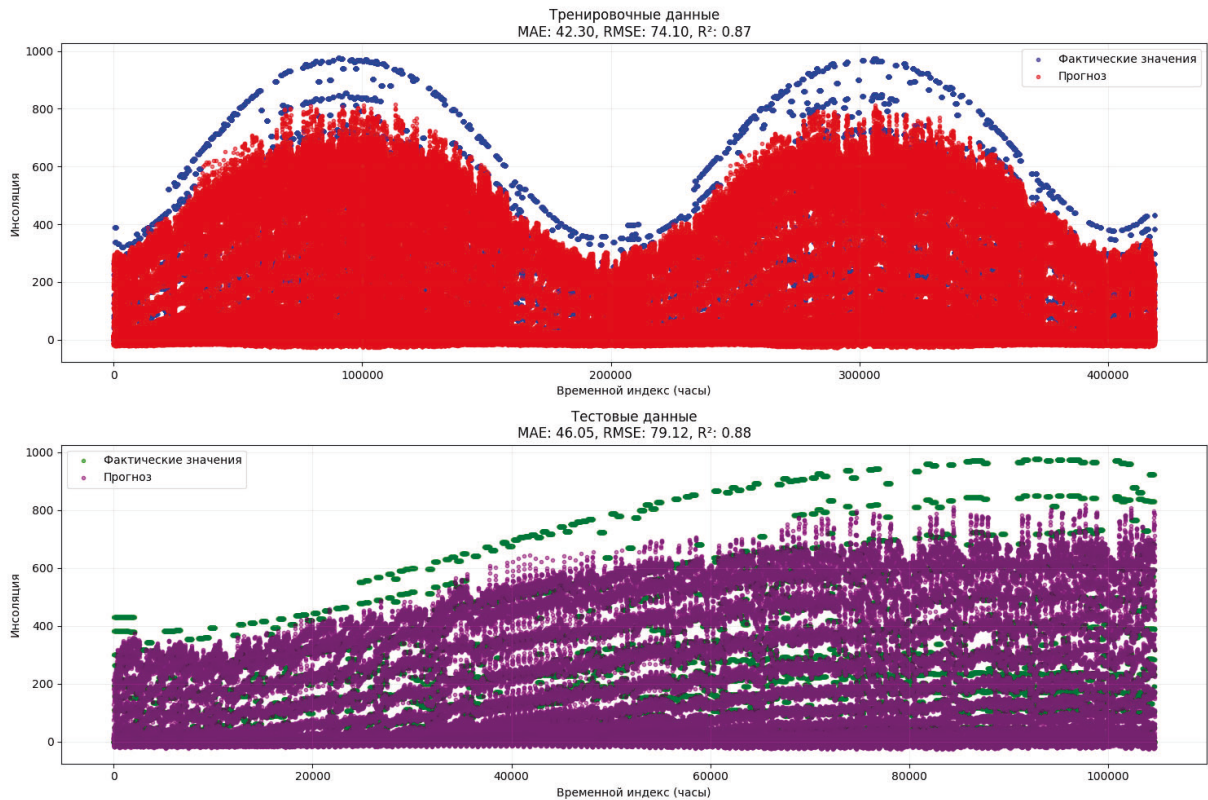


Рисунок 3.23 – Результат работы LSTM – сети

Рисунок 3.23 демонстрирует результаты работы LSTM-сети на сокращенном датасете из 22000 записей. Модель сохранила архитектуру из трех слоев по 128 нейронов и обучалась в течение 12 эпох. На графике отчетливо видно, что несмотря на уменьшение объема данных, сеть способна воспроизводить суточные колебания и повышение общих прогнозируемых значений на фоне сезонности.

На Рисунке 3.24 представлены результаты GRU-сети на том же сокращенном датасете. Архитектура из трех слоев по 150 нейронов показала сравнимые с LSTM результаты. При этом время обучения составило всего 300 секунд, что подтверждает вычислительную эффективность GRU-архитектуры. Однако

визуальный анализ графика показывает, что модель иногда ошибается в предсказании амплитуды пиковых значений, особенно в тестовой выборке.

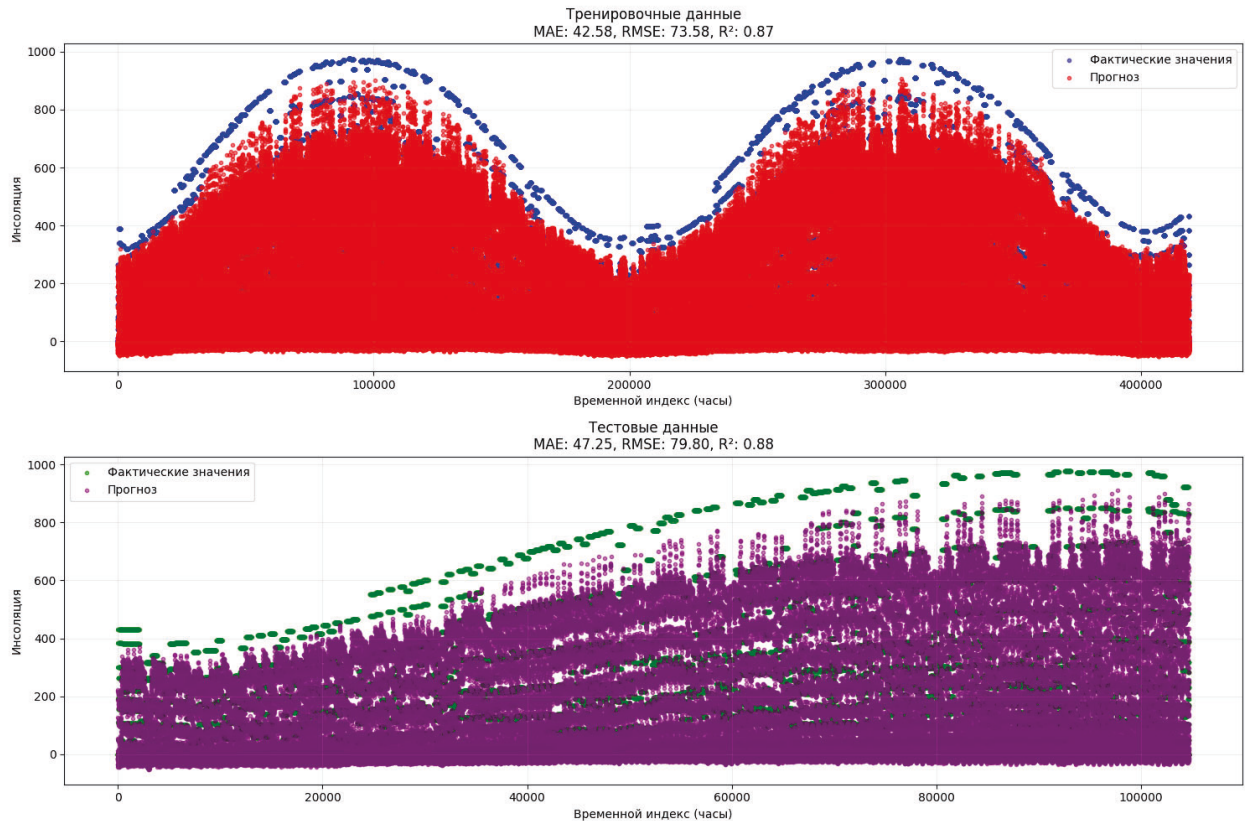


Рисунок 3.24 – Результат работы GRU – сети

Рисунок 3.25 иллюстрирует эффективность XGBoost на малых данных. Несмотря на сокращение объема выборки, модель демонстрирует результаты: RMSE 26,23 (обучение) и 36,16 (тест) при R^2 0,98. Это значительно превосходит показатели нейросетевых подходов. График предсказаний почти идеально совпадает с фактическими значениями, за исключением отдельных моментов резких изменений суммарной солнечной радиации. Такой результат объясняется тем, что алгоритмы на основе деревьев решений менее чувствительны к объему обучающих данных и лучше справляются с малыми выборками.

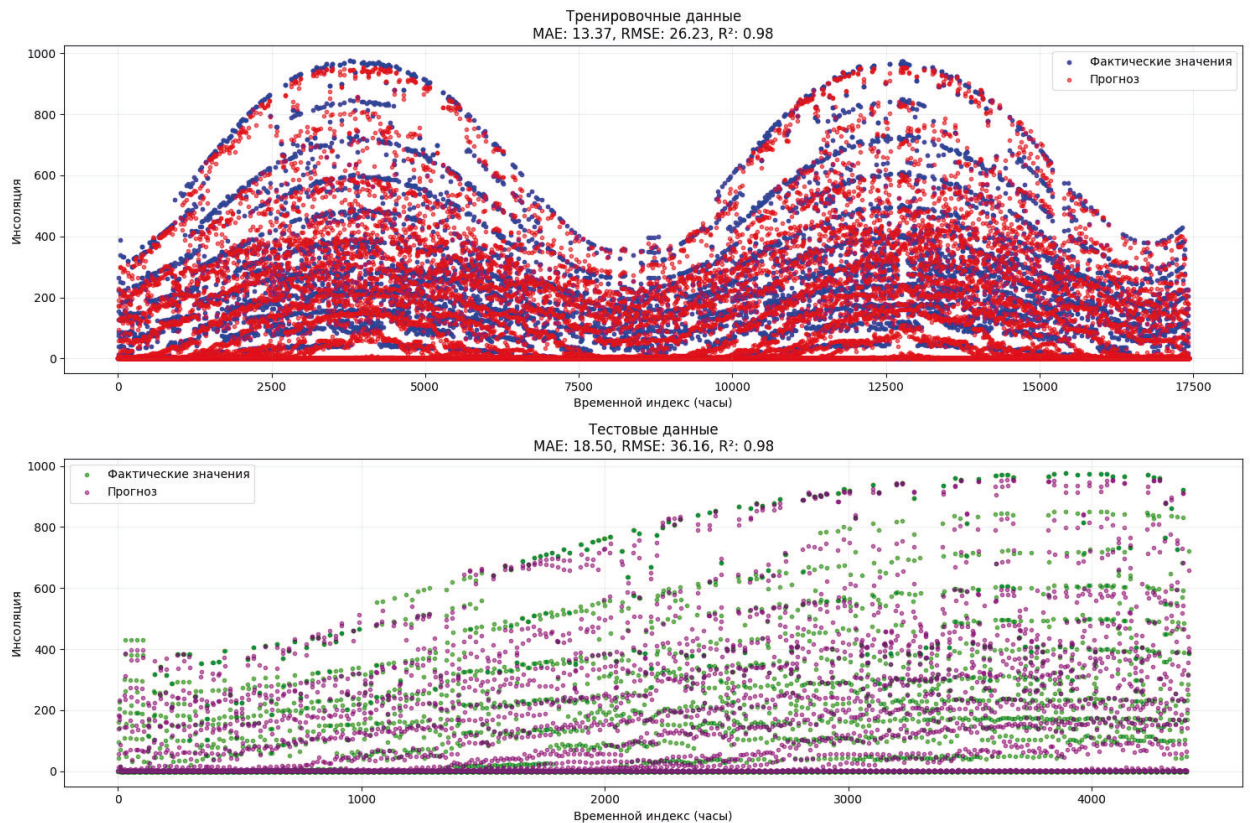


Рисунок 3.25 – Результат работы XGBoost – сети

На Рисунке 3.26 представлены результаты применения алгоритма EnvSwitch на сокращённом наборе данных. Уменьшение объёма обучающей выборки приводит к некоторому снижению точности прогнозирования для всех моделей. Тем не менее адаптивный характер EnvSwitch позволяет частично компенсировать данный эффект. На графике видно, что алгоритм сохраняет способность адекватно воспроизводить общую динамику временного ряда и корректно реагировать на изменения уровня суммарной солнечной радиации. Использование различных моделей в зависимости от текущего режима позволяет избежать существенного накопления ошибки на отдельных участках.

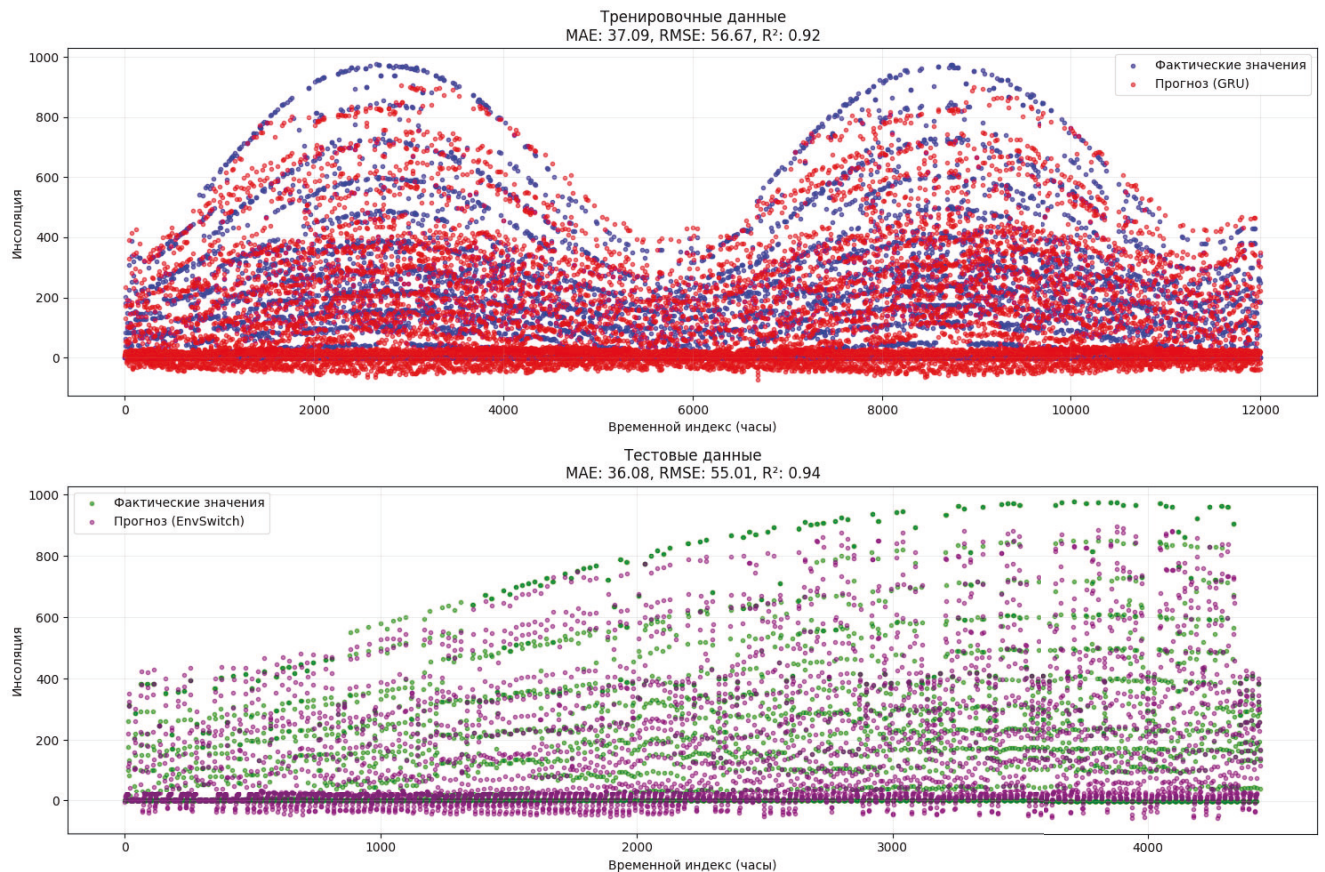


Рисунок 3.26 – Результат работы EnvSwitch – сети

На Таблице 3.5 продемонстрировано сравнение численных значений метрик, которое подтверждает, что при уменьшении объёма данных точность всех моделей снижается. Однако EnvSwitch демонстрирует устойчивые показатели качества прогнозирования, что свидетельствует о целесообразности применения адаптивного подхода в условиях ограниченного объёма обучающих данных.

Таблица 3.5 – Метрики для оценки для датасета из 20 000 значений

	LSTM			GRU			XGBoost			EnvSwitch		
	RMSE	MAE	R ²	RMSE	MAE	R ²	RMSE	MAE	R ²	RMSE	MAE	R ²
Train	74,1	42,3	0.87	73,58	42,58	0,87	26,23	13,37	0,98	56,67	37,09	0,92
Test	79,12	46,05	0.88	79,8	47,25	0,88	36,16	18,5	0,98	55,01	36,08	0,94

Полученные результаты имеют практическое значение для разработки цифрового двойника солнечной электростанции, поскольку показывают, что выбор

модели прогнозирования не может основываться только на её потенциальной точности или объёме доступных обучающих данных. В зависимости от текущего состояния временного ряда и объёма информации, используемой для обучения, эффективность различных моделей существенно различается. В этих условиях целесообразным является применение адаптивного подхода, реализованного в алгоритме EnvSwitch, который обеспечивает динамический выбор модели прогнозирования. Это позволяет использовать алгоритмы на основе деревьев решений в режимах с ограниченной информативностью данных и рекуррентные нейронные сети при работе с полными временными рядами, тем самым повышая устойчивость и точность прогнозирования в составе цифрового двойника.

3.5. Выводы по главе 3

Глава 3 посвящена практической реализации умного цифрового двойника солнечной электростанции (СЭС).

Предложена агентно-сервисная архитектура цифрового двойника, где агенты (серверы, клиенты, оборудование) взаимодействуют через стандартизированные интерфейсы. Агентный подход основан на использовании автономных сущностей (агентов), способных воспринимать среду, принимать решения и взаимодействовать для достижения общих целей, что обеспечивает распределенность, масштабируемость и гибкость системы. Сервис-ориентированная архитектура (SOA) дополняет этот подход, предлагая модульность, слабую связанность и стандартизированные интерфейсы, что упрощает интеграцию компонентов и повышает адаптивность системы к изменениям. Агенты-решатели (математические модели) и агенты-предсказатели (модели ИИ) координируются для оптимизации работы солнечной электростанции, что демонстрирует гибкость и масштабируемость предложенного решения.

Реализация компонентов цифровой тени основана на онтологическом представлении предметной области, в рамках которого формализованы сущности солнечной электростанции и их взаимосвязи. Полученная структура данных

отображена в виде логической схемы и реализована с использованием СУБД, что обеспечило системную организацию цифрового представления объекта и возможность его последующего расширения. Программная реализация научного прототипа цифрового двойника осуществлена в виде web-приложения, интегрирующего компоненты цифровой тени, вычислительное ядро и модули визуализации результатов моделирования. Использование web-архитектуры обеспечило возможность выполнения вычислительных экспериментов, конфигурирования параметров модели и анализа результатов прогнозирования в интерактивном режиме, что повышает практическую применимость разработанной системы.

Разведочный анализ данных суммарной солнечной радиации показал, что имеются: суточная и годовая сезонность, стабильный тренд, большое количество нулевых значений (ночные показатели), слабая корреляция между остальными погодными значениями и нормальное распределение остатков с редкими выбросами.

Сравнение моделей LSTM, GRU и XGBoost показало, что эффективность их применения определяется как объёмом обучающих данных, так и характером временного ряда. Модель XGBoost демонстрирует высокую точность при работе с ограниченными наборами данных (значения R^2 достигают 0.98), тогда как рекуррентные нейронные сети LSTM и GRU более эффективны при использовании протяжённых временных рядов, несмотря на более высокие вычислительные затраты. Полученные результаты обосновывают целесообразность применения адаптивного подхода, реализованного в алгоритме EnvSwitch, который позволяет динамически выбирать модель прогнозирования в зависимости от текущих условий и тем самым сочетать преимущества различных методов.

Заключение

Достигнута цель диссертационной работы: разработан методический подход к построению цифровых двойников, включающий построение агрегированной математической модели СЭС, формализованной модели умного цифрового двойника; разработку численного метода ансамблирования моделей машинного обучения с выбором на основе состояния внешней среды; разработку типовой методики построения ЦД и ее конкретизацию для ЦД СЭС; реализацию программного комплекса – научного прототипа ЦД СЭС.

В ходе работы автором получены следующие результаты:

1. Выполнен анализ существующих математических моделей и методов машинного обучения и возможностей их применения при разработке умных цифровых двойников энергетических объектов.
2. Построена агрегированная модель солнечной электростанции и предложена формализованная модель умного цифрового двойника на основе модификации базовой модели ЦД.
3. Выполнена адаптация численного метода ансамблирования моделей с динамическим выбором на основе состояния внешней среды для повышения точности прогнозирования временных рядов в составе виртуальной среды цифрового двойника солнечной электростанции.
4. Разработаны методики построения умного цифрового двойника с применением онтологического и математического моделирования и моделей машинного обучения и виртуальной среды для эмулирования данных цифрового двойника солнечной электростанции на этапе научного прототипирования.
5. Разработана агентно-сервисная архитектура цифрового двойника солнечной электростанции и реализован программный комплекс (научный прототип умного цифрового двойника солнечной электростанции), интегрирующий основные базовые компоненты ЦД.

Результаты диссертационной работы применены при выполнении базового проекта по госзаказу ИСЭМ СО РАН (АААА-А21-121012090007-7) “Методология

построения ИТ-инфраструктуры для разработки интеллектуальных систем управления развитием и функционированием систем энергетики” (2021-2025 гг.) и проекта по гранту РФФИ (№ 23-21-00382) “Разработка и интеграция методов и технологий построения цифровых двойников энергетических объектов (на примере возобновляемых источников энергии)” (2023 -2024 гг.).

Результаты диссертационной работы отмечены именной стипендией Губернатора Иркутской области (2023 г.). В 2024 г. автор в составе молодежного коллектива стал лауреатом Иркутского областного конкурса в сфере науки и техники, в номинации «За лучшие научные, научно – технические и инновационные разработки молодых ученых».

Список сокращений и условных обозначений

- БД – база данных;
- ВИЭ – возобновляемые источники энергии;
- ИИ – искусственный интеллект;
- ИТ – информационные технологии;
- КПД – коэффициент полезного действия;
- МО – машинное обучение;
- РНС – рекуррентная нейронная сеть;
- СДЭС – солнечно-дизельная электростанция;
- СУБД – система управления базами данных;
- СЭС – солнечная электростанция;
- ТИТТ – технологически изолированные и труднодоступные территории;
- ТЭК – топливно-энергетический комплекс;
- ФС – фотоэлектрические системы;
- ЦД – цифровой двойник;
- CRISP-DM – Cross-Industry Standard Process for Data Mining (методология по исследованию данных);
- GRU – Gated Recurrent Unit (управляемый рекуррентный блок);
- IDEF0 – Integrated Definition (методология функционального моделирования);
- LCOE – Levelized Cost of Electricity (нормированная стоимость единицы энергии);
- LSTM – Long Short-Term Memory (долгая краткосрочная память);
- MAE – Mean Absolute Error (средняя абсолютная ошибка);
- MAPE – Mean Absolute Percentage Error (средняя абсолютная процентная ошибка);
- NREL – National Renewable Energy Laboratory (Национальная лаборатория по изучению возобновляемой энергии);
- R² – Коэффициент детерминации;

RMSE – Root Mean Square Error (среднеквадратическая ошибка);

RNN – Recurrent Neural Network (рекуррентная нейронная сеть);

SOA – Service-Oriented Architecture (сервис-ориентированная архитектура);

TPS – Thermal Protection System (система тепловой защиты).

Список литературы

1. ГОСТ Р 57700.37-2021 Компьютерные модели и моделирование. Цифровые двойники изделий. Общие положения.
2. Top Trends In The Gartner Hype Cycle For Emerging Technologies 2017 [Электронный ресурс]. – URL: <https://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017>. (дата обращения: 10.04.2025).
3. 10 прорывных идей в энергетике на следующие 10 лет // Глобальная энергия [Электронный ресурс] // Глобальная энергия. – URL: <https://globalenergyprize.org/ru/regional-to-global/10ideas/>. (дата обращения: 10.04.2025).
4. Что такое возобновляемая энергия? | Организация Объединенных Наций [Электронный ресурс]. – URL: <https://www.un.org/ru/climatechange/what-is-renewable-energy>. (дата обращения: 10.04.2025).
5. Возобновляемая энергетика в России и мире | Российское энергетическое агентство [Электронный ресурс] // Российское энергетическое агентство. – URL: <https://rosenergo.gov.ru/upload/iblock/e04/3xtm87iv99x76b23c6wjul3as5pzz8zj.pdf>. (дата обращения: 10.04.2025).
6. Best Research-Cell Efficiency Chart | Photovoltaic Research | NREL [Электронный ресурс]. – URL: <https://www.nrel.gov/pv/cell-efficiency>. (дата обращения: 10.04.2025).
7. Luo, W. Potential-induced degradation in photovoltaic modules: a critical review / W. Luo, Y. S. Khoo, P. Hacke, V. Naumann, D. Lausch, S. P. Harvey, J. P. Singh, J. Chai, Y. Wang, A. G. Aberle, S. Ramakrishna // Energy & Environmental Science. – 2017. – Т. 10. – № 1. – С. 43–68.
8. Кирпичникова, И. М. Анализ видов деградации фотоэлектрических модулей и их вольтамперных характеристик / И. М. Кирпичникова, В. А. Заварухин, В. А. Серов // Вестник Пермского национального исследовательского

политехнического университета. Электротехника, информационные технологии, системы управления. – 2024. – № 51. – С. 176–200.

9. Климат Иркутской области как компонент природы [Электронный ресурс]. – URL: <https://geoirkobl.irkutsk.ru/041NaturaKlimat.html>. (дата обращения: 13.04.2025).

10. Зиновьев, В. В. Контроль и диагностика состояния солнечных модулей в условиях неоднородного освещения / В. В. Зиновьев. – PhD Thesis. – Ижевск: ФГБУН «Удмуртский федеральный исследовательский центр Уральского отделения Российской академии наук». – 2025. – 124 с.

11. Зиновьев, В. В. Математическая модель фотоэлектрического преобразователя с использованием W-функции Ламберта / В. В. Зиновьев, А. П. Бельтюков, О. А. Бартенев // Известия Института математики и информатики Удмуртского государственного университета. – 2016. – № 2(48). – С. 22–30.

12. Сазонов, Вас. В. Математическое моделирование работы солнечных батарей космического аппарата / Вас. В. Сазонов // Математическое моделирование. – 2021. – Т. 33. – № 9. – С. 87–107.

13. Ковалев, В. З. Математическое моделирование фотоэлектрических панелей как составляющей комплекса распределенной генерации / В. З. Ковалев, А. О. Парамзин, О. В. Архипова // Инженерный вестник Дона. – 2022. – № 12. – С. 1–15.

14. Герра, Д. Д. Математическое моделирование солнечной электростанции в климатических условиях Республики Куба / Д. Д. Герра, Э. В. Яковлева // Вопросы электротехнологии. – 2020. – № 1(26). – С. 52–60.

15. Kratochvil, J. Photovoltaic array performance model. / J. Kratochvil, W. Boyson, D. King. – 2004.

16. Dobos, A. PVWatts Version 5 Manual / A. Dobos. – National Renewable Energy Laboratory. – 2014.

17. Grieves, M. Digital Twin: Manufacturing Excellence through Virtual Factory Replication / M. Grieves. – 2015.

18. Grieves, M. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems / M. Grieves, J. Vickers // *Transdisciplinary Perspectives on Complex Systems*. – Cham: Springer International Publishing. – 2017. – С. 85–113.
19. Прохоров, А. Цифровой двойник. Анализ, тренды, мировой опыт / А. Прохоров, М. Лысачев. – Первое изд. – ООО «Альянс Принт». – 2020. – 401 с.
20. Digital Shadows in the Internet of Production [Электронный ресурс]. – URL: <https://ercim-news.ercim.eu/en115/special/digital-shadows-in-the-internet-of-production>. (дата обращения: 10.04.2025).
21. Shafto, M. Draft modeling, simulation, information technology & processing roadmap / M. Shafto, M. Conroy, R. Doyle, E. Glaessgen, C. Kemp, J. LeMoigne, L. Wang.
22. Kritzinger, W. Digital Twin in manufacturing: A categorical literature review and classification / W. Kritzinger, M. Karner, G. Traar, J. Henjes, W. Sihn // *IFAC-PapersOnLine*. – 2018. – Т. 51. – № 11. – С. 1016–1022.
23. Tao, F. Digital twin-driven product design, manufacturing and service with big data / F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, F. Sui // *The International Journal of Advanced Manufacturing Technology*. – 2018. – Т. 94. – № 9–12. – С. 3563–3576.
24. Fuller, A. Digital Twin: Enabling Technologies, Challenges and Open Research / A. Fuller, Z. Fan, C. Day, C. Barlow // *IEEE Access*. – 2020. – Т. 8. – С. 108952–108971.
25. Barricelli, B. R. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications / B. R. Barricelli, E. Casiraghi, D. Fogli // *IEEE Access*. – 2019. – Т. 7. – С. 167653–167671.
26. Boschert, S. Digital Twin—The Simulation Aspect / S. Boschert, R. Rosen // *Mechatronic Futures*. – Cham: Springer International Publishing. – 2016. – С. 59–74.
27. Новая парадигма цифрового проектирования и моделирования глобально конкурентоспособной продукции нового поколения - FEA.RU | CompMechLab - разработка и применение цифровых двойников (digital twin), цифровое проектирование, CAD/CAE/CAM/CAO/НПС [Электронный ресурс]. – URL: <https://fea.ru/news/6721>. (дата обращения: 10.04.2025).

28. Боровков, А. И. Цифровые двойники: определение, подходы и методы разработки / А. И. Боровков, Ю. А. Рябов // Цифровая трансформация экономики и промышленности: Сборник трудов научно-практической конференции с зарубежным участием. – Санкт-Петербург: Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого». – 2019. – С. 234–245.
29. Боровков, А. И. Цифровые двойники и цифровая трансформация предприятий ОПК / А. И. Боровков, Ю. А. Рябов, К. В. Кукушкин // Вестник Восточно-Сибирской Открытой Академии. – № 32. – С. 2.
30. Ключевые инструменты и форматы развития цифровой экономики: «умные» цифровые двойники и Центр НТИ СПбПУ | Центр НТИ СПбПУ «Новые производственные технологии» [Электронный ресурс]. – URL: <https://nticenter.spbstu.ru/news/6722>. (дата обращения: 10.06.2025).
31. Arora, N. Introduction to Big Data Analytics / N. Arora, A. Singh, V. Shahare, G. Datta // Towards the Integration of IoT, Cloud and Big Data. – Singapore: Springer Nature Singapore. – 2023. – С. 1–18.
32. Xie, R. Digital Twin Technologies for Turbomachinery in a Life Cycle Perspective: A Review / R. Xie, M. Chen, W. Liu, H. Jian, Y. Shi // Sustainability. – 2021. – Т. 13. – № 5. – С. 2495.
33. Бахтадзе, Н. Н. Методы синтеза цифровых двойников на основе цифровых идентификационных моделей производственных процессов / Н. Н. Бахтадзе, А. Е. Коньков, Д. В. Елпашев, В. Н. Кушнарев, К. С. Мухтаров, А. Пуртов В., В. Е. Пятецкий, А. А. Черешко // Информационные технологии и вычислительные системы. – 2024. – № 4. – С. 100–111.
34. How digital twins are transforming wind operations [Электронный ресурс]. – URL: <https://www.windpowerengineering.com/how-digital-twins-are-transforming-wind-operations/>. (дата обращения: 10.04.2025).

35. Alao, K. T. A review on digital twin application in photovoltaic energy systems: challenges and opportunities / K. T. Alao, S. I. U. H. Gilani, K. Sopian, T. O. Alao // *JMST Advances*. – 2024. – Т. 6. – № 3. – С. 257–282.
36. Yassin, M. A. M. Digital twin in power system research and development: Principle, scope, and challenges / M. A. M. Yassin, A. Shrestha, S. Rabie // *Energy Reviews*. – 2023. – Т. 2. – № 3. – С. 100039.
37. Sifat, Md. M. H. Towards electric digital twin grid: Technology and framework review / Md. M. H. Sifat, S. M. Choudhury, S. K. Das, Md. H. Ahamed, S. M. Mueen, Md. M. Hasan, Md. F. Ali, Z. Tasneem, Md. M. Islam, Md. R. Islam, Md. F. R. Badal, S. H. Abhi, S. K. Sarker, P. Das // *Energy and AI*. – 2023. – Т. 11. – С. 100213.
38. Sleiti, A. K. Digital twin in energy industry: Proposed robust digital twin for power plant and other complex capital-intensive large engineering systems / A. K. Sleiti, J. S. Kapat, L. Vesely // *Energy Reports*. – 2022. – Т. 8. – С. 3704–3726.
39. Javaid, M. Digital Twin applications toward Industry 4.0: A Review / M. Javaid, A. Haleem, R. Suman // *Cognitive Robotics*. – 2023. – Т. 3. – С. 71–92.
40. Kumari, N. A Comprehensive Review of Digital Twin Technology for Grid-Connected Microgrid Systems: State of the Art, Potential and Challenges Faced / N. Kumari, A. Sharma, B. Tran, N. Chilamkurti, D. Alahakoon // *Energies*. – 2023. – Т. 16. – № 14. – С. 5525.
41. Attaran, M. Digital Twin: Benefits, use cases, challenges, and opportunities / M. Attaran, B. G. Celik // *Decision Analytics Journal*. – 2023. – Т. 6. – С. 100165.
42. Ведомственный проект «Цифровая энергетика» [Электронный ресурс]. – URL: <https://minenergo.gov.ru>. (дата обращения: 10.04.2025).
43. АО «ВНИИАЭС» разрабатывает цифровые двойники атомных станций малой мощности [Электронный ресурс]. – URL: <https://www.tek-all.ru/news/id7230-ao-vniiaes-razrabativaet-tsifrovie-dvoyniki-atomnih-stantsiy-maloy-moschnosti/>. (дата обращения: 10.04.2025).
44. Во ВНИИАЭС началась разработка цифровых двойников атомных станций малой мощности с реакторами РИТМ-200Н и «Шельф-М» | Атомная

энергия 2.0 [Электронный ресурс]. – URL: <https://www.atomic-energy.ru/news/2021/02/11/111440>. (дата обращения: 10.04.2025).

45. Цифровой двойник энергосистемы. В НИУ «МЭИ» разработана перспективная отечественная программная платформа | Новости энергетики «ЭЛЕКТРОЭНЕРГИЯ. Передача и распределение» [Электронный ресурс]. – URL: <https://eepir.ru/new/tsds-niu-mpei-310323/>. (дата обращения: 10.06.2025).

46. Ковалёв, С. П. Проектирование информационного обеспечения цифровых двойников энергетических систем / С. П. Ковалёв // Системы и средства информатики. – 2020. – Т. 30. – № 1. – С. 66–81.

47. Стенников, В. А. Автоматизация вычислений при проектировании интегрированной энергетической системы на основе ее цифрового двойника / В. А. Стенников, Е. А. Барахтенко, Д. В. Соколов, Майоров, Г. С. // Известия Российской академии наук. Энергетика. – 2024. – № 2. – С. 3–20.

48. Стенников, В. А. Построение цифровых двойников для проектирования интегрированных энергетических систем / В. А. Стенников, Е. А. Барахтенко, Д. В. Соколов, Г. С. Майоров // Программная инженерия. – 2023. – Т. 14. – № 12. – С. 592–602.

49. Olayiwola, O. Enhanced Solar Photovoltaic System Management and Integration: The Digital Twin Concept / O. Olayiwola, U. Cali, M. Elsdén, P. Yadav // Solar. – 2025. – Т. 5. – № 1. – С. 7.

50. Yalçın, T. Exploiting Digitalization of Solar PV Plants Using Machine Learning: Digital Twin Concept for Operation / T. Yalçın, P. Paradell Solà, P. Stefanidou-Voziki, J. L. Domínguez-García, T. Demirdelen // Energies. – 2023. – Т. 16. – № 13. – С. 5044.

51. Karunanidhi, B. Development of a Digital Twin Framework for a PV System to Resolve Partial Shading / B. Karunanidhi, L. Ramasamy, A. Alexander Stonier, C. R. Sathiasamuel // Mathematical Problems in Engineering. – 2024. – Т. 2024. – С. 1–15.

52. Толковый словарь по искусственному интеллекту [Электронный ресурс]. – URL: <https://www.raai.org/pages/UGFnZVR5cGU6MTAwMw==#L208>. (дата обращения: 10.04.2025).

53. Flach, P. Machine Learning: The Art and Science of Algorithms that Make Sense of Data / P. Flach. – 1 изд. – Cambridge University Press. – 2012.
54. Машинное обучение [Электронный ресурс]. – URL: http://www.machinelearning.ru/wiki/index.php?title=%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5_%D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D0%B5. (дата обращения: 10.04.2025).
55. Understanding LSTM Networks -- colah's blog [Электронный ресурс]. – URL: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. (дата обращения: 10.04.2025).
56. Hochreiter, S. Long Short-Term Memory / S. Hochreiter, J. Schmidhuber // Neural Computation. – 1997. – Т. 9. – № 8. – С. 1735–1780.
57. Chung, J. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling / J. Chung, C. Gulcehre, K. Cho, Y. Bengio. – arXiv. – 2014.
58. Friedman, J. H. Greedy function approximation: A gradient boosting machine. / J. H. Friedman // The Annals of Statistics. – 2001. – Т. 29. – № 5.
59. Kumari, P. Deep learning models for solar irradiance forecasting: A comprehensive review / P. Kumari, D. Toshniwal // Journal of Cleaner Production. – 2021. – Т. 318. – С. 128566.
60. Ahmed, R. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization / R. Ahmed, V. Sreeram, Y. Mishra, M. D. Arif // Renewable and Sustainable Energy Reviews. – 2020. – Т. 124. – С. 109792.
61. Das, U. K. Forecasting of photovoltaic power generation and model optimization: A review / U. K. Das, K. S. Tey, M. Seyedmahmoudian, S. Mekhilef, M. Y. I. Idris, W. Van Deventer, B. Horan, A. Stojcevski // Renewable and Sustainable Energy Reviews. – 2018. – Т. 81. – С. 912–928.
62. Guo, X. Study on short-term photovoltaic power prediction model based on the Stacking ensemble learning / X. Guo, Y. Gao, D. Zheng, Y. Ning, Q. Zhao // Energy Reports. – 2020. – Т. 6. – С. 1424–1431.

63. Ren, Y. Ensemble methods for wind and solar power forecasting—A state-of-the-art review / Y. Ren, P. N. Suganthan, N. Srikanth // *Renewable and Sustainable Energy Reviews*. – 2015. – Т. 50. – С. 82–91.
64. Husein, M. Day-Ahead Solar Irradiance Forecasting for Microgrids Using a Long Short-Term Memory Recurrent Neural Network: A Deep Learning Approach / M. Husein, I.-Y. Chung // *Energies*. – 2019. – Т. 12. – № 10. – С. 1856.
65. Breiman, L. Bagging predictors / L. Breiman // *Machine Learning*. – 1996. – Т. 24. – № 2. – С. 123–140.
66. Wolpert, D. H. Stacked generalization / D. H. Wolpert // *Neural Networks*. – 1992. – Т. 5. – № 2. – С. 241–259.
67. Dietterich, T. G. Ensemble Methods in Machine Learning / T. G. Dietterich // *Multiple Classifier Systems*. – Berlin, Heidelberg: Springer Berlin Heidelberg. – 2000. – С. 1–15.
68. Rahimi, N. A Comprehensive Review on Ensemble Solar Power Forecasting Algorithms / N. Rahimi, S. Park, W. Choi, B. Oh, S. Kim, Y. Cho, S. Ahn, C. Chong, D. Kim, C. Jin, D. Lee // *Journal of Electrical Engineering & Technology*. – 2023. – Т. 18. – № 2. – С. 719–733.
69. Массель, Л. В. Разработка цифрового двойника ветровой электростанции: постановка задачи и проектирование / Л. В. Массель, А. Г. Массель, Н. И. Щукин, А. Р. Цыбиков // *Информационные и математические технологии в науке и управлении*. – 2022. – № 1. – С. 79–90.
70. Цыбиков, А. Р. Построение цифровых двойников возобновляемых источников энергии / А. Р. Цыбиков, Л. В. Массель, А. Г. Массель, Н. И. Щукин // *Труды Международной научно-практической конференции «Проблемы и решения эффективного использования альтернативных источников энергии»*. – 2023. – С. 68–75.
71. Massel, L. Digital twin development of a solar power plant / L. Massel, N. Shchukin, A. Cybikov // *E3S Web of Conferences*. – 2021. – Т. 289. – С. 03002.

72. Ruiz, J. C. S. Smart Digital Twin for ZDM-based job-shop scheduling / J. C. S. Ruiz, J. M. Bru, R. P. Escoto // 2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT). – Rome, Italy: IEEE. – 2021. – С. 510–515.
73. Массель, Л. В. Семантическое моделирование при построении цифровых двойников энергетических объектов и систем / Л. В. Массель, Массель, А.Г. // Онтология проектирования. – 2023. – Т. 13. – № 1. – С. 44–54.
74. Массель, Л. Фрактальный подход к построению информационных технологий / Л. Массель // Информационная технология исследований развития энергетики. – Новосибирск: Наука. – 1995. – С. 40–67.
75. Массель, Л. Фрактальная модель структурирования знаний / Л. Массель // Сб. науч. трудов Национальной конференции с международным участием «Искусственный интеллект-94». – Рыбинск. – 1994. – С. 46–49.
76. Массель, Л. Фрактальный подход к структурированию знаний и примеры его применения / Л. Массель // *Ontology of Designing*. – 2016. – Т. 6. – № 2. – С. 149–161.
77. Saba, D. Optimization of a Multi-source System with Renewable Energy Based on Ontology / D. Saba, F. Z. Laallam, A. E. Hadidi, B. Berbaoui // *Energy Procedia*. – 2015. – Т. 74. – С. 608–615.
78. Notton, G. Hybrid Photovoltaic/Wind Energy Systems For Remote Locations / G. Notton, S. Diaf, L. Stoyanov // *Energy Procedia*. – 2011. – Т. 6. – С. 666–677.
79. Duffie, J. A. Solar Engineering of Thermal Processes / J. A. Duffie, W. A. Beckman. – 1 изд. – Wiley. – 2013.
80. Jordan, D. C. Photovoltaic Degradation Rates—an Analytical Review / D. C. Jordan, S. R. Kurtz // *Progress in Photovoltaics: Research and Applications*. – 2013. – Т. 21. – № 1. – С. 12–29.
81. Карамов, Д. Н. Моделирование солнечной электростанции с учётом изменения параметров окружающей среды / Д. Н. Карамов, И. В. Наумов // *Электрические станции*. – 2020. – № 6(1067). – С. 21–28.

82. Koehl, M. Modeling of the nominal operating cell temperature based on outdoor weathering / M. Koehl, M. Heck, S. Wiesmeier, J. Wirth // *Solar Energy Materials and Solar Cells*. – 2011. – Т. 95. – № 7. – С. 1638–1646.
83. ГОСТ Р 57721-2017 Информационно-коммуникационные технологии в образовании. Эксперимент виртуальный. Общие положения.
84. Wen, Q. Time Series Data Augmentation for Deep Learning: A Survey / Q. Wen, L. Sun, F. Yang, X. Song, J. Gao, X. Wang, H. Xu. – 2020.
85. Chatterjee, S. GAN-based synthetic time-series data generation for improving prediction of demand for electric vehicles / S. Chatterjee, D. Hazra, Y.-C. Byun // *Expert Systems with Applications*. – 2025. – Т. 264. – С. 125838.
86. Oh, C. Time-Series Data Augmentation based on Interpolation / C. Oh, S. Han, J. Jeong // *Procedia Computer Science*. – 2020. – Т. 175. – С. 64–71.
87. Chapman, P. Step-by-step data mining guide / P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer, R. Wirth.
88. Hyndman, R. J. Another look at measures of forecast accuracy / R. J. Hyndman, A. B. Koehler // *International Journal of Forecasting*. – 2006. – Т. 22. – № 4. – С. 679–688.
89. Armstrong, J. S. Error measures for generalizing about forecasting methods: Empirical comparisons / J. S. Armstrong, F. Collopy // *International Journal of Forecasting*. – 1992. – Т. 8. – № 1. – С. 69–80.
90. Nagelkerke, N. J. D. A note on a general definition of the coefficient of determination / N. J. D. Nagelkerke // *Biometrika*. – 1991. – Т. 78. – № 3. – С. 691–692.
91. Гудфеллоу Ян. Глубокое обучение / Гудфеллоу Ян, Бенджио Йошуа, Курвилль Аарон. – 2 изд. – ДМК Пресс. – 2018. – 652 с.
92. Bergstra, J. Random Search for Hyper-Parameter Optimization / J. Bergstra, Y. Bengio // *The Journal of Machine Learning Research*. – 2012. – Т. 13. – С. 281–305.
93. Brochu, E. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning / E. Brochu, V. Cora, N. Freitas // *CoRR*. – 2010. – Т. abs/1012.2599.

94. Lu, J. Learning under Concept Drift: A Review / J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, G. Zhang // IEEE Transactions on Knowledge and Data Engineering. – 2018. – С. 1–1.
95. Barros, R. S. M. D. An overview and comprehensive comparison of ensembles for concept drift / R. S. M. D. Barros, S. G. T. D. C. Santos // Information Fusion. – 2019. – Т. 52. – С. 213–244.
96. Song, Z. Digital Twins for the Future Power System: An Overview and a Future Perspective / Z. Song, C. M. Hackl, A. Anand, A. Thommessen, J. Petzschmann, O. Kamel, R. Braunbehrens, A. Kaifel, C. Roos, S. Hauptmann // Sustainability. – 2023. – Т. 15. – № 6. – С. 5259.
97. Массель, Л. Методология построения ИТ-инфраструктуры для разработки интеллектуальных систем управления развитием и функционированием систем энергетики / Л. Массель.
98. Массель Л. В. Построение цифровых двойников ветровой и солнечной электростанций на основе онтологического подхода / Массель Л. В., Массель А. Г., Щукин Н. И., Цыбиков А. Р., Лосев А. С. // Автоматизация в промышленности. – 2022. – Т. 7. – С. 28–32.
99. Р 50.1.028-2001 Информационные технологии поддержки жизненного цикла продукции. Методология функционального моделирования.
100. Цыбиков, А. Р. Разработка цифрового двойника солнечной электростанции с использованием онтологий / А. Р. Цыбиков, Н. И. Щукин, Л. В. Массель // Знания - Онтологии - Теории (ЗОНТ-2021): Материалы VIII Международной конференции, Новосибирск, 08–12 ноября 2021 года. – Новосибирск: Институт математики им. С.Л. Соболева СО РАН. – 2021. – С. 268–274.
101. Цыбиков, А. Р. Разработка цифрового двойника солнечной электростанции с использованием технологии Big Data / А. Р. Цыбиков, Н. И. Щукин, Л. В. Массель // Труды Международного научно-технического конгресса «Интеллектуальные системы и информационные технологии - 2021». – 2021. – С. 598–605.

102. Massel, L. V. Agent-Service Approach to Building Digital Twins / L. V. Massel, A. G. Massel, A. R. Tsybikov. – Sochi, Russian Federation: IEEE. – 2024. – С. 411–415.

103. Massel, L. Designing a Digital Twin of a Wind Farm / L. Massel, A. Massel, N. Shchukin, A. Tsybikov // 15th International Conference “Intelligent Systems” (INTELS’22). – MDPI. – 2023. – С. 30.

104. Цыбиков, А. Р. Программа прогнозирования временных рядов для цифрового двойника солнечной электростанции с использованием LSTM-сетей / А. Р. Цыбиков, Л. В. Массель, А. Г. Массель, Н. И. Щукин // Свидетельство о гос. регистрации программы для ЭВМ № 2022683763. Правообладатель: ФГБУН ИСЭМ СО РАН. № 2022681766; заявл. 16.11.2022; зарегистр. 08.12.2022; опубл. 08.12.2022, 1 с.

105. Цыбиков, А. Р. Оценка устойчивости энергетических систем с применением методов машинного обучения / А. Р. Цыбиков, Л. В. Массель, А. Г. Массель, Т. Г. Мамедов, Д. А. Гаськова, Н. И. Щукин // Информационные и математические технологии в науке и управлении. – 2022. – № 4. – С. 248–261.

106. Городецкий, В. Современное состояние и перспективы индустриальных применений многоагентных систем / В. Городецкий, О. Бухвалов, П. Скобелев // Управление большими системами: сборник трудов. – 2017. – № 66. – С. 94–157.

107. Russell, S. J. Artificial intelligence: a modern approach / S. J. Russell, P. Norvig. – Fourth Edition изд. – Hoboken, NJ: Pearson. – 2021. – 1115 с.

108. What Is SOA (Service-Oriented Architecture)? | Oracle United Kingdom [Электронный ресурс]. – URL: <https://www.oracle.com/uk/service-oriented-architecture-soa/>. (дата обращения: 10.04.2025).

109. What is SOA? - Service-Oriented Architecture Explained - AWS [Электронный ресурс] // Amazon Web Services, Inc. – URL: <https://aws.amazon.com/what-is/service-oriented-architecture/>. (дата обращения: 10.04.2025).

110. Цыбиков, А. Р. Вычислительное ядро цифрового двойника солнечной электростанции / А. Р. Цыбиков, Л. В. Массель, А. Г. Массель, Н. И. Щукин, Д. Н. Карамов // Свидетельство о гос. регистрации программы для ЭВМ № 2021680672. Правообладатель: ФГБУН ИСЭМ СО РАН. № 2021681401; заявл. 15.12.2021; зарегистр. 21.12.2021; опубл. 21.12.2021, 1 с.

111. Цыбиков, А. Р. Программа для обработки математических формул и вычислений с использованием наборов данных цифрового двойника / А. Р. Цыбиков, Л. В. Массель, А. Г. Массель, Н. И. Щукин // Свидетельство о гос. регистрации программы для ЭВМ № 2024686190. Правообладатель: ФГБУН ИСЭМ СО РАН. № 2024685833; заявл. 06.11.2024; зарегистр. 06.11.2024; опубл. 06.11.2024.

112. Цыбиков, А. Р. Вычислительное ядро цифрового двойника ветровой электростанции / А. Р. Цыбиков, Л. В. Массель, А. Г. Массель, Н. И. Щукин // Свидетельство о гос. регистрации программы для ЭВМ № 2022682299. Правообладатель: ФГБУН ИСЭМ СО РАН. № 2022681789; заявл. 16.11.2022; зарегистр. 21.11.2022; опубл. 21.11.2022, 1 с.

113. Цыбиков, А. Р. Компонент визуализации цифрового двойника ветровой электростанции / А. Р. Цыбиков, Л. В. Массель, А. Г. Массель, Н. И. Щукин // Свидетельство о гос. регистрации программы для ЭВМ № 2024686216. Правообладатель: ФГБУН ИСЭМ СО РАН. № 2024685836; заявл. 06.11.2024; зарегистр. 06.11.2024; опубл. 06.11.2024, 1 с.

114. Цыбиков, А. Р. База данных цифрового двойника солнечной электростанции (DB DTSP) / А. Р. Цыбиков, Л. В. Массель, А. Г. Массель, Н. И. Щукин // Свидетельство о гос. регистрации базы данных № 2022620268. Правообладатель: ФГБУН ИСЭМ СО РАН. № 2021623049; заявл. 15.12.2021; зарегистр. 02.02.2022; опубл. 02.02.2022, 1 с.

115. Светуньков, И. С. Методы социально-экономического прогнозирования: учебник и практикум для вузов / И. С. Светуньков, С. Г. Светуньков. – Москва: Издательство Юрайт. – 256 с.

116. Introduction — statsmodels [Электронный ресурс]. – URL: <https://www.statsmodels.org/v0.13.5/>. (дата обращения: 27.03.2025).

Приложение А. Свидетельства о регистрации программ и баз данных



Рисунок А.1 – Свидетельство о государственной регистрации базы данных

№2022620268

Авторский вклад: реализация базы данных, тестирование



Рисунок А.2 – Свидетельство о государственной регистрации программы
для ЭВМ №2021681401

Авторский вклад: Проектирование интерфейсов, тестирование объектной модели, реализация программы



Рисунок А.3 – Свидетельство о государственной регистрации программы для
ЭВМ №2022263763

Авторский вклад: Проектирование интересов, реализация алгоритмов
программы



Рисунок А.4 – Свидетельство о государственной регистрации программы
для ЭВМ №2022682299

Авторский вклад: Проектирование пользовательских интерфейсов,
тестирование объектной модели



Рисунок А.5 – Свидетельство о государственной регистрации программы
для ЭВМ №2024686216

Авторский вклад: проектирование пользовательских интерфейсов,
реализация пользовательских интерфейсов и интеграционное тестирование



Рисунок А.6 – Свидетельство о государственной регистрации программы
для ЭВМ №2024686190

Авторский вклад: проектирование пользовательских интерфейсов,
реализация программы