

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

На правах рукописи



СНЕГИРЕВ ДЕНИС АЛЕКСЕЕВИЧ

**ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ
КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ ГЕНЕРАЦИИ
ВЕТРОЭЛЕКТРИЧЕСКИХ СТАНЦИЙ**

2.4.3. Электроэнергетика

АВТОРЕФЕРАТ

диссертации на соискание учёной степени
кандидата технических наук

Екатеринбург – 2025

Работа выполнена на кафедре «Автоматизированные электрические системы», Уральского энергетического института, Федерального государственного автономного образовательного учреждения высшего образования «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина».

Научный руководитель: доктор технических наук, профессор
Паздерин Андрей Владимирович

Официальные оппоненты: **Соломин Евгений Викторович**,
доктор технических наук, профессор, Федеральное государственное автономное образовательное учреждение высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)», г. Челябинск, профессор кафедры «Электрические станции, сети и системы электроснабжения»;

Чемборисова Наиля Шавкатовна,
доктор технических наук, профессор, Федеральное государственное бюджетное образовательное учреждение высшего образования «Национальный исследовательский университет «МЭИ» г. Москва, профессор кафедры электроэнергетических систем;

Суворов Алексей Александрович,
кандидат технических наук, доцент, Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский Томский политехнический университет», доцент Отделения электроэнергетики и электротехники Инженерной школы энергетики.

Защита состоится 11 ноября 2025 года в 13:00 ч на заседании диссертационного совета УрФУ 2.4.10.26 по адресу: 620062, г. Екатеринбург, ул. Мира, д. 19, ауд. И-420 (зал Ученого совета).

С диссертацией можно ознакомиться в библиотеке и на сайте ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б. Н. Ельцина», по адресу:
<https://dissovet2.urfu.ru/mod/data/view.php?d=12&rid=7547>

Автореферат разослан «__» октября 2025 г.

Ученый секретарь
диссертационного совета

Самойленко Владислав Олегович



ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования.

Развитие возобновляемых источников энергии (ВИЭ) во всем мире обосновывается соображениями экологии, энергетической безопасности, энергетической эффективности, всеобщей доступности, а также интенсификацией функционирования розничных и оптовых рынков электроэнергии с вовлечением большого количества субъектов. Одним из наиболее быстро растущих секторов возобновляемой энергетики является использование ветроэнергетических ресурсов. Всего в мире на конец 2024 года установленная мощность ветроэлектрических станций (ВЭС) составила 1 136 ГВт. За 2024 год прирост составил рекордные 114,9 ГВт. Практически половина установленной мощности ВЭС приходится на Китай, около 521 ГВт.

Установленная мощность ВЭС в России достигла 2 556 МВт на конец 2024 года, чуть более 1% от установленной мощности всех электростанций Единой энергетической системы (ЕЭС) России. За последние 5 лет было введено более 2 ГВт установленной мощности ВЭС в рамках программы поддержки развития возобновляемой энергетики с помощью договоров о предоставлении мощности (ДПМ) ВИЭ. По разным оценкам, до 2035 года может быть введено до 5 ГВт ВЭС для достижения целевого показателя объема производства электрической энергии с использованием возобновляемых источников энергии (кроме гидроэлектростанций установленной мощностью более 25 МВт) в 4,5 %, закрепленного в распоряжении Правительства Российской Федерации от 08.01.2009 №1-р.

В объединенной энергосистеме Юга доля установленной мощности ВЭС уже достигла практически 8 %, а в отдельных регионах, например Республике Калмыкия, около 45 %. При такой значительной доле стохастической генерации традиционный подход к планированию балансов мощности, в том числе при краткосрочном планировании режимов работы электроэнергетических энергосистем (ЭЭС), должен дополняться прогнозированием выработки ВЭС. Чем больше доля ВЭС, тем большее влияние будет оказывать точность прогноза на надежную и экономически эффективную работу ЭЭС.

Под краткосрочным прогнозированием генерации ВЭС понимается определение средней выдаваемой мощности на каждый час следующих суток, что эквивалентно выработке электрической энергии за соответствующий час суток. Основной областью применения таких прогнозов являются расчеты планов выработки для участия ВЭС в оптовом рынке электроэнергии «на сутки вперед» (РСВ) в соответствии с регламентами оптового рынка. На сегодняшний день основную часть дохода собственники ВЭС получают на рынке мощности, при этом требуется обеспечивать определенный коэффициент использования установленной мощности за отчетный период. При работе на РСВ ВЭС подают ценопринимающие заявки, поэтому при выполнении оптимизационного расчета для формирования планового диспетчерского графика данный тип генерации оказывается более приоритетным, чем генерация традиционного типа, которая подает ступенчатые ценовые заявки. Отклонение фактической выработки электроэнергии (ЭЭ) от

плановой приводит к возникновению убытков, в том числе упущенной выгоды, при работе ВЭС на оптовом рынке. Отклонения от планового почасового производства могут приводить к убыткам и для покупателей ЭЭ из-за возникновения незапланированного стоимостного небаланса между суммарными требованиями и обязательствами поставщиков и потребителей. Он распределяется между всеми участниками рынка по достаточно сложным правилам.

В случае, если фактическая выработка ВЭС отклоняется выше заявленного на РСВ объема, то за отклонения ВЭС получает оплату по индикаторам балансирующего рынка (БР) вместо цены РСВ. Причем за значительные отклонения более 10% от установленной мощности оплата может быть снижена до 1 руб/МВт·ч, кроме того увеличивается доля стоимостного небаланса, который будет распределен на ВЭС.

Если же фактическая выработка ВЭС отклоняется ниже заявленного на РСВ объема, то для компенсации отклонений ВЭС вынуждена покупать электроэнергию по индикаторам БР. Значительные отклонения также увеличивают долю стоимостного небаланса.

Таким образом, разработка и внедрение методик повышения точности краткосрочного прогнозирования генерации ВЭС, обеспечивает для:

- собственников ВЭС – снижение убытков и упущенной выгоды, связанных с отклонением фактической выработки от планового почасового производства;
- участников оптового рынка, в том числе потребителей – снижение затрат на компенсацию стоимостного небаланса;
- системного оператора – повышение надежности функционирования энергосистем с ВЭС.

Степень научной разработанности темы исследования.

Анализ мирового опыта в области краткосрочного прогнозирования выработки ВЭС показывает, что единого универсального способа для решения данной задачи на текущий момент нет, но наблюдается повсеместный интерес к методам машинного обучения. Используемые методики прогнозирования во многом носят индивидуальный характер, который определяется составом и качеством доступных исходных данных, полученных от метеопровайдеров и собранных с помощью SCADA на самой ВЭС. Значительное внимание в научных публикациях по данной теме уделяется выбору моделей и алгоритмов машинного обучения. В меньшей степени уделяется внимание трансформации и декомпозиции исходных данных, отбору информативных признаков, идентификации выбросов. Ещё меньшее количество исследований посвящено изучению стратегий агрегации мощности отдельных ветроустановок (ВЭУ), а также оптимизации гиперпараметров моделей машинного обучения применительно к задаче краткосрочного прогнозирования выработки ВЭС. Практически отсутствуют исследования, связанные с учетом состава работающих ветроустановок.

Объект исследования представляет собой энергосистемы, к которым присоединены ВЭС, а также временные ряды, используемые для обучения и тестирования моделей прогнозирования: измерений генерации отдельных ВЭУ, групп ВЭУ и всей ВЭС, прогнозных и фактических метео данных.

Предметом исследования является повышение точности краткосрочного прогнозирования выработки ветроэлектрических станций с помощью оптимизации гиперпараметров моделей машинного обучения, объединения турбин в группы, исключения недостоверных наблюдений и учета состава работающих ВЭУ.

Целью работы является совершенствование методик краткосрочного прогнозирования выработки ветроэлектрических станций для снижения погрешности прогноза.

Для достижения данной цели в работе решаются следующие **задачи**:

1. Анализ существующих подходов к оценке и повышению эффективности краткосрочного прогнозирования генерации ВЭС.
2. Сравнение существующих и разработка новых методик оптимизации гиперпараметров моделей машинного обучения применительно к задаче краткосрочного прогнозирования выработки ВЭС.
3. Сравнение стратегий агрегации мощности отдельных ветроустановок и разработка нового способа объединения ветроустановок в группы.
4. Повышение эффективности методов предварительной обработки данных (исключение недостоверных наблюдений).
5. Повышение эффективности учета состава работающих ВЭУ.

Научная новизна работы состоит в следующем:

1. Предложено два новых критерия для оценки эффективности краткосрочного прогнозирования генерации ВЭС: нормализованная суммарная абсолютная ошибка, показывающая суммарную ошибку прогноза в выработке электрической энергии, и вероятность охвата интервала ошибки в заданном диапазоне, которая может использоваться для оценки надежности получаемых прогнозов.
2. Предложены новые методики подбора гиперпараметров моделей машинного обучения, основанные на использовании байесовской оптимизации с функциями выбора «вероятность улучшения» и «нижний доверительный интервал», обеспечивающие снижение погрешности краткосрочного прогнозирования генерации ВЭС.
3. Разработан альтернативный способ объединения турбин в группы с помощью метода иерархической агломеративной кластеризации по расстоянию, определенному с помощью алгоритма динамической трансформации временной шкалы. Предложенный способ позволил снизить погрешность прогнозирования выработки ВЭС на уровне групп турбин.
4. Предложена новая более эффективная с точки зрения снижения погрешности прогноза генерации ВЭС методика идентификации выбросов по характеристикам мощности с использованием алгоритма глобально-локальной оценки выбросов.
5. Разработана новая методика учета состава работающих ВЭУ, основанная на трансформации отклика регрессионных моделей, обеспечивающая снижение погрешности прогнозирования генерации на уровне ВЭС и групп турбин.

Теоретическая значимость работы состоит в развитии методик повышения эффективности краткосрочного прогнозирования выработки ВЭС в области оптимизации гиперпараметров моделей машинного обучения, определения

способа агрегации мощности отдельных ВЭУ, идентификации выбросов по характеристике мощности и учета состава работающих ВЭУ.

Практическая значимость работы подтверждается положительными результатами оценки эффективности предложенных методик в задаче краткосрочного прогнозирования генерации ВЭС при использовании реальных данных, полученных с находящейся в промышленной эксплуатации ВЭС. Исследования по данной теме выполнялась в рамках совместной научно-исследовательской работы УрФУ и компании «Прософт-Системы» (договор № 21-53810 и договор 22-59892).

Практическая ценность исследования также подтверждается актом внедрения. Результаты работы были использованы в ООО "Прософт-Системы" при разработке модуля краткосрочного прогнозирования выработки ветроэлектрических станций для программного комплекса «Энергосфера».

Положения, выносимые на защиту.

1. Методики подбора гиперпараметров моделей машинного обучения для краткосрочного прогнозирования выработки ВЭС, основанные на использовании байесовской оптимизации с функциями выбора «вероятность улучшения» и «нижний доверительный интервал».

2. Стратегия прогнозирования суммарной мощности ВЭС по группам ВЭУ, при этом объединение в группы выполняется методом иерархической агломеративной кластеризации по расстоянию, определенному с помощью алгоритма динамической трансформации временной шкалы.

3. Методика идентификации выбросов по характеристикам мощности ВЭС, групп турбин и отдельных турбин с использованием алгоритма глобально-локальной оценки выбросов по иерархической основанной на плотности пространственной кластеризации для приложений с шумами.

4. Методика учета состава работающих ветроустановок с помощью трансформации отклика регрессионных моделей.

Методология и методы исследования.

При выполнении настоящей работы были применены методы математического моделирования, математического анализа, математической статистики и машинного обучения. Для обработки данных, проведения вычислительных экспериментов (ВЭ) и визуализации результатов используется программное обеспечение, разработанное автором на языках *MATLAB* с использованием пакетов *MATLAB*, *Statistics and Machine Learning Toolbox*, *Signal Processing Toolbox*, и *Python* с использованием библиотек *matplotlib*, *kneed*, *hdbscan*.

Соответствие диссертации паспорту научной специальности.

Диссертационная работа соответствует специальности 2.4.3. Электроэнергетика, так как область исследования соответствует пункту:

«20. Разработка методов использования информационных и телекоммуникационных технологий и систем, искусственного интеллекта в электроэнергетике, включая проблемы разработки и применения информационно-измерительных, геоинформационных и управляющих систем для оперативного и ретроспективного мониторинга, анализа, прогнозирования и

управления электропотреблением, режимами, надежностью, уровнем потерь энергии и качеством электроэнергии», паспорта специальности.

Обоснованность и достоверность полученных в работе результатов обеспечена корректной реализацией и применением методов машинного обучения, а также методологией проведения вычислительных экспериментов.

Апробация диссертационной работы.

Основные положения диссертационной работы докладывались и обсуждались на конференциях:

1. II Международная научно-техническая конференция «*Smart Energy Systems 2021*» (*SES-2021*), Казань – 2021;

2. Международная научно-практическая конференция студентов, аспирантов и молодых ученых – Даниловские чтения, Екатеринбург – 2022, 2023;

3. Международная научно-техническая конференция *Belarusian-Ural-Siberian Smart Energy Conference*, Екатеринбург – 2023;

4. VI Международная научно-техническая конференция «Релейная защита и автоматика», Москва – 2023;

5. *22nd International Conference on Renewable Energy and Power Quality (ICREPQ'24)*, Бильбао, Испания – 2024.

Публикации по теме исследования.

Основное содержание диссертационной работы отражено в 8 научных трудах, из них 5 статей – в рецензируемых научных журналах, определенных ВАК РФ и Аттестационным советом УрФУ, включая 2 статьи в изданиях, входящих в международную базу цитирования *Scopus*; получено 2 свидетельства о регистрации программы для ЭВМ.

Личный вклад автора заключается в разработке новых методик повышения эффективности краткосрочного прогнозирования выработки ветроэлектрических станций, а также разработке программного обеспечения, с помощью которого были проведены вычислительные эксперименты и подтверждена эффективность предложенных методик.

Объем и структура диссертации

Диссертация состоит из введения, четырёх глав, заключения, списка литературы и приложений. Объём работы включает в себя 227 страниц печатного текста, 21 таблицу, 48 рисунков и 3 приложения.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы, сформулированы цели, задачи и методы исследования, а также определены научная новизна работы, ее теоретическая и практическая значимость.

В первой главе приведены данные о динамике развития ВЭС в России и в мире. Показано, что прогнозирование выработки ВЭС является одной из наименее затратных мер, позволяющих повысить эффективность планирования режимов в энергосистемах со значительной долей установленной мощности ВЭС. Рассмотрены основные особенности краткосрочного прогнозирования выработки ВЭС, в том числе постановка задачи, этапы и основные источники погрешности

прогнозов. Проанализированы критерии для оценки эффективности прогнозов. Выполнен обзор основных подходов к повышению эффективности прогнозирования генерации ВЭС.

Ввиду того, что регламенты оптового рынка ЭЭ предусматривают работу с детерминированными данными, в настоящей работе задача прогнозирования генерации ВЭС осуществляется в детерминистической постановке:

$$\hat{P}(t + k|t) = f(x(t + k|t), \beta), \quad (1)$$

где $\hat{P}(t + k|t)$ – прогнозная мощность на момент времени $t + k$, рассчитанная в момент времени t ; $x(t + k|t)$ – вектор прогнозных значений признаков (объясняющих переменных) на момент времени $t + k$, определенных в момент времени t ; f – модель прогнозирования, β – параметры модели, k – горизонт прогнозирования (интервал упреждения), в данной работе рассматриваются краткосрочные прогнозы на сутки вперед.

Основные этапы процесса прогнозирования представлены на рисунке 1. Эти этапы включают в себя создание и применение модели.

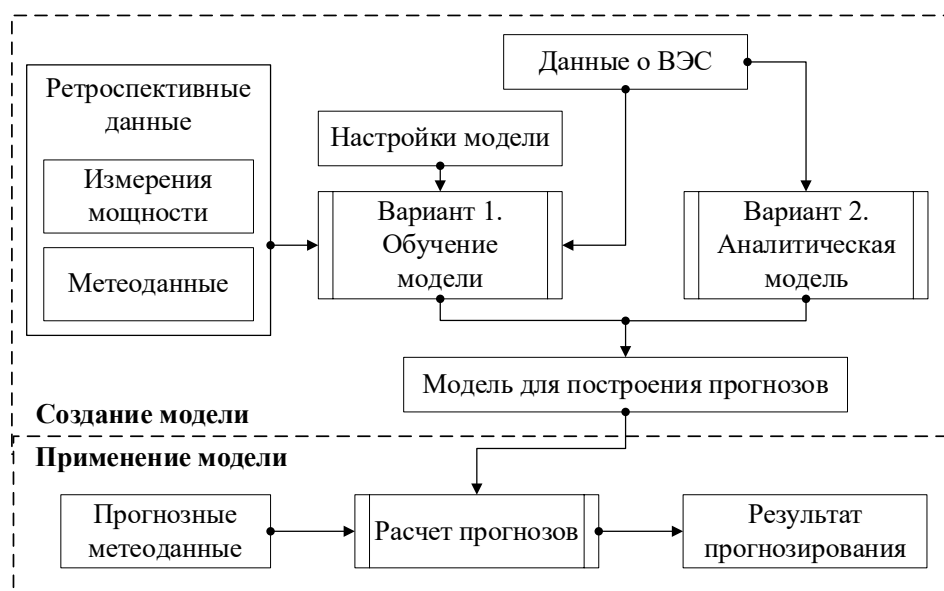


Рисунок 1 – Основные этапы процесса прогнозирования генерации ВЭС

На этапе создания модели главной задачей является определить модель прогнозирования f . Существуют две основных категории: модели, требующие ретроспективной информации для обучения (статистические модели, модели машинного обучения, гибридные модели), и аналитические модели, для которых требуются лишь данные о ВЭС: оборудовании, расположении, высоте турбин, топографических данных и т.д. В ходе обучения модели по ретроспективным данным выполняется определение параметров модели β , позволяющих наилучшим образом с точки зрения функции потерь L описать связь между вектором зависимых переменных P и матрицей независимых переменных X :

$$\beta^* = \underset{\beta}{\operatorname{argmin}} L(P, f_{\beta}(X)), \quad (2)$$

где β^* – искомые значения параметров модели f ; L – функция потерь, обычно квадратичная; P – ретроспективные наблюдения мощности, представляющие собой вектор откликов; X – ретроспективные наблюдения метеоданных,

полученные, как правило, с помощью численного прогноза погоды (ЧПП), представляющие собой матрицу признаков.

На этапе применения модели определяется значение $\hat{P}(t + k|t)$ с помощью прогнозных метеоданных $x(t + k|t)$, также полученных с помощью ЧПП.

Аналитическая модель представляет собой кубическую зависимость мощности, выдаваемой ВЭУ, от скорости ветра:

$$P = \frac{1}{2} \cdot A \cdot \rho \cdot C_p \cdot V^3, \quad (3)$$

где P – мощность турбины; A – ометаемая площадь; ρ – плотность воздуха; C_p – коэффициент аэродинамической эффективности ВЭУ; V – скорость ветра.

На первый взгляд может показаться, что рассматриваемую задачу можно свести к прогнозированию скорости ветра и пересчета полученного значения в мощность по уравнению (3). Но фактически выдаваемая мощность может значительно отличаться от теоретических значений, как показано на рисунке 2. Лишь малое количество точек оказывается вблизи аналитической модели.

Среди основных источников погрешности прогноза выделяют:

1. Погрешность численного прогноза погоды.
2. Характерные для конкретного расположения ВЭС факторы, резко изменяющие скорость или направление воздушного потока, попадающего на лопасти турбин: взаимное влияние ВЭУ друг на друга, наличие препятствий, наклон поверхности и другие особенности ландшафта.
3. Обледенение лопастей ВЭУ и другие экстремальные погодные условия.
4. Погрешности в работе систем управления ВЭУ: системы ориентации, системы отслеживания точки максимальной мощности и т.д.
5. Низкая информационная обеспеченность.



Рисунок 2 – Сравнение аналитической модели и реальных наблюдений

Для оценки эффективности получаемых прогнозов используется ряд общепринятых для данной задачи критериев. Нормализованный квадратный

корень из средней квадратичной ошибки ($NRMSE$) позволяет оценить наличие грубых ошибок прогноза, является квадратичной мерой. Чем больше значение ошибки, тем больше окажется вклад квадрата этой ошибки в величину показателя. Нормализованная средняя абсолютная ошибка ($NMAE$) используется для оценки общей погрешности прогноза, является линейной мерой. Все рассматриваемые ошибки вносят пропорциональный своей величине вклад в значение показателя. Нормализованное смещение ($NBias$) позволяет выявить систематическую ошибку, когда в среднем результаты прогноза завышаются или занижаются. Также было предложено два новых критерия. Нормализованная сумма абсолютных ошибок ($NSAE$) показывает суммарную ошибку прогноза в выработке электрической энергии, поскольку в числителе и знаменателе используется сумма мощностей, усредненных на одном и том же интервале времени. Вероятность охвата интервала ошибки в диапазоне ($EICP_{20}$) может использоваться для оценки надежности получаемых прогнозов. Эта величина показывает долю ошибок, находящихся в интервале $\pm 20\%$ от установленной мощности ВЭС. В качестве основного критерия для сравнения эффективности методик прогнозирования генерации ВЭС в данной работе используется фактор улучшения (SS).

Выражения для расчета данных критериев приведены в таблице 1. При описании критериев для оценки точности и надежности прогноза используются следующие условные обозначения: P_{inst} – установленная мощность ВЭС, кВт; N – количество наблюдений в тестовой выборке, использованных для определения значений критерия; P_i – фактическая измеренная мощность для i -го наблюдения, кВт; $\hat{P}_i$ – прогнозная мощность для i -го наблюдения, кВт; e_i – ошибка прогноза для i -го наблюдения, кВт, определяемая как $e_i = P_i - \hat{P}_i$; $NRMSE_{eval}$ – показатель $NRMSE$ для рассматриваемой методики, %; $NRMSE_{ref}$ – показатель $NRMSE$ для референсной методики, %. Референсная методика определяется для каждой серии вычислительных экспериментов.

Таблица 1 – Критерии для оценки эффективности и надежности прогнозов

$NRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{e_i}{P_{inst}} \right)^2} \cdot 100\%$	$NMAE = \frac{1}{N} \sum_{i=1}^N \frac{ e_i }{P_{inst}} \cdot 100\%$
$NBias = \frac{1}{N} \sum_{i=1}^N \frac{e_i}{P_{inst}} \cdot 100\%$	$NSAE = \frac{\sum_{i=1}^N e_i }{\sum_{i=1}^N P_i} \cdot 100\%$
$EICP_{20} = \frac{1}{N} \sum_{i=1}^N c_i \cdot 100\%,$ $c_i = \begin{cases} 1, & \text{if } e_i < 0,2 \cdot P_{inst} \\ 0, & \text{в остальных случаях.} \end{cases}$	$SS = \left(1 - \frac{NRMSE_{eval}}{NRMSE_{ref}} \right) \cdot 100\%$

При проведении исследований в данной работе использовались реальные данные с ВЭС установленной мощностью 98,8 МВт (26 турбин *Vestas V126-3.80*). Набор данных включал в себя временные ряды измерений мощности каждой турбины и скорости ветра на высоте гондолы, полученные из *SCADA* системы.

Также набор данных содержал результаты ЧПП «на сутки вперед», полученные от метеопровайдера, включающие информацию о скорости и направлении ветра, температуре, влажности воздуха и атмосферном давлении. Общее количество наблюдений в рассматриваемом наборе данных составляет 6343. В тестовую выборку отнесено 20% наблюдений (1268 шт.). Оставшиеся наблюдения (5075 шт.) использовались для обучения и валидации моделей.

Во второй главе приведена постановка задачи оптимизации гиперпараметров. Кратко описаны рассматриваемые модели машинного обучения, а также существующие подходы к оптимизации гиперпараметров. Более подробно рассмотрен способ определения гиперпараметров с использованием байесовской оптимизации. Предложены две новые методики подбора гиперпараметров, основанные на использовании байесовской оптимизации с функциями выбора «вероятность улучшения» и «нижний доверительный интервал» применительно к задаче краткосрочного прогнозирования выработки ВЭС. Проведены вычислительные эксперименты, по результатам которых определено, что модель градиентного бустинга над деревьями решений обеспечивает наименьшую погрешность. Показано, что для ряда моделей машинного обучения предложенные методики оптимизации гиперпараметров позволяют снизить время оптимизации и значение целевой функции (ЦФ), а также получить набор гиперпараметров, обеспечивающий снижение ошибки прогноза по сравнению с другими рассмотренными подходами.

При решении практических задач, связанных с прогнозированием, в том числе и генерации ВЭС, широчайшее распространение получили модели машинного обучения. При обучении модели (2) основной задачей является определение параметров модели β^* . При этом часть параметров определяется непосредственно в процессе восстановления регрессии (значение весов в нейронной сети, значения коэффициентов в полиномиальной модели и т.д.), а часть задаются заранее (количество слоев в нейронной сети, порядок полиномиальной модели и т.д.). Такого рода параметры называются гиперпараметрами.

Единого подхода к тому, как определить подходящую для конкретной задачи модель и ее гиперпараметры, пока что нет. Выделяют следующие способы.

Поиск по решетке (вариация параметров) (GS). Суть данного подхода заключается в полном переборе всех возможных наборов гиперпараметров по заданному вручную подмножеству пространства гиперпараметров. Успешность результатов поиска в значительной степени зависит от заданных границ изменения и дискретизации каждого гиперпараметра. Эффективность данного подхода растет с расширением границ и уменьшением дискретизации, при этом соответственно в значительной степени увеличивается время, необходимое на полный перебор.

Случайный поиск (RS). В данном подходе полный перебор заменяется на оценку эффективности некоторой выборки из случайных комбинаций гиперпараметров. При этом частота дискретизации непрерывных величин гиперпараметров не задается, в результате, можно найти точку, которая не была бы найдена при поиске по решетке. Данный метод позволяет добиться большей эффективности при подборе гиперпараметров при меньших временных затратах.

Байесовская оптимизация (БО, БО). Основной идеей данного способа является направленный перебор гиперпараметров для минимизации ожидаемой ошибки прогноза модели регрессии. В результате, обеспечивается снижение количества проверяемых наборов и ускорение процесса выбора гиперпараметров.

Важной составляющей алгоритма байесовской оптимизации является функция выбора (*acquisition function*), определяющая следующий проверяемый набор гиперпараметров λ из пространства Λ . Данная функция должна обеспечивать баланс между исследованием новых областей, в которых стандартное отклонение суррогатной модели σ_M высоко, и уточнением тех наборов гиперпараметров, для которых среднее суррогатной модели μ_M имеет низкие значения.

Среди опубликованных исследований по краткосрочному прогнозированию генерации ВЭС при использовании байесовской оптимизации гиперпараметров применяется только функция «ожидаемого улучшения» (*EI*):

$$EI(\lambda) = \begin{cases} \mu_M(\lambda) - f_B(\lambda^+) \cdot \Phi(Z) + \sigma_M(\lambda) \cdot \phi(Z) & \text{при } \sigma_M(\lambda) > 0 \\ 0 & \text{при } \sigma_M(\lambda) = 0 \end{cases}, \quad (14)$$

где $\mu_M(\lambda)$ – среднее суррогатной модели для набора гиперпараметров λ ; $\sigma_M(\lambda)$ – стандартное отклонение суррогатной модели для набора гиперпараметров λ ; λ^+ – набор гиперпараметров, который на i -й итерации обеспечивает наименьшее значение ЦФ $f_B(\lambda^+)$, Φ и ϕ – функция распределения и плотность вероятности стандартного нормального распределения; параметр Z определяется как:

$$Z = \frac{\mu_M(\lambda) - f_B(\lambda^+)}{\sigma_M(\lambda)}. \quad (15)$$

В представленной работе впервые применительно к задаче оптимизации гиперпараметров моделей машинного обучения для краткосрочного прогнозирования выработки ВЭС предложено использование функций выбора «вероятность улучшения» (*PI*) и «нижний доверительный интервал» (*LCB*).

Функции «вероятность улучшения» и «нижний доверительный интервал» соответствуют выражениям:

$$PI(\lambda) = \Phi\left(\frac{\mu_M(\lambda) - f_B(\lambda^+) - m}{\sigma_M(\lambda)}\right), \quad (16)$$

$$LCB(\lambda) = 2 \cdot \sigma_M(\lambda) - \mu_M(\lambda), \quad (17)$$

где m – параметр, контролирующий исследование новых областей.

Рассматриваемые функции выбора различаются подходом к уточнению уже проверенных наборов гиперпараметров и исследованию новых областей. *EI* – сбалансированный поиск, *PI* – в большей степени уточняет, *LCB* – быстрее исследует. В ряде работ в смежных областях было выявлено, что функции *PI* и *LCB* иногда позволяют обеспечить определение искомого набора гиперпараметров за меньшее количество итераций и с меньшим значением целевой функции.

Для проверки предложенных методик и сравнения способов оптимизации гиперпараметров (СОГ) проведены вычислительные эксперименты с использованием моделей машинного обучения, который часто применяются для прогнозирования генерации ВЭС: регрессионное дерево решений (*RDT*), нейронная сеть прямого распространения (*FFNN*), градиентный бустинг над

деревьями решений (*GBDT*), регрессия на основе гауссовских процессов (*GPR*), случайный лес (*RF*), метод опорных векторов (*SVM*). Также приведены результаты расчета прогнозов с использованием аналитической модели (*AM*). Количество итераций оптимизации гиперпараметров во всех экспериментах равно 60. Результаты для моделей *GBDT* и *AM* и приведены в таблице 1. Фактор улучшения *SS* рассчитывался по отношению к эксперименту, в котором оптимизация гиперпараметров выполнялась с помощью метода поиска по решетке.

Таблица 1 – Оценка эффективности прогнозов на тестовом наборе данных для модели аналитической модели и модели *GBDT*

Модель	COF	SS, %	NRMSE, %	NMAE, %	NBias, %	NSAE, %	EICP 20, %	T, с
AM	-	0,0	33,0	24,1	22,5	67,5	54,3	0,1
GBDT	GS	0,0	14,6	10,3	-0,2	28,8	85,1	17,9
	RS	2,8	14,2	10,2	1,1	28,6	85,0	17,7
	EI	4,3	14,0	9,9	-0,3	27,7	86,2	19,3
	PI	6,8	13,6	9,6	-0,3	26,8	86,2	42,0
	LCB	8,7	13,3	9,4	-0,4	26,3	86,5	62,1

На рисунке 3 представлено сравнение итерационного процесса поиска минимума ЦФ для модели *GBDT*, для которой было получено наименьшее значение *NRMSE*, и различных способов оптимизации гиперпараметров. На рисунке 4 представлено сравнение рассмотренных моделей машинного обучения по основным критериям для оценки точности и надежности прогноза.

При применении предложенных функций выбора «вероятность улучшения» *PI* и «нижний доверительный интервал» *LCB* удалось добиться повышения эффективности прогнозирования для моделей *GBDT*, *RF* и *FFNN* по сравнению с другими рассмотренными способами оптимизации гиперпараметров. Для моделей *RDT*, *GPR* и *SVM* положительного результата добиться не удалось. При этом время оптимизации гиперпараметров для рассмотренных моделей машинного обучения для экспериментов с *PI* и *LCB*, как правило, оказывается выше.

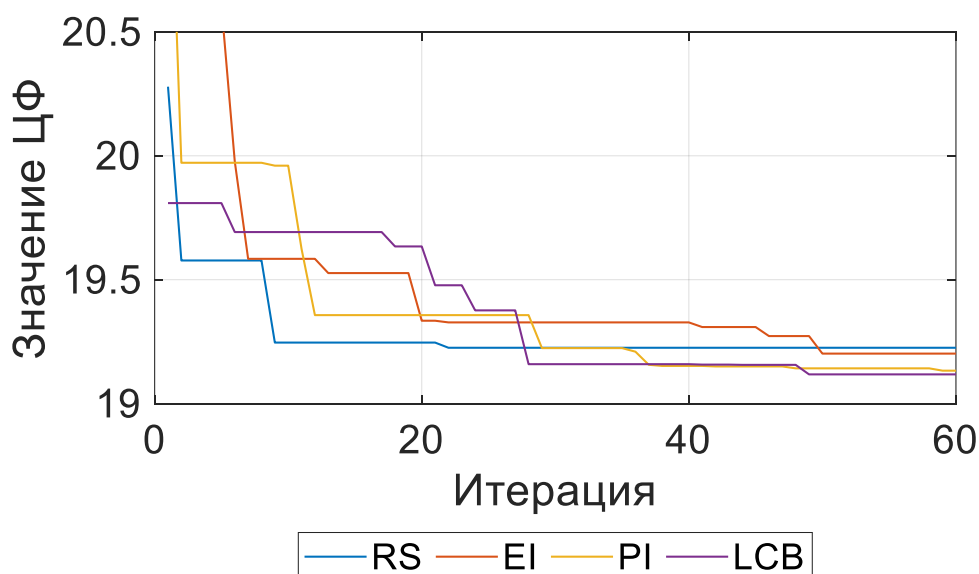


Рисунок 3 – Изменение значений целевой функции для различных способов оптимизации гиперпараметров: *RS* – случайный поиск; *EI* – БО с функцией выбора «ожидаемое улучшение», *PI* – «вероятность улучшения»; *LCB* – «нижний доверительный интервал».

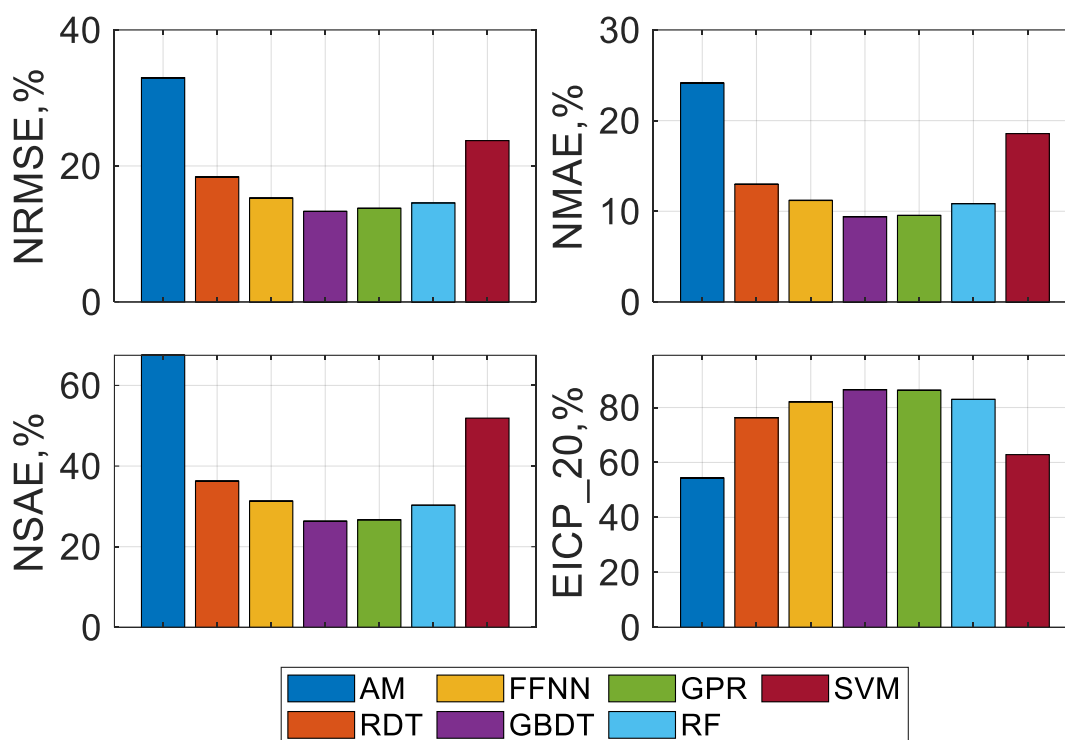


Рисунок 4 – Основные метрики для оценки точности и надежности прогнозов: *AM* – аналитическая модель; *RDT* – регрессионное дерево решений; *FFNN* – нейронная сеть прямого распространения; *GBDT* – градиентный бустинг над деревьями решений; *GPR* – регрессия на основе гауссовских процессов; *RF* – случайный лес; *SVM* – метод опорных векторов.

Среди рассмотренных моделей машинного обучения наибольшую эффективность показала модель *GBDT*. Сопоставимой эффективности удалось добиться при использовании модели *GPR*. При этом обучение модели *GPR* с оптимизацией гиперпараметров занимало в 10-60 раз большее количество времени в зависимости от способа оптимизации и проверяемых наборов гиперпараметров. При проведении последующих вычислительных экспериментов было принято решение использовать модель *GBDT*. Также по результатам проведенных вычислительных экспериментов и оценке точности результатов прогнозов можно сделать вывод, что предложенная методика оптимизации гиперпараметров, основанная на использовании байесовской оптимизации с функцией выбора «нижний доверительный интервал», обеспечивает наибольший фактор улучшения для модели *GBDT*. Эта методика также будет использоваться далее.

В третьей главе приведены основные причины различия в характере выработки мощности ветроустановок внутри одной ВЭС. Рассмотрены существующие стратегии агрегации мощности отдельных ВЭУ при получении суммарного прогноза по ВЭС: прогноз на уровне ВЭС, на уровне групп турбин и на уровне отдельных турбин. Показано, что в существующих исследованиях нет единого мнения о том, какой из подходов обеспечивает наибольшее снижение погрешности прогноза. Описана существующая методика объединения турбин в группы с использованием кластеризации *k*-средних по усредненным данным. Предложены три новых методики разделения турбин на группы только по временным рядам мощности ветроустановок. Проведены вычислительные эксперименты, показавшие, что наибольшей точности прогноза позволяет добиться

стратегия прогнозирования на уровне групп турбин. Наибольшего снижения погрешности удалось добиться с использованием предложенной методики объединения турбин в группы методом иерархической агломеративной кластеризации по расстоянию, определенному с помощью алгоритма динамической трансформации временной шкалы.

Как было отмечено ранее, мощность, выдаваемая отдельными турбинами ВЭС, может значительно отличаться, прежде всего, из-за эффекта «затенения» ветроустановок друг другом и элементами ландшафта. Сопоставление мощности, выдаваемых турбинами за 48 часов, представлено на рисунке 5.

С точки зрения агрегирования мощности отдельных турбин при выполнении прогнозов генерации ВЭС можно выделить три основных подхода:

- прогноз на уровне ВЭС – суммирование мощности всех ВЭУ для получения отклика и прогнозирование суммарной мощности ВЭС;
- прогноз на уровне групп турбин – суммирование мощности ВЭУ по группам, прогнозирование мощности каждой группы и суммирование прогнозов по группам для получения суммарной мощности ВЭС;
- прогноз на уровне турбин – прогнозирование мощности каждой ВЭУ и суммирование их прогнозов для получения мощности всей ВЭС.

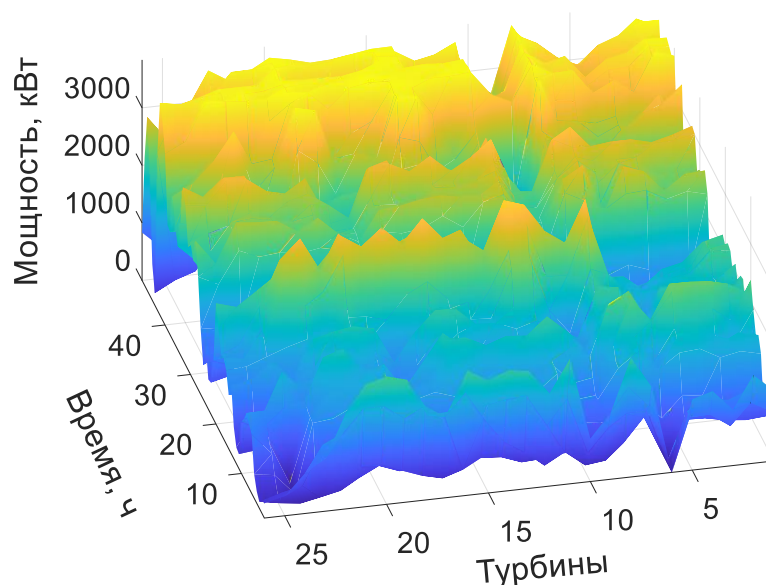


Рисунок 5 – Мощности турбин за 48 часов

Хотя в большинстве исследований для упрощения задачи используется прогноз на уровне ВЭС, в последнее время все чаще появляются работы, посвященные исследованию других способов агрегации мощности ВЭУ. Объединение в группы может выполняться по данным о средних значениях и среднеквадратичных отклонениях (СКО) измеренных мощностей и скоростей ветра каждой турбины с помощью метода кластеризации k -средних (K - MS). Суть данного метода заключается в итеративном повторении двух шагов: распределении объектов (ВЭУ) по кластерам, исходя из расстояния между объектом и центром кластера, и пересчете центров кластеров. Итерационный процесс продолжается до тех пор, пока распределение объектов по кластерам не перестанет меняться, либо при достижении заданного количества итераций.

В данной работе предложено три новых методики распределения турбин по группам. Сравнение существующей и предложенных методик приведено в таблице 2. Отличительной особенностью всех трех предложенных методик является использование в качестве признаков при кластеризации только временных рядов измерений мощности ВЭУ.

Первая методика, как и существующая, предполагает использование метода k -средних, вторая и третья – алгоритма иерархической агломеративной кластеризации (HAC). Данный алгоритм основан на выделении каждого объекта в отдельный кластер, последовательном расчете расстояний между кластерами и объединении ближайших кластеров до достижения их заданного количества. Различие между второй и третьей методикой заключается в способе определения расстояния между объектами: классическое евклидово и DTW -расстояние, основанное на алгоритме динамической трансформации временной шкалы.

Исследования в смежных областях показывают, что применение данного расстояния позволяет получить более корректные результаты кластеризации с точки зрения определения похожести временных рядов в случае, если они несколько смещены друг относительно друга по оси времени. Эффект смещения может быть связан с влиянием резко меняющихся скорости и направления ветра, «затенения» ВЭУ, запаздыванием в работе систем управления и т.д.

Таблица 2 – Методики объединения ВЭУ в группы

Методика	Метод	Признаки	Размерность матрицы признаков*	Расстояние между объектами
$K-MS$ (сущ.)	k -средних	Средние значения и СКО скорости ветра и мощности	$T_n \times 4$	евклидово
$K-P$ (предл.)	k -средних	Измерения мощности	$T_n \times N$	евклидово
$H-PE$ (предл.)	HAC	Измерения мощности	$T_n \times N$	евклидово
$H-PD$ (предл.)	HAC	Измерения мощности	$T_n \times N$	DTW -расстояние

* T_n – количество ВЭУ, N – количество наблюдений.

Для определения наиболее эффективной методики агрегации мощности турбин был проведен ряд вычислительных экспериментов. Было выполнено сравнение стратегии прогнозирования на уровне ВЭС (методика P), на уровне групп турбин с упрощённым способом распределения по группам исходя из принадлежности к фидерам (GF) и на уровне отдельных турбин (T).

Для всех четырёх рассматриваемых методик количество кластеров, на которые необходимо разделить ВЭУ, является неизвестной величиной. Для ее определения были проведены вычислительные эксперименты, результаты которых представлены на рисунке 6. В качестве критерия для определения количества кластеров использовался фактор улучшения SS по отношению к методике P .

Как видно из рисунка 6, наибольшего фактора улучшения удалось добиться при использовании методики $H-PD$ с количеством кластеров, равным 6. Результаты кластеризации для эксперимента с использованием методики $H-PD$ представлены на рисунке 7. Результаты вычислительных экспериментов по определению наиболее эффективной методики агрегации мощности отдельных ветроустановок и количества групп турбин N сведены в таблицу 3. Как видно из таблицы, выполнение прогнозов на уровне групп турбин и на уровне отдельных турбин

позволяет снизить ошибку прогноза. При этом на обучение отдельных регрессионных моделей тратится большее количество времени.

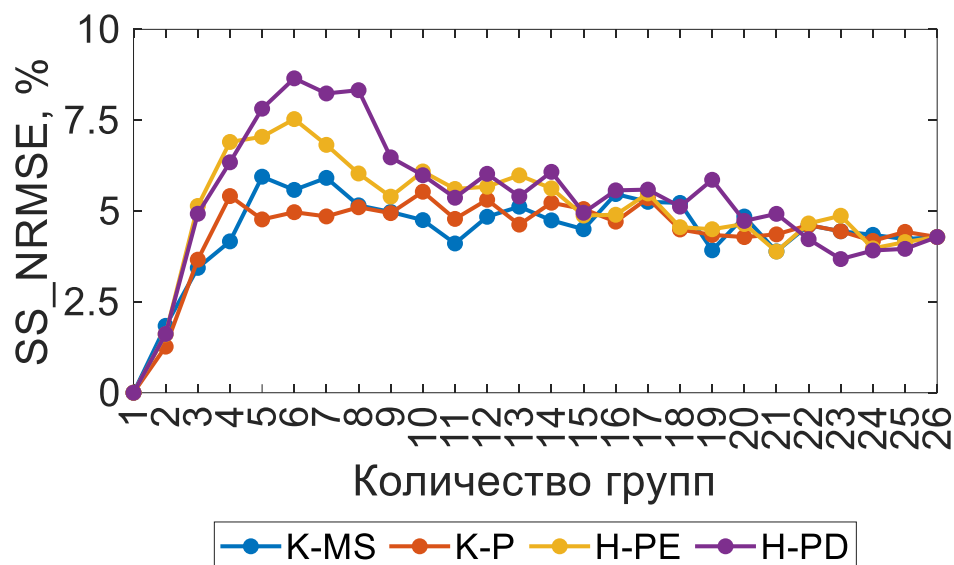


Рисунок 6 – Определение наиболее удачного распределения турбин по группам: *K-MS* – кластеризация *k*-средних, евклидово расстояние, среднее и СКО мощности и скорости ветра; *K-P* – кластеризация *k*-средних, евклидово расстояние, временные ряды мощности турбин, *H-PE* – иерархическая кластеризация, евклидово расстояние, временные ряды мощности, *H-PD* – иерархическая кластеризация, *DTW*-расстояние, временные ряды мощности турбин.

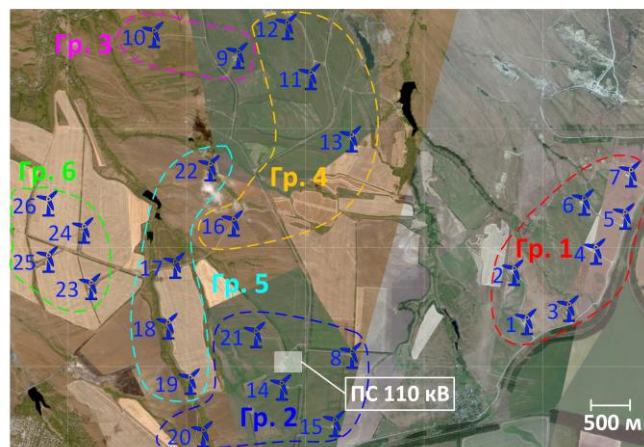
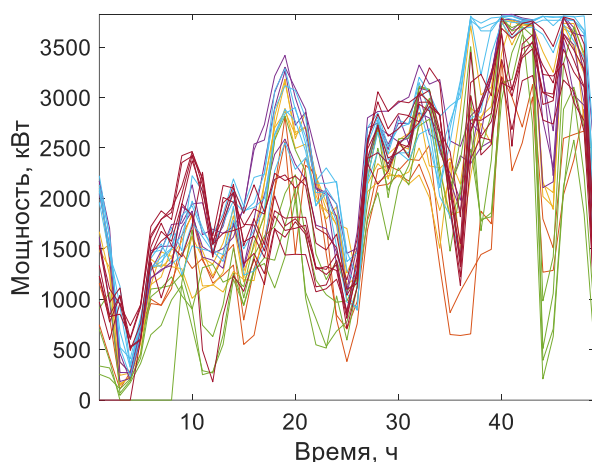


Рисунок 7 – Результаты распределения ВЭУ по группам с помощью методики *H-PD*: временные ряды мощности турбин – слева (одинаковым цветом обозначены временные ряды в одном кластере), группы ВЭУ на местности – справа

Таблица 3 – Оценка эффективности прогнозов на тестовом наборе данных для рассматриваемых подходов к агрегации мощности отдельных установок

ВЭ	<i>N</i>	<i>SS</i> , %	<i>NRMSE</i> , %	<i>NMAE</i> , %	<i>NBias</i> , %	<i>NSAE</i> , %	<i>EICP_20</i> , %	<i>T</i> , с
<i>P</i>	1	0,0	13,31	9,4	-0,4	26,3	86,5	63,6
<i>GF</i>	4	5,2	12,61	9,2	-0,4	25,6	87,5	189,2
<i>T</i>	26	4,3	12,74	9,3	-0,4	26,0	87,9	757,7
<i>K-MS</i>	6	5,6	12,57	9,2	-0,3	25,6	87,8	179,3
<i>K-P</i>	6	5,0	12,65	9,3	-0,4	25,9	88,6	185,9
<i>H-PE</i>	6	7,5	12,31	9,1	-0,4	25,4	89,1	213,2
<i>H-PD</i>	6	8,7	12,16	9,0	-0,5	25,1	89,5	203,9

Прогнозирование как на уровне отдельных турбин, так и на уровне групп турбин, даже при упрощенном разделении турбин на группы по принадлежности к фидерам, позволяет обеспечить снижение как средней, так и среднеквадратичной ошибки прогноза. Было определено, что наибольшей эффективностью с точки зрения снижения среднеквадратичной ошибки прогноза обладает методика *H-PD* при разделении турбин на 6 групп. В дальнейших экспериментах для разделения турбин на группы будет использоваться данная методика с обозначением *GS*.

В четвертой главе приведена существующая классификация типов выбросов на характеристике мощности. Описана существующая методика идентификации выбросов по суммарной характеристике мощности с использованием метода межквартильного размаха. Предложена новая методика, основанная на использовании алгоритма глобально-локальной оценки выбросов по иерархической основанной на плотности пространственной кластеризации для приложений с шумами. Также предложена новая методика для учета состава работающих ВЭУ с помощью трансформации откликов. Проведены вычислительные эксперименты, результаты которых позволяют судить об эффективности предложенных методик. Приведены итоговые результаты исследования, позволяющие оценить частный и суммарный эффект от предложенных в данном исследовании методик повышения точности прогнозирования генерации ВЭС.

При классификации выбросов по характеристике мощности обычно выделяют пять основных типов, определяющих области, в которых могут наиболее часто возникают выбросы. Основные причины выбросов – отказы ВЭУ, сбои в работе измерительных устройств и каналов связи, искусственные ограничения выходной мощности, эффекты «затенения» ВЭУ и усреднения скорости ветра и выходной мощности при порывах ветра, запаздывание в работе системы управления ориентации ВЭУ, экстремальные погодные условия.

Среди существующих подходов к идентификации выбросов с помощью характеристики мощности наиболее распространенным является метод межквартильного размаха (методика *QM*). Данный метод основан на разделении наблюдений мощности, находящихся в диапазоне скорости ветра 0,5 м/с, на подвыборки. Для наблюдений в подвыборке рассчитываются границы Q_L и Q_U . Значения мощности, находящиеся вне данных границ, идентифицируются как выбросы. Границы определяются в соответствии с формулой:

$$[Q_L, Q_U] = [q_1 - 1,5 \cdot Q_R, q_3 + 1,5 \cdot Q_R], \quad (18)$$

где q_1 – первый квартиль; q_3 – третий квартиль; Q_R – межквартильный размах, определяемый выражением $Q_R = q_3 - q_1$.

Также в одном из существующих исследований был применен подход, заключающийся в использовании алгоритма основанной на плотности пространственной кластеризации для приложений с шумами (*DBSCAN*) для идентификации выбросов по суммарной характеристике мощности (методика *DB*).

Упрощенно алгоритм *DBSCAN* можно представить как последовательность следующих этапов: определение для каждой точки количества точек внутри максимального радиуса соседства ε ; если таких точек больше минимального числа

точек, которые должны образовывать плотную область N_{min} , то точка считается основной; основные точки, находящиеся в ε -окрестности друг друга, соединяются ребром (являются напрямую достижимыми); выделяются компоненты связности, образующие кластеры; точки, в окрестности которых есть основные, но количество точек в окрестности меньше N_{min} , являются граничными точками и относятся к кластеру ближайшей основной точки (являются достижимыми); точки, в окрестности которых нет основных точек и в которых содержится менее N_{min} наблюдений, считаются выбросами (являются недостижимыми). В данной работе предложена модификация существующей методики идентификации выбросов для автоматического определения максимального радиуса соседства с помощью алгоритма *KNeedle*.

Также в данной работе предлагается новая методика идентификации выбросов, основанная на алгоритме глобально-локальной оценки выбросов (*GLOSH*) по иерархической основанной на плотности пространственной кластеризации для приложений с шумами (*HDBSCAN**) (методика *GL*). Данный алгоритм является усовершенствованной версией *DBSCAN*. Его отличительной особенностью является возможность поиска кластеров наблюдений с различной плотностью, а также идентификации локальных выбросов вблизи областей с высокой плотностью. Результаты выявления выбросов на суммарной характеристике мощности представлены на рисунке 8.

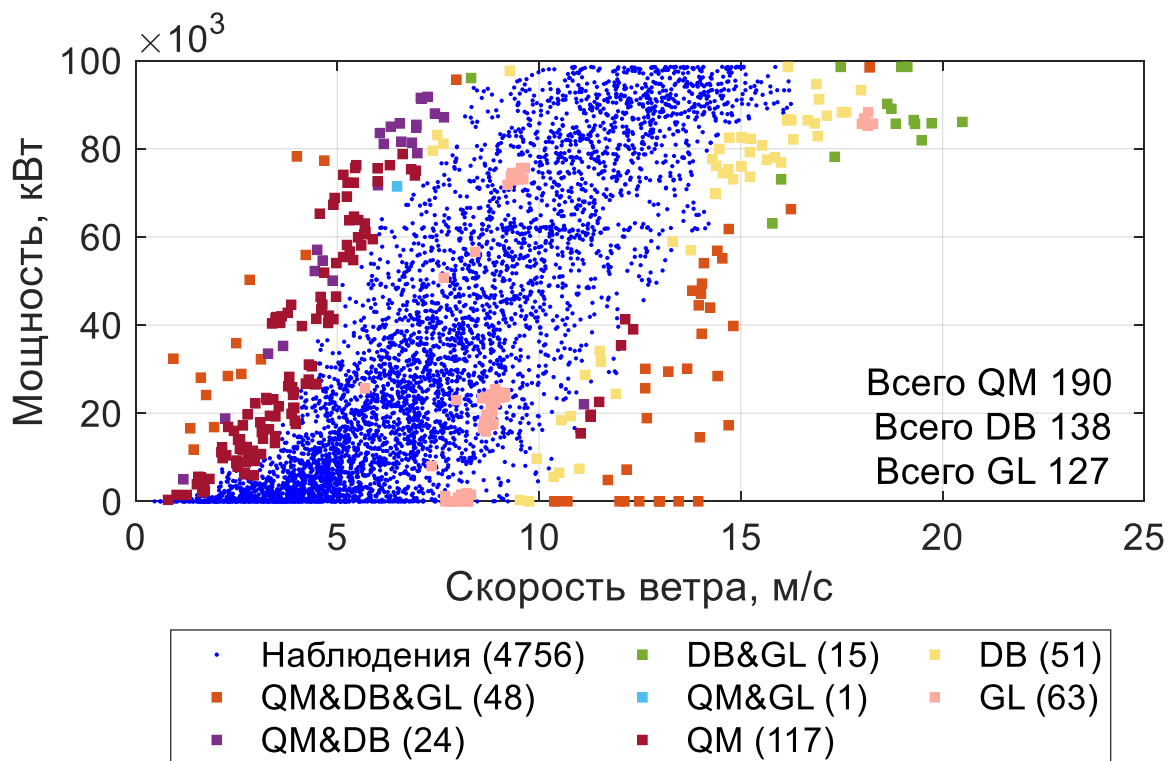


Рисунок 8 – Результаты идентификации выбросов: *QM* – метод межквартильного диапазона; *DB* – кластеризация *DBSCAN*; *GL* – глобально-локальная оценка выбросов *GLOSH*; символ «&» – пересечение множеств.

Результаты вычислительных экспериментов по определению наиболее эффективной методики идентификации выбросов сведены в таблицу 4.

Таблица 4 – Результаты оценки эффективности прогнозов для рассматриваемых подходов к поиску и исключению выбросов по характеристике мощности

ВЭ	SS, %	NRMSE, %	NMAE, %	NBias, %	NSAE, %	EICP_20, %	T, с
<i>P</i>	0,0	13,31	9,4	-0,4	26,3	86,5	63,6
<i>GC</i>	8,7	12,16	9,0	-0,5	25,1	89,5	203,9
<i>GC-QM</i>	8,7	12,15	8,9	-0,1	24,9	90,3	175,6
<i>GC-DB</i>	13,7	11,49	8,4	-0,3	23,5	92,0	190,4
<i>GC-GL</i>	15,9	11,20	8,4	-0,5	23,5	92,7	172,0

Таблица 4 показывает, что с помощью идентификации выбросов было обеспечено снижение ошибки прогноза. Кроме того, за счет сокращения количества наблюдений обеспечивается снижение времени обучения регрессионных моделей. Наибольшего повышения эффективности прогноза с точки зрения фактора улучшения удалось добиться при использовании идентификации выбросов на уровне групп турбин с помощью методики *GL*.

В опубликованных исследованиях вопрос учета состояния турбин не освещен в должной степени. Наблюдения, соответствующие отключенному состоянию установок, обычно исключаются на этапе предварительной обработки данных. Тем не менее, существует подход, заключающийся в учете состояния турбин в качестве признаков при обучении моделей (методика *A*).

В данной работе предлагается новая методика учета состава работающих ВЭУ, основанная на трансформации отклика (методика *RT*). При этом выполняется пересчет фактической измеренной мощности турбин в возможную доступную мощность, исходя из количества включенных установок:

$$P_{av} = P \cdot \frac{K}{K_{av}}, \quad (19)$$

где P_{av} – доступная мощность, кВт; P – измеренная мощность, кВт; K – количество турбин в группе или ВЭС; K_{av} – количество включенных турбин.

Трансформация отклика по данному соотношению выполняется для обучающей выборки для групп турбин или ВЭС. Для отдельных турбин наблюдения с $K_{av} = 0$ исключаются. Для корректировки прогнозных значений используется обратное соотношение, получаемая прогнозная доступная мощность снижается для получения прогнозной фактической мощности группы турбин, отдельных турбин или ВЭС.

Результаты вычислительных экспериментов по определению наиболее эффективной методики учета состояния ветроустановок представлены в таблице 5. Таблица 5 – Результаты оценки эффективности прогнозов для рассматриваемых подходов к учету состава работающих ветроустановок

ВЭ	SS, %	NRMSE, %	NMAE, %	NBias, %	NSAE, %	EICP_20, %	T, с
<i>P</i>	0,0	13,31	9,4	-0,4	26,3	86,5	63,6
<i>GC</i>	8,7	12,16	9,0	-0,5	25,1	89,5	203,9
<i>GC-A</i>	9,6	12,03	8,6	-0,5	24,0	89,1	270,1
<i>GC-RT</i>	13,5	11,51	8,5	-0,2	23,9	90,5	186,0

Таблица 5 показывает, что учет состава работающих ветроустановок как в виде признака, так и с помощью трансформации отклика, позволяет снизить

погрешность прогноза. При этом меньшая погрешность прогноза была получена при использовании трансформации отклика.

Итоговые результаты исследований представлены на рисунке 9, а также сведены в таблицу 6. В дополнительном эксперименте комбинировались исключение выбросов по методике *GL* и учет состояния включенных установок с помощью трансформации отклика *RT*.

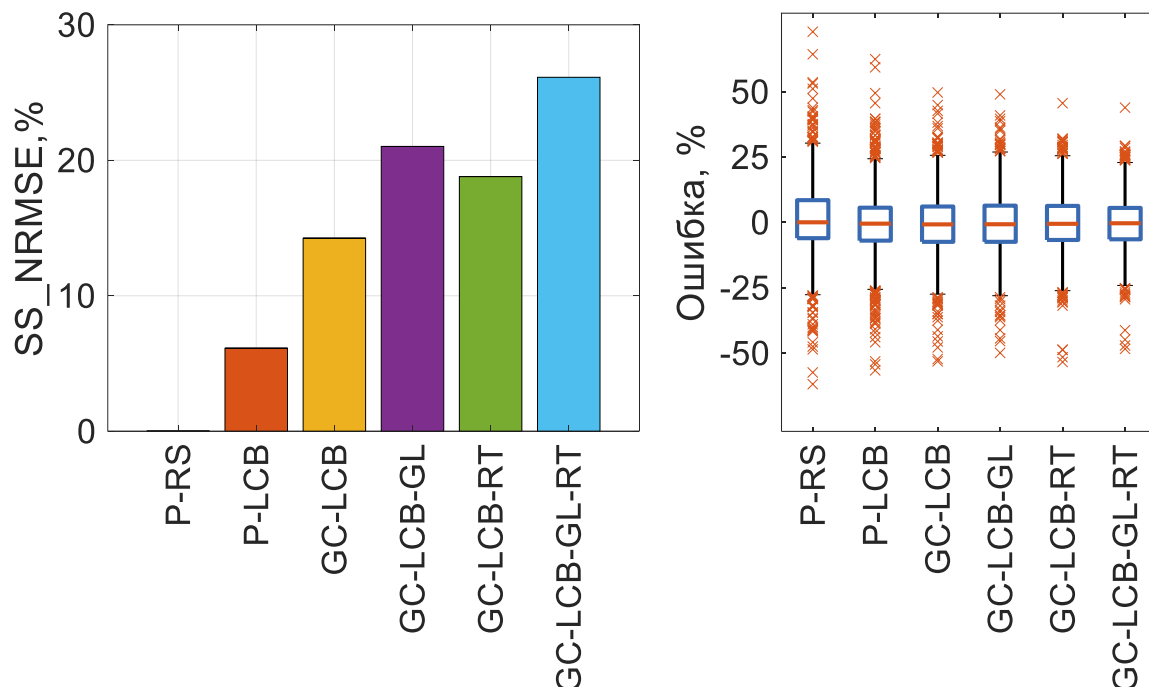


Рисунок 9 – Сравнение предложенных методик повышения эффективности прогнозирования генерации ВЭС (факторы улучшения – слева, диаграмма размаха – справа):

P – прогноз на уровне ВЭС; *GC* – прогноз на уровне групп турбин, определенных иерархической кластеризацией по DTW-расстоянию; *RS* – оптимизация гиперпараметров методом случайного поиска; *LCB* – байесовская оптимизация гиперпараметров с функцией выбора «нижний доверительный интервал»; *GL* – глобально-локальная оценка выбросов; *RT* – учет состава работающих установок с помощью трансформации отклика.

Таблица 6 – Результаты исследования по повышению эффективности краткосрочного прогнозирования выработки ВЭС

ВЭ	SS, %	NRMSE, %	NMAE, %	NBias, %	NSAE, %	EICP_20, %	T, с
<i>P-RS</i>	0,0	14,18	10,2	1,1	28,6	85,0	13,5
<i>P-LCB</i>	6,1	13,31	9,4	-0,4	26,3	86,5	63,6
<i>GC-LCB</i>	14,2	12,16	9,0	-0,5	25,1	89,5	203,9
<i>GC-LCB-GL</i>	21,0	11,20	8,4	-0,5	23,5	92,7	222,0
<i>GC-LCB-RT</i>	18,8	11,51	8,5	-0,2	23,9	90,5	186,0
<i>GC-LCB-GL-RT</i>	26,1	10,47	7,8	-0,2	21,9	93,1	254,6

В заключении приведены основные выводы и обобщены полученные результаты.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИОННОЙ РАБОТЫ

1. По результатам обзора литературы были определены перспективные направления повышения эффективности краткосрочного прогнозирования генерации ВЭС: оптимизация гиперпараметров моделей машинного обучения, изменение стратегии агрегации мощности отдельных ВЭУ, идентификация выбросов, учет состава работающих ВЭУ.

2. Предложено два новых критерия для оценки эффективности краткосрочного прогнозирования выработки ВЭС: нормализованная суммарная абсолютная ошибка, позволяющая оценить погрешность прогноза с точки зрения вырабатываемой электроэнергии, и вероятность охвата интервала ошибки в заданном диапазоне, обеспечивающая возможность оценить надежность прогнозов.

3. С использованием разработанного программного обеспечения выполнено сравнение эффективности различных способов оптимизации гиперпараметров моделей машинного обучения. Предложена новая методика, основанная на использовании байесовской оптимизации с функцией выбора «нижний доверительный интервал», которая позволила снизить среднеквадратичную ошибку прогноза на 6,1 % по отношению к подходу, основанному на подборе гиперпараметров методом случайного поиска.

4. Проведена оценка эффективности основных стратегий агрегирования мощности турбин при прогнозировании суммарной мощности ВЭС: прогноз на уровне всей станции, на уровне групп турбин и на уровне отдельных ВЭУ. Разработан новый способ объединения турбин в группы с помощью метода иерархической агломеративной кластеризации по *DTW*-расстоянию, который позволил снизить среднеквадратичную ошибку прогноза на 8,7 % по отношению к прогнозу на уровне всей ВЭС.

5. Предложена новая методика идентификации выбросов по характеристике мощности с использованием алгоритма глобально-локальной оценки выбросов по иерархической кластеризации, которая позволила снизить среднеквадратичную ошибку прогноза на 15,9 % по отношению к прогнозу по ВЭС без исключения выбросов.

6. Разработана новая методика учета состава работающих ветроустановок с помощью трансформации отклика регрессионных моделей, которая позволила снизить среднеквадратичную ошибку прогноза на 13,5 % по отношению к прогнозу без учета состояния турбин.

7. Совместное использование предложенных методов позволило снизить среднеквадратичную ошибку прогноза генерации ВЭС на 26,1 %.

8. Практическая ценность исследования подтверждается актом внедрения. Результаты работы были использованы в ООО «Прософт-Системы» при разработке модуля краткосрочного прогнозирования выработки ветроэлектрических станций для программного комплекса «Энергосфера».

Публикации по теме диссертации

Статьи, опубликованные в рецензируемых научных журналах и изданиях, определенных ВАК РФ и Аттестационным советом УрФУ:

1. Снегирев, Д. А. Оптимизация краткосрочного прогнозирования выработки ветроэлектрических станций с помощью машинного обучения / **Д. А. Снегирев**, А. В. Паздерин, В. О. Самойленко, П. А. Крючков // *Электрические станции*. – 2025. – № 6(1127). – С. 40-49; 0,82 п.л./0,66 п.л.; K1.

2. Снегирев, Д. А. Оценка влияния способов агрегации ветроэнергетических установок на эффективность краткосрочного прогнозирования выработки электроэнергии / **Д. А. Снегирев**, А. В. Паздерин, В. О. Самойленко // *Электричество*. – 2025. – № 5. – С. 62-73; 1,41 п.л./1,13 п.л.; K1.

3. Снегирев, Д. А. Идентификация выбросов на характеристиках мощности для повышения эффективности краткосрочного прогнозирования генерации ветроэлектростанций / **Д. А. Снегирев**, А. В. Паздерин, В. О. Самойленко, П. И. Бартоломей // *Электротехнические системы и комплексы*. – 2025. – № 2(67). – С. 25-34; 1,19 п.л./0,95 п.л.; K2.

4. Snegirev D. A. The Selection of Machine Learning Model and Its Hyperparameters Using Bayesian Optimization for Short-Term Wind Power Forecasting / **D. A. Snegirev**, V. O. Samoylenko, A. V. Pazderin, P. I. Bartolomey // 2023 Belarusian-Ural-Siberian Smart Energy Conference (BUSSEC). – Yekateringburg, Russia. – IEEE, 2023. – P. 18-23; 0,71 п.л./0,57 п.л. (*Scopus*).

5. Snegirev D. A. Short-Term Wind Power Forecasting Based on Gaussian Process Regression / **D. A. Snegirev**, V. O. Samoylenko, A. V. Pazderin, A. S. Berdin // 2023 6th International Scientific and Technical Conference on Relay Protection and Automation (RPA). – Moscow, Russia. – IEEE, 2023. – Vol. 1. – P. 1-13; 0,67 п.л./0,53 п.л. (*Scopus*).

Патенты и программы:

6. WindPowerForecast: свидетельство об официальной регистрации программы для ЭВМ № 2022615338 Российская Федерация / **Снегирев Д.А.**, Паздерин А.В., Самойленко В.О.; заявитель и правообладатель Снегирев Д.А. – заявл. 29.03.2022; опубл. 15.04.2022.

7. WindPowerForecastML: свидетельство об официальной регистрации программы для ЭВМ № 2023614895 Российская Федерация / **Снегирев Д.А.**, Паздерин А.В., Самойленко В.О.; заявитель и правообладатель Снегирев Д.А. – заявл. 22.02.2023; опубл. 07.03.2023.

Другие публикации:

8. Snegirev D. Short-term wind power forecasting with responses transformation for taking into account wind turbine operational state / **D. Snegirev**, A. Pazderin, V. Samoylenko, P. Bartolomey // 22th International Conference on Renewable Energies and Power Quality (ICREPQ'24). – Bilbao, Spain. – 2024. – P. 1-6; 0,65 п.л./0,52 п.л.