УДК 621.321

Томин Никита Викторович,
к.т.н., старший научный сотрудник, Институт систем энергетики им. Л.А. Мелентьева СО РАН,
e-mail: tomin.nv@gmail.com
Куприков Александр Александрович,
магистрант, Иркутский научно-исследовательский технический университет,
sanyakuprik54@gmail.com

КРАТКОСРОЧНОЕ ПРОГНОЗИРОВАНИЕ ВЫРАБОТКИ МОЩНОСТИ ВЕТРОЭНЕРГЕТИЧЕСКИХ УСТАНОВОК НА БАЗЕ АНСАМБЛЕВЫХ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ[1]

Tomin N.V., Kuprikov A.A.

SHORT-TERM WIND POWER FORECASTING BASED ON ENSEMBLE MACHINE LEARNING METHODS

Аннотация. Рассмотрены регрессионные модели для краткосрочного прогнозирования выходных мощностей ветроэнергетических установок, работающих в составе ветропарков, с применением алгоритмов машинного обучения.

Ключевые слова: регрессивные модели, ветростанции, прогнозирование

Annotation. Regression models for short-term forecasting of output capacities of wind power plants operating as part of wind farms using machine learning algorithms are considered.

Keywords: regression models, wind farm, modeling.

Проблема энергосбережения сегодня стоит достаточно остро, так как промышленность и технологии постоянно развиваются, а это ведет к существенному увеличению расходов электроэнергии, увеличению затрат предприятия и росту стоимости продукции, выпускаемой предприятием. В условиях растущей конкуренции предприятия вынуждены искать пути снижения энергопотребления и увеличивать энергоэффективность отдельных производств.

О важности вопросов энергосбережения на государственном уровне свидетельствует законодательная база. В 2009 году выпущен закон «Об энергосбережении и повышении энергетической эффективности и о внесении изменений в отдельные законодательные акты Российской Федерации» [1], в соответствие, со всеми изменениями, а именно последними от 14.07.2022 № 253-ФЗ, в соответствии с которым все предприятия, а также мелкие потребители, должны постоянно снижать потребление электроэнергии. Этот шаг в масштабах государства способен улучшить ситуацию в экономической сфере. Для создания соответствующего механизма 27 декабря 2010 г. Распоряжением Правительства Российской Федерации № 2446-р принята к действию программа «Энергосбережение и повышение энергетической эффективности на период до 2020

[1] Данная работа была выполнена при поддержке проекта государственного задания фундаментальных исследований РАН «Теоретические основы, модели и методы управления развитием и функционированием интеллектуальных электроэнергетических систем», FWEU-2021-0001 (рег. №: АААА-А21-121012190027-4).

года» [2]. В соответствии с данной программой все крупные предприятия обязаны разрабатывать и внедрять мероприятия по уменьшению потребления энергопотребления [3].

Ветер и полученная энергия, производимая ветроэнергетической установкой (далее ВЭУ), не стабильны и их невозможно запланировать. Существует ряд трудностей, с одной из которых сталкиваются диспетчеры электроэнергетических систем (ЭЭС) во всех странах, непредсказуемость и резкая изменчивость выходной мощности ветровой электростанции (далее ВЭС).

Таким образом, проблема исследования заключается в определении оптимальных моделей краткосрочного прогнозирования и моделирования выходной мощности ВЭУ с применением алгоритмов машинного обучения.

Целью статьи является рассмотрение регрессионных моделей для краткосрочного прогнозирования выходных мощностей ветроэнергетических установок, работающих в составе ветропарков, с применением алгоритмов машинного обучения.

ВЭУ или ветрогенератор – устройство для преобразования кинетической энергии воздушного потока в электрическую энергию [3]. Это альтернативный источник энергии. Его выработка, а также выходные характеристики тока связаны кубической зависимостью со скоростью ветра. Из-за этого и в виду непостоянства ветра самостоятельное его применение невозможно. Требуются также аккумуляторы для накопления электроэнергии, а также оборудование для зарядки батарей. В среднем ВЭУ вырабатывает 150 кВт·ч/мес. электроэнергии.

Краткосрочное прогнозирование выходной мощности ветроэнергетических установок осуществляется с целью расчета режима энергосистемы, планирования поставок электроэнергии и оптимизации режимов работы традиционных электростанций. Интервал прогнозирования при этом составляет от 1 часа до 3 суток [4].

Качественное прогнозирование ветровой мощности необходимо как для успешной интеграции ветроэнергетики в энергосистеме, так и для работы рынков электроэнергии. Многие рынки электроэнергии во всем мире перешли от регулируемого варианта, реализуемого по вертикально-интегрированным монопольным моделям производства, передачи и распределения электроэнергии, к дерегулируемым моделям с дезагрегацией производства и распределения электроэнергии. Учет внеплановой мощности (к примеру, мощности от ВЭС) на рынках электроэнергии производится различными способами. В некоторых случаях государственное регулирование цен на поставку устанавливается для переменных по генерации возобновляемых источников энергии. На других же рынках все участники должны соревноваться в равных условиях. К примеру, системный оператор рынка электроэнергии в Австралии NEMMCO, для обеспе-

чения режимной надежности ЭЭС и распределения генерируемой мощности при удовлетворении спроса рассматривает ветровую энергию как «негативный спрос».

От предполагаемого применения прогнозирование выработки мощности ВЭС выполняется в разных временных интервалах (табл. 1).

При диапазонах от миллисекунд до нескольких минут могут быть использованы прогнозы для активного управления турбиной и для управления режимом работы всей системы в реальном времени. Такой тип прогнозов обычно называют сверхкраткосрочным. Краткосрочные прогнозы (от 1 ч до 72 ч) необходимы при расчете режима и планировании поставок электроэнергии, управлении режимом энергосистемы, а также при торговле электроэнергией на рынке. Помимо этого, такие прогнозы крайне важны для принятия решения об использовании традиционных электростанций и оптимизации режимов работы этих станций. Среднесрочные прогнозы от 3 до 7 дней необходимы для общего технического обслуживания ВЭС, сетей всей энергосистемы и для планирования хранения энергии.

Таблица 1

Классификация временных интервалов прогнозирования скорости ветра и выработки мощности ВЭС

Временной диапазон	Интервал прогнозирования	Применение
Сверхкраткосрочный	Миллисекунды-несколько минут	• Управление турбиной • Управление электрической сетью в реальном времени
Краткосрочный	1-72 ч.	• Расчет режима энергосистемы • Планирование поставок электроэнергии • Торговля электроэнергией на рынке • Планирование и оптимизация режимов работы традиционных электростанций
Среднесрочный	3-7 дней	• Техническое обслуживание ВЭС • Планирование хранения электроэнергии • Планирование эксплуатации ВЭС совместно с традиционными станциями

Для более длительных временных интервалов (более 5—7 дней вперед) прогнозы предназначены для планирования эксплуатации ветропарков или традиционных электростанций. При этом текущая эксплуатация морских ветропарков является дорогостоящим мероприятием. Отметим, что ошибки прогно-

зирования обычно возрастают по мере увеличения временного горизонта. За последние 20 лет проведены исследования в области предсказания мощности ВЭУ, которые осуществлялись в рамках двух основных подходов — физического и статистического. Первый изучает процесс, основываясь на физических законах. Второй пытается извлечь информацию из ретроспективных исходных данных. У каждого из них есть свои сильные стороны, но часто используется комплексный подход.

Ряд исследований также показывает [6], что модели на основе деревьев решений (ДР), представляющие собой перспективную и более «прозрачную» технологию машинного обучения, достаточно эффективные для прогнозирования сложных нестационарных реализаций, не всегда могут быть обработаны на базе нейросетевых моделей. По сути, ДР представляют собой так называемые «жадные» алгоритмы построения модели принятия решений, использующиеся в статистике и анализе данных для прогнозных моделей. Структура ДР представляет собой «листья» и «ветки». На ребрах («ветках») ДР записаны атрибуты, от которых зависит целевая функция, в «листьях» записаны значения целевой функции, а в остальных узлах — атрибуты, по которым различаются случаи. Модель ДР позволяет получать стабильные решения, сравнимые, например, с машинами опорных векторов (МОВ) и нейросетями, не используя при этом больших вычислительных мощностей, требуемых предыдущими моделями.

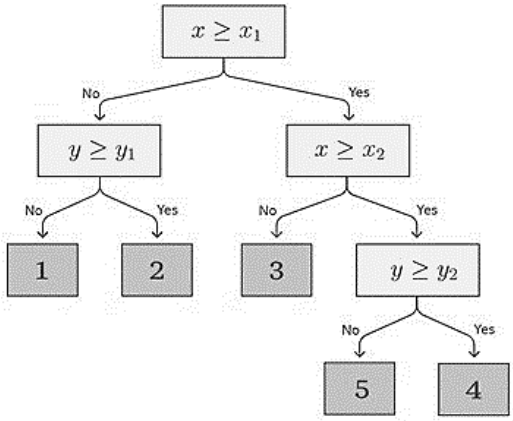

Рисунок 1 – Типичная структура ДР

Хотя обобщающая способность ДР невысока, их предсказания вычисляются довольно просто, из-за чего ДР часто используют как кирпичики для построения ансамблей — моделей, делающих предсказания на основе агрегации предсказаний других моделей.

Наиболее популярной модификацией алгоритма ансамблей ДР, применяемой для задач прогнозирования, стали модели случайного леса (от англ.

Random Forest) [7] и градиентного бустинга [8], которые позволяют строить множество деревьев на различных подмножествах обучающей выборки и, вследствие закона больших чисел, достигают лучших конечных результатов прогнозов путем выбора среднего из всех прогнозов деревьев. Данный пример приходит из жизни древних греков. Если много древних греков соберутся на одном холме и проголосуют независимо друг от друга, то вместе они придут к разумному решению несмотря на то, что опыт каждого из них субъективен.

Аналогом голосования в мире машинного обучения является бэггинг, которое и используется в основе алгоритма случайного леса. Идея бэггинга заключается в следующем. Пусть обучающая выборка состояла из n объектов. Выберем из неё n примеров равновероятно, с возвращением. Получим новую выборку X^1, в которой некоторых элементов исходной выборки не будет, а какие-то могут войти несколько раз. С помощью некоторого алгоритма обучим на этой выборке модель $b_1(x) = b(x, X^1)$. Повторим процедуру: сформируем вторую выборку X^2 из n элементов с возвращением и с помощью того же алгоритма обучим на ней модель $b_2(x) = b(x, X^2)$. Повторив процедуру k раз, получим k моделей, обученных на k выборках. Чтобы получить одно предсказание, усредним предсказания всех моделей:

$$a(x) = \frac{1}{k}\left(b_1(x) + \cdots + b_k(x)\right) \quad (1)$$

Процесс генерации подвыборок с помощью семплирования с возвращением называется бутстрепом (bootstrap), а модели $b_1(x) + \cdots + b_k(x)$ часто называют базовыми моделями. Модель $a(x)$ называется ансамблем этих моделей.

Бустинг - это ансамблевый метод, в котором так же, как и в методах выше, строится множество базовых алгоритмов из одного семейства, объединяющихся затем в более сильную модель. Отличие состоит в том, что в бэггинге и случайном лесе базовые алгоритмы учатся независимо и параллельно, а в бустинге — последовательно. Каждый следующий базовый алгоритм в бустинге обучается так, чтобы уменьшить общую ошибку всех своих предшественников. Как следствие, итоговая композиция будет иметь меньшее смещение, чем каждый отдельный базовый алгоритм (хотя уменьшение разброса также может происходить).

Математически это можно представить следующим образом. Алгоритм учится предсказывать разницу между правильным ответом и текущим предсказанием композиции из $k-1$ алгоритмов

$$b_k(x) = \mathrm{a}rg\min_{b\in\mathcal{B}}, \mathcal{L}(s^{k-1}, b(x)) \quad (2)$$

а композиция в целом обновляется по формуле

$$a_k(x) = a_{k-1}(x) - b_k(x) \quad (3)$$

Обучение K базовых алгоритмов завершает построение композиции.

На текущий момент основным видом бустинга с точки зрения применения на практике является градиентный бустинг. Хотя случайный лес — мощный и достаточно простой для понимания и реализации алгоритм, на практике он чаще всего уступает градиентному бустингу.

В [9] представлено широкомасштабное исследование моделей ансамблей ДР — случайный лес и градиентный бустинг для внутрисуточных прогнозов скорости ветра для ВЭС «Сотавенто» (Испания) и ветрогенерации всей Испании (по данным Системного оператора Испании). Эффективность моделей на основе ДР сравнивалась с МОВ и многослойным персептроном. Однако полученные результаты не выявили явного «победителя». На уровне прогноза для одной ВЭС модели на основе ансамблей ДР превосходят модель МОВ. В то же время прогнозы для всей ЭЭС Испании показали, что результаты ансамблей ДР близки к показателям точности более сложных в настройке МОВ. И наконец, наиболее низкие точности прогнозов для этих случаев зафиксированы при использовании моделей многослойного персептрона.

Для тестирования ансамблевых алгоритмов был использован реальный ветропарк, состоящий из 10 ВЭУ, исходные данные о которых содержали в себе почасовые измерения мощности и скорости ветра. На рисунке 2 показана гистограмма выходной мощности ветропарка, суммирующая мощности всех ВЭУ. Результаты расчетов показали, что большая часть измерений мощности находится в диапазоне от 0 до 1 – 126941 из 138710 измерений, нулевая мощность была зафиксирована в 11762 измерениях, единичная мощность – в 7. При этом доля измерений ВЭУ с нулевой мощностью для 3 зоны составила 900, а для 8 – 1562. Таким образом, ветропарк в целом работает достаточно эффективно.

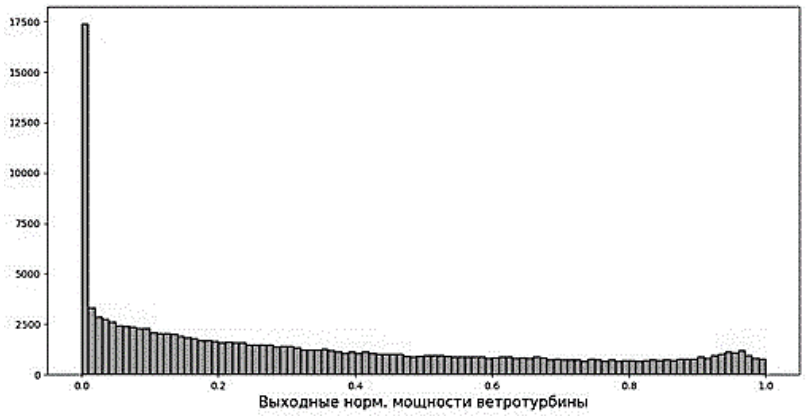

Рисунок 2 – Гистограмма выходной мощности ВЭУ, суммированная по всем ВЭУ ветропарка

В данном экспериментальном примере рассматривались данные только одной из ВЭУ ветропарка. Согласно графику, изображенному на рисунке 3, ну-

левая выходная мощность этой ВЭУ наблюдается при скорости ветра около 1,5–2 м/с на высоте 10 м и 2–4 м/с на высоте 100 м. Ненулевая выходная мощность в более широком диапазоне скоростей ветра: 2–6 м/с на высоте 10 м и 4–10 м/с на высоте 100 м.

Рассмотрим прогноз с использованием алгоритма градиентного бустинга (метод CatBoost) [10], случайного леса и традиционной линейной регрессии (метод регуляризации Тихонова). Для каждого интервала прогнозирования были рассчитана среднеквадратическая (RMSE) и средняя абсолютная ошибка (MAE) для обучающего и тестового набора данных. Результаты прогнозирования выходной мощности ВЭУ на 24 часа вперед представлены на рисунке 4.

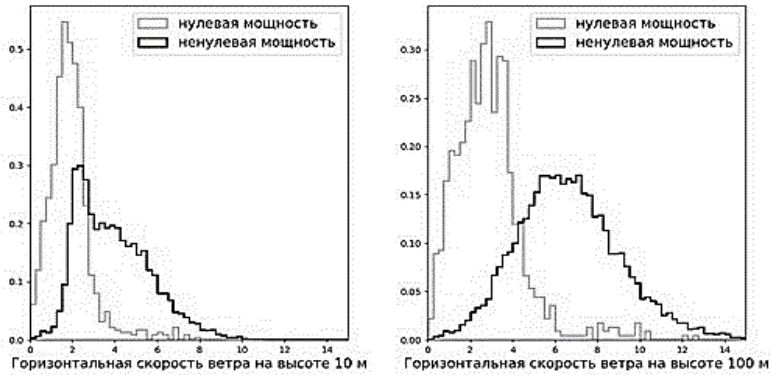

Рисунок 3 – График распределения зональных скоростей ветра для уровней 10 и 100 м для случаев нулевой и ненулевой выходной мощности рассматриваемой ветроустановки

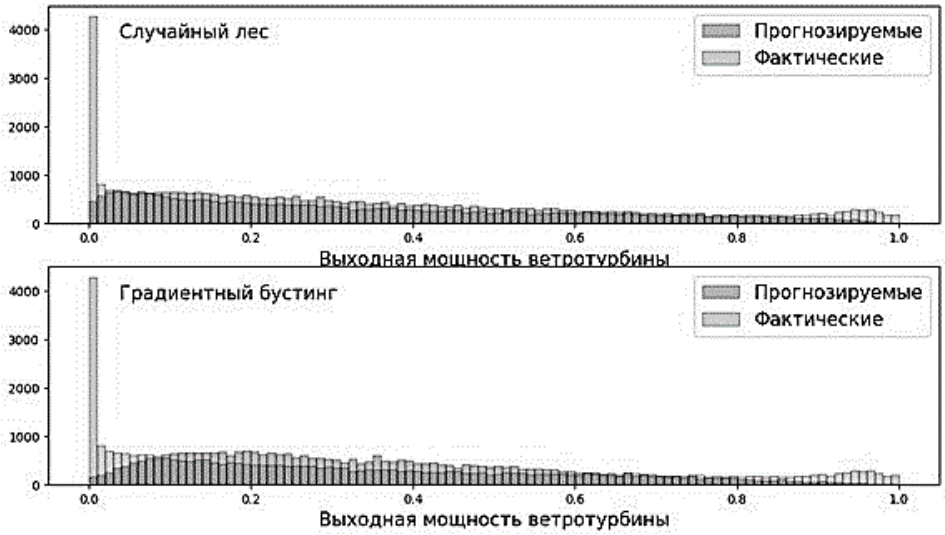

Рисунок 4 – Гистограмма с результатами прогнозирования выходной мощности ВЭУ на базе алгоритмов случайного леса и градиентного бустинга на 24 часа вперед

Как видно из таблицы 2, алгоритм случайного леса наиболее эффективен для прогнозирования на 24 часа вперёд, т.к. для данного интервала RMSE и MAE имеют наименьшие значения.

Таблица 2

Ошибки прогноза выработки мощности ВЭУ на 24 часа вперёд для сравниваемых алгоритмов

Алгоритм прогнозирования	Ошибка	
	RMSE, %	MAE, МВт
Случайный лес	13.6	0.09
Градиентный бустинг	18.2	0.14
Линейная регрессия	27.8	0.16

Для прогнозирования мощности ВЭУ требуется сложный математический аппарат, в настоящее время для этих целей широко используется методология машинного обучения. В данной статьей были исследованы наиболее популярные ансамблевые методы машинного обучения, такие как случайный лес и градиентный бустинг, для суточного прогнозирования выработки мощности ВЭУ. Данные методы были протестированы для прогноза ветрогенерации реального ветропарка на 24 часа вперед. Их эффективность сравнивалась с моделью традиционной линейной регрессии.

ЛИТЕРАТУРА

1. **Бубенчиков, А.А.** Применение ветроэнергетических установок с концентраторами ветровой энергии в регионах с малой ветровой нагрузкой // Международный научно-исследовательский журнал. Бубенчиков А. А., Артамонова Е. Ю., Дайчман Р. А., Файфер Л. А., Катеров Ф. В., Бубенчикова А. А. 2015. № 5– 2. 35с.

2. **Жолудева, Ю.К.** Математическая модель ветрогенератора // Научный альманах. Жолудева Ю. К., Мальцев М. В. 2017. №4-3(30). С. 63 – 67.

3. **Российская Федерация. Законы.** Об энергосбережении и о повышении энергетической эффективности, и о внесении изменений в отдельные законодательные акты Российской Федерации [Электронный ресурс]: Федеральный закон от 23.11.2009 № 261. URL: https://rg.ru/2009/11/27/energo-dok.html (дата обращения: 15.05.19)

4. **Стенникова, В.А.** Современные методы и модели прогнозирования скорости ветра и выработки мощности ветроэнергетических установок // Глава в монографии «Региональные аспекты ветроэнергетики»: под. ред. Стенникова В.А., Курбацкого В.Г., Томин Н.В., 2020.

5. **Безруких, П.П.** Ветроэнергетика. Вымыслы и факты. Ответы на 100 вопросов. М.: Институт устойчивого развития, Безруких П. П. 2014. 74 с. 2.

6. **Tomin N.,** Zhukov A., Sidorov D., Kurbatsky V., Panasetsky D., Spiryaev V. Random Forest Based Model for Preventing Large-Scale Emergencies in Power Systems // International Journal of Artificial Intelligence, 2015 Spring (March), Vol. 13, No. 1, pp. 211-228.

7. **Breiman,** “Random Forests”, Machine Learning, 45(1), 5-32, 2001.

8. **Friedman, J.H.** (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29, 1189-1232.

9. **Torres-Barrána,** A. Regression tree ensembles for wind energy and solar radiation prediction / A. Torres-Barrána, Á. Alonsoa, J. R. Dorronsoro // Neurocomputing. – 2019. – Vol. 326-327. – P. 151-160.

10. **Anna Veronika Dorogush,** Andrey Gulin, Gleb Gusev, Nikita Kazeev, Liudmila Ostroumova Prokhorenkova, Aleksandr Vorobev "Fighting biases with dynamic boosting". arXiv:1706.09516, 2017.