Министерство науки и высшего образования Российской Федерации
Федеральное государственное
автономное образовательное учреждение высшего образования
«Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»
Уральский энергетический институт
Кафедра «Электротехника»

ДОПУСТИТЬ К ЗАЩИТЕ ПЕРЕД ГЭК
Зав. кафедрой электротехники
_____ В.Э. Фризен
«13» июня 2024 г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

ПРОГНОЗИРОВАНИЕ ВЫРАБОТКИ ВЕТРОЭНЕРГЕТИЧЕСКИХ УСТАНОВОК НА ОСНОВЕ МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Научный руководитель: А.И. Хальясмаа _____
канд. техн. наук, доц. подпись

Нормоконтролер: С.А. Ерошенко _____
доц., канд. техн. наук подпись

Студент группы ЭНМ-221201 Д.А. Харлашкин _____
подпись

Екатеринбург
2024

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное
учреждение высшего образования
«Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»

Институт *Уральский энергетический*
Кафедра *«Электротехника»*
Направление подготовки *01.04.04 – Прикладная математика*
Образовательная программа *Искусственный интеллект в электроэнергетике*

УТВЕРЖДАЮ
Зав. кафедрой В.Э. Фризен
«15» января 2024 г.

ЗАДАНИЕ

на выполнение выпускной квалификационной работы

студента Харлашкина Дмитрия Александровича группы *ЭНМ-221201*
(фамилия, имя, отчество)

1. Тема выпускной квалификационной работы *Прогнозирование выработки ветроэнергетических установок на основе методов искусственного интеллекта*
Утверждена распоряжением по институту от «10» января 2024 г. № 33.07-05/003

2. Руководитель Хальясмаа Александра Ильмаровна, заведующий научной лабораторией ЦДЭЭ, доцент, канд. техн. наук
(Ф.И.О., должность, ученое звание, ученая степень)

3. Исходные данные к работе
3.1 Метеорологические данные для прогнозирования генерации ветроэнергетической установки 3.2 Измерения потока ветра, падающего на плоскость лопастей ветроэнергетической установки

4. Перечень демонстрационных материалов Презентация (15 слайдов)

5. Календарный план

№ п/п	Наименование этапов выполнения работы	Срок выполнения этапов работы	Отметка о выполнении
1.	*1 раздел (глава Введение)*	до 15 февраля 2024 г.	
2.	*2 раздел (глава Ветроэнергетические установки)*	до 15 марта 2024 г.	
3.	*3 раздел (глава Анализ выборки данных по генерации ветроэнергетической установки)*	до 08 мая 2024 г.	
4.	*4 раздел (глава Разработка модели прогнозирования генерации ветроэнергетической установки на основе машинного обучения)*	до 01 июня 2024 г.	
5.	*ВКР в целом*	до 06 июня 2024 г.	

Руководитель _____ А.И. Хальясмаа
(подпись) Ф.И.О.
Задание принял к исполнению 15.01.2024 _____
дата (подпись)

6. Выпускная квалификационная работа закончена «06» июня 2024 г. считаю возможным допустить Харлашкина Дмитрия Александровича к защите выпускной квалификационной работы в Государственной экзаменационной комиссии.

Руководитель_________ А.И. Хальясмаа
(подпись) Ф.И.О.

7. Допустить Харлашкина Дмитрия Александровича к защите магистерской диссертации в Государственной экзаменационной комиссии (протокол заседания кафедры №___ от «___» _____ 202_г.).

Зав. кафедрой _____ В.Э. Фризен
(подпись) Ф.И.О.

СОДЕРЖАНИЕ

РЕФЕРАТ 5
ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ 6
ПЕРЕЧЕНЬ СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ 7
ВВЕДЕНИЕ 8
1 ОБЗОР ОБЩЕГО СОСТОЯНИЯ И ПЕРСПЕКТИВЫ РАЗВИТИЯ ВИЭ 12
1.1 Обзор развития ВИЭ в мире 12
1.2 Обзор развития ВИЭ в России 17
1.3 Обзор методов прогнозирования выработки 25
2 ВЕТРОЭНЕРГЕТИЧЕСКИЕ УСТАНОВКИ. 29
2.1 Общая информация 29
2.2 Технические решения и перспективы развития 33
3 АНАЛИЗ ВЫБОРКИ ДАННЫХ ПО ГЕНЕРАЦИИ ВЕТРОЭНЕРГЕТИЧЕСКОЙ УСТАНОВКИ 42
3.1 Используемая выборка данных 42
3.2 Предварительная обработка и визуализация данных 42
4 РАЗРАБОТКА МОДЕЛИ ПРОГНОЗИРОВАНИЯ ГЕНЕРАЦИИ ВЕТРОЭНЕРГЕТИЧЕСКОЙ УСТАНОВКИ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ 61
4.1 Подключение библиотек обработки данных и машинного обучения 61
4.2 Выделение целевой переменной и признаков 61
4.3 Преобразование данных 62
4.4 Оценка значимости признаков 62
4.5 Построение моделей машинного обучения 64
4.6 Результаты тестирования моделей 68
ЗАКЛЮЧЕНИЕ 72
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 74
ПРИЛОЖЕНИЕ А ФРАГМЕНТ КОДА ДЛЯ ПРЕОБРАЗОВАНИЯ ДАННЫХ 80

РЕФЕРАТ

Отчет 87 с., 50 рис., 3 табл., 44 источн., 1 прил.

Ключевые слова: возобновляемые источники энергии, ветроэнергетические установки, искусственный интеллект, машинное обучение, прогнозирование генерации электрической энергии ветроэнергетическими установками, ансамбли деревьев решений.

Объектом исследования данной работы являются системы искусственного интеллекта для прогнозирования генерации ветроэнергетических установок.

Предметом исследования являются математические модели и алгоритмы машинного обучения для краткосрочного прогнозирования генерации ветроэнергетических установок.

Цель работы – оценка существующих подходов к краткосрочному прогнозированию генерации ветроэнергетических станций, разработка и апробация методов прогнозирования выработки ветроэнергетических установок на основе алгоритмов искусственного интеллекта, повышающих точность прогноза.

Результатом работы стала разработка алгоритма для прогнозирования генерации ветроэнергетической установки с использованием технологий искусственного интеллекта.

Установлено значительное улучшение точности расчета генерации ветровой турбины с помощью ансамблей регрессионных деревьев решений по сравнению с текущей моделью, основанной на зависимости выходной мощности ветроэнергетической установки от скорости ветра.

ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ

В настоящей работе применяют следующие термины с соответствующими определениями.

Возобновляемые источники энергии – это природные ресурсы или процессы, из которых получают энергию и которые по человеческим меркам являются неистощаемыми или возобновляются быстрее, чем расходуются.

Выброс данных – это значения, которые сильно отличаются от большинства других значений в наборе данных.

Нормализация данных – это процесс приведения разных масштабов и единиц измерения к единому виду.

Обучающая выборка данных – выборка представляет собой набор данных, на которых модель обучается, т.е. на основе которых она настраивает свои веса и определяет зависимости между входными данными и целевой переменной.

Тестовая выборка данных – представляет собой независимый набор данных, на которых модель не обучалась, но который используется для оценки ее качества и точности предсказаний.

Оншорный ветропарк – это группы ветровых турбин, установленных на суше.

Офшорный ветропарк – это группы ветровых турбин, установленных в море или океане, с фиксированным фундаментом или имеющие плавающее основание, генерирующие электроэнергию из ветра.

ПЕРЕЧЕНЬ СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

В настоящем ВКР применяют следующие сокращения и обозначения.

Ветровая электростанция	– ВЭС
Ветроэнергетические установки	– ВЭУ
Возобновляемые источники энергии	– ВИЭ
Гигаватт	– ГВт
Искусственный интеллект	– ИИ
Линии электропередачи	– ЛЭП
Малая гидроэлектростанция	– МГЭС
Мегаватт	– МВт
Международное энергетическое агентство	– МЭА
Солнечная электростанция	– СЭС
Тераватт-час	– ТВтч
Технологически изолированные и труднодоступные территории	– ТИТТ
Традиционные источники энергии	– ТЭР
Электроэнергетические системы	– ЭЭС

ВВЕДЕНИЕ

Эффективное функционирование электроэнергетического сектора напрямую влияет на социальное и экономическое развитие. Поэтому обеспечение надежного электроснабжения населения является приоритетной задачей.

С начала промышленной революции ископаемое топливо стало основным источником энергии для промышленного производства, что привело к увеличению загрязнения атмосферы.

В связи с этим, разработка экологически чистых и энергоэффективных технологий является критически важной для решения сложных экологических проблем. Одним из способов решения этого вопроса является развитие возобновляемых источников энергии.

Функционирование электроэнергетических систем с высокой долей генерации на основе ВИЭ отличается от принципов работы традиционных ЭЭС. Одним из основных отличий является невозможность управления объемами генерации на ВИЭ-объектах. Это требует решения проблемы резервирования ВИЭ-генерации в периоды снижения генерирующей способности.

Существует два основных подхода к решению этой проблемы: использование систем накопления электрической энергии и подключение безуглеродных ЭЭС к традиционным ЭЭС для параллельной работы. Оба этих подхода позволяют обеспечить стабильность энергоснабжения при высокой доле ВИЭ-генерации в электроэнергетической системе.

Одной из основных проблем при эксплуатации ВИЭ является стохастический характер их генерации, что создает высокую степень неопределенности при планировании режимов электроэнергетических систем. Изменение времени суток и погодных условий влияет на режим работы ВИЭ, что в свою очередь влияет на перетоки мощности в сети.

Рассмотрим ветроэнергетику, как часть ВИЭ и одну из наиболее перспективных и экологически чистых форм генерации электроэнергии,

активно развивающуюся по всему миру. Однако, в отличие от традиционных источников энергии, ветроэнергетические установки подвержены значительным колебаниям в производительности, обусловленным изменчивостью скорости ветра и других метеорологических условий.

Из-за этого остро встает вопрос прогнозирования выработки электроэнергии ВЭУ.

Точное прогнозирование выработки ветроэнергетических установок позволяет оптимизировать работу энергосистем, улучшая их стабильность и надежность, особенно в условиях увеличения доли возобновляемых источников в общем энергобалансе.

Актуальность данного исследования обусловлена несколькими факторами: необходимостью повышения надежности и эффективности ветроэнергетики, развитием методов ИИ, экономическим и экологическим аспектами.

В данной диссертации будет проведено исследование нескольких алгоритмов искусственного интеллекта для прогнозирования выработки ветроэнергетических установок.

Для каждого алгоритма будут проведены следующие этапы исследования: подготовка данных, включающая предварительную обработку данных, удаление выбросов, нормализация и разделение на обучающую и тестовую выборки; обучение модели, включающее настройку параметров алгоритма и обучение модели на обучающих данных; оценка производительности, включающая оценку точности прогнозирования на тестовых данных и сравнение с другими алгоритмами; анализ результатов, включающий интерпретацию полученных результатов, выявление преимуществ и недостатков каждого алгоритма.

В результате исследования будет определен наиболее эффективный алгоритм ИИ для прогнозирования выработки ВЭУ, а также будут предложены рекомендации по его использованию в практической деятельности.

Объектом исследования данной работы являются системы искусственного интеллекта для прогнозирования генерации ветроэнергетических установок.

Предметом исследования являются математические модели и алгоритмы машинного обучения для краткосрочного прогнозирования генерации ветроэнергетических установок.

Цель работы – оценка существующих подходов к краткосрочному прогнозированию генерации ветроэнергетических станций, разработка и апробация методов прогнозирования выработки ветроэнергетических установок на основе алгоритмов искусственного интеллекта, повышающих точность прогноза.

Задачи, поставленные для достижения цели:

- анализ существующих подходов к прогнозированию генерации ветроэнергетических станций и данных, используемых в прогнозных моделях. Оценка их преимуществ и недостатков в контексте применения к ветроэнергетике;
- оценка влияния предварительной обработки исходных данных на точность прогноза генерации ветроэнергетических станций с использованием методов машинного обучения, с целью определения необходимости применения алгоритма предварительной обработки данных для улучшения качества прогноза.;
- адаптация ансамблевых методов машинного обучения, основанных на деревьях решений, для краткосрочного прогнозирования генерации ветроэнергетических станций с точностью не менее 75%.

Научная новизна работы заключается в следующем:

- проведена адаптация регрессионных моделей, основанных на ансамблевых методах машинного обучения, для краткосрочного прогнозирования генерации ветроэнергетических станций на основе метеорологических данных из открытых источников;

- был разработан алгоритм для прогнозирования генерации ветроэнергетической установки с использованием технологий искусственного интеллекта. Установлено значительное улучшение точности расчета генерации ветровой турбины с помощью ансамблей регрессионных деревьев решений по сравнению с текущей моделью, основанной на зависимости выходной мощности ветроэнергетической установки от скорости ветра.

1 ОБЗОР ОБЩЕГО СОСТОЯНИЯ И ПЕРСПЕКТИВЫ РАЗВИТИЯ ВИЭ

1.1 Обзор развития ВИЭ в мире

Актуальность данной темы обусловлена значительным ростом установленной мощности объектов генерации электроэнергии на основе возобновляемых источников энергии в мировом масштабе с 2017 по 2022 годы. За этот период общая мощность таких установок удвоилась, возрастя с 2 250 ГВт до 3 372 ГВт [1]. Основной вклад в этот рост внесли ветровая и солнечная энергетика, на долю которых пришлось 90% от общего прироста мощности в 2022 году.

В 2022 году мощности возобновляемых источников энергии увеличились на 295 ГВт (+9,6%). Солнечная энергетика продолжала лидировать с ростом мощностей на 192 ГВт (+22%), за ней следовала ветровая энергетика с увеличением на 75 ГВт (+9%). Гидроэнергетика также показала рост мощностей на 21 ГВт (+2%), а биоэнергетика - на 8 ГВт (+5%). Геотермальная энергия, однако, увеличилась лишь на незначительные 181 МВт [1, 2].

При этом согласно отчету Международного энергетического агентства в 2023 году произошел резкий рост мощности возобновляемых источников энергии, в основном благодаря рынку ВИЭ в Китае.

Годовые приросты мощностей возобновляемых источников энергии в мире увеличились почти на 50% и достигли почти 510 гигаватт (ГВт) в 2023 году, что является самым быстрым темпом роста за последние два десятилетия. В то время как рост возобновляемых мощностей в Европе, США и Бразилии достиг рекордного уровня, ускорение роста в Китае было необычайным. В 2023 году Китай ввел в эксплуатацию столько же солнечных фотоэлектрических систем, сколько весь мир в 2022 году, при этом количество ветровых электростанций также выросло на 66% в годовом исчислении. В глобальном масштабе около 75% прироста возобновляемых мощностей во всем мире приходится на солнечные фотоэлектрические системы [3].

Согласно прогнозам, при существующей политике и рыночных условиях к 2028 году глобальная мощность возобновляемых источников энергии достигнет 7 300 ГВт. Такая траектория роста приведет к увеличению глобальной мощности в 2,5 раза по сравнению с нынешним уровнем к 2030 году. Рост ввода мощностей ВИЭ в динамике и прогнозе до 2028 года можно посмотреть на рисунке 1.1 [3].

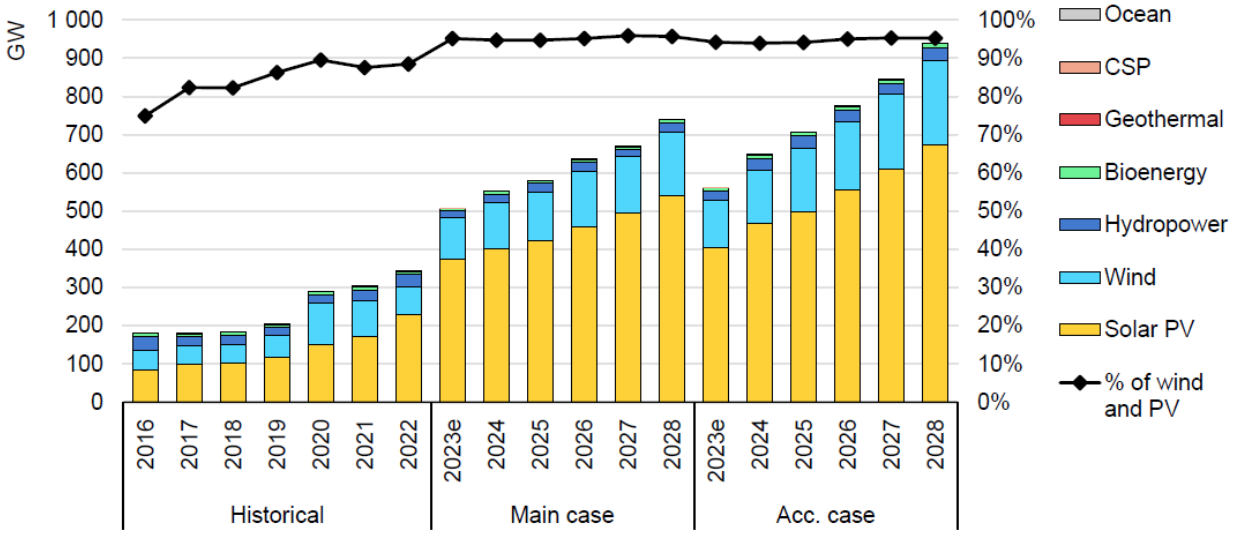


Рисунок 1.1 – Увеличение мощностей возобновляемой электроэнергии по технологиям и сегментам

На долю Китая приходится почти 60% новых возобновляемых мощностей, которые, как ожидается, будут введены в эксплуатацию во всем мире к 2028 году.

К 2028 году ожидается увеличение количества наземных ВЭС и СЭС более чем вдвое в таких странах как США, Европейский Союз, Бразилия Индия в сравнении в предыдущим периодом в пять лет.

По оценкам, в 2023 году 96% вновь установленных солнечных фотоэлектрических и береговых ветроэнергетических мощностей имели более низкие затраты на выработку, чем новые угольные и газовые электростанции.

Ожидается, что к 2028 году потенциальная выработка электроэнергии из возобновляемых источников достигнет примерно 14 400 ТВтч, что почти на

70% больше, чем в 2022 году. В течение следующих пяти лет можно будет достичь нескольких важных показателей в области возобновляемой энергетики [3]:

- в 2024 году переменная генерация из возобновляемых источников превзойдет гидроэнергетику;
- в 2025 году возобновляемые источники энергии превзойдут производство электроэнергии на угле;
- в 2025 году ветер превзойдет производство ядерной электроэнергии;
- в 2026 году солнечная фотоэлектрическая энергия превзойдет производство ядерной электроэнергии;
- в 2028 году солнечная фотоэлектрическая энергия превзойдет выработку ветровой электроэнергии.

В течение прогнозируемого периода потенциальный рост производства электроэнергии из возобновляемых источников превысит рост глобального спроса, что указывает на медленное снижение производства электроэнергии на основе угля, в то время как производство электроэнергии на основе природного газа остается стабильным. К 2028 году ВИЭ будут обеспечивать 42% мирового производства электроэнергии, при этом процент производства солнечной и ветровой и энергии составит порядка 25%.

При этом в целом, ветроэнергетика во всем мире сталкивается с финансовыми проблемами, большинство крупных западных производителей сообщили об убытках за последние два года. Поскольку в большинстве стран реализация проектов развития электроэнергетики за пределами Китая шло медленнее, чем ожидалось, прогноз ветровой генерации на суше был пересмотрен в сторону понижения.

На крупных рынках (исключая Китай), таких как Индия, расширение наземных ветровых электростанций существенно не продвигается, а прогноз был пересмотрен в сторону понижения для региона АСЕАН, Африки и Ближнего Востока из-за медленного прогресса проекта и сохраняющейся политической неопределенности. В Европейском Союзе длительное время

ожидания разрешений, проблемы с цепочкой поставок, а также более высокие затраты на оборудование и финансовые затраты также сокращают ожидаемое развертывание береговых ветровых электростанций.

Однако, наибольший рост затрат был отмечен в сфере развития морской ветроэнергетики, что привело к пересмотру прогноза на этот год на 16% за пределами Китая.

После значительного роста мощностей, обусловленного благоприятной политикой и снижением затрат в течение последнего десятилетия, стоимость материалов, используемых для производства морской ветровой энергии, снизилась.

С начала 2020 года стоимость компонентов начала расти из-за ограничений в цепочках поставок и инфляции, что привело к увеличению затрат на инвестиции в морскую ветроэнергетику более чем на 20%. Кроме того, более высокие затраты на финансирование оказывают дополнительное давление на застройщиков, которые заключили соглашения о покупке электроэнергии при низкой стоимости оборудования и займов до начала глобального энергетического кризиса.

Две крупнейшие страны по расширению морских ветряных электростанций за пределами Китая - США и Великобритания - испытывают значительное влияние изменяющихся экономических условий. В Соединенных Штатах более 10 ГВт мощности, ранее предоставленных в рамках государственных тендеров, были отложены или находятся под угрозой задержки или отмены, что привело к пересмотру прогноза развития морской ветроэнергетики в США на более чем 60% в сторону понижения. Экономика любого проекта является основным фактором. Некоторые разработчики, отменившие выигранные проекты, планируют подать заявку на строительство дополнительных мощностей на будущих аукционах, предположительно по более высокой цене.

Высокий рост производства электроэнергии из возобновляемых источников, в сочетании с сильной политической поддержкой, привел к

снижению затрат на производство ветровой и солнечной фотоэлектрической энергии более чем на 80% с 2010 года. Однако значительная часть этого роста в странах с развитой экономикой произошла в период количественного смягчения, когда базовые процентные ставки центральных банков (без учета премий за проект и риска компании) были ниже 1% в Европейском Союзе, Великобритании и США. В развивающихся странах процентные ставки были значительно выше в тот же период, но это было частично компенсировано более низкими инвестиционными затратами, особенно в Китае и Индии. Доступность недорогого финансирования, поддерживаемого правительствами и международными финансовыми институтами, также способствовала росту [4].

В 2023 году новые мощности возобновляемой энергетики, финансируемые в странах с развитой экономикой, столкнулись с более высокими базовыми процентными ставками, чем в Китае и в среднем по миру, главным образом потому, что китайское правительство решило снизить процентные ставки, а не следовать мировой тенденции. На развивающихся рынках и в развивающихся странах разработчики возобновляемых источников энергии всегда получали инвестиции по более высоким процентным ставкам, что приводило к более высоким затратам, препятствующим более быстрому расширению мощностей возобновляемых источников энергии. Отсутствие доступного финансирования остается самой серьезной проблемой для развития проектов возобновляемой энергетики в большинстве стран.

Последствия этой новой макроэкономической среды многообразны. Во-первых, инфляция привела к увеличению стоимости оборудования для наземных и морских ветряных и, частично, солнечных фотоэлектрических систем. Во-вторых, более высокие процентные ставки увеличивают стоимость финансирования капиталоемких технологий с использованием возобновляемых источников энергии. В-третьих, запоздалая политическая реакция (поскольку правительства ожидали продолжения снижения затрат в последнее десятилетие) и проблемы с выдачей разрешений привели к тому,

что несколько аукционов в странах с развитой экономикой, особенно в Европе, остались без реализации. Кроме того, девелоперам, чьи контракты на покупку электроэнергии были подписаны до этих макроэкономических изменений, пришлось отказаться от своих проектов или пересмотреть их финансовые показатели. Это может привести к снижению объемов инвестиций в возобновляемые источники энергии и замедлению их развития.

Для дальнейшего развития необходимо продолжать проведение исследований и разработок в области технологий возобновляемой энергии, чтобы снизить затраты на производство и улучшить эффективность систем. Это позволит сделать возобновляемые источники энергии более конкурентоспособными на рынке и ускорить их внедрение.

Тем не менее, развитие ВИЭ-генерации является одним из ключевых направлений мировых программ по борьбе с климатическими изменениями и формированию «зеленой» энергетики.

Сокращение использования ископаемого топлива и выбросов в энергетическом секторе считается жизненно важным для достижения глобальных климатических целей. Более 100 стран на климатическом саммите COP28 в Дубае в 2023 году согласились утроить мощности возобновляемых источников энергии к 2030 году [5].

ВИЭ-объекты могут служить альтернативой централизованному электроснабжению для удаленных потребителей, при условии, что местные ВИЭ могут обеспечивать электроснабжение прилегающих территорий.

В таких случаях строительство ВИЭ-объекта вместо дорогостоящих протяженных линий электропередачи является экономически целесообразным. Государственная поддержка ВИЭ включает частичное субсидирование строительства таких объектов на конкурсной основе.

1.2 Обзор развития ВИЭ в России

Хотя Россия и располагает значительными запасами традиционных источников энергии, развитие ВИЭ является важным стратегическим курсом для будущего всего энергетического сектора страны.

В России существует разнообразие видов возобновляемой энергии в различных регионах, которое определяется географическими особенностями, местным климатом и рельефом. В каждом регионе страны существует один или два типа ВИЭ, которые экономически оправданы конкретно для этой местности, но в некоторых местах можно найти все виды возобновляемой энергии, как, например, в Республике Дагестан. Это отличает Россию от других стран, где из-за однородности географических условий преобладает или используется только один вид возобновляемой энергии. [6].

Несмотря на огромные ресурсы ветровой, геотермальной, солнечной энергии, гидроэнергетические ресурсы и энергии биомассы, доля ВИЭ в энергобалансе страны на сегодняшний день составляет скромные 2-3%.

Динамика совокупной установленной мощности объектов ВИЭ-генерации в России приведена на рисунке 1.2[7].

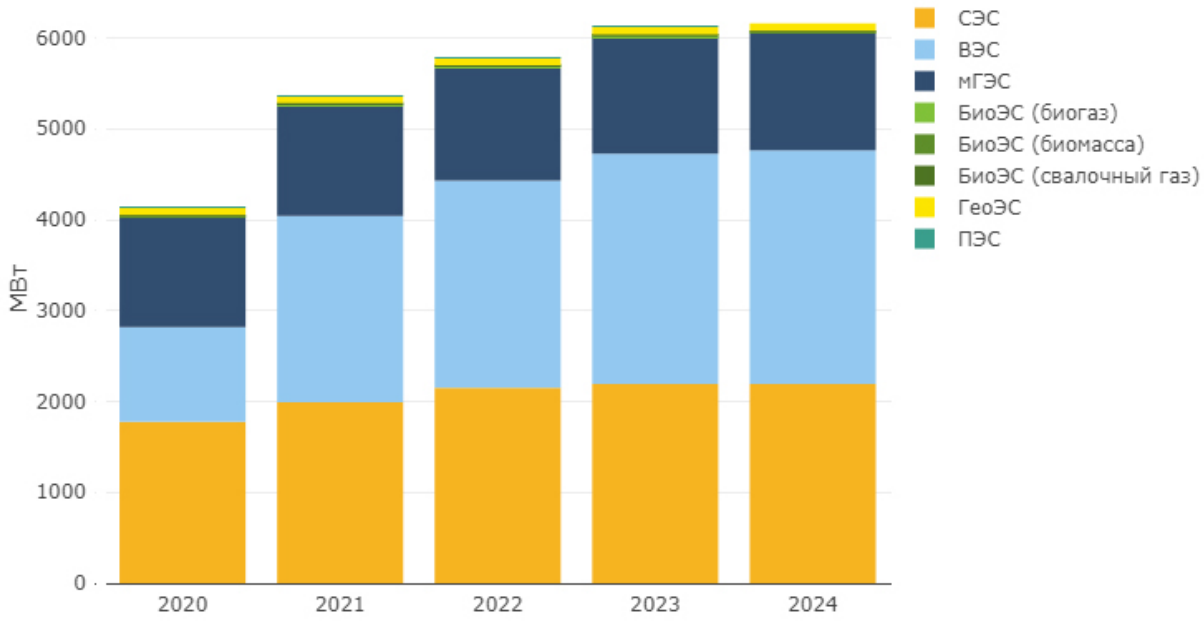


Рисунок 1.2 – Динамика совокупной установленной мощности объектов ВИЭ-генерации в России

По мнению многих специалистов, недостаточное внимание развитию ВИЭ в России связано с рядом объективных и субъективных факторов, включая невысокую экономическую конкурентоспособность ВИЭ по отношению к топливной энергетике (цены и тарифы на электрическую и

тепловую энергию в районах централизованного энергоснабжения России существенно ниже, чем в других странах), огромные запасы ТЭР и большой резерв мощности объединенных энергосистем (от 3 до 16 ГВт) при низких темпах роста электропотребления (прогнозный рост – 1 – 1,3 % ежегодно) .

Ускоренное развитие возобновляемых источников энергии в России необходимо по нескольким причинам.

Во-первых, это связано с необходимостью обеспечения энергетической безопасности технологически изолированных и труднодоступных территорий, которые находятся вне систем централизованного энергоснабжения. В таких регионах технологии использования ВИЭ могут быть конкурентоспособными и обеспечивать стабильное энергоснабжение.

Во-вторых, развитие ВИЭ в России является важным фактором модернизации экономики. Это обусловлено появлением инновационных предприятий, разработкой передовых новых технологий, поддержкой малого и среднего бизнеса, формированием новых рабочих мест, улучшением социальной сферы и экологической обстановки.

Кроме того, ускоренное развитие ВИЭ в России необходимо для создания надежного задела в инновационном развитии энергетики страны для будущих поколений и снижения негативного воздействия традиционных источников энергии на окружающую среду [6].

Если в ТИТТ использование возобновляемых источников энергии уже может конкурировать на рынке, то для централизованных систем энергоснабжения часто необходимы меры государственной поддержки, аналогичные тем, которые применяются в других странах.

Таким образом, ускоренное развитие ВИЭ в России является необходимым для обеспечения энергетической безопасности, модернизации экономики и создания надежного задела в инновационном развитии энергетики страны для будущих поколений.

Технологически изолированные и труднодоступные территории России — это регионы, не подключенные к Общероссийской энергетической системе,

где электроснабжение осуществляется за счет автономных, децентрализованных источников. Тарифы на электроэнергию в этих районах устанавливаются местными властями исходя из фактических затрат на ее производство. Большая часть таких территорий находится в северных и приравненных к ним регионах, включая некоторые участки Мурманской, Архангельской, Магаданской, Амурской и Сахалинской областей, а также Республики Саха (Якутия), Карелия, Чукотского, Ханты-Мансийского и Ямало-Ненецкого автономных округов. Эти места можно охарактеризовать такими особенностями в области энергоснабжения как:

- малая плотность жителей и их неравномерное расселение охватывают примерно две трети территории страны, где проживает более 15 миллионов человек, что составляет приблизительно 10% от общей численности населения страны;
- энергоснабжение недостаточно надежно из-за износа генерирующего оборудования, что требует больших резервов и увеличивает расход топлива. Высокий износ также повышает риск аварий, особенно критичных в условиях, когда эти объекты часто являются единственным источником электроэнергии;
- высокие затраты на доставку дизельного топлива для примерно 900 дизельных электростанций, общая годовая выработка которых составляет около 2,54 миллиарда кВт·ч.

Международный и отечественный опыт показывает, что оптимизация и модернизация систем энергоснабжения в ТИТТ России может быть эффективно достигнута через развитие распределенной генерации с использованием возобновляемых источников энергии. Этот подход позволит значительно сократить потребление дорогостоящих топливных ресурсов, увеличить надежность энергоснабжения, снизить стоимость производимой электроэнергии и улучшить экологическую ситуацию в этих регионах. Развитие распределенной генерации, включая использование ВИЭ, в ТИТТ

имеет важное стратегическое значение для всей страны и отражено в ряде программных и стратегических документов развития энергетики РФ [6].

В некоторых регионах России (например, Приморский край) предлагается развитие распределительного электросетевого комплекса, чтобы объединить изолированные энергорайоны с зонами централизованного энергоснабжения и заменить дизельные генераторы. Однако такой подход ограничен экономическими, климатическими и территориальными факторами и возможен только в небольшом количестве случаев.

В других регионах России предпочитают модернизацию генерации на привозном топливе, включая проекты возобновляемых источников энергии, с использованием существующих механизмов привлечения инвестиций [8].

Для осуществления модернизации объектов производства электрической энергии в ТИТТ России можно воспользоваться энергосервисным договором, который предусмотрен действующим законодательством. Энергосервисный договор позволяет возвратить вложенные средства за счет экономии топливных затрат, достигнутой благодаря повышению энергетической эффективности объекта. В случае объектов электроэнергетики это может быть достигнуто с помощью современного генерирующего оборудования с низким удельным расходом топлива, автоматизации, ВИЭ и систем накопления энергии.

В настоящее время проекты ВИЭ в ТИТТ РФ осуществляются децентрализованно и пока не наблюдается общей тенденции в данных регионах, несмотря на принятие ряда механизмов поддержки на федеральном уровне.

К концу 2021 года в 8 регионах России, относящихся к ТИТТ, эксплуатировались 43 объекта генерации ВИЭ общая установленная электрическая мощность которых составил 9,68 МВт. Суммарный объем выработки этих генерирующих объектов за 2021 год составил 23,65 ГВт·ч, что эквивалентно сжиганию около 5 675 тонн дизельного топлива [9].

В конце 2019 года в России был принят Федеральный закон № 471-ФЗ (внесен поправками в действующий ФЗ № 35 [10]), который установил право для любого физического или юридического лица, владеющего объектом микрогенерации, продавать избыток выработанной электрической энергии в сеть, а сбытовая организация обязана ее приобрести. Законодательный акт дает определение объекту микрогенерации и устанавливает условия, выполнение которых относит генерирующий объект к этой категории.

Объект микрогенерации - это объект по производству электрической энергии, принадлежащий на праве собственности или ином законном основании потребителю электрической энергии. Его энергопринимающие устройства технологически присоединены к объектам электросетевого хозяйства с уровнем напряжения до 1 000 вольт. Он может функционировать на основе использования возобновляемых источников энергии и используется указанным потребителем для производства электрической энергии в целях удовлетворения собственных бытовых и (или) производственных нужд, а также для продажи в порядке, установленном основными положениями функционирования розничных рынков. Объем выдачи электрической энергии таким объектом по производству электрической энергии в электрическую сеть не превышает величину максимальной присоединенной мощности энергопринимающих устройств указанного потребителя и составляет не более 15 кВт. Для подключения данного объекта к электрической сети не применяется электрическое оборудование, предназначенное для эксплуатации в более чем одном помещении здания, включая оборудование, входящее в состав общего имущества многоквартирного дома.

Согласно указанным документам, для начала подачи электроэнергии в сеть требуется выполнить ряд действий.

При осуществлении технологического присоединения региональная сетевая компания обязана безвозмездно установить специальный двунаправленный смарт-счетчик, который обеспечивает почасовые измерения потребляемой из сети и отпускаемой в сеть электрической энергии. Анализ

потребления и производства энергии объектом микрогенерации осуществляется энергосбытовой компанией в удаленном режиме. [11, 12].

Механизм работы объекта микрогенерации заключается в следующем: сначала вырабатываемая электроэнергия используется для собственных нужд частного потребителя по приоритетной схеме. Если же электроэнергии оказывается в избытке, она направляется во внешнюю сеть, которая в данном случае выступает в роли идеального аккумулятора энергии, способного беспрепятственно накапливать и выдавать неограниченную мощность без потерь и инерции. При необходимости потребитель может получить недостающую энергию из внешней сети на условиях взаимозачета (сальдирования) или продать излишки и получить за них оплату.

В качестве дополнительного стимула, помимо возможности сокращения покупки электроэнергии и продажи избытка, до 2029 года доход, получаемый владельцем объекта микрогенерации, не будет рассматриваться как предпринимательская деятельность и, соответственно, не будет подлежать налогообложению.

Прогнозы многих экспертов были весьма оптимистичными, но рынок микрогенерации в России не получил широкого распространения на данный момент. Так ПАО «Россети» получили всего 118 заявок на технологическое присоединение объектов микрогенерации с начала ввода механизма поддержки. Общая мощность заявок была всего 1,23 МВт. По оценкам Минэнерго России, в ближайшее время объемы ввода микрогенерации на основе ВИЭ составят 15-30 МВт в год. При этом эксперты считают вполне достижимым результатом рынка микрогенерации на основе возобновляемых источников энергии без негативного воздействия на энергосистему до 15 ГВт. [13].

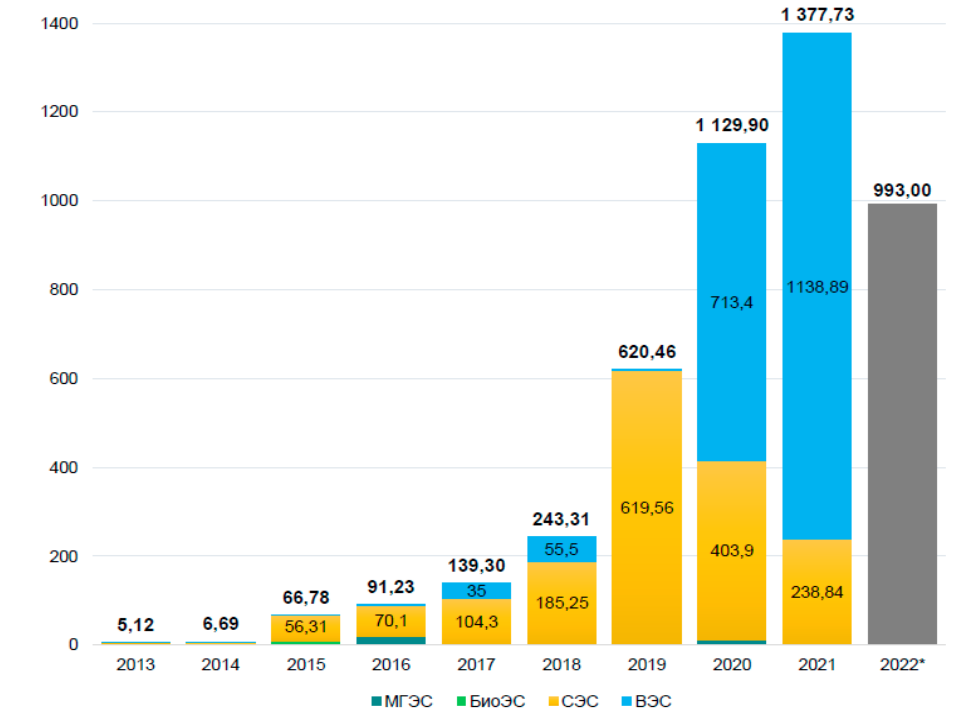
За время функционирования механизма поддержки и стимулирования развития ВИЭ с 2013 по 2021 гг. в России введены в эксплуатацию 121 электростанция ВИЭ, в том числе:

- 83 СЭС мощностью 1 678,6 МВт;

- 27 ВЭС мощностью 1 946,1 МВт;
- 11 МГЭС мощностью 58,1 МВт.

Так, например, дивизион государственной корпорации «Росатом» АО «НоваВинд» [14] запустила в работу Кочубеевскую ВЭС установленная электрическая мощность которой составила 210 МВт, что делает ее самой крупной действующей ВЭС в России. Станция расположена в Ставропольском крае, в ее состав входят 84 ветроэнергетических установок мощностью по 2,5 МВт каждая, размещенных на площади 75 гектаров. Планируемая среднегодовая выработка электроэнергии на Кочубеевской ВЭС составляет 597 млн кВт·ч. Однако фактически выработка электроэнергии за 2021 год составила 528,5 млн кВт·ч. В строительство Кочубеевской ВЭС было инвестировано более 28 млрд рублей.

Динамика ввода генерирующих мощностей ВИЭ в России показана на рисунке 1.3.



* – прогноз ввода мощностей ВИЭ в 2022 году по данным АРВЭ: СЭС – 138 МВт, ВЭС – 731 МВт, МГЭС – 124 МВт

Рисунок 1.3 – Ввод генерирующих мощностей ВИЭ в России в 2013 – 2021 гг., МВт

1.3 Обзор методов прогнозирования выработки

Однако развитие возобновляемых источников энергии невозможно без точного прогнозирования производства электроэнергии. Точные прогнозы работы ветрогенераторов позволяют оптимизировать функционирование энергосистем, улучшая их стабильность и надежность, особенно в условиях увеличения доли возобновляемых источников энергии в общем энергетическом балансе.

Современные алгоритмы машинного обучения и глубокого обучения показывают высокую эффективность в решении задач прогнозирования, включая анализ сложных и нелинейных зависимостей в больших объемах данных. Применение этих методов к прогнозированию выработки ВЭУ может значительно улучшить качество прогнозов по сравнению с традиционными статистическими подходами. Точные прогнозы помогают энергопредприятиям минимизировать затраты на хранение энергии и покупку дефицитной электроэнергии на рынке в периоды пикового спроса, а также оптимизировать расходы на поддержание и обслуживание ветроустановок, управляя нагрузкой в зависимости от предсказываемой выработки. Увеличение доли ветроэнергии в энергобалансе способствует снижению выбросов парниковых газов и других вредных веществ. Точное прогнозирование помогает максимизировать использование ветроэнергии, минимизируя при этом зависимость от ископаемых топлив. [15, 16]

Например, алгоритмы машинного обучения могут использовать исторические данные о скорости ветра, направлении ветра, температуре воздуха и других факторах для построения модели прогнозирования выработки электроэнергии. Эти модели могут учитывать не только линейные, но и нелинейные зависимости между факторами и выработкой электроэнергии [17].

Кроме того, использование методов искусственного интеллекта позволяет учитывать сложные взаимосвязи между факторами и адаптироваться к изменяющимся условиям. Например, алгоритмы могут

автоматически обнаруживать и учитывать изменения в топографии местности или изменения в техническом состоянии ветротурбин [18].

Существует несколько методов прогнозирования выработки ветроэнергии, которые используют различные подходы и данные для создания прогнозов.

Статистические модели, такие как авторегрессионные модели (AR), интегрированные скользящие средние (ARIMA) и экспоненциальное сглаживание, используют исторические данные о скорости ветра и других метеорологических параметрах для создания прогнозов. Эти модели основаны на статистических закономерностях и трендах в данных.

Физические модели, такие как численное моделирование атмосферы (NWP), используют физические законы и уравнения для моделирования атмосферных процессов и прогнозирования скорости ветра. Эти модели требуют большого объема данных и вычислительных ресурсов, но могут предоставить более точные прогнозы на большие временные горизонты [19, 20].

Методы машинного обучения, такие как регрессионные модели, нейронные сети и случайные леса, используют исторические данные о скорости ветра, метеорологических параметрах и других факторах для создания прогнозов. Эти модели могут учитывать нелинейные зависимости и сложные взаимосвязи между факторами.

Гибридные модели объединяют различные методы, такие как статистические модели, физические модели и машинное обучение, для создания более точных прогнозов. Например, гибридная модель может использовать физическую модель для создания базового прогноза, а затем применить машинное обучение для уточнения прогноза на основе исторических данных.

Интеллектуальные системы, такие как экспертные системы и системы искусственного интеллекта, используют знания и опыт экспертов, а также данные и модели для создания прогнозов. Эти системы могут учитывать не

только метеорологические данные, но и другие факторы, такие как состояние ветротурбин и энергетической системы.

Каждый метод имеет свои преимущества и ограничения, и выбор метода зависит от доступных данных, требуемой точности прогноза и временного горизонта прогнозирования. Важно отметить, что комбинирование различных методов может привести к более точным и надежным прогнозам выработки ветроэнергии.

Традиционные подходы к прогнозированию выработки ветроэнергии, такие как статистические модели и физические модели, имеют некоторые недостатки, которые могут быть преодолены с помощью применения искусственного интеллекта.

Так, подобные подходы часто сталкиваются с проблемой недостаточной точности прогнозов, особенно на коротких временных горизонтах. ИИ, в частности методы машинного обучения, могут обучаться на большом объеме исторических данных и учитывать сложные нелинейные зависимости между факторами, что позволяет достичь более точных прогнозов.

Традиционные подходы часто ограничены в учете множества факторов, которые могут влиять на выработку ветроэнергии. ИИ позволяет учитывать большое количество переменных, включая метеорологические данные, состояние ветротурбин и энергетической системы, что позволяет получить более полную картину и улучшить прогнозы.

Также эти подходы могут быть неэффективными при адаптации к изменяющимся условиям, таким как изменение топографии местности или изменение состояния ветротурбин. ИИ позволяет автоматически обнаруживать и учитывать эти изменения, что позволяет создавать более адаптивные прогнозы.

Традиционные подходы часто ограничены в использовании только одного типа данных, такого как скорость ветра. ИИ позволяет интегрировать различные источники данных, включая метеорологические данные, данные

сенсоров ветротурбин и данные энергетической системы, что позволяет получить более полную информацию для прогнозирования [21].

ИИ позволяет автоматизировать процесс прогнозирования и оптимизировать его на основе обратной связи. Это позволяет сократить время и ресурсы, необходимые для создания прогнозов, и повысить эффективность работы ветроэнергетических установок [22].

Таким образом, вышеописанные факторы подчеркивают важность разработки и внедрения эффективных методов прогнозирования выработки электроэнергии ВЭУ, особенно с использованием современных технологий искусственного интеллекта. Такие методы могут значительно улучшить планирование производства и распределения электроэнергии, а также способствовать дальнейшему расширению использования возобновляемых источников энергии.

2 ВЕТРОЭНЕРГЕТИЧЕСКИЕ УСТАНОВКИ.

2.1 Общая информация

Использование энергии ветра началось еще в древние времена. Первые упоминания о ветряных мельницах датируются VII веком в Сасанидском Иране, где они использовались для помола зерна. В средние века ветряные мельницы распространились по всей Европе и стали неотъемлемой частью сельской местности, служа не только для помола зерна, но и для откачки воды.

В XIX веке ветряные двигатели начали использоваться в США для электрификации отдаленных ферм. Эти устройства, известные как ветрогенераторы, позволяли получать электроэнергию для освещения и работы небольших электроприборов. Однако с развитием централизованной энергетики интерес к ветроэнергетике временно снизился.

Возрождение интереса к ветроэнергетике началось в 1970-х годах в связи с энергетическим кризисом и необходимостью поиска альтернативных источников энергии. С тех пор технологии ветроэнергетики значительно продвинулись: увеличилась мощность ветротурбин, повысилась их эффективность и надежность, снизились затраты на производство электроэнергии.

Современные ветроэлектростанции могут генерировать сотни мегаватт электроэнергии, что делает их важным компонентом энергетической системы многих стран. Ветроэнергетика активно развивается не только в Европе и Северной Америке, но и в Азии, Латинской Америке и Африке.

Перспективы развития ветроэнергетики выглядят очень обнадеживающе. Согласно прогнозам Международного энергетического агентства, ветроэнергетика может стать крупнейшим источником электроэнергии в мире к 2050 году. Это будет способствовать сокращению выбросов парниковых газов и поможет в достижении целей по борьбе с климатическим изменением.

Одним из ключевых направлений развития является увеличение масштабов использования офшорной ветроэнергетики, где ветротурбины

устанавливаются в море. Это позволяет использовать более сильные и стабильные ветры, что повышает эффективность генерации электроэнергии. Также активно разрабатываются новые технологии, например, ветротурбины вертикальной оси, которые могут работать в условиях слабого ветра и более компактны, что делает их подходящим для использования в городских условиях и на ограниченных территориях.

Кроме того, важным аспектом развития ветроэнергетики является улучшение хранения и передачи электроэнергии. Ветровая энергия является переменной и зависит от погодных условий, поэтому необходимы эффективные системы аккумулирования энергии для обеспечения стабильного энергоснабжения. Развитие технологий, таких как литий-ионные батареи и гидроаккумулирующие системы, позволяет улучшить гибкость и надежность ветроэнергетических систем.

Также важным направлением является интеграция ветроэнергетики в существующие энергетические системы. Для этого необходимо разрабатывать более умные и гибкие сети, способные эффективно управлять переменным потоком энергии от ветротурбин. Применение искусственного интеллекта и машинного обучения в управлении энергетическими системами позволяет оптимизировать процессы генерации, передачи и потребления электроэнергии, уменьшая потери и повышая эффективность [23].

Ветроэнергетика - это процесс преобразования кинетической энергии ветра в электрическую энергию с помощью ветроустановок. Ветроустановки состоят из ветротурбин, генератора и системы передачи энергии. Ветроустановки работают на принципе преобразования энергии ветра в механическую. Лопасти ветротурбины, под действием ветра, начинают вращаться. Это вращение передается генератору, который в результате производит электричество.

Первоисточником возникновения ветра, а также многих других динамических процессов, происходящих на земной поверхности и в её атмосфере, является энергия Солнца. Прямое поглощение солнечного

излучения атмосферой минимально, так как активно его поглощают лишь водяной пар, озон, углекислый газ и пыльные частицы. В отличие от этого, основные компоненты воздуха — азот и кислород — прозрачны для солнечных лучей. Таким образом, основная масса солнечной энергии передается атмосферному воздуху от нагретой земной поверхности через инфракрасное излучение, молекулярную диффузию, конвекцию и другие физические процессы.

Когда воздух нагревается у земной поверхности, он поднимается вверх, так как его удельный вес уменьшается при нагреве. Это вертикальное движение воздуха является первичным и основным в ряду последовательных процессов, которые приводят к возникновению ветра. Последовательность этих процессов включает нагревание, охлаждение и снова нагревание воздуха во время его движения относительно вращающейся вокруг своей оси земной поверхности во времени и пространстве, что приводит к формированию атмосферной циркуляции.

Под воздействием этих циркуляций перемещение воздушных масс происходит как во времени, так и в пространстве. И формы, и размеры морей, океанов, материков меняют картину циркуляций вдоль побережий. С удалением от побережий циркуляция стабилизируется, но влияние местных особенностей на скорость ветра иногда остается значительным [24].

При ветроэнергетических расчетах исходные данные об общем уровне интенсивности ветра определяются на основе метеорологических данных за длительный период времени, равный примерно 20-25 годам, соответствующий сроку эксплуатации ветроэнергетической установки.

Так, средняя скорость ветра рассчитывается как среднее арифметическое значение, полученное из ряда измерений скорости ветра, выполненных через равные интервалы времени в течение заданного периода.

Еще один рассчитываемый параметр фоновая скорость ветра – это скорость ветра, рассчитанная на условия открытой (без влияния крупных водных объектов) ровной местности в приземном слое на высоте 10 метров от

поверхности земли, средняя годовая (сезонная или месячная) скорость ветра за 20-25-летний период, оцененная на основе многолетних исследований государственных метеорологических станций и постов.

Для оценки эффективности использования ветроэнергетической установки необходимо знать среднегодовую скорость ветра на высоте опоры. Для принятия обоснованных решений о финансировании строительства ветроустановок необходимо проводить практические измерения скорости ветра на высоте установки ветротурбины (80-120 метров) в течение года. Для этого используются мачтовые конструкции с установленным оборудованием для дистанционной передачи данных о скорости ветра и других параметрах в режиме реального времени. Только расчетных значений скорости ветра при принятии подобных решений недостаточно.

Для использования энергии ветрового потока необходимо знать удельную кинетическую энергию.

Основным показателем эффективности ветроустановок служит коэффициент использования энергии ветрового потока ξ (или коэффициент использования мощности ветрового потока), который представляет собой отношение механической энергии, генерируемой ветродвигателем, к полной энергии ветра, проходящей через ометаемую площадь ветродвигателя. Идеальное значение этого коэффициента составляет 0,593. Ометаемая площадь - это площадь круга, создаваемого ветроколесом при вращении.

Эффективность использования энергии ветра определяется количеством дней в году, когда скорость ветра достаточна для работы ветроустановок. Оценка этой эффективности производится с помощью коэффициента использования мощности (capacity factor), который представляет собой отношение фактически выработанной энергии к теоретически возможной.

Типичное значение данного коэффициента - 15-40% (максимальное значение достигается для площадок с высокими ветровыми нагрузками) [24].

2.2 Технические решения и перспективы развития

Ветроэнергетические установки укрупненно можно классифицировать по мощности (маломощные 30-5000Вт; средней мощности 5-100кВт; большой мощности 100кВт-5Мвт, гигантские более 5 МВт), месту расположения (на суше – оншорные) и в шельфовых зонах водных бассейнов (оффшорные), а также по конструкции (с горизонтальной и вертикальной осью вращения ветроколеса).

Отсюда ветрогенераторы можно разделить на 3 категории:

- промышленные,
- коммерческие,
- бытовые (для частного использования).

Промышленные ВЭУ строятся, как правило, государством или крупными энергетическими корпорациями. Обычно они объединяются в сети, создавая крупные ветроэнергетические комплексы. Главное отличие ВЭС от традиционных электростанций (тепловых, атомных) заключается в том, что они не требуют сырья и не производят отходов [23]. Пример ВЭС приведен на рисунке 2.1



Рисунок 2.1 – Пример ВЭС

Как было сказано ранее существует несколько типов и конструкций ветротурбин, которые используются в ветроустановках.

На сегодняшний день наиболее распространенными типами ветрогенераторов являются горизонтально-осевые ветротурбины (HAWT) и вертикально-осевые ветротурбины (VAWT).

Горизонтально-осевые ветротурбины (HAWT) имеют три или более лопасти, расположенных на вертикальной оси, и являются наиболее распространенным типом ветрогенераторов. Они обладают высокой эффективностью и широко используются в крупных ветроэнергетических парках [24, 25].

Примеры данного типа ВЭУ показаны на рисунке 2.2



Рисунок 2.2 – Примеры горизонтально-осевых ветротурбин

Рассмотрим данный тип ВЭУ подробнее.

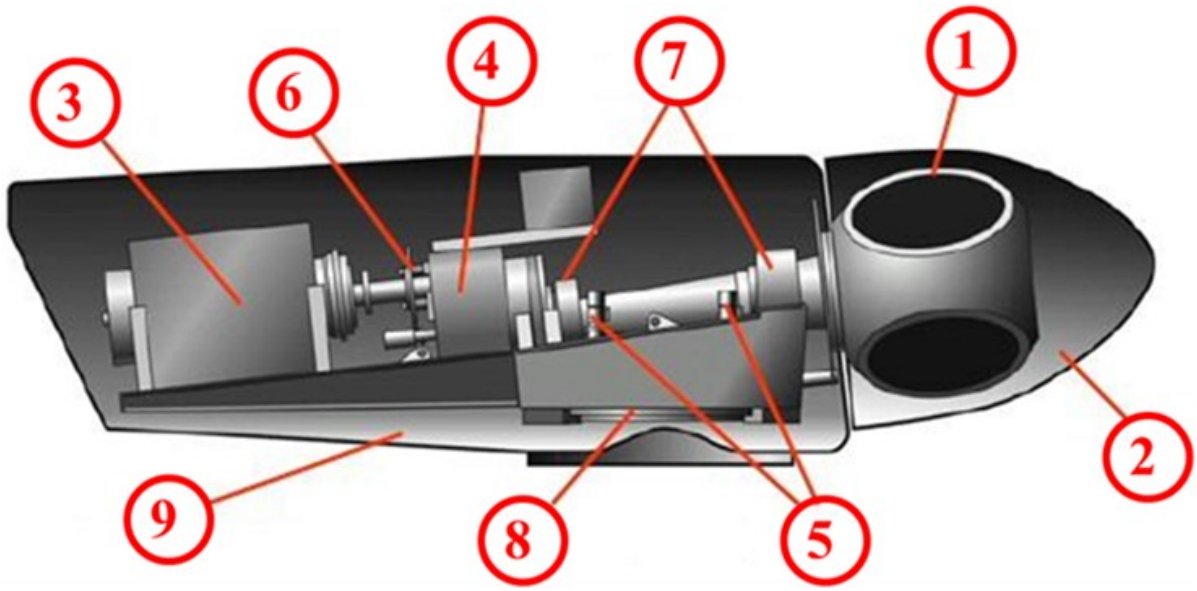
Гандола с ветроколесом называется ветрогенератором. Ветрогенератор размещается (устанавливается) в верхней части башни. Башни современных ВЭУ имеют телескопическую конструкцию и собираются на фундаменте установки из нескольких секций с использованием прочных резьбовых соединений.

Элементы конструкции, расположенные в гандоле можно увидеть на рисунке 2.3

Ротор ветрогенератора состоит из ветроколеса и первичного вала, с которого энергия ветра передается на рабочие механизмы ВЭУ.

Ветроколесо представляет собой втулку, к которой прикреплены лопасти.

Специальное устройство мультипликатор (редуктор) предназначего для повышения оборотов первичного вала до рабочего значения оборотов генератора (с 15-20 до 1800 об/мин).



1 – втулка ветроколеса; 2 – обтекатель; 3 – генератор;
4 – мультипликатор; 5 – следящая система; 6 – дисковый тормоз;
7 – основной вал; 8 – подшипник азимута; 9 – рама гондолы

Рисунок 2.3 – Элементы конструкции, расположенные в гандоле

Редуктор - это сложное механическое устройство, содержащее большое количество деталей, таких как шестерни, оси, червячные механизмы и другие. Из-за этого редуктор считается одним из наименее надежных компонентов ветрогенераторов. При его выходе из строя замена или ремонт становятся дорогостоящими операциями, особенно учитывая высоту современных ветрогенераторов (80-120 м). Фирма Enercon разработала безредукторную конструкцию ВЭУ, в которой ось (ступица) ветроколеса соединяется непосредственно с осью генератора. В данной конструкции нет

необходимости повышать обороты на валу генератора, увеличение происходит за счет использования многосекционной конструкции статора и ротора генератора [24].

Такие ВЭУ являются более надежными и долговечными, что выражается в более высокой их стоимости (на 10-20%).

Маломощные ВЭУ имеют более простую конструкцию и используются в качестве зарядных устройств для аккумуляторных систем, создания локальных электрических сетей в местах, где нет доступа к общей энергосистеме и других целей.

Горизонтально-осевые ветротурбины наиболее эффективны в местностях с высокой средней скоростью ветра, обычно в районах с открытыми пространствами, таких как равнины, побережья, вершины холмов и гор.

Но также широкое распространение получило направление строительства оффшорных ветропарков и, как следствие, появились ВЭУ, предназначенные для использования в шельфовых зонах водных бассейнов

Достоинством офшорных ветропарков является возможность размещения большого количества ВЭУ высокой мощности, меньшее влияние на здоровье населения (шум, оптические эффекты и др.). При этом стоимость строительства и дальнейшие эксплуатационные расходы несколько выше, чем для ВЭУ находящихся на суше [24, 25].

Как и у любой сложной технической системы у ВЭУ с горизонтальной осью вращения существуют свои достоинства и недостатки.

Ветроэнергетические установки с горизонтальной осью вращения имеют несколько преимуществ. Одно из них заключается в возможности изменения угла атаки лопастей в зависимости от времени суток и сезона, что позволяет эффективно использовать энергию ветра. Кроме того, высокая мачта способствует повышению эффективности установки за счет использования энергии ветра на больших высотах над землей. Фактически, с увеличением высоты каждые 10 м скорость ветра увеличивается на 20%.

Однако, у ВЭУ с горизонтальной осью вращения также есть и недостатки:

- необходимость применения высоких мачт (башен) с высотой более 100 м и длинных лопастей затрудняет их транспортировку и монтаж (стоимость транспортировки и монтажа составляет до 20 % от общей стоимости проекта);
- для монтажа ветроустановок большой мощности требуется специальное оборудование (подъемное) и высококвалифицированные кадры, поэтому производство таких ВЭУ осуществляется только небольшим количеством стран;
- из-за больших размеров ВЭУ они оказывают помехи на радиосвязь и работу радиолокационных установок;
- необходимость в специальных системах ориентации ветротурбины на ветер.

Вертикально-осевые ветротурбины (VAWT) имеют лопасти, расположенные вокруг вертикальной оси, и обладают рядом преимуществ, таких как более компактный дизайн, возможность работы при низких скоростях ветра и меньшая зависимость от направления ветра [24, 25].

Сегодня используются различные конструкции роторов ветроэнергетических установок с вертикальной осью вращения, включая роторы Савониуса, Савониуса спиральной конструкции, Дарье, Гиромила, Геликоидный и другие [23]. Примеры подобных установок приведены на рисунке 2.4.



Рисунок 2.4 – Примеры вертикально-осевых ветротурбин

Ротор Савониуса был создан финским инженером в 1922 году. Его принцип работы основан на разнице сил сопротивления воздушного потока при обтекании вогнутых и выпуклых диафрагм полукруглого или логарифмического профиля ветроротора. Коэффициент сопротивления ветровому потоку при обтекании вогнутой диафрагмы составляет около 1,3, а при обтекании в противоположном направлении - 0,35. Разница сил сопротивления создает крутящий момент на приводном валу.

Сейчас существует несколько конструктивных вариантов ВЭУ на основе ротора Савониуса, отличающихся компоновкой роторов, их количеством в установке и используемыми материалами.

Однако высокая материалоемкость и сложности с балансировкой препятствуют массовому производству мощных ветророторов Савониуса. Данные ВЭУ используются для питания небольших автономных технологических объектов с невысокими требованиями к мощности и качеству электроэнергии.

Для удешевления в настоящее время диафрагмы роторов изготавливаются из пластика, что также упрощает балансировку.

Преимущества таких ветроэнергетических установок включают в себя низкий уровень шума, компактное размещение и эффективную работу при низких скоростях ветра (3-5 м/с). Благодаря вертикальной оси вращения, ротор не требует устройств для ориентации на ветер, что значительно упрощает конструкцию и делает ветряную установку изысканно простой. Однако данная турбина характеризуется невысоким коэффициентом использования энергии ветра, всего 0,18-0,24, и коэффициентом использования мощности - 17-18%. [24].

Ротор Дарье, создал французский авиаинженер в 1931 году. Он представляет собой симметричную конструкцию, состоящую из двух или более аэродинамических крыльев, закрепленных на радиальных балках.

Подъемная сила действует на каждое крыло, движущееся относительно потока и зависящая от угла между векторами скорости потока и мгновенной

скорости крыла. Максимальное значение подъемной силы достигается при ортогональности данных векторов. Так как вектор мгновенной скорости крыла циклически изменяется в процессе вращения ротора, момент силы, развиваемый ротором, также является переменным.

Из-за необходимости движения крыльев для создания подъемной силы, ротор Дарье имеет ограничения в самозапуске. Улучшение самозапуска достигается при использовании трех и более лопастей, однако все же требуется начальное ускорение ротора. [25].

Турбина Гиромила представляет собой вариацию турбины Дарье с прямой конструкцией лопастей и рядом преимуществ. Среди них высокий стартовый момент, широкая кривая вращательного момента, относительно низкая скорость вращения лопастей, меньшие напряжения на лопастях, повышенная эффективность и более надежная работа в условиях ветровых турбулентностей.

Данный вид конструкций ВЭУ также имеет свои плюсы и минусы.

К достоинствам можно отнести отсутствие необходимости ориентации на ветер (они могут работать в любом направлении ветра), гораздо меньший шум при работе, гораздо более компактные габариты, что делает их более доступными для использования в ограниченных пространствах, лучшую работу на низких скоростях ветра и в условиях турбулентности [25].

К недостаткам можно отнести низкую эффективность на высоких скоростях ветра, большие колебания в выходной мощности, большие потери на трение, ограниченную мощность.

Также существуют и иные конструкции и инженерные решения.

Двухроторные ветротурбины: Эти ветротурбины имеют два ротора, расположенных на одной вертикальной оси. Они обладают высокой эффективностью и могут использоваться в различных условиях ветра.

Диффузорные ветротурбины: Эти ветротурбины имеют диффузорную конструкцию, которая увеличивает скорость ветра и повышает эффективность ветротурбины.

Конструкция ветротурбин может варьироваться в зависимости от требований и условий эксплуатации. Некоторые ветротурбины имеют фиксированные лопасти, в то время как другие имеют регулируемые лопасти, которые могут изменять угол наклона для оптимального использования ветра.

Современные ветротурбины постоянно развиваются и совершенствуются. Некоторые инновации в конструкции ветротурбин включают улучшенные лопасти, изготовленные из новых материалов, таких как композитные материалы, что позволяет создавать лопасти с более эффективной формой и повышенной прочностью, увеличивая эффективность ветротурбин и улучшая их работу в различных условиях ветра.

Также используются регулируемые лопасти, которые позволяют изменять угол наклона лопастей в зависимости от скорости ветра, что оптимизирует работу ветротурбины в различных условиях [24].

Улучшенные системы передачи, такие как безредукторные системы, уменьшают потери энергии и повышают эффективность ветротурбин.

Применение систем искусственного интеллекта и машинного обучения в управлении ветротурбинами позволяет оптимизировать их работу, учитывая метеорологические условия и изменения ветра. Развитие систем мониторинга и диагностики позволяет оперативно обнаруживать и устранять возможные неисправности ветротурбин, что повышает их надежность и снижает затраты на обслуживание.

Инновации в конструкции ветротурбин позволяют повысить эффективность, надежность и экономическую эффективность ветроэнергетики.

Что касается стран с самой большой долей выборки электроэнергии ветрогенераторами, то наиболее выделяются следующие: Дания (один из лидеров, около 50% общего потребления электроэнергии), Испания (около 20% общего потребления электроэнергии), Германия (20% общего потребления электроэнергии), США и Китай (около 7% общего потребления электроэнергии).

Эти страны активно инвестируют в развитие ветроэнергетики, строят крупные ветроэнергетические парки и разрабатывают новые технологии для повышения эффективности и надежности ветрогенераторов. Ветроэнергетика становится все более значимым источником электроэнергии во всем мире, и ожидается, что ее доля в общем потреблении электроэнергии будет продолжать расти в ближайшие годы.

Выработка электроэнергии ветроустановками зависит от ряда факторов, влияющих на эффективность и производительность ветротурбин. Скорость ветра является ключевым фактором, влияющим на выработку электроэнергии, при этом оптимальный диапазон скорости ветра обеспечивает максимальную эффективность ветротурбины. Направление ветра также влияет на выработку электроэнергии, поскольку ветротурбины могут поворачивать лопасти для оптимального использования ветра в различных направлениях. Топография местности, где расположены ветроустановки, может влиять на скорость и направление ветра, создавая ветровые тени, что может снижать скорость ветра и, следовательно, выработку электроэнергии. Высота ветротурбины также играет роль, поскольку чем выше ветротурбина, тем больше вероятность, что она будет работать в зоне с более высокой скоростью ветра, что приведет к увеличению выработки электроэнергии. Техническое состояние ветротурбины, включая состояние лопастей, генератора и системы передачи, также влияет на выработку электроэнергии. Регулярное обслуживание и ремонт ветротурбин помогают обеспечить их оптимальную работу и максимальную выработку электроэнергии. Метеорологические условия, такие как температура воздуха, влажность и атмосферное давление, также могут влиять на работу ветротурбин и, следовательно, на выработку электроэнергии. Учет всех этих факторов и их взаимосвязей является важным аспектом для точного прогнозирования выработки электроэнергии ветроустановками.

3 АНАЛИЗ ВЫБОРКИ ДАННЫХ ПО ГЕНЕРАЦИИ ВЕТРОЭНЕРГЕТИЧЕСКОЙ УСТАНОВКИ

3.1 Используемая выборка данных

Большинство существующих работ по прогнозированию генерации ветроэнергетических установок рассматривают скорость ветра как целевую переменную [26, 27, 28], поскольку считается, что на ее основе можно получить значение генерируемой мощности с низкой погрешностью при интервале дискретизации по времени до 1 часа [29, 30, 31]. Это утверждение как правило основано на том, что для всех ветроэнергетических установок производителями предоставляются функциональные зависимости выходной мощности от скорости ветра. Однако, физика процесса выработки электроэнергии ветровой турбиной требует более внимательного изучения возможности вычисления генерации по скорости ветра с учетом таких факторов как инерция, плотность воздуха, термические процессы [32, 33].

Для выполнения такого исследования необходима выборка данных с параметрами ветрового потока и значениями генерации. Рассматривается выборка данных за год, содержащая 50530 значений следующих параметров, записанных SCADA-системой с реальной ветроэнергетической установки, с 10-минутным шагом:

- дата и время;
- вырабатываемая активная мощность (кВт);
- скорость ветра на высоте турбины (м/с);
- расчетное значение мощности согласно паспортным данным ветроэнергетической установки при текущей скорости ветра (кВ);
- направление ветра на высоте турбины (°), при этом турбина поворачивается так, чтобы ветроколесо было перпендикулярно потоку ветра.

3.2 Предварительная обработка и визуализация данных

На рисунке 3.1 приведен код для подключения библиотек для обработки и визуализации данных:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.simplefilter("ignore")
```

Рисунок 3.1 – Код для подключения необходимых библиотек

На рисунке 3.2 приведен код для чтения исходного файл и результаты чтения (фрагмент выборки данных).

```
df=pd.read_csv('T1.csv')
df.columns = ['Дата_Время', 'Факт.генерация',
              'Скорость_ветра', 'Расч.генерация', 'Направление_ветра']
df
```

	Дата_Время	Факт.генерация	Скорость_ветра	Расч.генерация	Направление_ветра
0	01 01 2018 00:00	380.047791	5.311336	416.328908	259.994904
1	01 01 2018 00:10	453.769196	5.672167	519.917511	268.641113
2	01 01 2018 00:20	306.376587	5.216037	390.900016	272.564789
3	01 01 2018 00:30	419.645905	5.659674	516.127569	271.258087
4	01 01 2018 00:40	380.650696	5.577941	491.702972	265.674286
...	...	...	...	...	...

Рисунок 3.2 – Код для чтения выборки данных и ее фрагмент

На рисунках 3.3–3.6 показаны распределения значений параметров.

На рисунках 3.7–3.9 приведены диаграммы совместного распределения параметров. Видно, что теоретическая (расчетная) зависимость генерации от скорости ветра имеет существенную погрешность, а определение фактической генерации на основе данных только расчетной модели будет иметь большее отклонение, чем на основе модели, учитывающей скорость ветра.

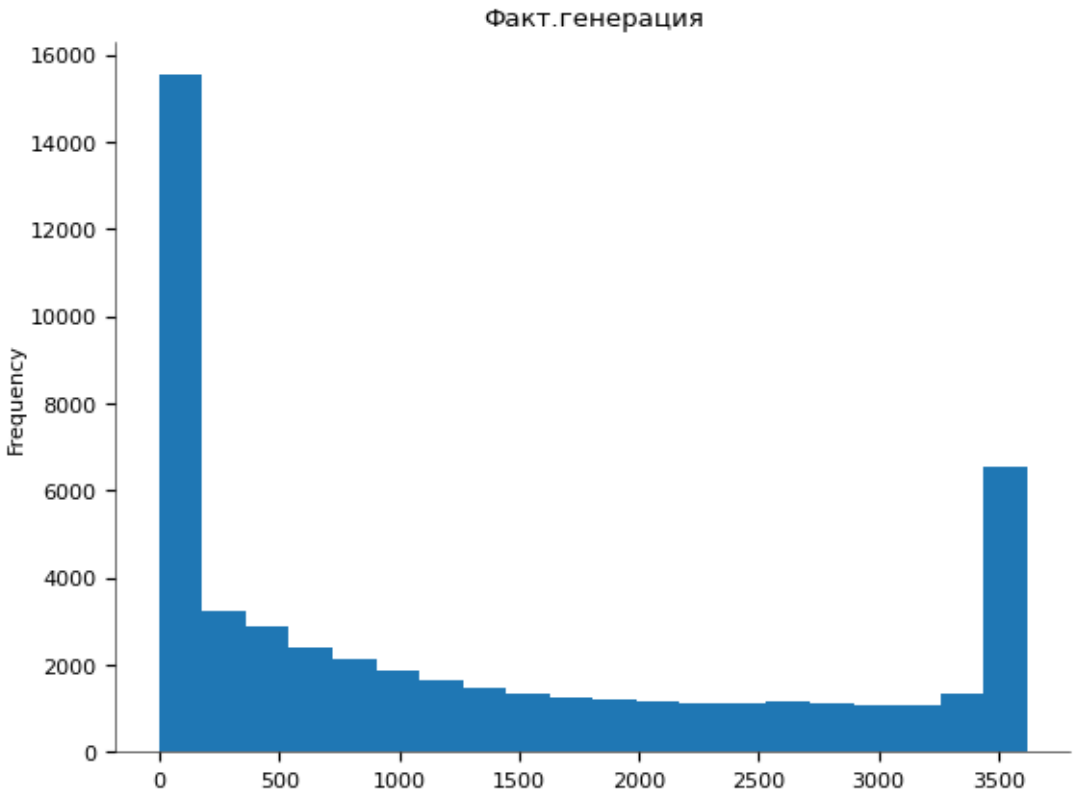


Рисунок 3.3 – Распределение значений генерации

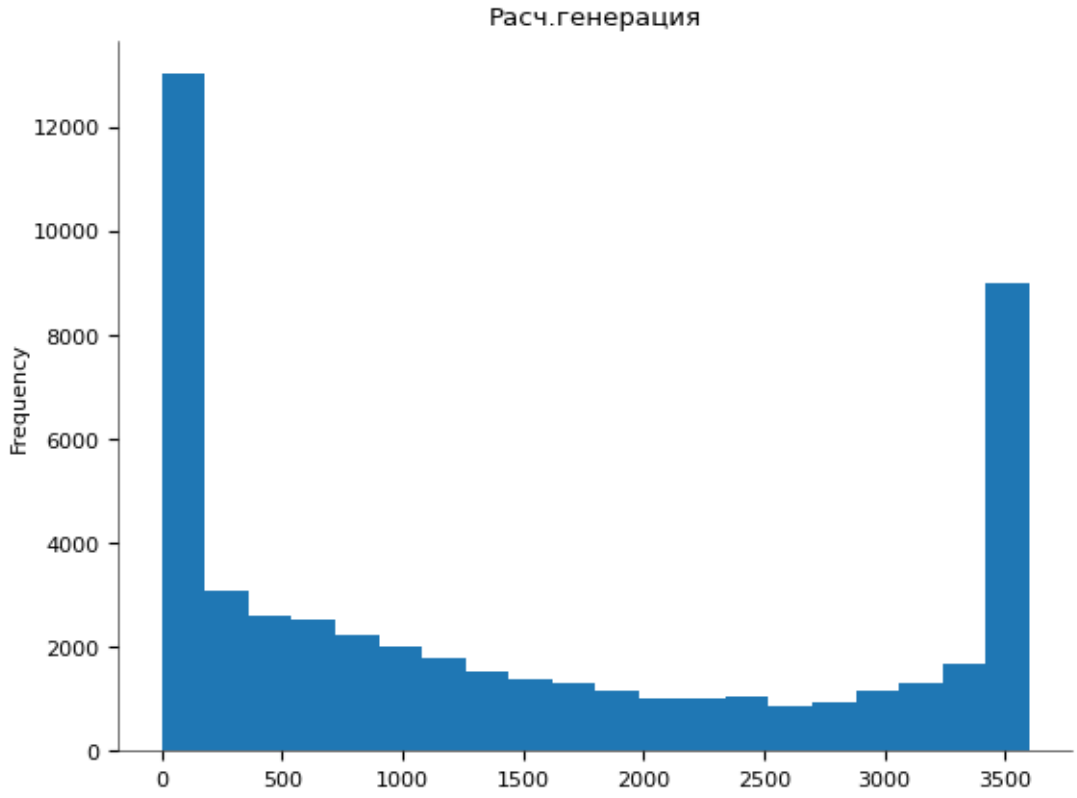


Рисунок 3.4 – Распределение значений расчетной генерации

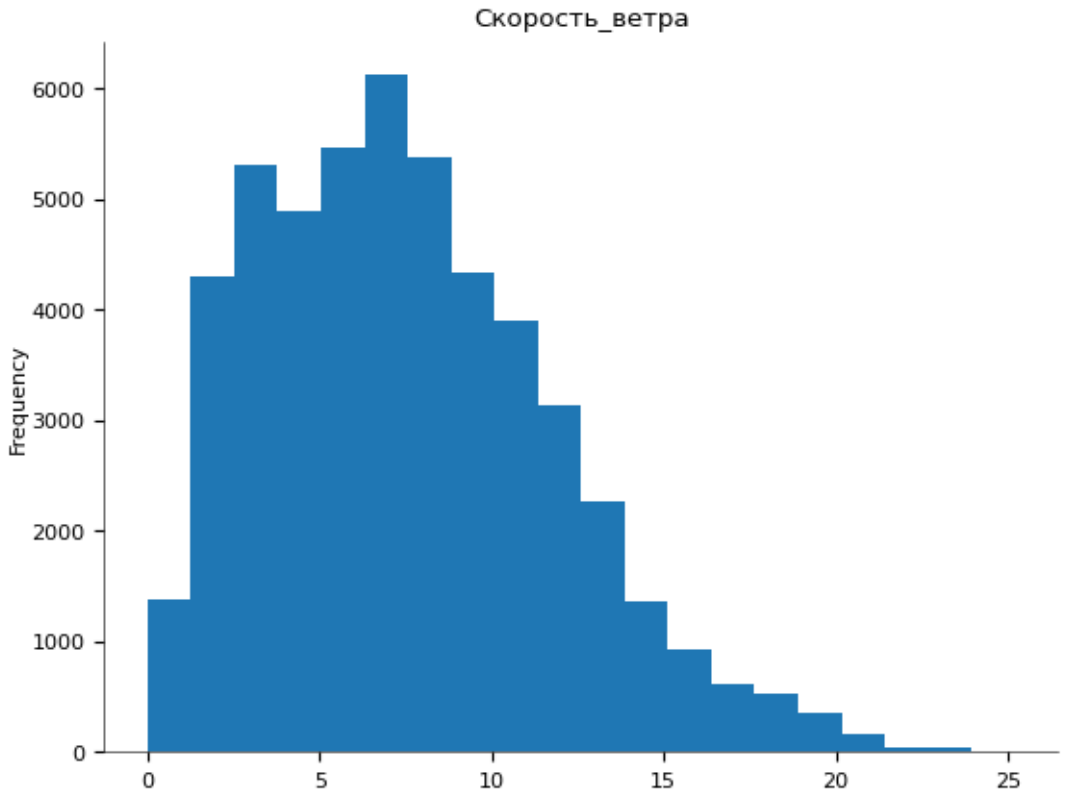


Рисунок 3.5 – Распределение значений скорости ветра

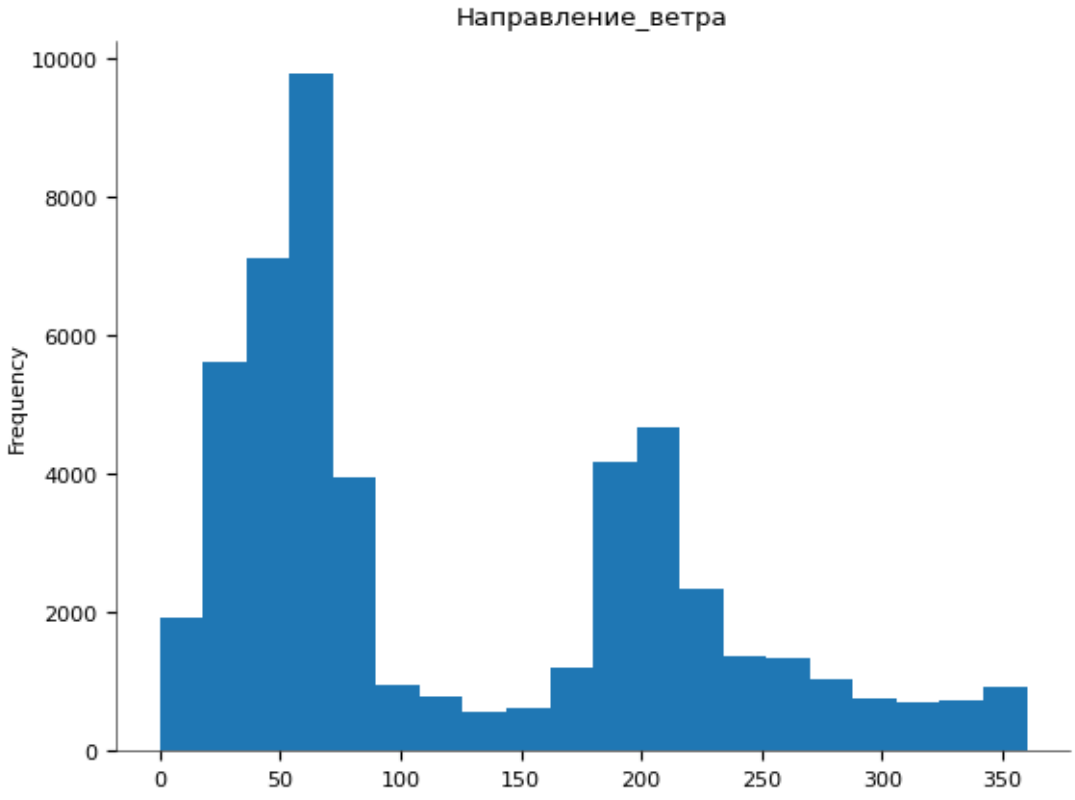


Рисунок 3.6 – Распределение значений направления ветра

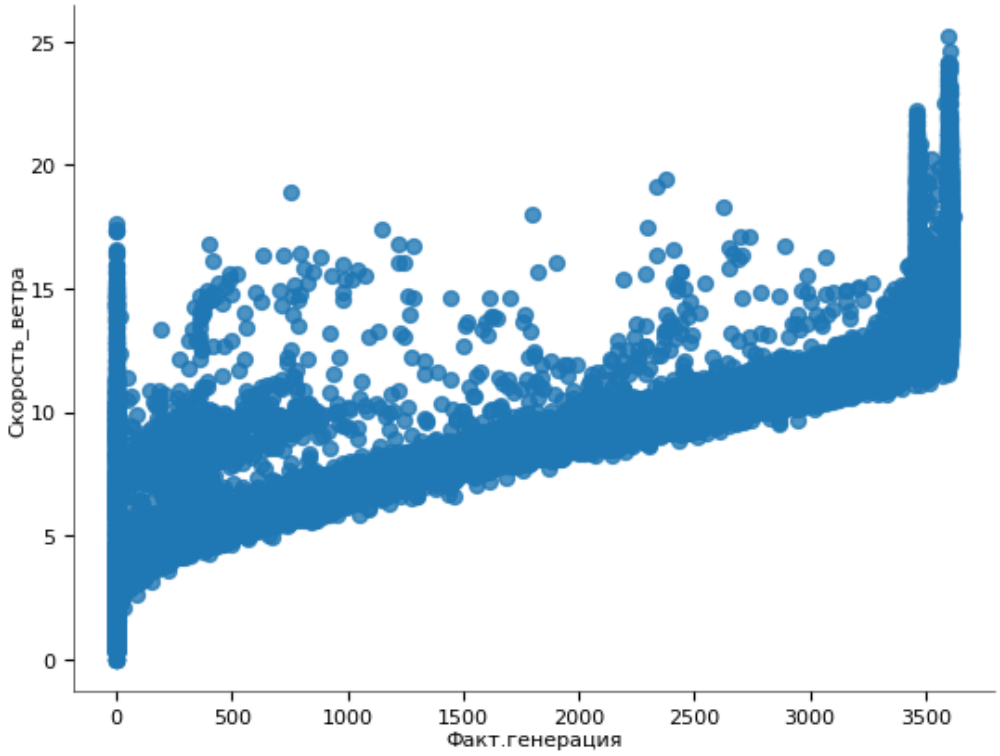


Рисунок 3.7 – Совместное распределение скорости ветра и фактической генерации

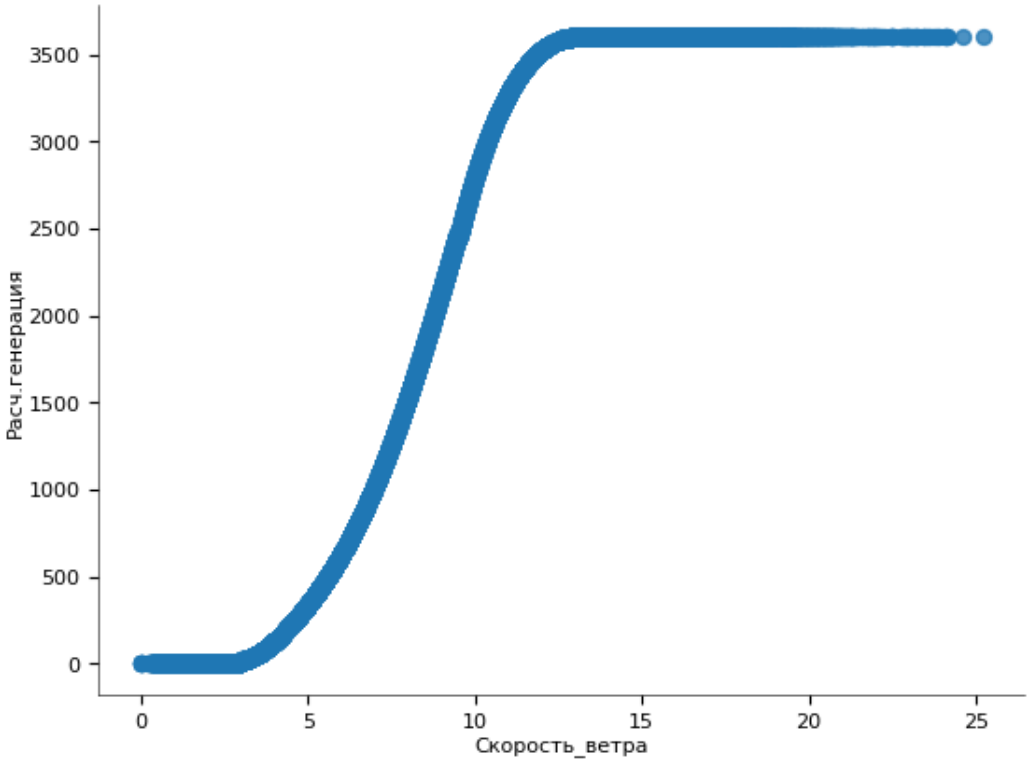


Рисунок 3.8 – Совместное распределение скорости ветра и расчетной генерации

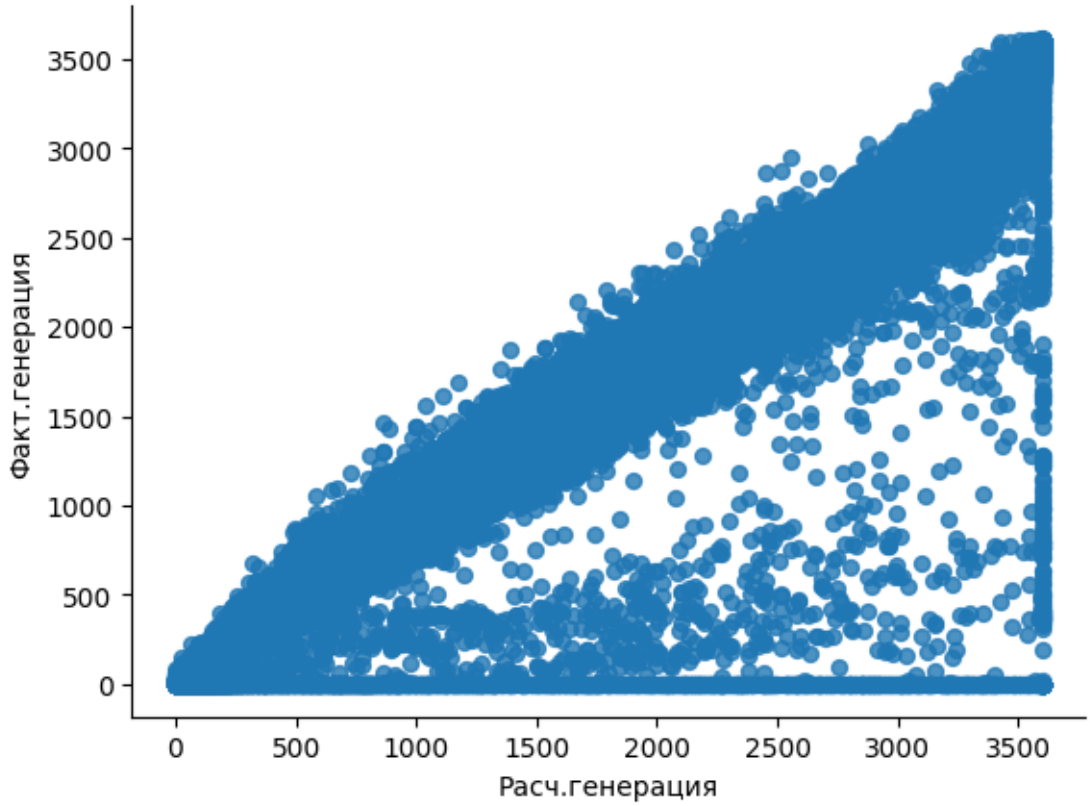


Рисунок 3.9 – Совместное распределение фактической и расчетной генерации

В таблице 3.1 приведены статистические параметры выборки данных.

Таблица 3.1 – Статистические параметры выборки данных

Параметр	Факт. генерация	Скор. ветра	Расч. генерация	Направл. ветра
Количество	50530	50530	50530	50530
Среднее	1307,68	7,55	1492,17	123,68
Среднеквадратическое отклонение	1312,45	4,22	1368,01	93,44
Минимальное	-2,471	0,00	0,00	0,00
25-й перцентиль	50,67	4,20	161,32	49,31
50-й перцентиль	825,83	7,10	1063,77	73,71

Продолжение таблицы 3.1

Параметр	Факт. генерация	Скор. ветра	Расч. генерация	Направл. ветра
75-й перцентиль	2482,50	10,30	2964,97	201,69
Максимальное	3618,73	25,20	3600,00	359,99

Видно, что расчетная генерация существенно завышена при малых скоростях ветра. Также можно заметить, что в данных фактической генерации есть отрицательные значения, хотя и очень близкие к нулю. Их можно приравнять к нулю, как показано на рисунке 3.10.

```
count_negative_values = (df['Факт.генерация'] < 0).sum()
print(count_negative_values)
df.loc[df['Факт.генерация'] < 0, 'Факт.генерация'] = 0
count_negative_values = (df['Факт.генерация'] < 0).sum()
print(count_negative_values)
```

```
57
0
```

Рисунок 3.10 – Код для устранения ошибочных отрицательных значений фактической генерации

Можно предположить, что для построения более точной зависимости генерации от скорости ветра являются полезными признаки, связанные со временем: месяц, сезон, час суток и т.п. Благодаря использования технологий искусственного интеллекта возможно построить более точную модель определения генерации ветроэнергетической установки с использованием множества признаков и выделения более сложных зависимостей, чем учитывается в расчетной модели [34, 35].

На рисунке 3.11 приведен код, выполняющий извлечение таких признаков, а также проверку, что в выборке отсутствуют пропуски.

Перед следующим этапом – построением моделей с использованием технологий искусственного интеллекта, выборку данных следует разделить на обучающую и тестовую. Код показан на рисунке 3.12.

Код для визуализации показан на рисунке 3.13. Результаты визуализации аналогичны показанным на рисунках 3.3–3.6.

```
df['Дата_Время']=pd.to_datetime(df['Дата_Время'],format='%d %m %Y %H:%M')
df['Неделя']=df['Дата_Время'].dt.day // 7 + 1

df['Неделя'].value_counts()
df['Месяц']=df['Дата_Время'].dt.month
seasons_dict = {1: 'Зима', 2: 'Зима', 3: 'Весна', 4: 'Весна', 5: 'Весна',
                6: 'Лето', 7: 'Лето', 8: 'Лето', 9:
                'Осень', 10: 'Осень', 11: 'Осень', 12: 'Осень'}
df['Сезон'] = df['Месяц'].map(seasons_dict)

df['День']=df['Дата_Время'].dt.day
df['Час']=df['Дата_Время'].dt.hour+1

df.drop(columns=['Дата_Время'],inplace=True)
df.isna().sum()
```

```
Факт.генерация      0
Скорость_ветра      0
Расч.генерация      0
Направление_ветра   0
Неделя              0
Месяц               0
Сезон               0
День                0
Час                 0
```

Рисунок 3.11 – Извлечение признаков, связанных со временем

```
from sklearn.model_selection import train_test_split

df_train, df_test=train_test_split(df,test_size=0.2,random_state=42)
print(df_train.shape)
print(df_test.shape)
```

```
(40424, 9)
(10106, 9)
```

Рисунок 3.12 – Разделение выборки на обучающую и тестовую

```
date_col=['Неделя', 'Месяц', 'Сезон', 'Час', 'День']
num_col=['Факт.генерация', 'Скорость_ветра', 'Расч.генерация', 'Направление_ветра',]
df_train[num_col].hist(bins=20,figsize=(12,8))
```

```
array([[<Axes: title={'center': 'Факт.генерация'}>,
        <Axes: title={'center': 'Скорость_ветра'}>],
       [<Axes: title={'center': 'Расч.генерация'}>,
        <Axes: title={'center': 'Направление_ветра'}>]], dtype=object)
```

Рисунок 3.13 – Запуск визуализации с применением новых признаков

На рисунке 3.14 показаны диаграммы рассеяния между всеми признаками выборки.

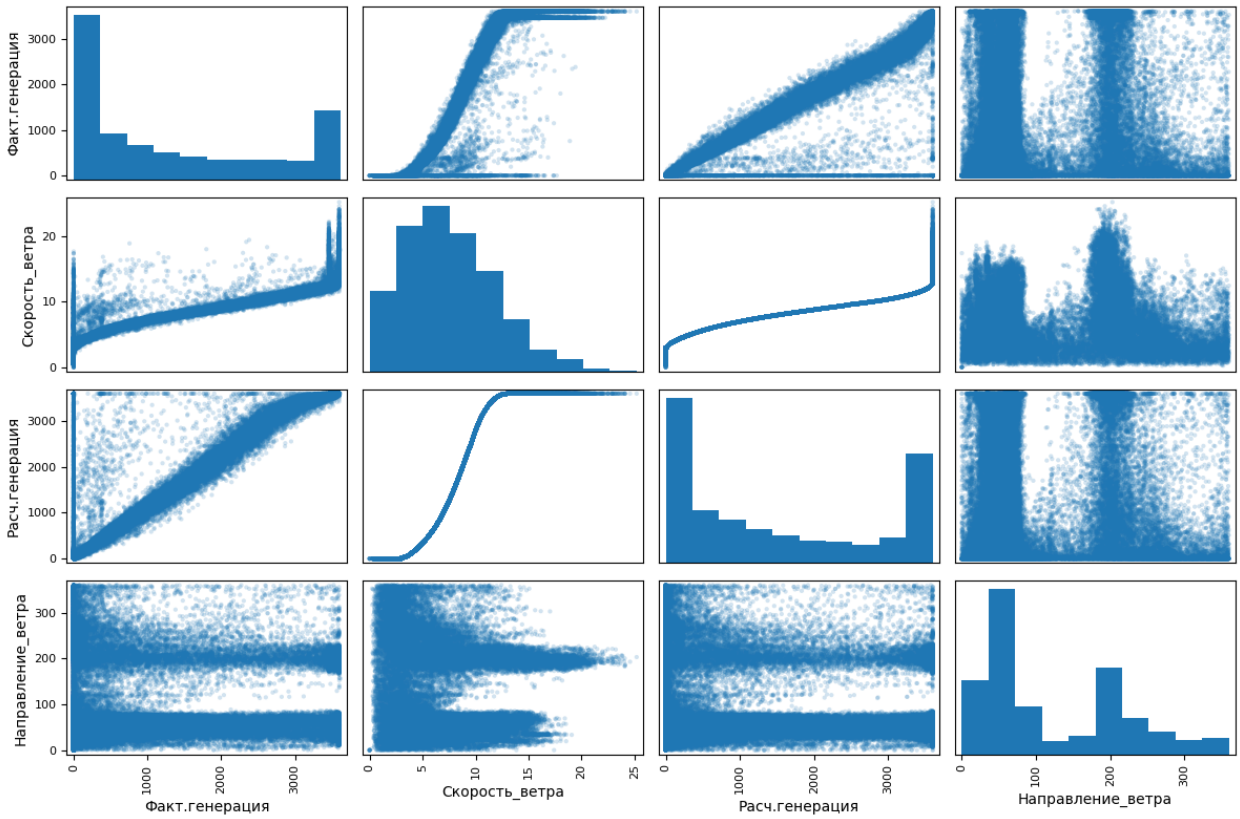


Рисунок 3.14 – Диаграммы рассеяния между признаками

Из диаграммы рассеяния между скоростью ветра и расчетной мощностью следует:

- турбина не может генерировать электроэнергию, если скорость ветра менее 4 м/с;
- когда скорость ветра превышает 4–11 м/с, соотношение между ними линейное, это означает, что увеличение скорости ветра позволяет турбинам генерировать больше энергии;

- после того, как скорость ветра превысит 11 м/с, вырабатываемая электроэнергия насыщается на 3600 кВтч.

Зависимость фактической генерации в целом подобна, но с очень большим количеством отклонений вверх, которые, вероятно, связаны с инерционностью ветроэнергетической установки [36].

На рисунке 3.15 приведен код для построения дополнительных результатов визуализации. На рисунках 3.16–3.19 показаны результаты визуализации.

```
for col in num_col:
    fig, ax = plt.subplots(1, 2, figsize=(12, 5))
    sns.distplot(df_train[col], ax=ax[0] ,color='green')
    sns.boxplot(x=df_train[col], ax=ax[1])
    plt.show()
```

Рисунок 3.15 – Код для визуализации плотностей распределения величин

На рисунке 3.20 показан результат построения матрицы коэффициентов корреляции.

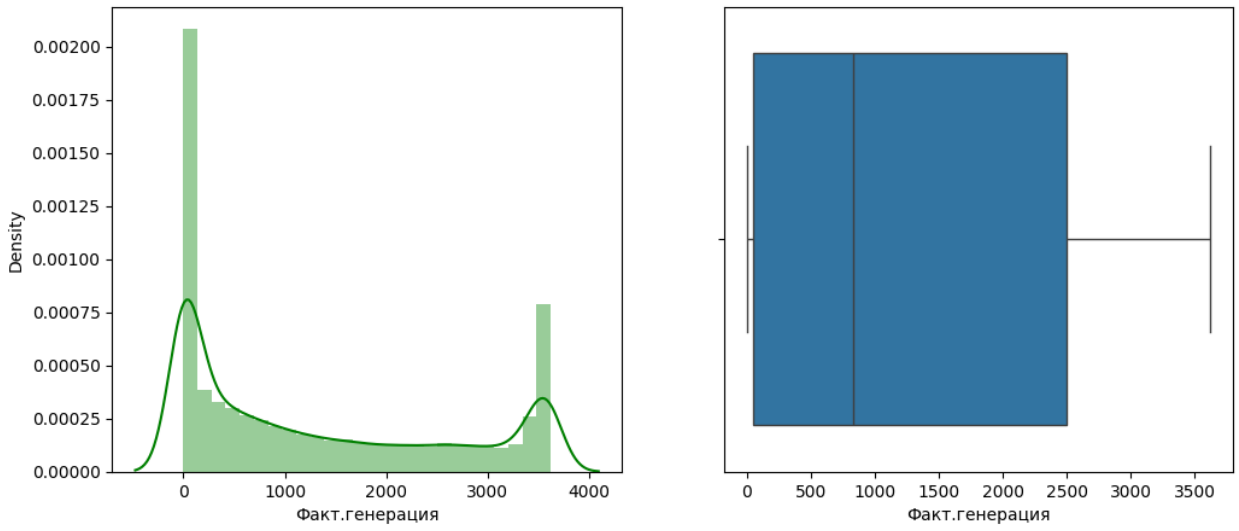


Рисунок 3.16 – Распределение фактической генерации

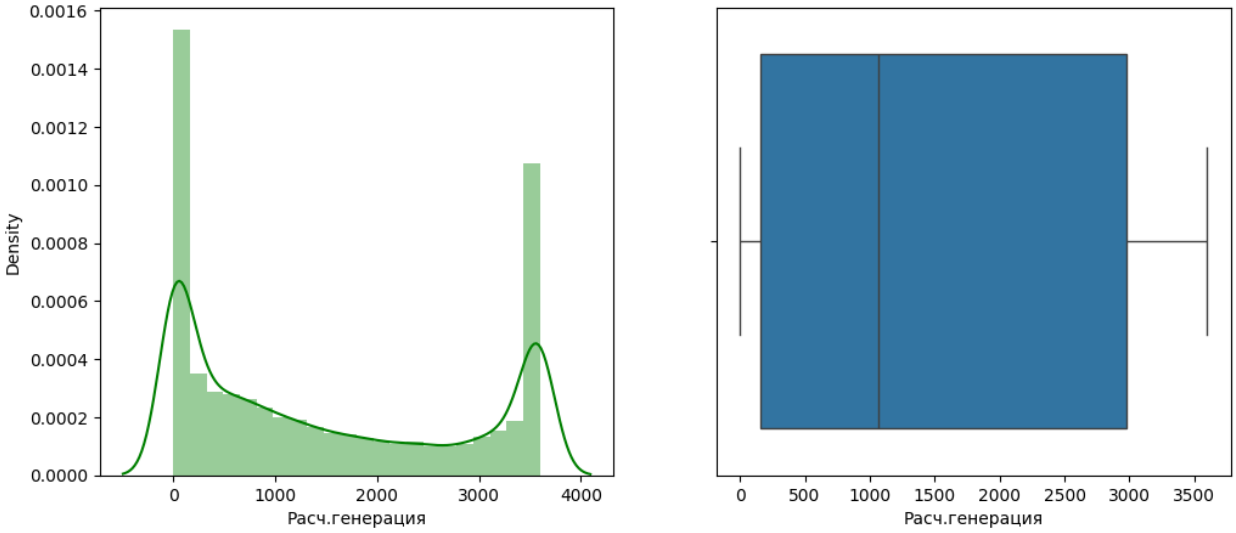


Рисунок 3.17 – Распределение расчетной генерации

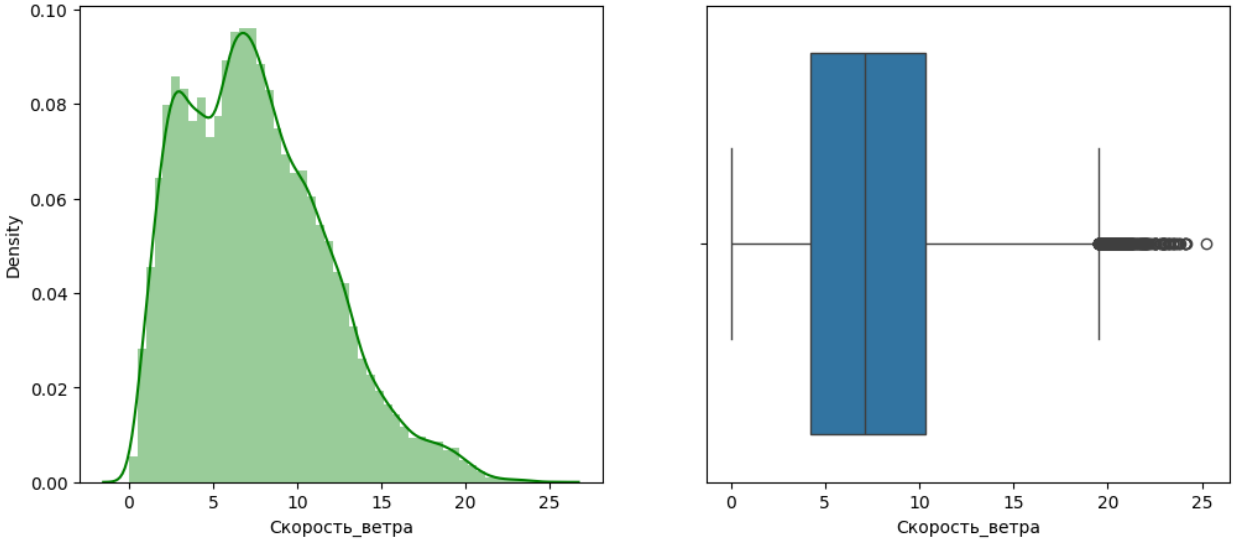


Рисунок 3.18 – Распределение скорости ветра

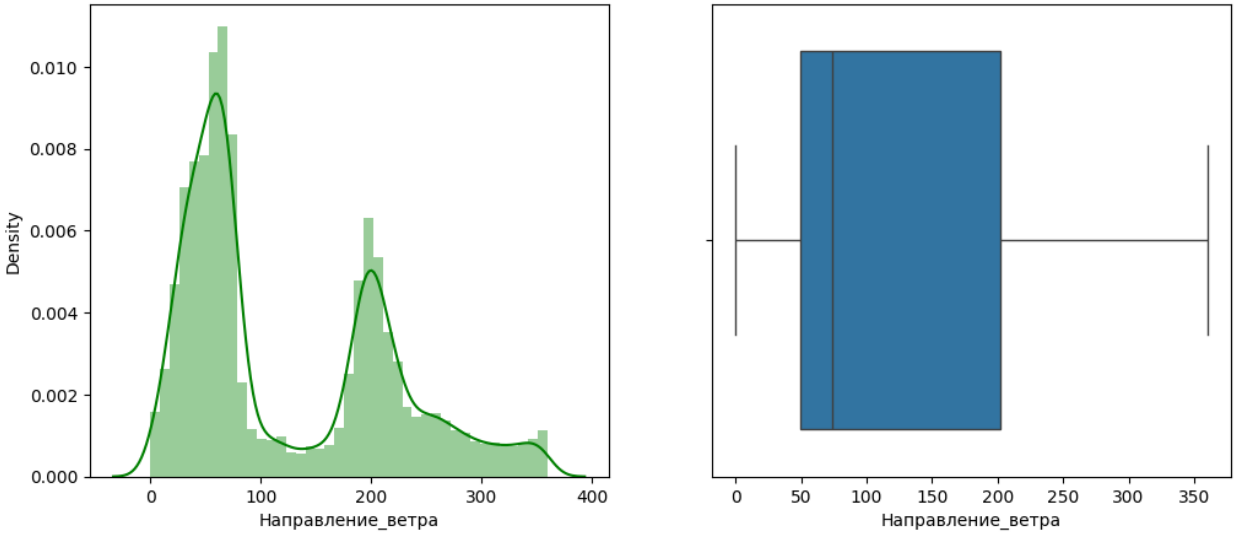


Рисунок 3.19 – Распределение направления ветра

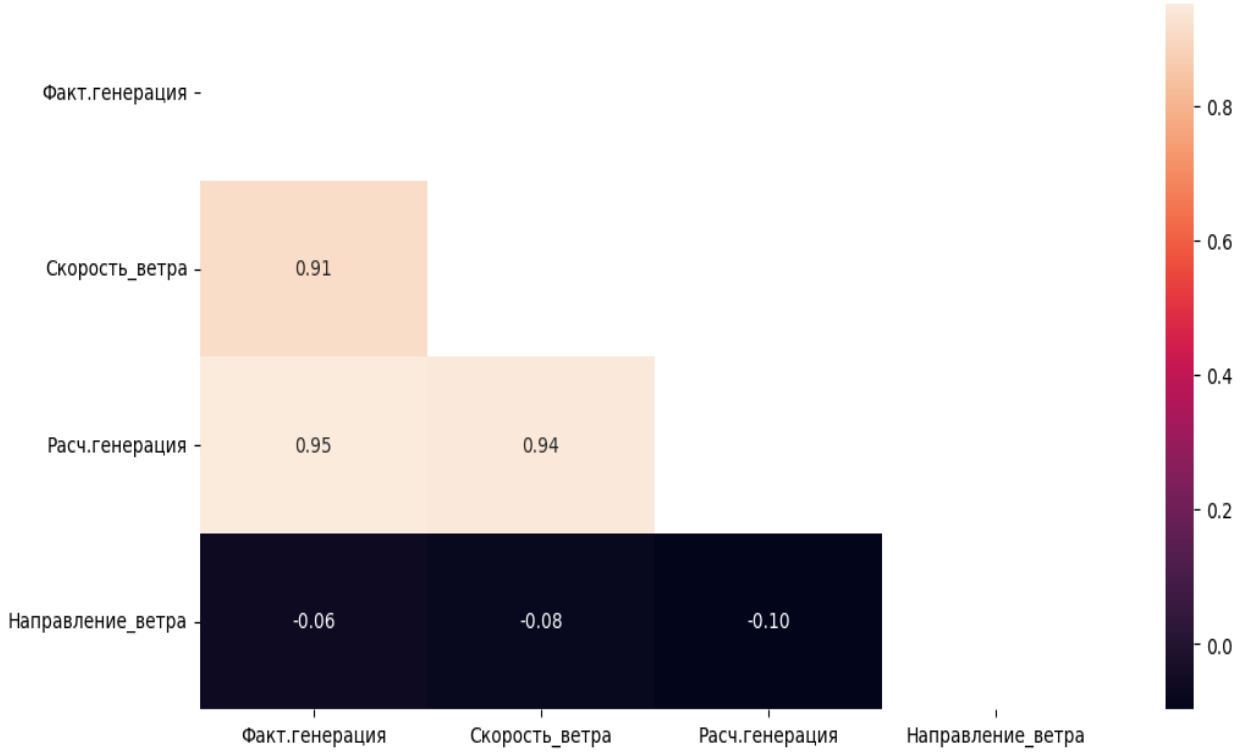


Рисунок 3.20 – Коэффициенты корреляции между параметрами

Видно, что коэффициент корреляции Пирсона между расчетной генерацией и фактической близок к единице. Поэтому нужно признать, что в целом использование расчетной генерации вместо фактической допустимо, особенно, если учесть, что в данной работе используется 10-минутный интервал, а во многих исследованиях, посвященных прогнозированию генерации ветроэнергетических установок, – часовой, и при суммировании погрешности снижаются.

Тем не менее, для систем управления электроэнергетическими установками с ветровыми турбинами отдельные отклонения в прогнозной выработке могут приводить к неэффективному оперативному управлению режимами работы. Поэтому задача прогнозирования именно выходной мощности, а не скорости ветра остается актуальной [37, 38].

На рисунке 3.21 показан код для построения визуализации с учетом месяцев года.

На рисунке 3.22 – показаны полученные результаты.

```python
fig,axes=plt.subplots(nrows=len(num_col)//2 ,ncols=2,figsize=(12,8))

for idx,col in enumerate(df_train[num_col]):
    row_idx=idx//2
    col_idx=idx%2
    sns.barplot(x=df_train['Месяц'],
                y=df_train[col],
                data=df_train,
                ax=axes[row_idx,col_idx],
                errorbar=None,
                palette='rocket')

plt.tight_layout()
plt.show()
```

Рисунок 3.21 – Код для визуализации с учетом месяцев года

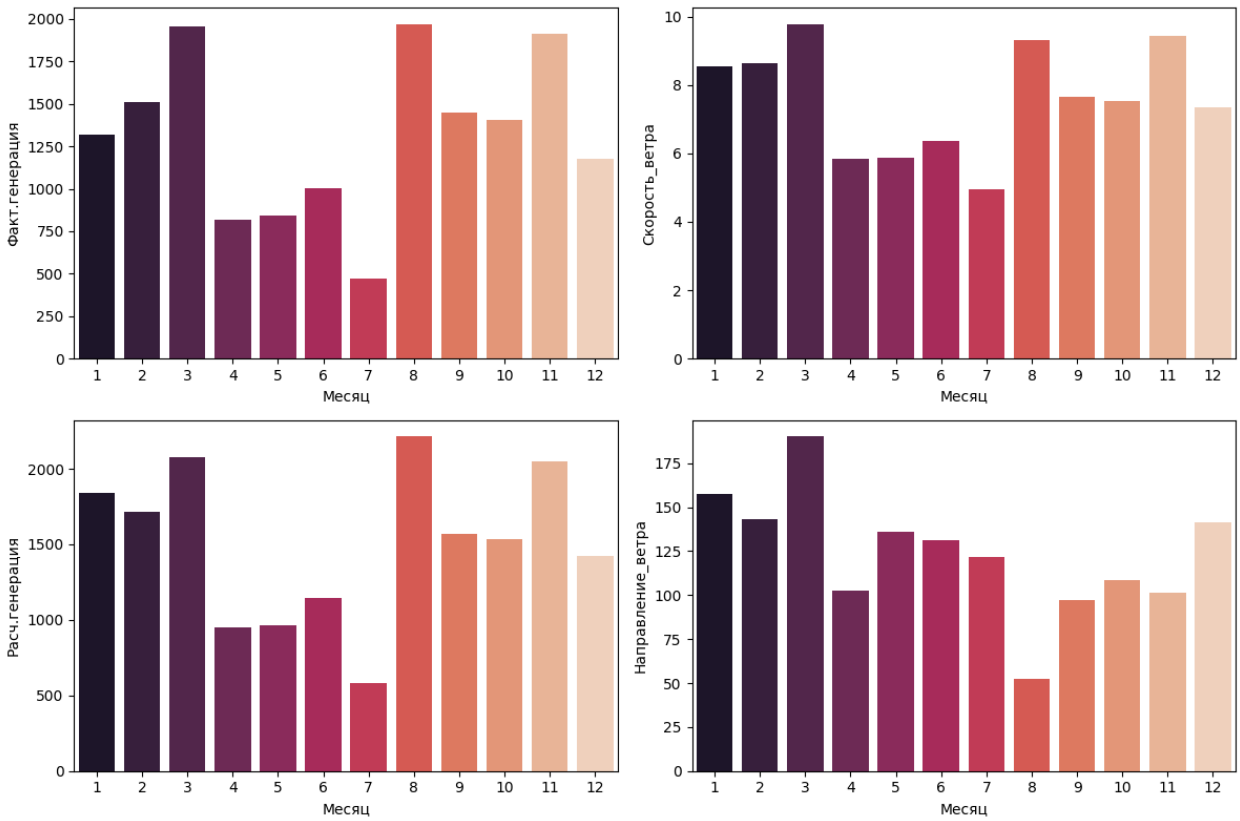


Рисунок 3.22 – Результаты визуализации по месяцам

Видна неравномерность распределения скорости ветра по месяцам года, особенно выделяются в меньшую сторону июль, в большую – август, ноябрь и март.

Аналогично на рисунке 3.23 приведены результаты распределения по сезонам года.

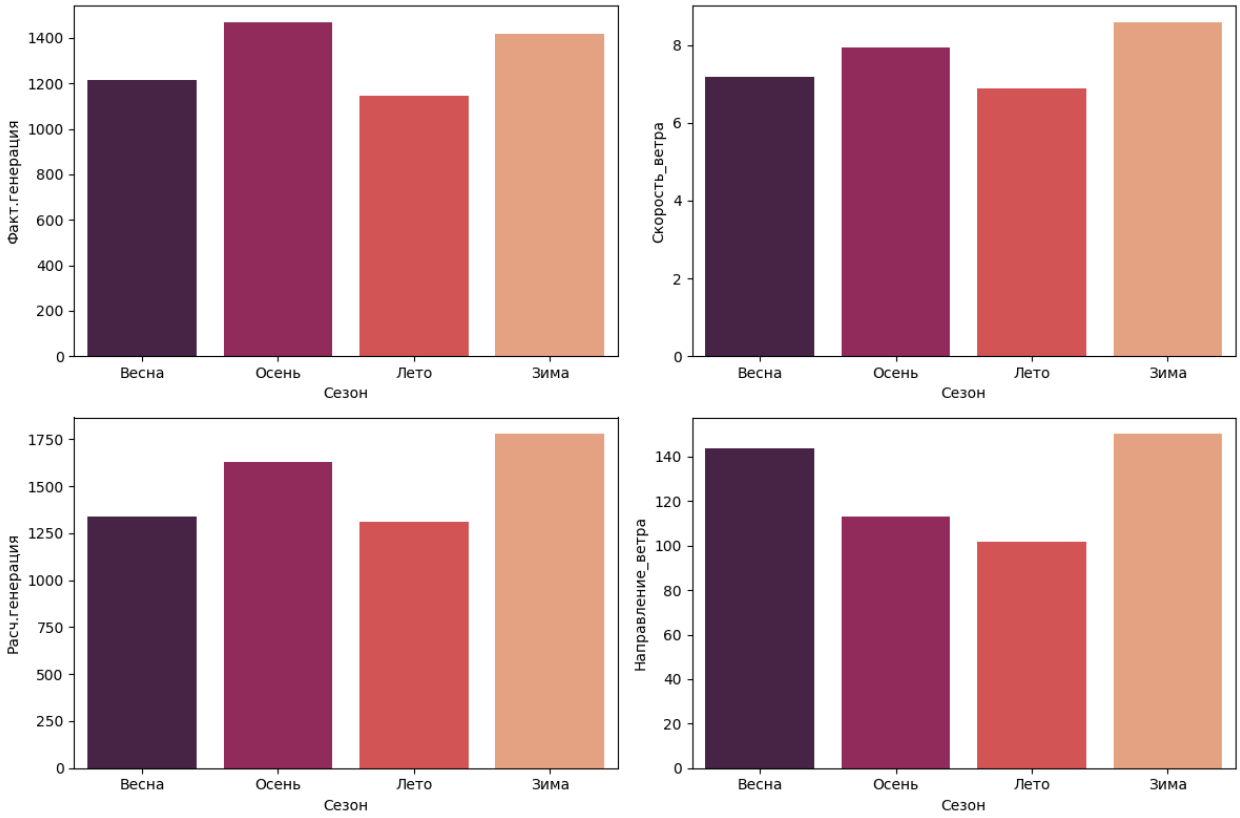


Рисунок 3.23 – Распределение параметров по сезонам года

На рисунках 3.24–3.30 показаны дополнительные результаты визуального анализа данных.

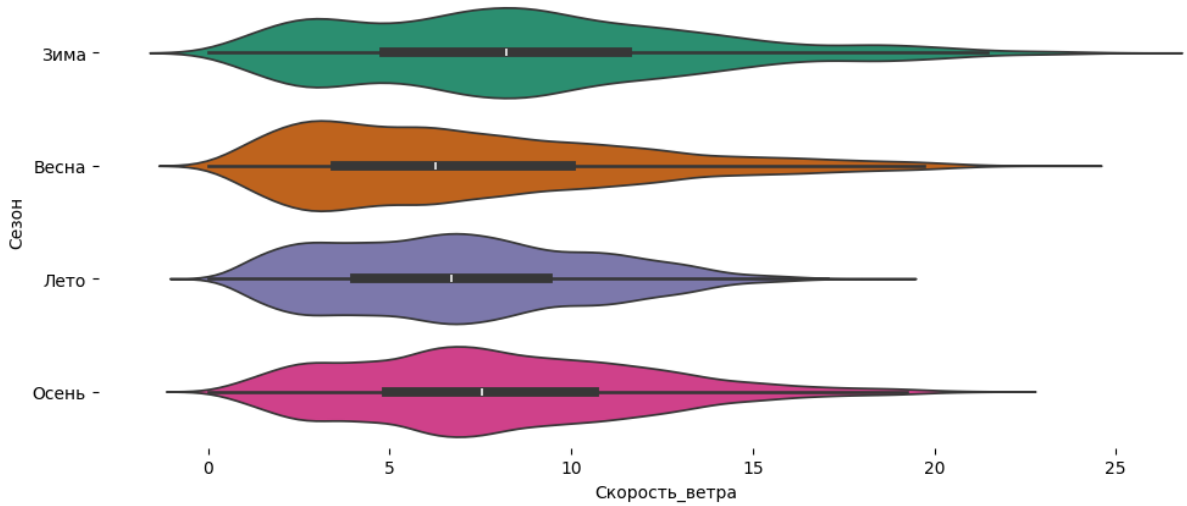


Рисунок 3.24 – Распределение скорости ветра по сезонам года

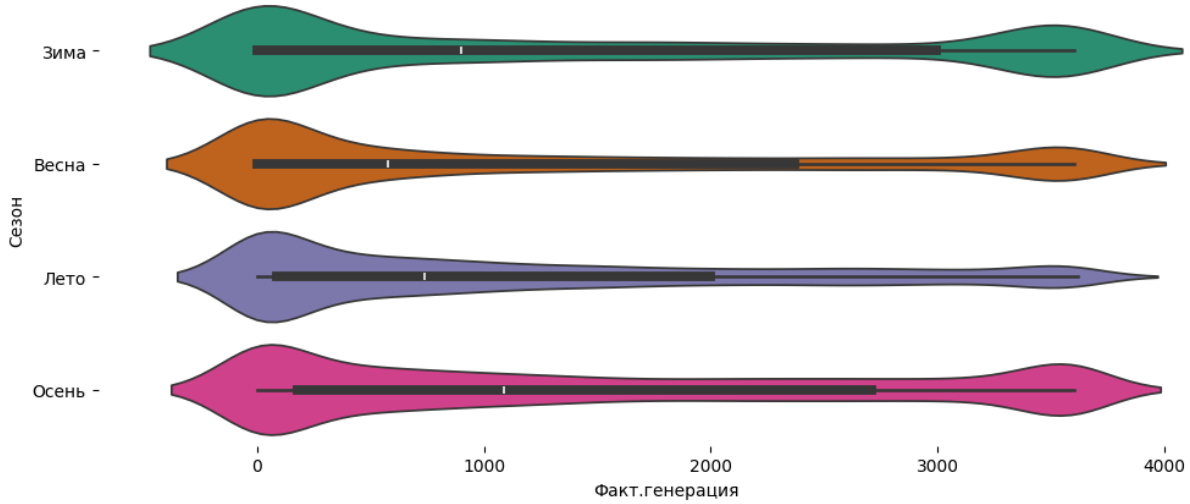


Рисунок 3.25 – Распределение фактической генерации по сезонам года

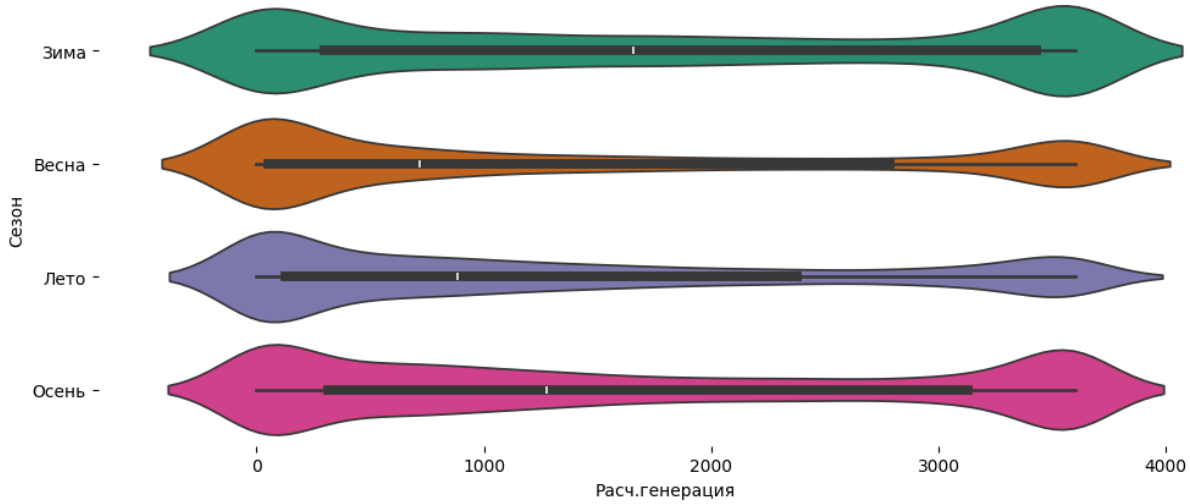


Рисунок 3.26 – Распределение расчетной генерации по сезонам года

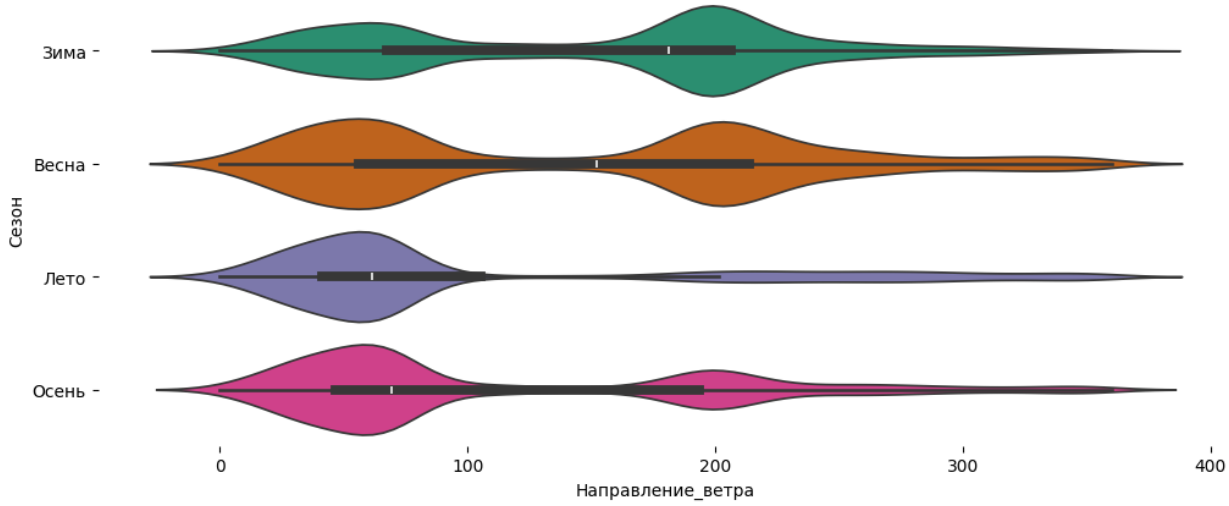


Рисунок 3.27 – Распределение направления ветра по сезонам года

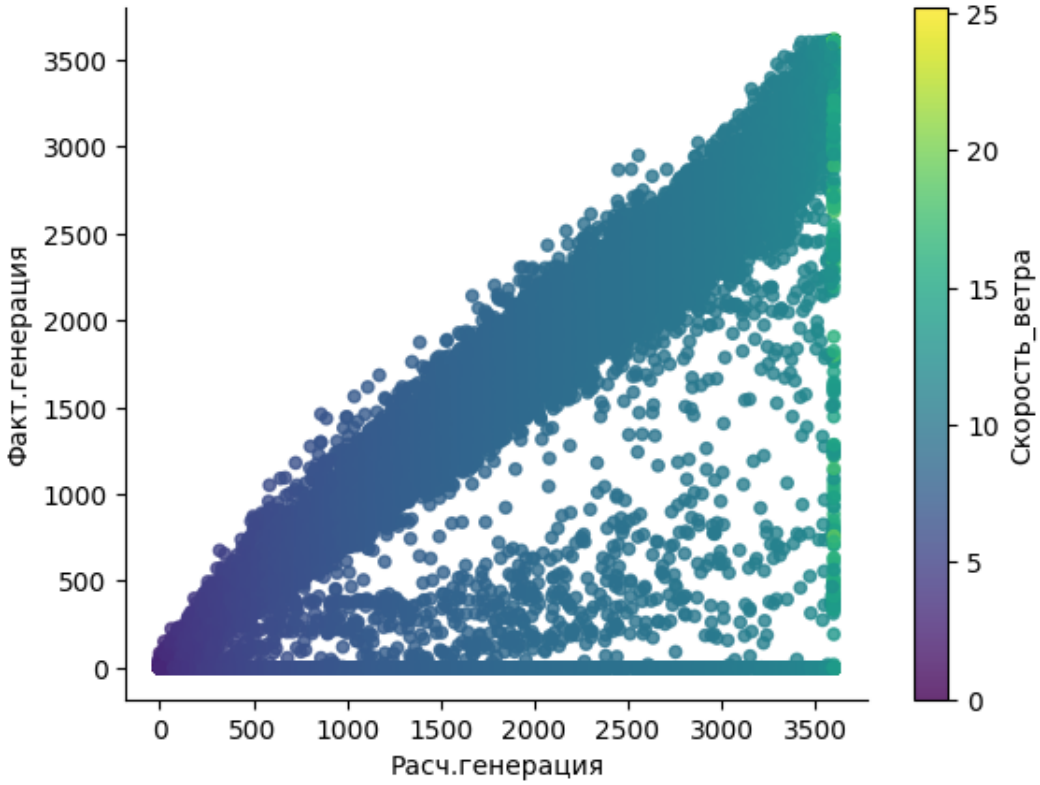


Рисунок 3.28 – Зависимость фактической генерации
от расчетной при различных скоростях ветра

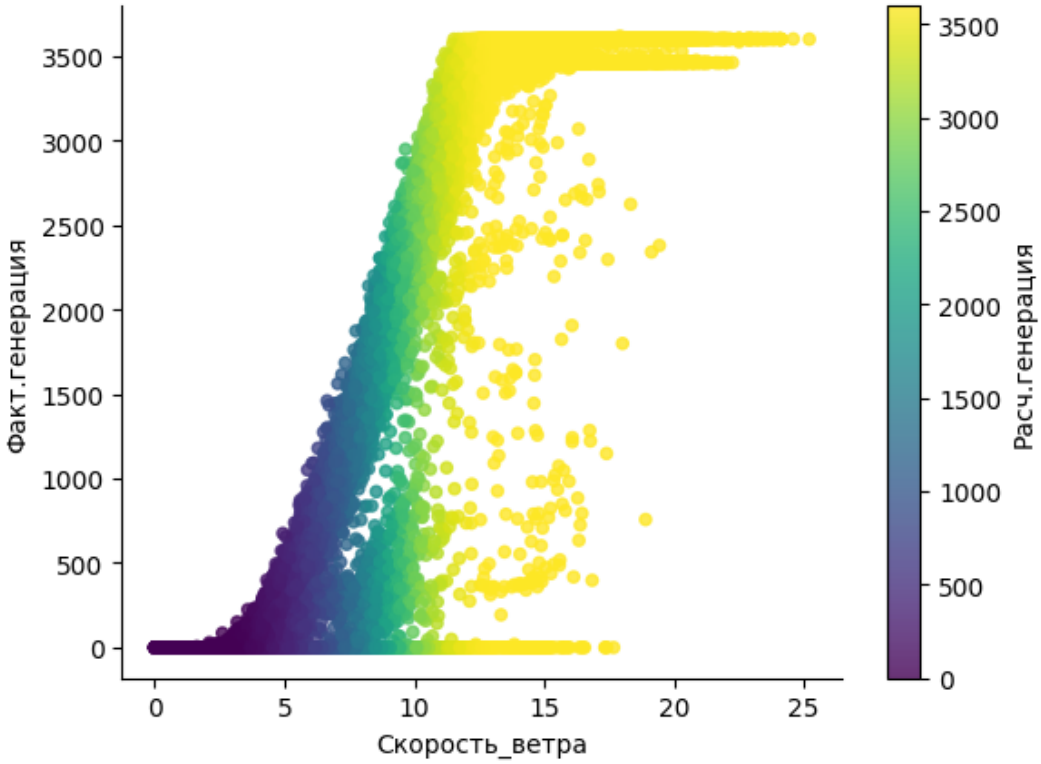


Рисунок 3.29 – Зависимость фактической генерации от скорости ветра
при различных значениях расчетной генерации

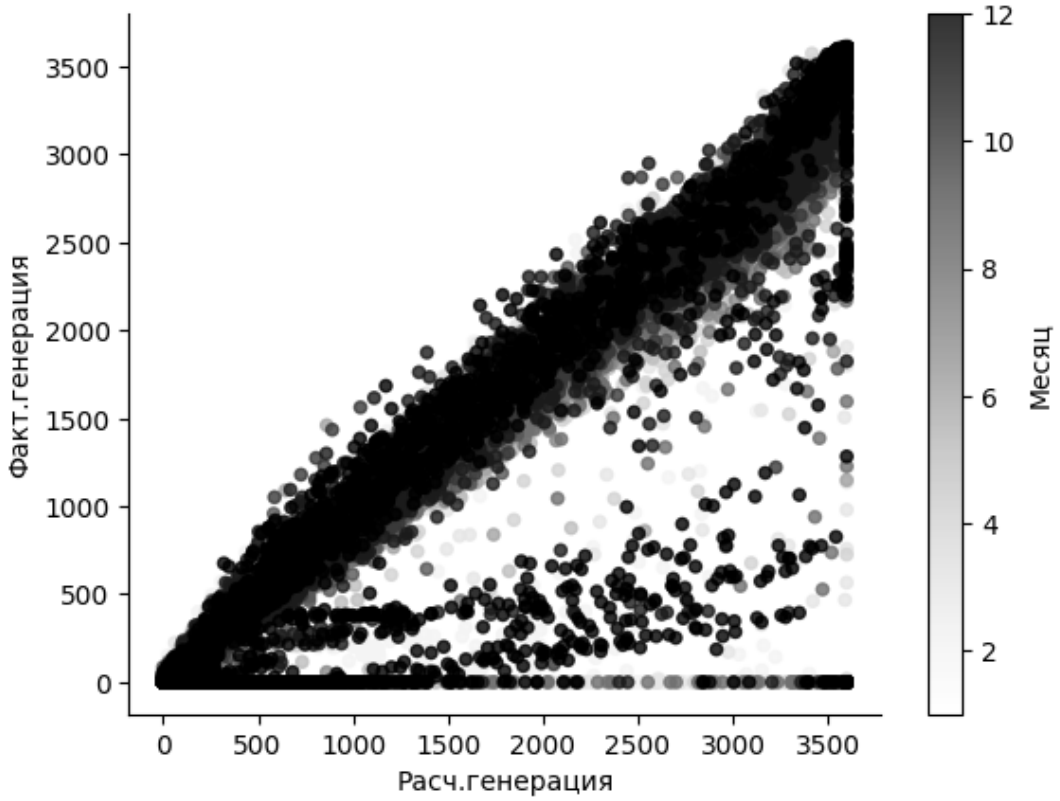


Рисунок 3.30 – Зависимость фактической генерации
от расчетной в различные месяцы года

На рисунках 3.31–3.33 показаны фрагменты временного ряда фактических и расчетных мощностей ветроэнергетической установки.

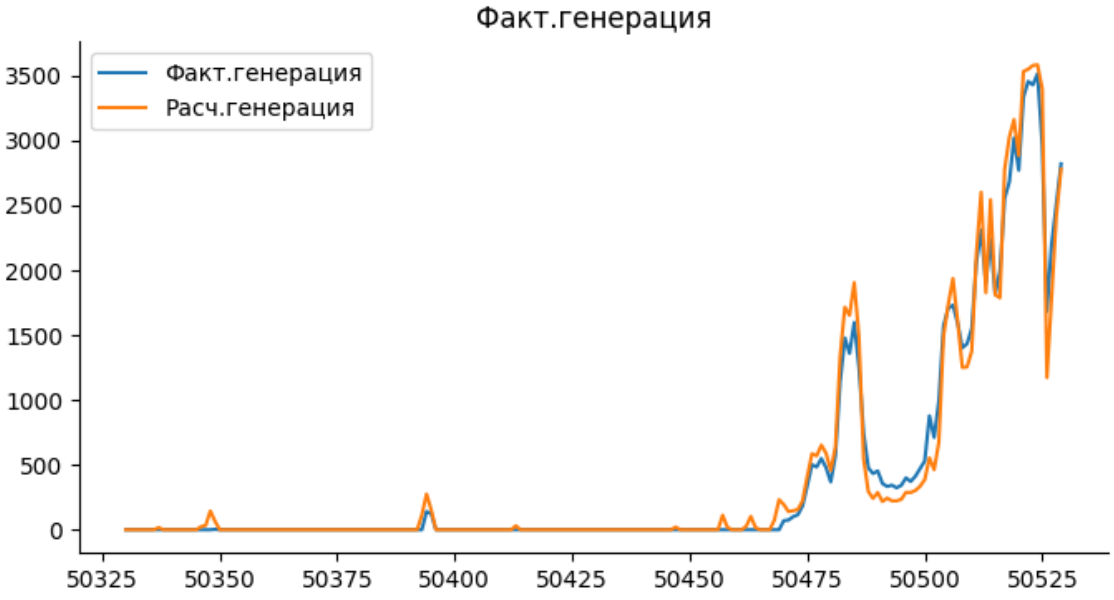


Рисунок 3.31 – Фрагмент временного ряда



Рисунок 3.32 – Фрагмент временного ряда



Рисунок 3.33 – Фрагмент временного ряда

На рисунке 3.34 показана визуализация розы ветров, выполненная в виде диаграммы рассеяния.

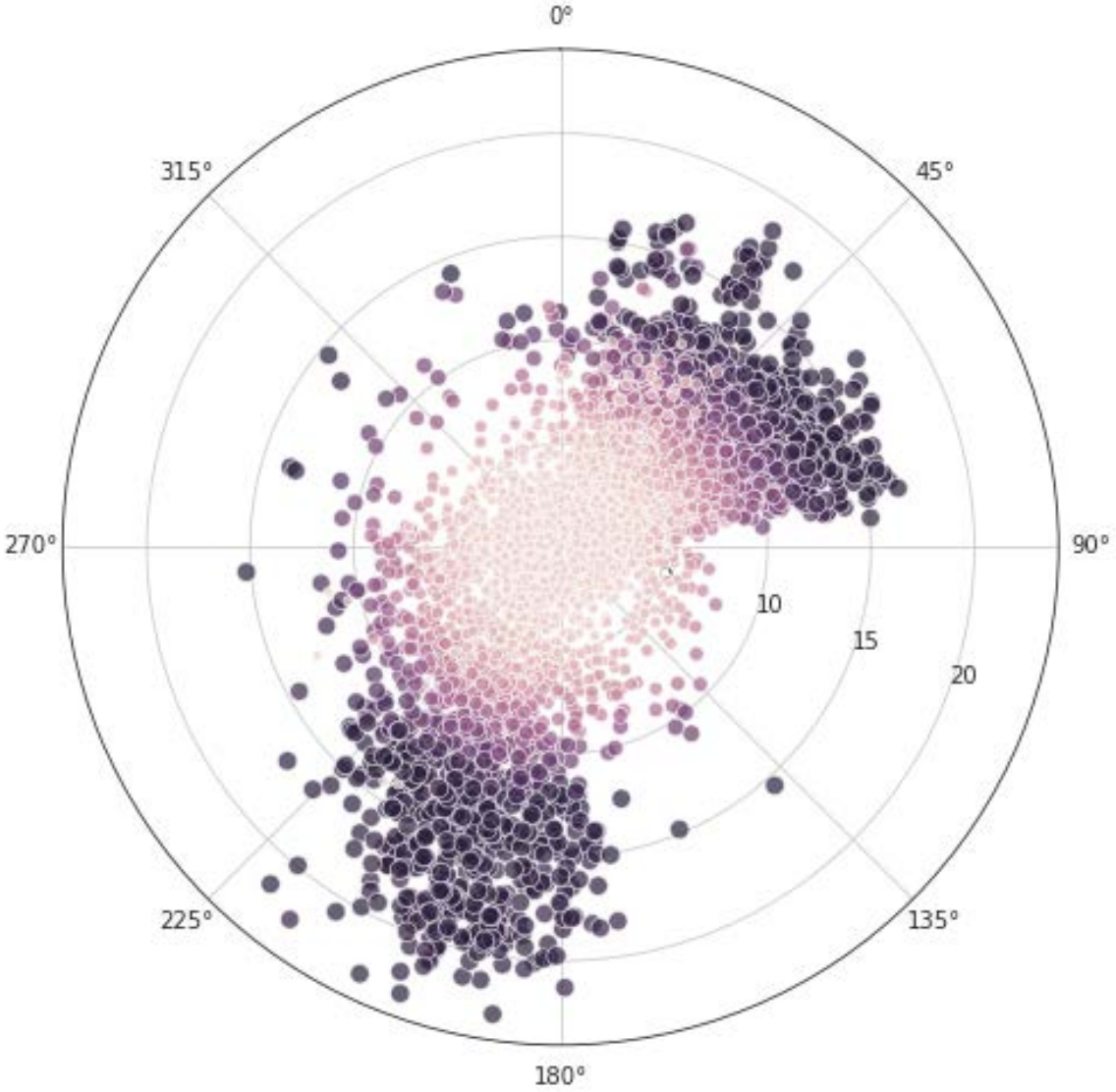


Рисунок 3.34 – Роза ветров

4 РАЗРАБОТКА МОДЕЛИ ПРОГНОЗИРОВАНИЯ ГЕНЕРАЦИИ ВЕТРОЭНЕРГЕТИЧЕСКОЙ УСТАНОВКИ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

4.1 Подключение библиотек обработки данных и машинного обучения

На рисунке 4.1 показан код для подключения необходимых для дальнейшей работы библиотек.

```
from sklearn.preprocessing import LabelEncoder, Normalizer, StandardScaler
from sklearn.preprocessing import MinMaxScaler, RobustScaler
from sklearn.preprocessing import OneHotEncoder, OrdinalEncoder
from sklearn.model_selection import cross_val_score
from sklearn.metrics import accuracy_score,mean_squared_error
from sklearn.compose import ColumnTransformer
from sklearn.impute import SimpleImputer,KNNImputer
from sklearn.pipeline import make_pipeline,Pipeline
from sklearn.base import BaseEstimator,TransformerMixin
from sklearn.ensemble import RandomForestRegressor
```

Рисунок 4.1 – Код подключения необходимых библиотек обработки данных и машинного обучения

4.2 Выделение целевой переменной и признаков

На рисунке 4.2 показан код для отделения целевой переменной от признаков.

```
X_train,y_train,X_test,y_test = df_train.drop(columns='Факт.генерация'), \
  df_train['Факт.генерация'], \
  df_test.drop(columns='Факт.генерация'),\
  df_test['Факт.генерация']

print(X_train.shape)
print(y_train.shape)
print(X_test.shape)
print(y_test.shape)
```

```
(40424, 8)
(40424,)
(10106, 8)
(10106,)
```

Рисунок 4.2 – Код для разделения выборки на целевую переменную и признаки

4.3 Преобразование данных

Для многих методов машинного обучения необходимо выполнять нормализацию или стандартизацию данных. Использована стандартизация следующего вида:

$$x' = (x - m) / s, \qquad (4.1)$$

где x – входное значение, m – математическое ожидание для соответствующего признака, s – его среднеквадратическое отклонение.

Стандартизация набора данных делает закон распределения ближе к нормальному, что положительно сказывается на процессе обучения моделей. Кроме того, в рамках работы созданы реализации кодирования категориальных признаков и заполнения пропущенных значений. Они приведены в Приложении А.

В результате помимо стандартизации выполняется преобразование категориальных признаков (сезон) в бинарный вектор. Список полученных признаков:

- скорость ветра, м/с;
- направление ветра, градусы;
- расчетная выходная мощность, кВ;
- месяц (1–12);
- час суток (0–23);
- зима (бинарный признак);
- весна (бинарный признак);
- лето (бинарный признак);
- осень (бинарный признак).

4.4 Оценка значимости признаков

Одним из способов оценки значимости признаков является применение деревьев решений (предпочтительнее использовать ансамбль), поскольку при построении деревьев решений значимость признаков определяется в процессе построения модели по критерию Джинни [39].

Код для оценки значимости признаков приведен на рисунке 4.3. Результаты показаны на рисунке 4.4.

```
model=RandomForestRegressor(random_state=42)
model.fit(X_1_train_f1,y_train)

feature_importance=model.feature_importances_
feature_importance_df=pd.DataFrame({'Feature':X_1_train_f1.columns,
                                    'Importance':feature_importance})
feature_importance_df=feature_importance_df.sort_values(
    by='Importance',ascending=False)
```

Рисунок 4.3 – Код для оценки значимости признаков

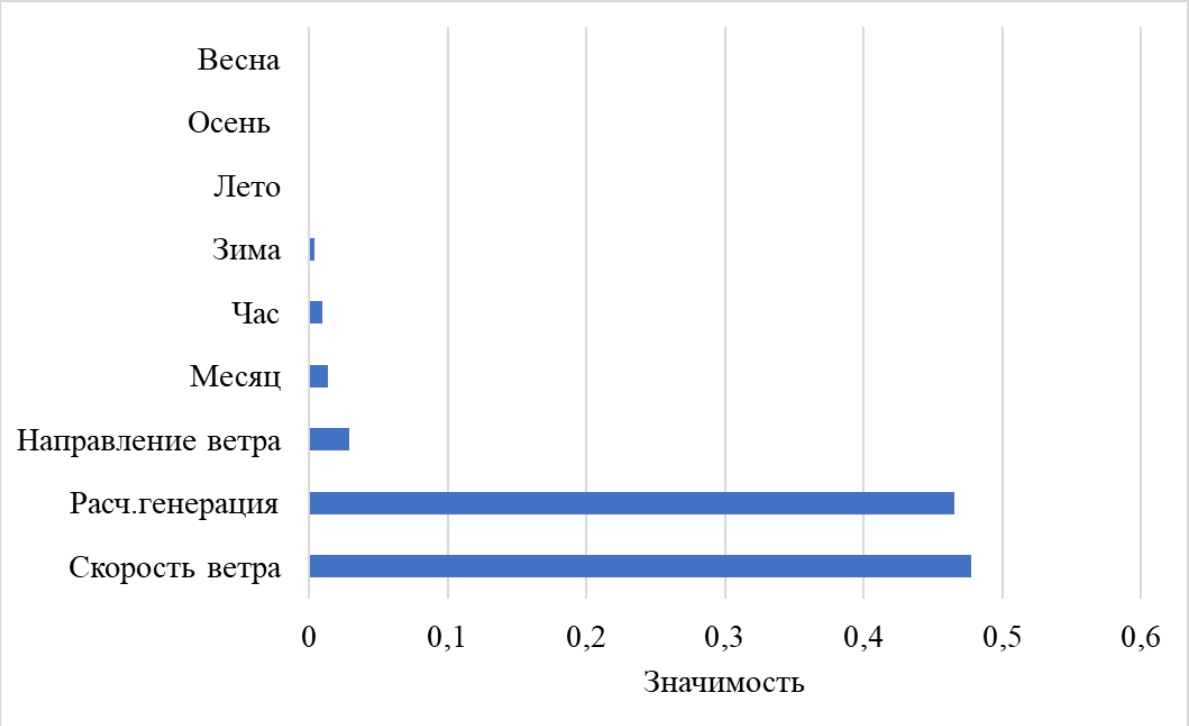


Рисунок 4.4 – Значимости признаков для прогноза выработки ветроэнергетической установки

Видно, что использование только скорости ветра или только расчетной мощности менее эффективно, чем их совместное использование. Кроме того, следует учитывать направление ветра и сезон года.

Таким образом, показано, что изначальная гипотеза о том, что возможно построить более точную модель расчета генерации ветроэнергетической установки, чем заложенная в технические данные установки, предварительно подтверждается.

4.5 Построение моделей машинного обучения

На рисунке 4.5 показан код для подключения моделей и методов машинного обучения.

```
from sklearn.metrics import r2_score,mean_squared_error
from sklearn.ensemble import GradientBoostingRegressor
from sklearn.svm import SVR
from sklearn.ensemble import RandomForestRegressor,\
  ExtraTreesRegressor,AdaBoostRegressor
from sklearn.linear_model import LinearRegression
from sklearn.tree import DecisionTreeRegressor
from xgboost import XGBRegressor,XGBRFRegressor
from sklearn.model_selection import RandomizedSearchCV
from catboost import CatBoostRegressor
```

Рисунок 4.5 – Код для подключения моделей и методов машинного обучения.

Используются следующие модели:

- линейная регрессия;
- метод опорных векторов;
- дерево решений;
- случайный лес;
- экстремально рандомизированные деревья;
- адаптивный бустинг над деревьями решений;
- градиентный бустинг над деревьями решений;
- экстремальный градиентный бустинг;
- категориальный бустинг.

Линейная регрессия – это статистическая модель, используемая для аппроксимации зависимости между набором признаков (независимых переменных) и целевой переменной (зависимой переменной). Линейная регрессия представляет собой математическую модель, где зависимость между признаками и целевой переменной описывается линейной функцией: взвешенной суммой входных признаков и свободного члена. Для определения параметров линейной регрессии используются различные методы, такие как метод наименьших квадратов, максимальное правдоподобие и другие.

Метод опорных векторов (SVM) является популярным алгоритмом машинного обучения, существенно модифицирующим модель логистической и линейной регрессии. SVM стремится найти гиперплоскость, которая максимально разделяет обучающие примеры разных классов с наибольшим зазором (маргинальным ядром) между ними. Это позволяет достичь высокой обобщающей способности модели. Обучающие примеры, которые находятся на границе маргинального ядра или внутри него, называются опорными векторами. Они играют ключевую роль в определении гиперплоскости разделения и определяют границы классов. SVM позволяет использовать ядерную функцию для нелинейного преобразования данных в пространство с более высокой размерностью, где линейное разделение может быть достигнуто. Это позволяет моделировать сложные нелинейные зависимости между признаками. Также SVM включает параметр регуляризации, который позволяет балансировать между точностью модели и ее обобщающей способностью. При переходе к задаче регрессии, вместо разделения классов в SVM стремится найти гиперплоскость, которая минимизирует сумму квадратов отклонений между предсказанными значениями и фактическими значениями обучающих примеров.

Регрессионное дерева решений – модель вида иерархической системы правил, каждое из которых делит подвыборку на две части, на основе сравнения одного признака с пороговым значением. Регрессионное дерево решений строится рекурсивно путем разделения данных на подмножества на основе значений признаков. На каждом узле дерева выбирается наилучший признак и значение разделения, которые минимизируют дисперсию целевой переменной в каждом подмножестве. Для выбора наилучшего разделения на каждом узле дерева используются различные критерии, такие как критерий Джини. Целью является минимизация дисперсии целевой переменной в каждом подмножестве [40].

Существует ряд методов, повышающих точность и стабильность (устойчивость к переобучению и выбросам) деревьев решений. Они основаны на объединении деревьев в ансамбль.

Случайный лес состоит из множества деревьев решений, каждое из которых обучается на случайной подвыборке данных и случайном подмножестве признаков. При обработке новых данных каждое обученное дерево «голосует», и окончательный результат определяется большинством голосов (с задаче классификации) или усреднением (в задаче регрессии). Преимущества: устойчивость к переобучению, возможность обработки больших наборов данных и многоклассовых задач, высокая скорость работы.

Метод экстремально рандомизированных деревьев во много близок к методу случайного леса, но, случайным образом выбирает подмножество признаков для каждого разделения узла. Это позволяет уменьшить корреляцию между деревьями в ансамбле и повысить обобщающую способность модели. Кроме того, на каждом узле дерева случайным образом выбирается порог разделения для каждого выбранного признака. Это также способствует уменьшению корреляции между деревьями в ансамбле.

Адаптивный бустинг последовательно обучает слабые модели (в данном случае деревья решений) на взвешенных выборках данных, где веса экземпляров обновляются на каждой итерации в зависимости от предыдущих ошибок модели. Преимущества: высокая точность, устойчивость к шуму в данных, возможность обработки неполных и несбалансированных данных.

Градиентный бустинг последовательно обучает слабые модели (в данном случае деревья решений) на взвешенных выборках данных, где веса экземпляров обновляются на каждой итерации в зависимости от градиента ошибки модели. Преимущества: высокая точность, возможность обработки больших наборов данных и многоклассовых задач, устойчивость к переобучению.

Экстремальный градиентный бустинг (XGBoost) является это оптимизированной реализацией градиентного бустинга, которая включает в

себя различные улучшения для повышения производительности и точности модели. Преимущества: высокая скорость работы, возможность обработки больших наборов данных, устойчивость к переобучению, возможность параллельной обработки.

Категориальный бустинг (CatBoost) – это оптимизированная реализация градиентного бустинга, которая специально разработана для обработки категориальных признаков. Она включает в себя различные улучшения для повышения производительности и точности модели. Преимущества: высокая скорость работы, возможность обработки больших наборов данных, устойчивость к переобучению, автоматическое обработка категориальных признаков.

Для оптимизации гиперпараметров моделей применен алгоритм случайного поиска. Код приведен на рисунке 4.6.

```
model_name=[]
r2score=[]
rmse=[]
models=[
    GradientBoostingRegressor(random_state=42),
    SVR(),
    RandomForestRegressor(random_state=42),
    LinearRegression(),
    ExtraTreesRegressor(random_state=42),
    AdaBoostRegressor(random_state=42),
    DecisionTreeRegressor(random_state=42),
    XGBRegressor(random_state=42),
    XGBRFRegressor(random_state=42),
    CatBoostRegressor(random_state=42,verbose=False)
]

for model in models:
    model.fit(X_1_train_f1 , y_train)
    y_pred = model.predict(X_1_test_f1)
    model_name.append(model.__class__.__name__)
    r2score.append(str(r2_score( y_test , y_pred ) * 100 ))
    rmse.append(str(mean_squared_error( y_test , y_pred,squared=False )))
models_df
```

Рисунок 4.6 – Код для обучения моделей

4.6 Результаты тестирования моделей

Результаты приведены в таблице 4.1 и на рисунке 4.7.

Использованы следующие показатели точности:

- среднеквадратическая ошибка – root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|y_i^* - y_i\right|^2}, \qquad (4.2)$$

- коэффициент детерминации (R2):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - y_i^*\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2}. \qquad (4.3)$$

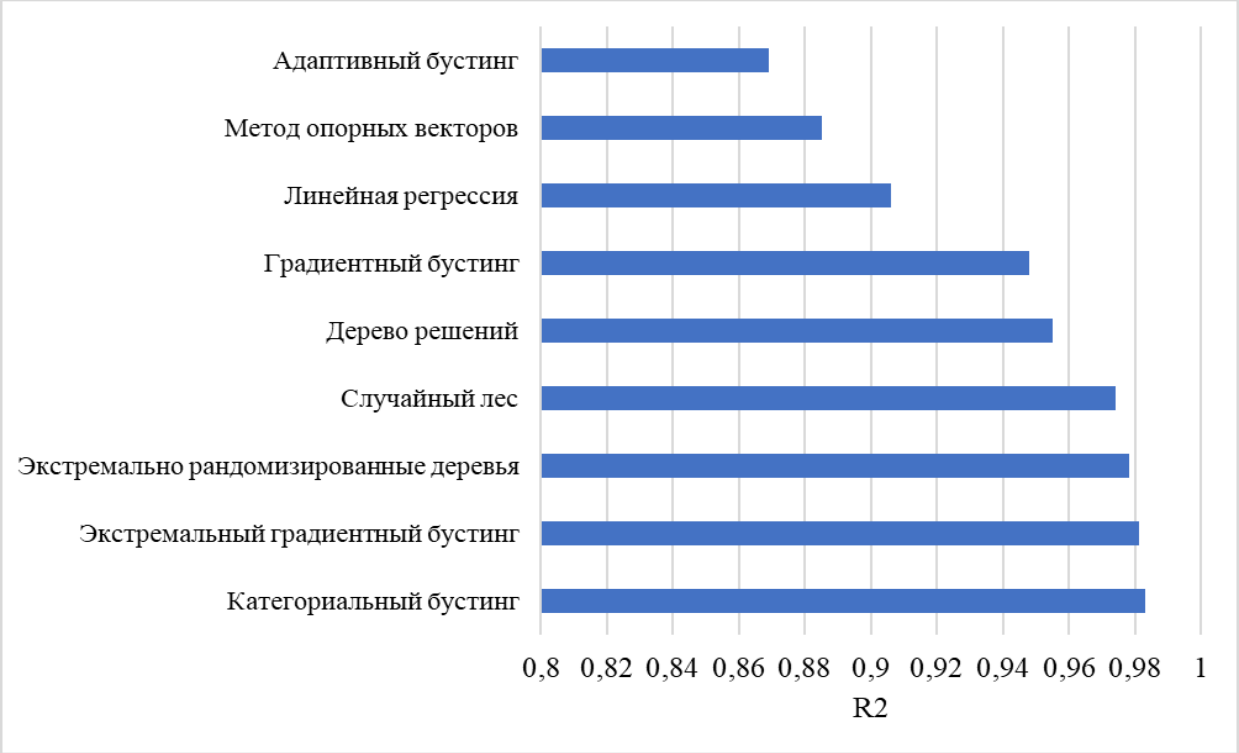


Рисунок 4.7 – Результаты моделей машинного обучения

Таблица 4.1 – Результаты применения моделей машинного обучения

Модель	R^2	RMSE
Категориальный бустинг	0,983	172,62
Экстремальный градиентный бустинг	0,981	177,89
Экстремально рандомизированные деревья	0,978	195,47

продолжение таблицы 4.1

Модель	R^2	RMSE
Случайный лес	0,974	210,66
Дерево решений	0,955	276,24
Градиентный бустинг	0,948	298,51
Линейная регрессия	0,906	400,34
Метод опорных векторов	0,885	443,14
Адаптивный бустинг	0,869	473,55

Наилучший результаты получен при использовании категориального бустинга со следующими значениями гиперпараметров:

- шаг обучения: 0.235;
- максимальная глубина деревьев: 9;
- число деревьев: 50.

Из проведенных экспериментов можно сделать ряд выводов:

- искомая модель не является линейной;
- важен учет категориальных признаков, в данном случае – сезонов года, что означает влияние сезонных факторов, возможно, связанных с температурой воздуха и осадками;
- ансамбли деревьев решений, за исключением адаптивного бустинга, показали близкую точность и превзошли остальные модели;
- слабый результат адаптивного бустинга можно свидетельствовать о наличии в данных выбросов, которые могут являться как ошибками в данных, так и редкими, но возможными значениями.

В таблице 4.2 приведены несколько примеров из выборки для сравнения расчетных значений мощности, прогнозных и фактических.

Таблица 4.2 Несколько примеров из выборки для сравнения расчетных значений мощности, прогнозных и фактических.

Фактическое значение, кВт	Прогнозное значение, кВт	Расчетное значение, кВт
3348,91	3224,844	3452,92
2647,51	2464,95	2989,12
0,00	0,00	779,12
1334,57	1301.16	1502,73
533,00	520.33	593,89

На рисунке 4.8 приведена визуализация расчетных, прогнозных и фактических значений для всей тестовой выборки данных.

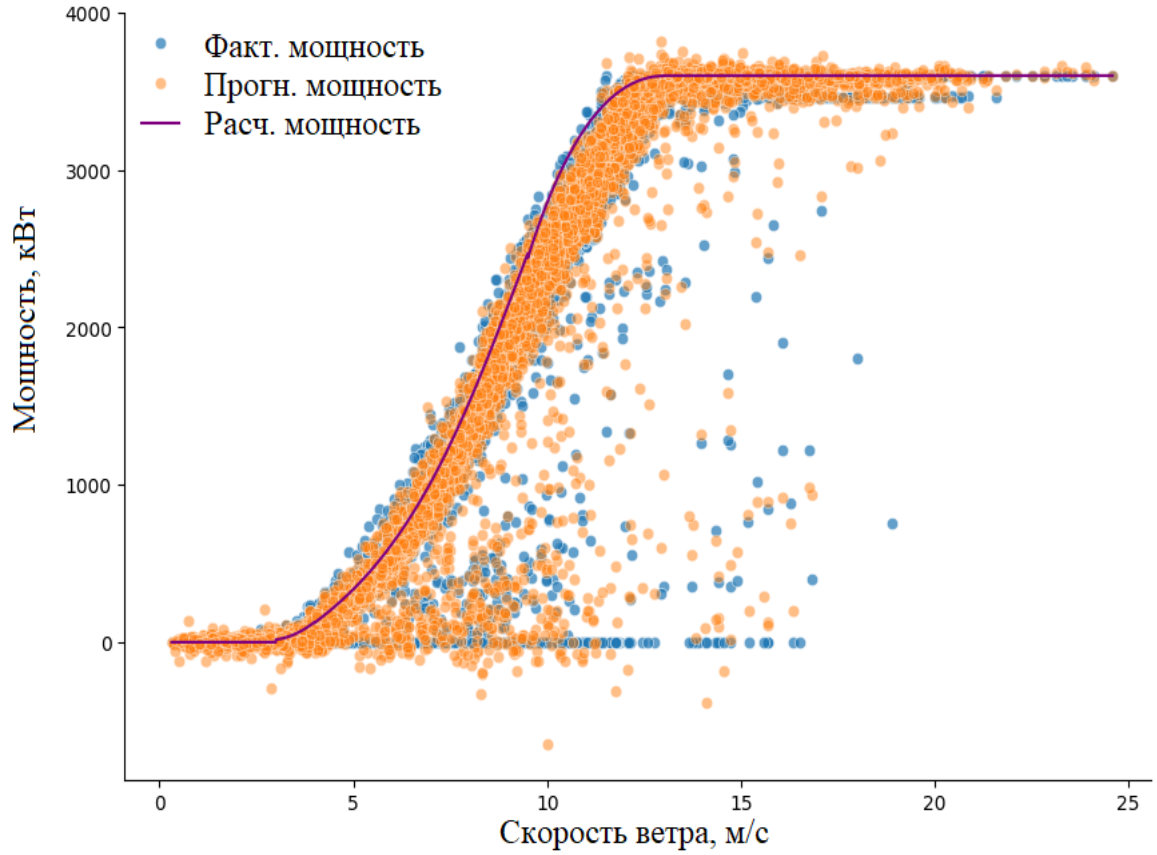


Рисунок 4.8 – Сопоставление расчетных, прогнозных и фактических значений для всей тестовой выборки данных.

Полученные результаты, в том числе высокое значение коэффициента детерминации, показывают, что модель успешно справляется с решением задачи прогнозирования выходной мощности ветроэнергетической установки и позволяет оценивать выходную мощность существенно точнее, чем широко используемая функциональна зависимость выработки от скорости ветра, определяемая с помощью расчетного выражения из паспортных данных установки.

Необходимо отметить, что точность прогнозирования мощности будет выше, чем точность прогнозирования с помощью чем широко используемой модели при любом горизонте прогнозирования. Точности прогноза на определенный горизонт зависит от точности прогноза скорости ветра и направления ветра.

На рисунке 4.9 приведен код для сохранения процесса обработки данных и модели машинного обучения.

```
best_model = CatBoostRegressor(**best_params)
best_model.fit(X_1_train_f1, y_train,verbose=False)
y_pred = best_model.predict(X_1_test_f1)

import joblib
joblib.dump(f1, 'one_hot_pipeline.pkl')

joblib.dump(best_model, 'one_hot_model.pkl')
```

Рисунок 4.9 – Код для сохранения полученного процесса обработки данных и модели машинного обучения

ЗАКЛЮЧЕНИЕ

В качестве объекта для анализа данных выбрана реальная ветроэнергетическая установка, для которой имеются данные из SCADA-системы, включающие фактическую и расчетную мощность, скорость и направление ветра за год с десятиминутным интервалом.

Проведен статистический и визуальный анализ данных, реализован код для подготовки выборки данных к применению методов машинного обучения.

Показано, что для повышения точности прогнозирования генерации ветроэнергетической установки необходимо либо прогнозировать величину вырабатываемой мощности напрямую, либо создавать модели, позволяющие вычислять ее по скорости ветра с меньшей погрешностью, чем заложена в расчетных зависимостях выработки от скорости ветра, указанных в паспортных данных.

Выполнена разработка алгоритма прогнозирования генерации ветроэнергетической установки на базе технологий искусственного интеллекта. Обучены и протестированы различные модели машинного обучения. Определено существенное улучшение точности вычисления генерации ветровой турбины с помощью ансамблей регрессионных деревьев решений по сравнению с используемой в настоящее время моделью зависимости выходной мощности ветроэнергетической установки от скорости ветра.

Среднеквадратичная ошибка составила 173 кВт при номинальной мощности рассматриваемой ветроэнергетической установки 3,6 МВт, таким образом, отношение среднеквадратичной ошибки к номинальной мощности составляет лишь 5 %. Значение коэффициента детерминации 0,98 также свидетельствует о высокой точности модели.

Определено, что наилучшая точность может быть получена с использованием ансамблей деревьев решений, что согласуется с

существующими работами по прогнозированию временных рядов в электроэнергетике [41, 42, 43, 44].

Установлено, что зависимость выходной мощности от скорости ветра не является линейной, а также меняется в зависимости от сезона года. Выявлено, что, в данных присутствуют статистические выбросы, которые могут являться как ошибками в данных, так и редкими, но возможными значениями.

В качестве направления дальнейших исследований можно выделить исследование моделей, использующих прогнозные значения скорости и направления ветра, а также метеорологические параметры и технологические параметры ветроэнергетических установок.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. International Renewable Energy Agency. Renewable capacity highlights // 2023, URL:https://www.irena.org/Publications/2023/Mar/Renewable-capacity-statistics-2023 (дата обращения: 09.03.2024).
2. Energy institute. Statistical Review of World Energy // 2023, 72nd edition URL:https://www.energyinst.org/__data/assets/pdf_file/0004/1055542/EI_Stat_Review_PDF_single_3.pdf (дата обращения: 09.03.2024).
3. International Energy Agency. Renewables 2024. Analysis and forecast to 2028 // 2024, URL:https://iea.blob.core.windows.net/assets/96d66a8b-d502-476b-ba94-54ffda84cf72/Renewables_2023.pdf (дата обращения: 09.03.2024).
4. Блог компании Сбер. Энергия ветра: как развивается возобновляемая энергетика, и когда она вытеснит уголь. // 2023, URL: https://habr.com/ru/companies/sberbank/articles/737268/ (дата обращения: 09.03.2024).
5. COP28 UAE // 2023, URL:https://www.cop28.com (дата обращения: 12.03.2024).
6. Минэнерго России Возобновляемая энергетика в России и мире // 2022, Москва, URL: https://rosenergo.gov.ru/upload/iblock/e04/3xtm87iv99x76b23c6wjul3as5pzz8zj.pdf (дата обращения: 12.03.2024)
7. Ассоциация развития возобновляемой энергетики (АРВЭ) // 2024, URL:https://rreda.ru/industry/statistics/#graph1 (дата обращения: 10.03.2024).
8. Возобновляемые источники энергии, зеленые инновации, энергосбережение и электротранспорт // Информационно-аналитическое издание выставки и форума «renwex. возобновляемая энергетика и электротранспорт», 2023, URL: https://www.renwex.ru/common/img/uploaded/exhibitions/renwex/doc/2023/digest-2023_01_ru.pdf

9. Российская Ассоциация Возобновляемых Источников и Электротранспорта // 2024, URL: https://rawi.ru (дата обращения: 09.03.2024).
10. Об электроэнергетике: Федер. закон Рос. Федерации от 26 марта 2003 г. № 35-ФЗ: принят Гос. Думой Федер. Собр. Рос. Федерации 21 февр. 2003 г.: одобр. Советом Федерации Федер. Собр. Рос. Федерации 12 марта 2003 г. // Рос. газ. 2003. 17 марта. URL: https://www.consultant.ru/document/cons_doc_LAW_41502/ (дата обращения: 15.03.2024).
11. Аликина В.С., Романова Т.Н. Перспективы развития ветроэнергетики // Международная научно-практическая конференция студентов, аспирантов и молодых ученых, посвященная памяти профессора Данилова Н. И. (1945–2015) — Даниловские чтения «Энерго- и ресурсосбережение. Энергообеспечение. Нетрадиционные и возобновляемые источники энергии. Атомная энергетика», 2019, URL: http://elar.urfu.ru/handle/10995/88128
12. Горячев С.В., Смолякова А.А. Проблемы и перспективы ветроэнергетических систем в России // Международный научно-исследовательский журнал 2022, Вып. № 5(119) Часть 1 DOI: https://doi.org/10.23670/IRJ.2022.119.5.048
13. Юрченко И.В. Обзор рынка ветроэнергетики России // Экономические науки, 2023, Вып. №3(220) С. 18–24 DOI: 10.14451/1.220.18
14. АО «НоваВинд» дивизион Росатома // 2024, URL: https://novawind.ru/company/ (дата обращения: 05.03.2024).
15. Doha B., Touria H., Mariam El J. Review of current artificial intelligence methods and metaheuristic algorithms for wind power prediction // Indonesian Journal of Electrical Engineering and Computer Science, 2023 Vol. 29, p.p. 626–634 DOI: 10.11591/ijeecs.v29.i2.pp626-634

16. Artificial Intelligence in Wind Speed Forecasting: A Review / S. Valdivia-Bautista, J. Domínguez-Navarro, M. Pérez-Cisneros [et al.] // 2023, Energies, URL:https://doi.org/10.3390/en16052457.
17. Grand challenges in the science of wind energy / Paul Veers, Katherine Dyke, Eric Lantz [et al.] // 2019, SCIENCE, Vol. 366, DOI: 10.1126/science.aau2027s
18. Котов А.С., Сивеев Т.М., Груздов А.Г. Обзор методов прогнозирования генерации энергии // 2022, Научный сетевой журнал «Столыпинский вестник», №8
19. Ali B.M., Wind Energy Prediction: Artificial Intelligence Perspective // 2023, 6th International Conference on Engineering Technology and its Applications (IICETA), Al-Najaf, Iraq, p.p. 885-891, DOI: 10.1109/IICETA57613.2023.10351296.
20. Юсупов К.Н., Тожибоев С.Т. Обзор методов прогнозирования в ветроэнергетике // 2022, СОК, №12, С. 70-73 URL: https://www.c-o-k.ru/get_issue_file.php?file_name=202212.pdf
21. Artificial Intelligence Methods Applied To Wind And Solar Energy Forecasting: A Comparative Study of Current Techniques / B. Doha, H. Touria, D. Mounir, E. Faissal [et al.] // 2023, International Conference on Digital Age & Technological Advances for Sustainable Development (ICDATA), Casablanca, Morocco, p.p. 94-99, DOI: 10.1109/ICDATA58816.2023.00026.
22. Ultra-short-term wind power forecasting techniques: comparative analysis and future trends / Y. Guangzheng, S. Lingxu, D. Qi, [et al.] // 2024, Frontiers in Energy Research, Vol. 11, DOI:10.3389/fenrg.2023.1345004.
23. Объединение лидеров нефтегазового сервиса и машиностроения россии национальная ассоциация нефтегазового сервиса // 2024, URL: https://nangs.org/news/renewables/wind (дата обращения: 15.03.2024).

24. Современное производство и техника // 2024, URL: https://itexn.com/8295_vetrojenergetika-vidy-ustrojstvo-ispolzovanie-vetrogeneratorov-vetrojenergeticheskie-parki.html (дата обращения: 15.03.2024).
25. Ветер Энергоперемен // 2018, Атомный эксперт. Технологии, Вып. №6, URL: https://atomicexpert.com/novawind_rosatom
26. Jamii J., Mansouri M., Mimouni M. F. Short Time Forecasting for Wind Power Generation Using Artificial Neural Network // 2022, 19th International Multi-Conference on Systems, Signals & Devices (SSD), Sétif, Algeria, p.p. 321-325, doi: 10.1109/SSD54932.2022.9955856.
27. Comparative analysis of approaches to short-term forecasting of electricity generation of wind turbine / N. Sergeev, P. Matrenin, N. Zubova. [et al.] // 2023, AIP Conference Proceedings, Vol. 2910(1), Art. 020144 (7 p.), DOI 10.1063/5.0175251.
28. Desai S., Sampat S., Vadodaria D., Pimpale M., Wind Energy Forecasting Using Artificial Intelligence // 2022, 4th International Conference on Artificial Intelligence and Speech Technology (AIST), Delhi, India, p.p. 1-6, doi: 10.1109/AIST55798.2022.10065273
29. Hydrogen energy storage systems to improve wind power plant efficiency considering electricity tariff dynamics / N. G. Kiryanova, P. V. Matrenin, S. V. Mitrofanov [et al.] // 2022, International Journal of Hydrogen Energy, Vol. 47, No. 18. P. 10156-10165, DOI 10.1016/j.ijhydene.2022.01.152.
30. Устойчивое краткосрочное прогнозирование скорости ветра с помощью адаптивных компактных нейронных сетей / П. В. Матренин, В. З. Манусов, Е. А. Игумнова [и др.] // 2020, Проблемы региональной энергетики, № 3(47), С. 69-80, DOI 10.5281/zenodo.4018960.
31. Куприков А.А., Томин Н.В., Краткосрочное прогнозирование выработки мощности ветроэнергетических установок на базе

ансамблевых методов машинного обучения // 2023, Сборник научных трудов ангарского государственного технического университета Том № 1, С. 235-243, URL:https://doi.org/10.36629/2686-7788-2023-1-235-243

32. Ateş K.T. Estimation of Short-Term Power of Wind Turbines Using Artificial Neural Network (ANN) and Swarm Intelligence. // 2023, Sustainability, №15, p. 13572, URL:https://doi.org/10.3390/su151813572
33. Donadio L., Fang J.; Porté-Agel F. Numerical Weather Prediction and Artificial Neural Network Coupling for Wind Energy Forecast // 2021, Energies, №14, p.338. https://doi.org/10.3390/en14020338
34. Improved EMD-Based Complex Prediction Model for Wind Power Forecasting / O. Abedinia, M. Lotfi, M. Bagheri, B. Sobhani, [et al.] // 2020, IEEE Transactions on Sustainable Energy, vol. №11, no. 4, pp. 2790-2802, DOI: 10.1109/TSTE.2020.2976038.
35. Buhan S., Çadırcı I. Multistage Wind-Electric Power Forecast by Using a Combination of Advanced Statistical Methods // 2015, IEEE Transactions on Industrial Informatics, vol. 11, no. 5, pp. 1231-1242, DOI: 10.1109/TII.2015.2431642.
36. Kim, Y., Hur, J. An Ensemble Forecasting Model of Wind Power Outputs Based on Improved Statistical Approaches // 2020, Energies, №13, p. 1071. URL:https://doi.org/10.3390/en13051071
37. Artificial Intelligence Based Hybrid Forecasting Approaches for Wind Power Generation: Progress, Challenges and Prospects / M. S. HOSSAIN LIPU, MD. SAZAL MIAH, M. A. HANNAN [et al.] // 2021, IEEE Access, vol. 9, pp. 102460-102489, DOI: 10.1109/ACCESS.2021.3097102
38. Effective artificial neural network-based wind power generation and load demand forecasting for optimum energy management / J. Jamii, M. Mansouri, M. Trabelsi, M. Fouazi Mimouni [et al.] // 2022, Energy Research, Vol. 10, URL: https://doi.org/10.3389/fenrg.2022.898413

39. Analysis of Artificial Intelligence based Forecasting Techniques for Renewable Wind Power Generation / J. Ranga, T. Arun Srinivas, S. Kumar, H. Anandaram [et al.] // 2022, Smart Technologies, Communication and Robotics (STCR), Sathyamangalam, India, pp. 1-5, DOI: 10.1109/STCR55312.2022.10009141
40. Региональные аспекты ветроэнергетики / В. А. Стенников, В. Г. Курбацкий, Б. Г. Санеев, А. М. Клер [и др.] по общ. ред. В. А. Стенникова, В. Г. Курбацкого: Книга, 2020, Новосибирск DOI 10.15372/REGIONAL2020SVA
41. Горшенин А.Ю., Васина Д.И. Сравнение используемых методов при прогнозировании выработки электроэнергии ветроэлектростанциями // 2023, Математические структуры и моделирование, №3(67). С. 36–42 DOI 10.24147/2222-8772.2023.3.36-42
42. Повышение точности прогнозирования генерации фотоэлектрических станций на основе алгоритмов k-средних и k-ближайших соседей / П.В. Матренин, А.И. Хальясмаа, В.В. Гамалей, С.А. Ерошенко [и др.] // 2023, Энергетика. Известия высших учебных заведений и энергетических объединений СНГ, №66(4), С. 305-321. URL:https://doi.org/10.21122/1029-7448-2023-66-4-305-321
43. Ансамблевая модель для прогнозирования выработки ветровых электростанций / А. Г. Русина, О. Тувшин, П. В. Матренин, Н. Н. Сергеев [и др.] // 2024, Известия высших учебных заведений. Проблемы энергетики, Том №26, № 1, С. 64–76. URL:https://doi.org/10.30724/1998-9903-2024-26-1-64-76.
44. Grey Wolf Optimizer for RES Capacity Factor Maximization at the Placement Planning Stage / A.M. Bramm, S.A. Eroshenko, A.I. Khalyasmaa, P.V. Matrenin [et al.] // 2023, Mathematics, Vol. 11, p. 2545, URL:https://doi.org/10.3390/math11112545.

ПРИЛОЖЕНИЕ А
ФРАГМЕНТ КОДА ДЛЯ ПРЕОБРАЗОВАНИЯ ДАННЫХ

```
class SSTransform(estimator, mixin):

    def __init__(self, cols):
        self.cols = cols
        self.scaler_ = None

    def fit(self, X, y=None):
        self.scaler_ = StandardScaler().fit(X.loc[:, self.cols])
        return self

    def transform(self, X):
        X_copy = X.copy()
        X_copy.loc[:, self.cols] = self.scaler_.transform(X_copy.loc[:,
self.cols])
        return X_copy

    def fit_transform(self, X, y=None):
        self.scaler_ = StandardScaler().fit(X.loc[:, self.cols])
        return self.transform(X)

class DFImputer(mixin, estimator):

    def __init__(self, mc=None, knnc=None):
        self.mc = mc
        self.knnc = knnc

    def fit(self, X, y=None):
```

```
            self.median_imputer = SimpleImputer(strategy='median')
            self.knn_imputer = KNNImputer()

            if self.mc is not None:
                self.median_imputer.fit(X[self.mc])
            if self.knnc is not None:
                self.knn_imputer.fit(X[self.knnc])
            return self

        def transform(self, X):
            X_imputed = X.copy()
            if self.mc is not None:
                X_median = pd.DataFrame(self.median_imputer.transform(X[self.mc]),
                                        columns=self.mc, index=X.index)
                X_imputed = pd.concat([X_imputed.drop(self.mc, axis=1), X_median], axis=1)
            if self.knnc is not None:
                X_knn = pd.DataFrame(self.knn_imputer.transform(X[self.knnc]),
                                     columns=self.knnc, index=X.index)
                X_imputed = pd.concat([X_imputed.drop(self.knnc, axis=1), X_knn], axis=1)
            return X_imputed

        def fit_transform(self, X, y=None):
            self.fit(X)
            return self.transform(X)

    class Extractor(estimator, mixin):
        def __init__(self, date_cols):
```

```
        self.date_cols = date_cols


    def fit(self, X, y=None):
        return self

    def transform(self, X):
        ef = []
        for col in self.date_cols:
            dates = pd.to_datetime(X[col], format='%d %m %Y %H:%M')
            for date in dates:
                month_val = date.month
                week_val = date.day // 7 + 1
                day_val = date.day
                hour_val = date.hour + 1
                # Determining season based on month
                if month_val in [3, 4, 5]:
                    season_val = 'Весна'
                elif month_val in [6, 7, 8]:
                    season_val = 'Лето'
                elif month_val in [9, 10, 11]:
                    season_val = 'Осень'
                else:
                    season_val = 'Зима'
                ef.append([month_val, week_val, day_val, season_val, hour_val])

        X_d = pd.DataFrame(ef, columns=['Месяц', 'Неделя', 'День', 'Сезон',
'Час'])
        return pd.concat([X.reset_index(drop=True),X_d],axis=1)
```

```
 def fit_transform(self, X, y=None):
        self.fit(X)
        return self.transform(X)

class OTransformer(estimator, mixin):

    def __init__(self, column, q1=0.25, q3=0.75):
        self.column = column
        self.q1 = q1
        self.q3 = q3

    def outlier_threshhold(self, dataframe, column):
        Q1 = dataframe[column].quantile(self.q1)
        Q3 = dataframe[column].quantile(self.q3)
        iqr = Q3 - Q1
        up_limit = Q3 + 1.5 * iqr
        low_limit = Q1 - 1.5 * iqr
        return low_limit, up_limit

    def fit(self, X, y=None):
        return self

    def transform(self, X):
        X_copy = X.copy()
        for col in self.column:
            low_limit, up_limit = self.outlier_threshhold(X_copy, col)
            X_copy.loc[(X_copy[col] < low_limit), col] = low_limit
            X_copy.loc[(X_copy[col] > up_limit), col] = up_limit
        return X_copy
```

```
    def fit_transform(self, X, y=None):
        return self.transform(X)


class ColDropper(estimator, mixin):

    def __init__(self, columns=None):
        self.columns = columns

    def fit(self, X, y=None):
        return self

    def transform(self, X):
        if self.columns is None:
            return X
        else:
            return X.drop(self.columns,axis=1)


class OHEncoder(estimator, mixin):

    def __init__(self, columns=None):
        self.columns = columns
        self.unique_values = {}
        self.feature_names_ = None

    def fit(self, X, y=None):
        if self.columns is None:
            self.columns = X.columns.tolist()
        self.unique_values = {col: X[col].unique() for col in self.columns}
```

```
        self.feature_names_ = self._get_feature_names()
        return self

    def _get_feature_names(self):
        feature_names = []
        for col in self.columns:
            for value in self.unique_values[col]:
                feature_names.append(f"{col}_{value}")
        return feature_names

    def transform(self, X):
        X_transformed = pd.DataFrame(index=X[self.columns].index)

        for col in self.columns:
            for value in self.unique_values[col]:
                X_transformed[f"{col}_{value}"] = (X[col] == value).astype(int)

        X = pd.concat([X, X_transformed], axis=1)
        return X.drop(columns=['Season'])

    def fit_transform(self, X, y=None):
        self.fit(X)
        return self.transform(X)


class LEColumns(estimator, mixin):

    def __init__(self, columns):
        self.columns = columns
        self.encoders_ = {}
```

```
        def fit(self, X, y=None):
            for col in self.columns:
                encoder = LabelEncoder()
                encoder.fit(X[col])
                self.encoders_[col] = encoder
            return self

        def transform(self, X):
            X_copy = X.copy()
            for col, encoder in self.encoders_.items():
                X_copy[col] = encoder.transform(X_copy[col])
            return X_copy

        def fit_transform(self, X, y=None):
            self.fit(X, y)
            return self.transform(X)

    class ResPipeline:
        def __init__(self) :
            self.date_cols=['Дата_Время']
            self.numerical_cols=['Скорость_ветра', 'Расч.генерация',
'Направление_ветра', 'Неделя', 'Месяц', 'Час', 'День']
            self.MLE_cols=['Сезон']
            self.full_pipeline=Pipeline([
                ('extract_date',Extractor(date_cols=self.date_cols)),
                ('label_encode', OHEncoder(columns=self.MLE_cols)),
                ('impute_num',DFImputer(knnc=self.numerical_cols)),
                ('remove_outlier',OTransformer(column=self.numerical_cols)),
                ('scale', SSTransform(cols=self.numerical_cols)),
```

```
        ('drop', ColDropper(columns=self.date_cols)),
    ])
def fit_transform(self, X_train):
    X_train = self.full_pipeline.fit_transform(X_train)
    return X_train
def transform(self, X_test):
    X_test = self.full_pipeline.transform(X_test)
    return X_test
```