

**ЯДЕРНАЯ, РАДИАЦИОННАЯ И
ЭКОЛОГИЧЕСКАЯ БЕЗОПАСНОСТЬ**
NUCLEAR, RADIATION AND
ENVIRONMENTAL SAFETY

<https://doi.org/10.26583/gns-2026-01-02>

УДК 621.311:620.92

EDN YVYQFO

Оригинальная статья / Original paper



**Алгоритм прогнозирования выработки объекта генерации
возобновляемой энергетики**

А.Н. Ленских¹ ✉, Е.Ю. Алтунина², А.А. Поволоцкая³, А.Е. Дембицкий³

¹Национальный исследовательский ядерный университет «МИФИ», г. Москва, Российская Федерация

²Национальный исследовательский университет ИТМО, г. Санкт-Петербург, Российская Федерация

³Волгодонский инженерно-технический институт – филиал Национального исследовательского ядерного университета «МИФИ», г. Волгодонск, Ростовская обл., Российская Федерация

✉ lenskikh.andrey51@mail.ru

Аннотация. В данной статье рассмотрено решение задачи определения плана выработки электроэнергии для объекта генерации, который представляет собой ветроэлектростанцию. Ветроэнергетика, как один из перспективных альтернативных способов получения энергии, сталкивается с высокой неопределенностью генерации из-за изменчивости метеорологических условий. Для решения задачи оптимизации плана выработки электроэнергии на ветроэлектростанции в данной работе применены методы машинного обучения, которые позволят анализировать большие объемы данных, получаемых от различных сенсоров и метеорологических станций. Использование моделей машинного обучения способствует точному прогнозированию выработки энергии, что в свою очередь позволяет оптимизировать работу ветроэлектростанции, в том числе за счет корректировки режимов для максимизации коэффициента использования установленной мощности. В работе рассматривалось несколько математических моделей – модель k-ближайших соседей, модель дерева принятия решений, модель случайного леса и градиентный бустинг. Данные модели были подобраны по критерию алгоритмической простоты – их обучение проходит относительно быстро, а также из-за независимости от типа данных. В ходе анализа полученных данных каждой модели была выбрана модель градиентного бустинга – за наименьшее время обработки данных получен наибольший коэффициент детерминации на валидационных данных. Также создан виртуальный интерфейс для более удобного ввода данных и визуализации результатов.

Ключевые слова: возобновляемые источники энергии, ветроэнергетика, ветрогенератор, регрессионная модель, моделирование, прогнозирование, градиентный бустинг, модель дерева принятия решений, модель случайного леса, модель k-ближайших соседей, машинное обучение

Для цитирования: Ленских А.Н., Алтунина Е.Ю., Поволоцкая А.А., Дембицкий А.Е. Алгоритм прогнозирования выработки объекта генерации возобновляемой энергетики. *Глобальная ядерная безопасность*. 2026;16(1):15–22. <https://doi.org/10.26583/gns-2026-01-02>

For citation: Lenskih A.N., Altunina E.Yu., Povolotskaya A.A., Dembitsky A.E. Forecasting algorithm of renewable energy generation facility production. *Nuclear Safety*. 2026;16(1):15–22. (In Russ.). <https://doi.org/10.26583/gns-2026-01-02>

Forecasting algorithm of renewable energy generation facility production

**Andrey N. Lenskih¹ ✉, Ekaterina Yu. Altunina², Anastasia A. Povolotskaya³,
Artyom E. Dembitsky³**

¹National Research Nuclear University «MEPhI», Moscow, Russian Federation

²National Research University «ITMO», St. Petersburg, Russian Federation

³Volgodonsk Engineering Technical Institute the branch of National Research Nuclear University «MEPhI»,
Volgodonsk, Rostov region, Russian Federation

✉ lenskikh.andrey51@mail.ru

Abstract. The article considers the solution to the problem of determining the power generation plan for the generating facility, which is a wind power plant. Wind power, as a key renewable energy source, faces high generation uncertainty due to variability of meteorological conditions. In order to solve the task of optimizing the power generation plan at wind power plants, this work applies machine-learning methods that allow for analyzing large amounts of data obtained from various sensors and meteorological stations. The use of machine learning models helps to accurately predict energy production, which in turn allows to optimize the operation of wind power plant, including by adjusting the modes of maximizing the installed capacity utilization factor (ICUF). Several mathematical models are considered – k-nearest neighbor model, decision tree model, random forest model and gradient boosting. These models are chosen because of their algorithmic simplicity, their learning is relatively fast, and also because of their independence from data type. As a result of the analysis of the data obtained from each model, the gradient-boosting model is chosen - the highest coefficient of determination on validation data is obtained for the shortest time of data processing. Also created a virtual interface for easier data entry and visualization of results.

Keywords: renewable energy, wind power, wind generator, regression model, modelling, forecasting, gradient boosting, decision tree model, random forest model, k-nearest neighbor model, machine learning

Введение

Возобновляемая электроэнергия – энергия, получаемая из неисчерпаемых источников энергии. К типам таких источников относят солнечную, ветровую, гидравлическую энергию, а также энергию морских течений, приливов, тепла земли, биомассы. В последние десятилетия доля возобновляемых источников в мировой экономике растет, что, в первую очередь, связано с экологической политикой многих стран и стремлением отказаться от сжигания традиционного топлива, в особенности угля.

Привлекательность возобновляемых источников энергии (далее – ВИЭ) заключена в их естественности, неограниченности, доступности. Также необходимо отметить простоту конструкций генерирующих объектов и стоимость производства электроэнергии из ВИЭ. Это дает возможность обеспечивать малонаселенные и отдаленные территории, где строительство тепловых и атомных электростанций может быть экономически не обоснованно, стабильной подачей энергии.

Недостатком почти всех видов объектов, генерирующих электроэнергию из ВИЭ (в особенности солнечных и ветроэлектростанций), является зависимость от условий окружающей среды [1], а необходимость использования генераторов и аккумуляторов в энергоустановках для обеспечения длительной бесперебойной работы обуславливает значительное удорожание получаемой энергии. Кроме того, невозможно точно определить какое количество энергии от ВИЭ перейдет в энергетическую сеть, поэтому часто

недостаток восполняется выработкой энергии на электростанциях, использующих традиционное топливо, в первую очередь на ТЭС, которые способны к выполнению этой задачи. Однако эксплуатация ТЭС в режиме маневрирования мощностью негативно влияет на оборудование. Для контроля указанного негативного влияния необходимо разработать план выработки объекта генерации ВИЭ. Это позволит снизить объемы дополнительных резервов, размещаемых на теплоэлектростанциях, и оптимизировать системные издержки на электроснабжение всех потребителей. Для создания такого плана предложено применить метод машинного обучения.

Целью данной работы является разработка алгоритма прогнозирования энерговыработки ветроэлектростанции на основе методов машинного обучения с возможностью учета метеорологических данных. Задачи работы:

- определение набора параметров и метрик для сравнения и выбора оптимальной модели прогноза;
- рассмотрение различных математических моделей для алгоритма;
- анализ полученных результатов;
- рассмотрение возможных дальнейших направлений улучшения алгоритма.

Выбор математической модели

Суть моделей регрессии заключается в моделировании отношений между зависимой переменной и двумя или более независимыми переменными путем выявления за-

висимостей по некоторым законам¹.

В данной работе сравнивались следующие модели:

– модель k-ближайших соседей заключается в том, что метка (исходная известная характеристика, на основе которой происходит обучение и предсказание модели), присвоенная точке запроса, будет вычисляться на основе среднего значения меток ее ближайших соседей;

– модель дерева принятия решений основана на предсказывании значения целевой переменной, изучая простые правила принятия решений, выведенные из характеристик данных. При визуализации структура имеет вид дерева с различными типами узлов – корневыми, внутренними и конечными;

– модель случайного леса основана на модели дерева принятия решений. Основная идея заключается в использовании ансамбля деревьев решения;

– градиентный бустинг – техника, при которой предсказания строятся не независимо, а последовательно. Бустинг воплощает идею линейного построения линейных моделей. Каждый последующий алгоритм старается уменьшить ошибку текущего ансамбля.

Данные модели были выбраны ввиду их алгоритмической простоты, относительно высокой скорости обучения и малого времени выполнения, а также ввиду их независимости от типа данных (т.е. все перечисленные модели могут работать с категориальными (номинальными) данными без предварительного преобразования в числовой формат). Кроме того, данные методы достаточно эффективны в задачах прогнозирования временных рядов и работы с табличными данными.

Для оценки модели будут рассматриваться следующие метрики:

а) средняя квадратичная ошибка – средний квадрат абсолютных разностей между целевыми значениями и значениями, предсказанными моделью на данном обучающем примере в процессе обучения:

$MSE = 1/N \sum_{i=1}^N |y_i - \hat{y}_i|^2$, где N – количество наблюдений; y_i – истинное значение для i-го наблюдения; \hat{y}_i – предсказанное моделью значение для i-го наблюдения;

б) коэффициент детерминации – доля дисперсии зависимой переменной, которая объясняется рассматриваемой моделью зависимости: $R^2 = 1 - RSS/TSS$, где RSS – сумма квадратов остатков регрессии, сумма TSS – сумма квадратов отклонений значений от среднего.

Предварительная обработка данных

Для прогнозирования мощности ветроэлектростанции (далее – ВЭС) авторами было принято решение использовать модель регрессии на основе машинного обучения. Это позволило получить ряд преимуществ: минимальные затраты времени на составление модели и достаточная точность прогноза модели при наличии большого количества входных данных [2]. Большинство моделей, используемых в машинном обучении, имеют опыт практического применения.

Для подготовки к рассмотрению нескольких математических моделей необходимо произвести предварительную обработку данных, т.к. модели машинного обучения уязвимы к разнообразию входных данных. Предварительная обработка данных начинается с исключения из рассмотрения пустых данных и отрицательных значений выработки электроэнергии, связанных с внеплановой недовыработкой из-за нестабильности ветра и неравномерности производства электроэнергии впоследствии. Рассмотрим их на диаграмме рассеивания выработки электроэнергии от скорости ветра на рисунке 1.

Можно предположить, что многие аномальные значения, наблюдаемые на графике (в основном под главной зависимостью), возникают из-за промежутков ремонта некоторых объектов ВЭС, преднамеренным снижением электропотребления на ВЭС для балансировки потребления электроэнергии (это особенно наблюдается в промежутках выработки наибольшего количества энергии), а также внешним воздействием на лопасти ветрогенераторов, к примеру, птиц [3].

¹ Мхитарян В.С. Анализ данных : учебник для вузов. – Москва: Издательство Юрайт, 2024. – 448 с.

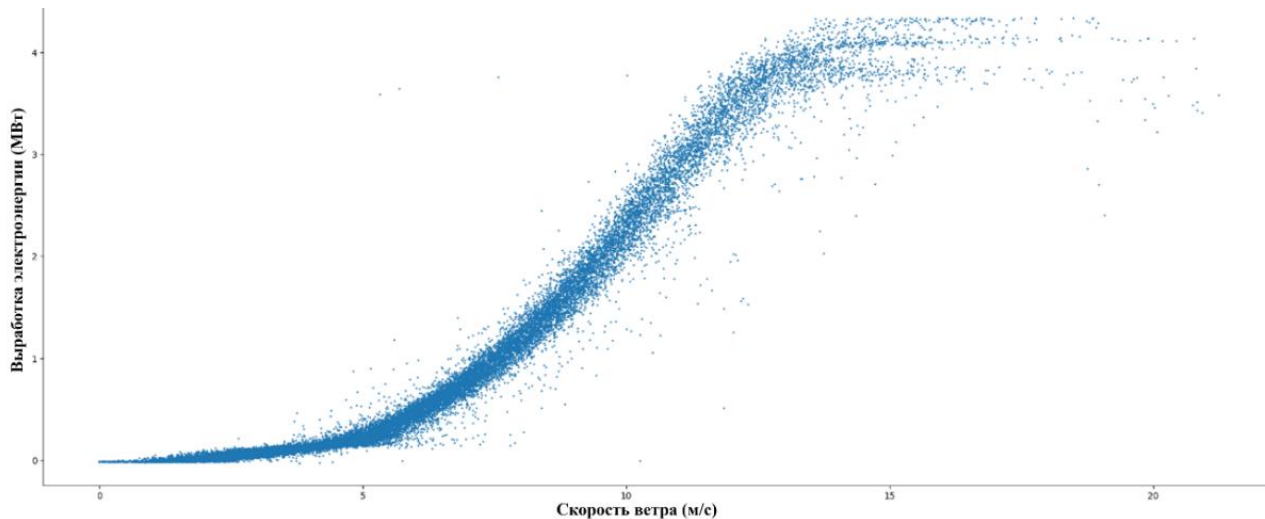


Рисунок 1. Диаграмма рассеивания выработки электроэнергии от скорости ветра
Figure 1. The scatter plot of electricity production on wind speed

Для удаления из рассмотрения ненужных в дальнейшей работе данных массив разделяется на несколько групп по скорости с шагом 0,1. В каждой из групп вычисляется среднее значение 0,05- и 0,75-квантиля ($q_{0,05}$, $q_{0,75}$). Разница этих двух квантилей будет определяться как $\Delta q = q_{0,75} - q_{0,05}$. В статистической фильтрации и очистке данных чаще всего используются множители около 1,4–1,5 для поиска интервалов, которые охватывают значимую часть данных,

исключая при этом редкие чрезмерные выбросы [4]. Выбор 1,4 является компромиссом между жесткостью фильтрации (чтобы не потерять ценные данные) и тщательностью очистки. В итоге отбрасываться будут те величины, которые не попадают в промежуток $[q_{0,05} - 1,4\Delta q; q_{0,05} + 1,4\Delta q]$. Результат предварительной обработки данных, полученный с помощью функций библиотек Pandas и NumPy на языке программирования Python, представлен на рисунке 2.

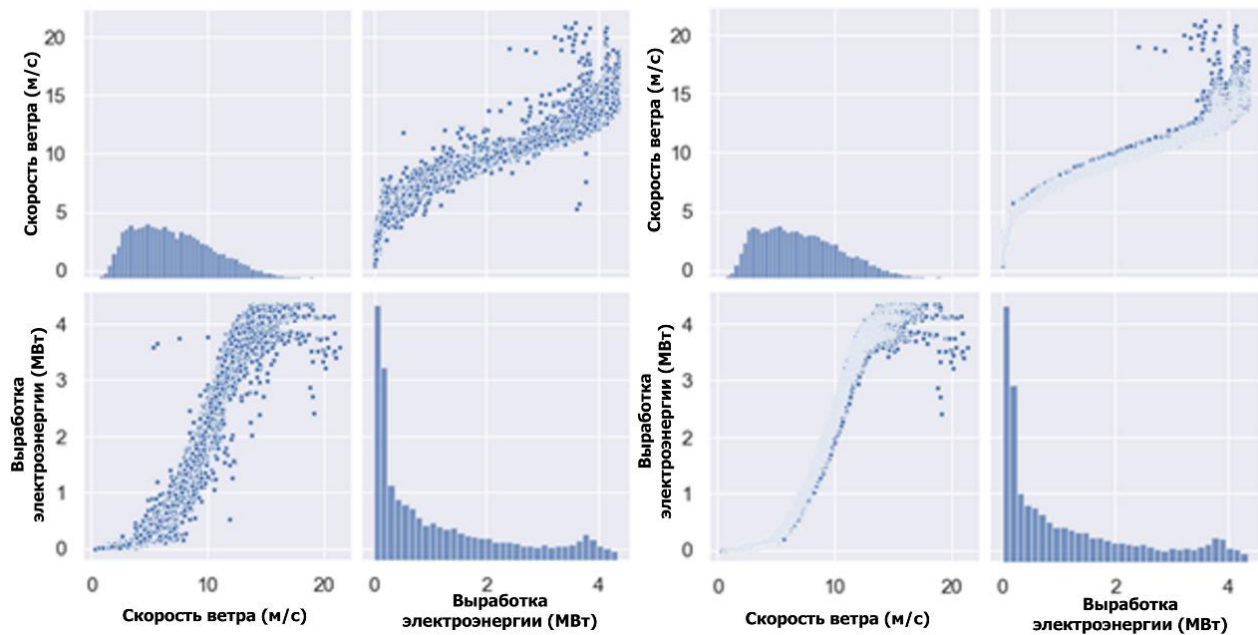


Рисунок 2. График зависимости выработки электроэнергии от скорости ветра до и после предварительной обработки

Figure 2. Graph of the dependence of before and after preprocessing

Перед обучением моделей необходимо применить стандартный масштабатор, использующийся при сильном различии между диапазонами или при различиях между единицами изменения свойств. Метод стандартизирует диапазон входных данных путем приведения их всех к значениям, при которых среднее значение наблюдаемых значений равно 0, а среднее отклонение – 1 [5].

Обучение моделей

На вход будет подаваться несколько свойств:

- скорость ветра;
- месяц;
- час рассматриваемого дня;

На выходе алгоритм формирует почасовой прогноз выработки на горизонт метеорологического прогноза (трое суток).

Обучение происходит на тестовой выборке, составляющей 40% от объема входного датасета с помощью функции `sklearn.model_selection.train_test_split` [6].

Полученные значения метрик представлены в таблице 1, результаты обучения моделей приведены на рисунке 3.

Таблица 1. Результаты анализа моделей

Table 1. Model analysis results

Модель	MSE, МВт	R ² тестовой выборки	R ² валид. Выборки	Время выполнения, с
KNeighborsRegressor	0.0178	0.9921	0.9883	0.6
DecisionTreeRegressor	0.0237	1.0000	0.98	0.6
RandomForestRegressor	0.0134	0.9988	0.9912	6.1
GradientBoostingRegressor	0.0121	0.9931	0.9921	1.7

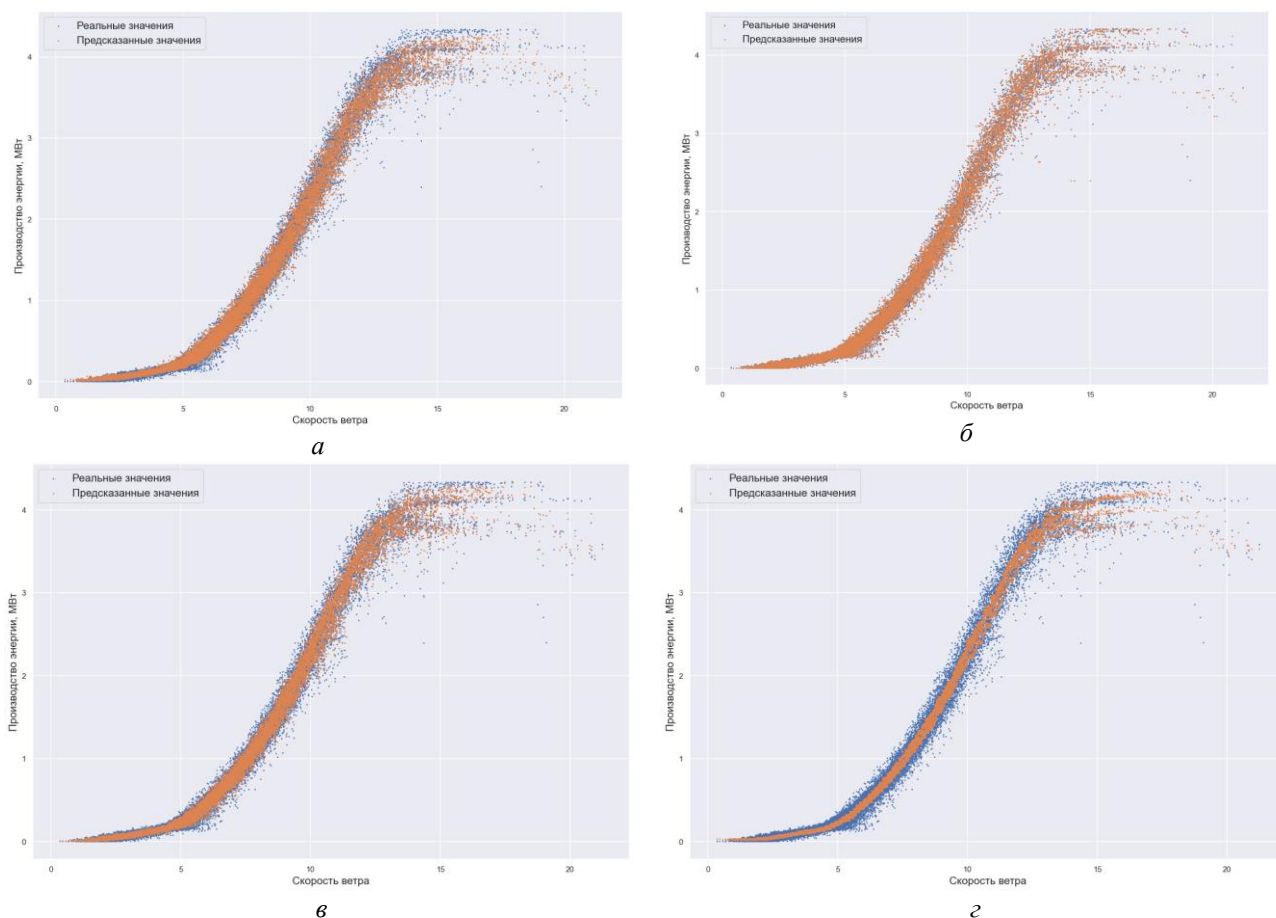


Рисунок 3. Прогнозирование с помощью: а – KNeighborsRegressor; б – DecisionTreeRegressor; в – RandomForestRegressor; г – GradientBoostingRegressor

Figure 3. Forecasting using а – KNeighborsRegressor; б – DecisionTreeRegressor; в – RandomForestRegressor; г – GradientBoostingRegressor

При анализе полученных графиков, метрик моделей и сравнении времени выполнения выбрана модель градиентного бустинга: при относительно малом времени обработки данных был получен наибольший коэффициент детерминации на валидационных данных. Также валидационные данные и метрики показывают, что модель RandomForest имеет большее переобучение, что привело к выбору именно градиентного

бустинга, который минимизирует ошибку последовательно, обеспечивая обобщение лучше, чем RandomForest.

Для удобного ввода и вывода информации создан веб-интерфейс на основе библиотеки React. Для интеграции python применяется микрофреймворк Flask, использующийся для чтения входных данных. Вид интерфейса на входе представлен на рисунке 4.

Рисунок 4. UI веб-приложения
Figure 4. UI web applications

На вход принимаются два файла – почасовые выработки конкретной ветроэлектростанции и метеорологический прогноз для местности, имеющие форматы csv, xls,.xlsx или txt. На выходе выгружается обрабо-

танный массив данных, а также график, отображающий почасовые данные выработки энергии, который можно увидеть на рисунке 5.



Рисунок 5. UI веб-приложения – вывод информации
Figure 5. UI web applications – information output

Выводы

В ходе работы проведено сравнение математических моделей для решения задачи определения плана выработки электроэнергии: модель k-ближайших соседей, модель дерева принятия решений, модель случайного леса и градиентный бустинг.

По результатам анализа выбран градиентный бустинг, на основе которого создана рабочая модель машинного обучения для получения величины выработки электроэнергии для ветроэлектростанции. Имея данные прогноза погоды, с помощью этой модели будет получен прогноз выработки электроэнергии, что поможет оптимизировать работу станции в отношении корректировки ее режимов.

Автоматизация прогноза выработки электроэнергии позволит повысить эффективность оперативного планирования, снизить трудозатраты, снизить влияние внешних факторов (в данном случае – погодных условий), ограничивающих выдачу мощности, увеличивая коэффициент использования

установленной мощности путем изменения режима работы ВЭС.

Среди математических моделей градиентный бустинг показал более низкую среднюю квадратичную ошибку и более высокий коэффициент детерминации по сравнению с другими моделями при относительно недолгом времени выполнения.

Одним из способов улучшения метрик модели является дополнение метеорологических данных, таких как температура и давление воздуха, направление ветра и высота расположения ветроэлектростанции.

Возможно дополнение данных минимальной скоростью ветра для работы турбины, скоростью для работы турбины в номинальном режиме и максимальной скорости ветра, после которой срабатывает тормоз турбины, и ее работа прекращается. Учет этих данных поможет в предварительной обработке данных, в первую очередь в отсеивании аномальных значений, наблюдаемых на высоких скоростях ветротурбины.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Захожий К.А. Возобновляемые источники энергии. *Colloquium-journal*. 2020;28(80):57-58. <https://doi.org/10.24412/2520-2480-2020-2880-57-58>
 2. Chen X., Zhang X., Dong M., Huang L., Guo Y., He S. Deep learning-based prediction of wind power for multi-turbines in a wind farm. *Frontiers in Energy research*. 2021;9:723775. <https://doi.org/10.3389/fenrg.2021.723775>
 3. Song D., Zheng S., Yang S., Yang J., Dong M., Su M., et al. Annual energy production estimation for variable-speed wind turbine at high-altitude site. *Journal of Modern Power Systems and Clean Energy*. 2021;9(3):684-687. <https://doi.org/10.35833/MPCE.2019.000240>
 4. Tukey J.W. *Exploratory data analysis*. Addison-Wesley Publishing Company, 1977. P. 688. Available at: https://archive.org/details/exploratorydataa0000tukey_7616/page/n3/mode/2up (accessed: 02.09.2025).
 5. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных. Москва: ДМК Пресс, 2015. 400 с. Режим доступа: https://rusneb.ru/catalog/000199_000009_008646425/?ysclid=mli2b67snu451963270 (дата обращения: 02.09.2025).
- Flah P. *Machine learning. The science and art of building algorithms that extract knowledge from data*. Moscow: HMB Press, 2015. 400 p. (In Russ.). Available at: https://rusneb.ru/catalog/000199_000009_008646425/?ysclid=mli2b67snu451963270 (accessed: 02.09.2025).

ВКЛАД АВТОРОВ:

Ленских А.Н. – постановка задачи, разработка математической модели;
Алтунина Е.Ю. – программная реализация математической модели, проведение численных исследований;
Поволоцкая А.А. – подготовка текста статьи;
Дембицкий А.Е. – методологический надзор за исследованием, общее руководство и координация работы.

ИСТОЧНИКИ ФИНАНСИРОВАНИЯ:

Авторы заявляют об отсутствии внешних источников финансирования.

AUTHORS' CONTRIBUTION:

Lenskih A.N. – problem statement, development of a mathematical model;
Altunina E.U. – software implementation of a mathematical model, conducting numerical studies;
Povolotskaia A.A. – preparation of the article text;
Dembitsky A.E. – study methodological oversight, overall work guidance and coordination.

FUNDING:

There are no sources of external funding.

КОНФЛИКТ ИНТЕРЕСОВ:

Авторы заявляют об отсутствии конфликта интересов.

CONFLICT OF INTEREST:

No conflict of interest.

ИНФОРМАЦИЯ ОБ АВТОРАХ:

Андрей Николаевич Ленских, аспирант, кафедра теоретической и экспериментальной физики ядерных реакторов, Национальный исследовательский ядерный университет «МИФИ», г. Москва, Российская Федерация.

e-mail: lenskikh.andrey51@mail.ru

Екатерина Юрьевна Алтунина, магистр, кафедра компьютерных технологий, Национальный исследовательский университет ИТМО, г. Санкт-Петербург, Российская Федерация.

e-mail: altunina0ekaterina@gmail.com

Анастасия Александровна Поволоцкая, кафедра атомной энергетики, Волгодонский инженерно-технический институт – филиал Национального исследовательского ядерного университета «МИФИ», г. Волгодонск, Ростовская обл., Российская Федерация.

e-mail: qqqt0@mail.ru

Артем Евгеньевич Дембицкий, кандидат технических наук, зав. кафедрой атомной энергетики, Волгодонский инженерно-технический институт – филиал Национального исследовательского ядерного университета «МИФИ», г. Волгодонск, Ростовская обл., Российская Федерация.

e-mail: demartev@yandex.ru

INFORMATION ABOUT THE AUTHORS:

Andrey N. Lenskih, Postgraduate, Department of theoretical and experimental physics of nuclear reactors, National Research Nuclear University «MEPhI», Moscow, Russian Federation.

e-mail: lenskikh.andrey51@mail.ru

Ekaterina Yu. Altunina, Master, Department of computer technology, National Research University ITMO, Saint-Petersburg, Russian Federation.

e-mail: altunina0ekaterina@gmail.com

Anastasia A. Povolotskaia, Department of atomic energy, Volgodonsk Engineering Technical Institute the branch of National Research Nuclear University «MEPhI», Volgodonsk, Rostov region, Russian Federation.

e-mail: qqqt0@mail.ru

Artyom E. Dembitsky, Cand. Sci (Eng), Head of the Department of Atomic Energy, Volgodonsk Engineering Technical Institute the branch of National Research Nuclear University «MEPhI», Volgodonsk, Rostov region, Russian Federation.

e-mail: demartev@yandex.ru

Поступила в редакцию / Received 05.11.2025

После доработки / Revision 09.02.2026

Принята к публикации / Accepted 12.02.2026