



АНСАМБЛЕВАЯ МОДЕЛЬ ДЛЯ ПРОГНОЗИРОВАНИЯ ВЫРАБОТКИ ВЕТРОВЫХ ЭЛЕКТРОСТАНЦИЙ

Русина А.Г., Осгонбаатар Тувшин, Матренин П.В., Сергеев Н.Н.

Новосибирский Государственный Технический Университет, г. Новосибирск, Россия

Резюме: Применение возобновляемых источников энергии является перспективным путем снижения вредных выбросов в процессе производства энергии и решения проблемы ухудшения экологической ситуации. Несмотря на то, что возобновляемая энергия считается чистой энергией, которую также называют «зеленой» энергией, вопрос ее использования в энергосистеме вызывает определенные трудности. Эффективное использование возобновляемых источников энергии требует актуальной информации о первичных энергоносителях, что особенно важно при крупномасштабной интеграции возобновляемых источников в систему. Таким образом, для обеспечения нормальных режимов работы энергосистемы необходимо прогнозировать выработку возобновляемых источников с допустимой погрешностью. **ЦЕЛЬ.** Прогнозирование суточного графика выработки ветровых электростанций. **МЕТОДЫ.** В работе использовались ансамблевые алгоритмы, основанные на деревьях решений и являющиеся одним из подходов машинного обучения. Программная реализация выполнена с помощью языка программирования Python. В качестве исходных данных использованы данные о скорости ветра и выработке определенных электростанций за период 2019-2021 гг. **РЕЗУЛЬТАТЫ.** С помощью предложенной методики был составлен прогноз суточных графиков выработки трех ветровых электростанций с погрешностью от 2,4 до 3,5 МВт или от 5,0 до 7,0 % установленной мощности соответствующих ветровых электростанций. Нормализованная средняя ошибка по модулю в процентах составила от 12,3 до 13,3 %. **ЗАКЛЮЧЕНИЕ.** Ансамблевые методы машинного обучения позволяют обнаруживать нелинейные и нестационарные зависимости во временных рядах, а также могут быть реализованы в задаче прогнозирования суточного графика выработки ветроустановок. Повышение точности прогнозирования выработки ветроустановок имеет высокую значимость для эффективного функционирования и планирования режимов энергосистемы.

Ключевые слова: прогнозирование; выработка; скорость ветра; ансамблевый метод; ветровая электростанция.

Для цитирования: Русина А.Г., Осгонбаатар Тувшин, Матренин П.В., Сергеев Н.Н. Ансамблевая модель для прогнозирования выработки ветровых электростанций // Известия высших учебных заведений. ПРОБЛЕМЫ ЭНЕРГЕТИКИ. 2024. Т.26. № 1. С. 64-76. doi:10.30724/1998-9903-2024-26-1-64-76.

ENSEMBLE MACHINE LEARNING MODEL FOR FORECASTING WIND FARM GENERATION

Rusina AG., Osgonbaatar Tuvshin, Matrenin PV., Sergeev NN.

Novosibirsk State Technical University, Novosibirsk, Russia

Abstract. The use of renewable energy sources is the only way to avoid emissions from energy production and to decide pollution of ecology. Despite the fact that renewable energy has become clean energy, which called green energy, the issue of using it is quite difficult for the control and regulate the energy system. Efficient use of renewable energy requires information on primary sources. If looking case of large-scale integration, this requirement will be significantly felt. Thus, to ensure the normal operating modes of the energy system, it is necessary to predict the

generation of renewable sources with an acceptable error. PURPOSE. To forecast the generation of wind farms. METHODS. This study is carried out by ensemble algorithms, such as Random Forest, AdaBoost and XGBoost, which are one of the machine learning approaches. The software implementation is made using the Python programming language. As initial inputs historical data on windspeed and generating of some windfarms in Mongolia by 2019-2021 were used. RESULTS. The proposed method predicted daily production schedules at three wind farms with an error of 2.4 to 3.4 MW or 5.0 to 7.0 percent of the installed capacity of the corresponding wind farm. Also, normalized MAE was 12,3 to 13.3 percent. CONCLUSIONS. Ensemble methods of machine learning made it possible to determine non-linear and non-stationary dependencies of the time series, and also can be implemented in the problem of predicting the daily production schedule. Increasing the accuracy of wind energy forecasting will affect positively the operation and planning of the power systems.

Keywords: forecasting; generate; windspeed; ensemble method; wind farm.

For citation: Rusina AG., Osgonbaatar Tuvshin, Matrenin PV., Sergeev NN. Ensemble machine learning model for forecasting wind farm generation. *Power engineering: research, equipment, technology.* 2024; 26 (1): 64-76. doi:10.30724/1998-9903-2024-26-1-64-76.

Введение (Introduction)

В настоящее время ресурсы традиционной энергетики значительно сокращаются и из-за их чрезмерного использования резко ухудшается состояние окружающей среды. Таким образом, возобновляемая энергия становится перспективным альтернативным источником электроэнергии, которая дает возможность снизить вредные выбросы от производства энергии. Основной целью производства электроэнергии с помощью возобновляемых источников энергии (ВИЭ) является снижение зависимости от ископаемого топлива, применение которого оказывает негативное воздействие на окружающую среду. Однако использование ВИЭ приводит к экономическим и техническим проблемам, в число которых входит усложнение процессов управления режимами, планирования электрической сети и функционирования оптового рынка электроэнергии. Задача обеспечения баланса генерируемой и потребляемой электроэнергии требует особого внимания при внедрении ВИЭ-генерации в систему. Растущая интеграция ВИЭ увеличивает нестабильность баланса мощности энергосистемы, поскольку появляется дополнительная неопределенность на стороне производства электроэнергии. Следовательно, точное прогнозирование выработки ВИЭ играет важную роль в обеспечении непрерывного функционирования энергосистемы.

На данный момент энергия ветра широко используется как один из ключевых ресурсов возобновляемой энергии, и ее доля на рынке электроэнергии стремительно растет во всем мире. С другой стороны, объем электроэнергии, вырабатываемой ветровой электростанцией, сильно меняется в зависимости от колебаний ветра и его направления на рассматриваемой территории. Факторы, которые часто влияют на выработку, включают метеорологические условия, турбулентность ветра, расположение электростанции, сезонность и т.д [1,2]. Прогнозирование скорости и направления ветра достаточно сильно затруднено из-за его резко переменного и стохастического характера изменения с течением времени [3].

В центральной энергосистеме Монголии функционируют 3 ветровые электростанции с установленной мощностью от 50 МВт до 55 МВт, расположенные в различных регионах страны. Общая установленная мощность данной энергосистемы составляет 1488 МВт, из которых 245 МВт вырабатывается за счет ВИЭ, в том числе 155 МВт – ветровыми электростанциями. Также правительством была поставлена цель увеличить объем мощности возобновляемых источников на 20 % к 2023 году и на 30 % к 2030 году для сокращения выбросов парниковых газов, связанных с энергетикой [4]. Видно, что ВИЭ, в частности энергия ветра, играют важную роль в производстве электроэнергии в Монголии. Иными словами, существует значительное влияние ВИЭ на функционирование энергосистемы, а также имеется вышеупомянутая необходимость прогнозирования их выработки с допустимой погрешностью. На данный момент при планировании режимов работы центральной энергосистемы на суточном интервале требуется совершенствование методики прогнозирования выработки ВИЭ.

Целью исследования является разработка алгоритмов прогнозирования суточного графика выработки ветровых электростанций. Практическая значимость исследования

заключается в снижении влияния неопределенности и случайного характера изменения ветра как первичного энергоносителя на функционирование энергосистемы и планирование ее режимов.

Литературный обзор (Literary review)

Имеется ряд публикаций, посвященных решению задач краткосрочного прогнозирования скорости ветра и выработки электроэнергии ветроустановок. В работах [5-8] предложенные методы разделены на физические, статистические и гибридные или ансамблевые.

Было установлено, что применение физических методов является одним из лучших подходов к долгосрочному прогнозированию. Эти методы используют метеорологические и физические данные, в том числе скорость ветра, температуру, давление и влажность наружного воздуха для создания математических моделей высокой сложности. В связи со сложностью определения метеорологической информации на близкие непродолжительные периоды времени, данные подходы не подходят для решения задачи краткосрочного прогнозирования. В целом, физические методы основаны на численном прогнозе погоды для расчета выработки ветровых электростанций [7].

Главное преимущество статистических методов заключается в том, что они не требуют физических данных для построения модели прогнозирования. Данные методы рассматривают линейную и нелинейную зависимость между исходными переменными временного ряда. Линейные модели обычно отслеживают устойчивое поведение ветра с помощью методов авторегрессии, включая *ARIMA*, *ARMA* и другие [9]. При прогнозировании скорости ветра в случае нелинейности и нестационарности рассматриваемого временного ряда применяются методы искусственного интеллекта, такие как нейронные сети [10], метод опорных векторов [11], нечеткая логика [12] и другие методы машинного обучения. В основном статистические методы используются для построения моделей краткосрочного прогнозирования [13].

Физические и статистические модели имеют свои характеристики и зачастую не могут полностью описать сложную закономерность временного ряда характеристик ветра, если они используются по отдельности. Применение ансамблевых методов, объединяющих несколько алгоритмов, является эффективным способом повышения достоверности одиночных моделей машинного обучения и других методов. Ниже приведены примеры исследований, направленных на решение рассматриваемой задачи.

В работе [14] авторы предложили ансамблевый метод, включающий алгоритм адаптивного бустинга (*AdaBoost*) для моделирования временных рядов скорости ветра и метод опорных векторов для преобразования скорости ветра в выработку ветровой электростанции. Предлагаемый метод был реализован на семи ветровых электростанциях и показал достаточную точность прогнозирования.

T. Peng и др. [15] разработали методику многоступенчатого прогнозирования скорости ветра с использованием ансамблевого алгоритма. В качестве исходных данных они использовали данные о скорости ветра с 10-минутными интервалами за весь 2018 год. Исходные данные были разделены по сезонности с помощью метода опорных векторов, а выработка электроэнергии за каждый сезон была рассчитана с использованием нейронных сетей.

В работе [16] показано применение ансамблевого метода для прогнозирования скорости ветра. Модель была построена с использованием вейвлет-пакетной декомпозиции, нейронных сетей и алгоритма бустинга. Полученная ансамблевая модель дала более точный результат по сравнению с одиночными моделями.

L. Wang с соавторами [17] сравнили результаты статистических методов и ансамблевого метода. В работе доказано, что адаптивный бустинг может применяться для прогнозирования скорости ветра с достаточной точностью.

В работе [18] гибридные методы, основанные на экстремальном градиентном бустинге (*XGBoost*), адаптивном бустинге и нейронных сетях, используются для прогнозирования почасовой скорости ветра.

В работе [19] авторы спрогнозировали скорость ветра используя ансамблевый метод, включающий методы опорных векторов и вейвлет-преобразования. Первичные данные временного ряда были декомпозированы с помощью вейвлет-преобразования, а модель прогнозирования была построена с помощью алгоритма опорных векторов путем обучения на этих декомпозированных данных.

Область применения ансамблевых моделей машинного обучения достаточно широкая. Вышеуказанные работы доказывают преимущества и возможности применения ансамблевых методов для построения многофакторной модели, соответствующей

стохастическому временному ряду. В данной работе предлагаются методы прогнозирования скорости ветра, основанные на алгоритмах деревьев решений [20], в том числе случайный лес [21,22], адаптивный бустинг [23], экстремальный градиентный бустинг [24]. В качестве исходных данных были использованы временные ряды скорости ветра и генерации за период 2019–2021 гг. на трех ветровых электростанциях, расположенных в различных регионах Монголии.

Методы исследования (Research Methods)

Дерево решения.

Дерево решения [20] представляет собой набор узлов и ветвей, построенных по иерархической архитектуре без цикла. Каждый узел выполняет функцию регрессии или классификации по одному признаку, которая применяется к исходным данным. Решение каждого узла передается ветвями к следующим узлам. Окончательные узлы (листовые) дают результат классификации или регрессии. Рассмотрим определенный временной ряд, выражающийся следующим образом:

$$S_n = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)\}$$

где X – исходные векторы, содержащие функции $f(X)$; Y – выходные скаляры или метки; S_n – исследуемые выборки временного ряда (X_n, Y_n) с номером наблюдения n .

Для разработки алгоритма необходимо разделить данные на обучающую и тестовую выборки. После обучения на данных алгоритм будет построен на основе модели, которая вычисляет зависимости между соответствующими переменными. Иными словами, в конце процесса обучения алгоритм выдает функцию $\hat{h}(X, S_n)$ модели временного ряда. На основе принципа деревьев решений разработаны некоторые ансамблевые алгоритмы для построения моделей регрессии и классификации, которые рассмотрены ниже.

Случайный лес.

Случайный лес – ансамблевый алгоритм, разработанный Брейменом [21]. Данный алгоритм является одним из ансамблевых методов, основанных на применении деревьев решений. На рисунке 1 показана структура алгоритма случайного леса.

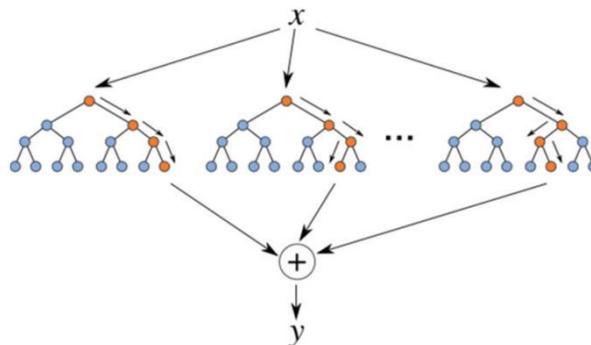


Рис. 1. Структура алгоритма случайного леса [22]

Fig. 1. Structure of the Random Forest algorithm [22]

Суть алгоритма случайного леса заключается в том, что выборки размером n из обучающего временного ряда S_n выбираются случайным образом и размещаются в узлах деревьев решений. На каждом дереве проводится регрессионный анализ и классификации случайных выборок и выводятся их модели, которыми выражаются зависимости между случайными переменными. В конце анализа каждое дерево дает выходную функцию $\hat{h}(X, S_n)$. Агрегирование результатов путем усреднения выходных функций в конечном итоге дает модель случайного леса:

$$\hat{Y} = \frac{1}{q} \sum_{i=1}^q \hat{h}(X, S_n^i)$$

где S_n^i – i -ая случайная выборка; q – число деревьев решений.

Кроме алгоритма случайного леса, еще одним методом повышения качества моделирования является бустинг (boosting), основная идея которого заключается в итеративном процессе построения ансамблевых моделей.

Адаптивный бустинг.

Адаптивный бустинг (AdaBoost) – алгоритм, который объединяет несколько слабых предикторов для создания сильного предиктора [23]. Слабыми предикторами в данном случае считаются отдельные регрессионные деревья.

В данном алгоритме выходу каждого слабого предиктора присваивается коэффициент, рассчитанный с учетом ошибок на предыдущих итерациях. В процессе

обучения алгоритма, весовые коэффициенты слабых предикторов получают большое значение в случае высокой погрешности и меньшее значение при низкой погрешности. Затем алгоритм снова обучается на основе обновленных весовых коэффициентов, чтобы улучшить прогнозируемый результат. Такой итерационный процесс продолжается до тех пор, пока все весовые коэффициенты не будут минимизированы. Процесс обучения алгоритма осуществляется в следующей последовательности. Предположим, что существует следующий набор выборок.

$$S = \{(x_i, y_i) | i = 1, 2, \dots, N\}. D_{q(i)}$$

где x_i – исходные переменные, y_i – выходные переменные, i – номер элемента выборки, $q = 1, 2, \dots, Q$ – номер итерации, D – весовые коэффициенты.

На первой итерации первичный весовой коэффициент определяется как

$$D_{1(i)} = 1/N$$

где N – количество элементов выборки набора S_n .

На основе весовых коэффициентов строится модель прогнозирования $\widehat{h}_q(x_i)$, а затем вычисляется погрешность модели:

$$e_{q(i)} = |\widehat{h}_q(x_i) - y_i|, \quad e_{q(i)} \in [0, 1]$$

где $e_{q(i)}$ – погрешность i -ого элемента выборки на q -ой итерации, y_i – фактическое значение целевой переменной i -ого элемента выборки.

Далее алгоритм вычисляет пропорциональную ошибку ε_q на каждой итерации по следующему выражению:

$$\varepsilon_q = \sum_{i=1}^N D_{q(i)} e_{q(i)}$$

где $e_{q(i)}$ – погрешность, $D_{q(i)}$ – распределение весов i -ого элемента выборки на q -ой итерации.

Следующим шагом является расчет весовых коэффициентов на каждой итерации:

$$w_q = 1/2 \log\left(1/\beta_q\right), \quad \text{при} \quad \beta_q = \varepsilon_q / (1 - \varepsilon_q)$$

где w_q – весовой коэффициент.

Затем необходимо обновить распределение весов, установленное на предыдущих итерациях:

$$D_{q+1,(i)} = (D_{q,(i)} * \beta_q^{-\varepsilon_q}) / V_{q+1}, \quad \text{при} \quad V_{q+1} = \sum_{i=1}^N D_{q+1,(i)}$$

где $D_{q+1,(i)}$ – новое распределение весов i -ого элемента выборки на $(q + 1)$ -ой итерации.

После итерации Q строится итоговая модель алгоритма адаптивного бустинга, которая выражается следующим образом:

$$\widehat{Y}' = \frac{1}{q} \sum_{q=1}^Q w_q \widehat{h}_q(x)$$

где \widehat{Y}' – выходная функция алгоритма.

С точки зрения структуры алгоритм бустинга, как и алгоритм случайного леса, заключается в том, что алгоритм объединяет несколько деревьев решений, называемых слабыми предикторами, для создания сильного предиктора. Отличие от случайного леса в том, что каждая последующая модель данного алгоритма стремится исправить ошибки, допущенные на предыдущей итерации. В итоге в зависимости от распределения весовых коэффициентов не все деревья участвуют в процессе принятия окончательного решения: деревья решений, часто допускающие ошибки, оказывают меньшее влияние на параметры выходной функции алгоритма.

Экстремальный бустинг.

Экстремальный бустинг (XGBoost) – один из самых распространенных алгоритмов ансамблевых методов, который основан на методе градиентного спуска [24]. Данный алгоритм имеет несколько отличий от предыдущих алгоритмов и учитывает отличие прогнозируемых значений от фактических. Целью алгоритма является минимизация погрешности с помощью градиентного спуска. Экстремальный бустинг можно рассматривать как модель, состоящую из Q деревьев решений. Предположим, что существует следующий набор выборок:

$$S = \{(x_i, y_i) | i = 1, 2, \dots, N\}$$

где x_i – исходные переменные, y_i – выходные переменные, i – номер элемента выборки.

Поскольку принцип алгоритма экстремального бустинга заключается в минимизации погрешности, необходимо определить целевую функцию, или функцию потерь:

$$L(y_i, \gamma) = \frac{1}{2} (y_i - \gamma)^2, \quad i = (1, 2, \dots, n)$$

где $L(y_i, \gamma)$ – функция потерь, определяющая отклонение прогнозируемых γ от фактических y_i значений, i – номер элемента выборки.

В начале процесса обучения необходимо определить начальное прогнозируемое значение следующим выражением:

$$\hat{h}_0(x) = \operatorname{argmin}_\gamma \sum_{i=1}^N L(y_i, \gamma)$$

где $\hat{h}_0(x)$ – начальное прогнозируемое значение.

В этом случае начальное значение $\hat{h}_0(x)$ равно среднему значению всех фактических y_i значений. То есть с этого значения начинается итерационный процесс в следующей последовательности.

1. Вычисление остатков:

$$r_{iq} = - \left[\frac{\partial L(y_i, \hat{h}(x_i))}{\partial \hat{h}(x_i)} \right]_{\hat{h}(x) = \hat{h}_{q-1}(x)}, \quad q = (1, 2, \dots, Q)$$

где q – номер дерева, r_{iq} – остатка i -ого элемента выборки на q -м дереве, y_i – фактическое значение i -ого элемента выборки, $\hat{h}(x_i)$ – прогнозируемое значение i -ого элемента выборки.

2. Повторное обучение на основе остатков r_{iq} .

3. Определение значений в конечных узлах деревьев решений:

$$\gamma_{jq} = \operatorname{argmin}_\gamma \sum_{x_i \in R_{ij}} L(y_i, \hat{h}_{q-1}(x_i) + \gamma), \quad j = (1, 2, \dots, J)$$

где j – номер конечного узла, γ_{jq} – значение j -го узла на q -м дереве.

4. Обновление прогнозируемых значений:

$$\hat{h}_q(x) = \hat{h}_{q-1}(x) + \mu \sum_{j=1}^J \gamma_{jq} I(x \in R_{jq}), \quad \mu \in [0, 1]$$

где $\hat{h}_q(x)$ – прогнозируемое значение q -ого дерева, $\hat{h}_{q-1}(x)$ – прогнозируемое значение $(q - 1)$ -го дерева, μ – коэффициент уровня обучения.

В каждом случае, когда добавляется дерево решений, алгоритм обучает новую функцию и определяет ее коэффициенты так, чтобы минимизировать погрешность предыдущих итераций. После итерации Q , алгоритм экстремального бустинга объединяет результаты всех итераций $\hat{h}_q(x)$ путем построения выходной модели \hat{Y}' .

$$\hat{Y}' = \frac{1}{q} \sum_{q=1}^Q L(y, \hat{h}_{q-1}(x)) + \hat{h}_q(x)$$

Ансамблевые алгоритмы, основанные на деревьях решений, способны обрабатывать значительный объем данных и не требуют сложных процедур оптимизации в отличие от нейронных сетей и метода опорных векторов [15].

Таким образом, модель случайного леса состоит из регрессоров, построенных независимо друг от друга. Алгоритмы адаптивного бустинга и экстремального бустинга отличаются от случайного леса тем, что состоят из регрессоров, построенных с учетом погрешности предыдущих регрессоров. Агрегирование результатов путем усреднения выходных данных в конечном итоге дает модель ансамблевого алгоритма.

Для построения модели с достаточной точностью с помощью данных алгоритмов необходимо указать число деревьев решений и их глубину. Чем больше деревьев решений, тем выше точность, но время построения случайного леса также пропорционально увеличивается. Точность модели также зависит от глубины деревьев решений. Несмотря на то, что увеличение глубины повышает качество ансамбля как на обучении, так и на тестировании, чем меньше глубина деревьев, тем быстрее выполняется данный алгоритм. Отсюда следует необходимость поиска оптимального количества деревьев решений и их глубины.

Зависимость выработки от скорости ветра.

По графику зависимости генерации от скорости ветра, представленному на рисунке 2, возможно проанализировать характер выработки ветровой турбины [25].

Данный график имеет три основные точки: минимальную скорость V_{cut-in} , номинальную скорость ветра V_{rated} и критическую скорость $V_{cut-out}$. Эти три точки делят график на четыре сегмента, каждый из которых имеет различные зависимости генерации от скорости ветра. Если скорость ветра V ниже минимальной V_{cut-in} или выше критической $V_{cut-out}$, генерация равна нулю. В случае $V_{rated} < V < V_{cut-out}$ мощность турбины достигает номинального значения. На графике те точки, которые находятся между минимальной и номинальной скоростью, могут быть описаны следующим выражением:

$$P = \frac{1}{2} C_p \rho \pi R^2 V^3 \quad \text{при } V_{cut-in} < V < V_{rated}$$

где P – выработка турбины, C_p – коэффициент ветровой турбины, ρ – плотность воздуха, R – радиус ротора, V – скорость ветра.

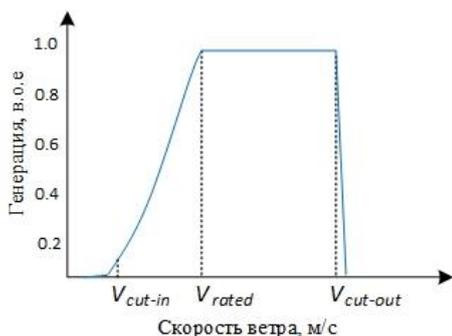


Рис. 2. График зависимости ветровой турбины

Fig. 2. Wind turbine power curve

*Источник: Составлено авторами Source: compiled by the author.

В данной статье рассматривается применение ансамблевых методов, описанных выше, для прогнозирования скорости ветра в трех районах, где построены ветровые электростанции с различной установленной мощностью. Также на основе полученных данных по скорости ветра были построены графики выработки ветровых электростанций.

Для оценки модели были использованы такие метрики, как средняя ошибка по модулю (англ. MAE – Mean Absolute Error) и нормализованная средняя ошибка по модулю в процентах (англ. nMAE – Normalized MAE).

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|,$$

$$nMAE = MAE / \frac{1}{N} \sum_{i=1}^N y_i * 100,$$

где N – размер выборки, y_i – фактическое значение i -ого элемента выборки, \hat{y}_i – прогнозируемое значение i -ого элемента выборки.

Результаты (Results)

Данные о скорости ветра и выработке регистрировались с часовым интервалом для трёх ветровых электростанций, расположенных в различных регионах Монголии, и были собраны за период с 2019 г по 2021 г. Весь набор данных был поделен на обучающую и тестовую выборки в отношении 70 на 30 процентов соответственно. В таблице 1 перечислены ветровые электростанции, которые функционируют в центральной энергосистеме Монголии.

Таблица 1

Table 1

Установленная мощность ветровой энергии
Installed capacity of wind energy

Наименование электростанций	Наименование в расчетах	Количество турбин	Установленная мощность, МВт
Салхит	ВЭС-1	31	50
Цэций	ВЭС-2	25	50
Сайншанд	ВЭС-3	25	55

*Источник: Составлено авторами Source: compiled by the author

В данной работе перечисленные выше ансамблевые алгоритмы были использованы для решения задачи прогнозирования скорости ветра. При построении модели важнейшей задачей является определение глубины деревьев решений, так как от этого существенно зависит достоверность модели R^2 . Зависимость достоверности модели R^2 от глубины деревьев решений (max_depth) показана на рисунке 3.

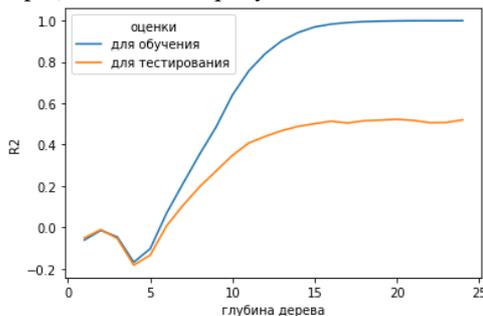


Рис. 3. Зависимость достоверности модели R^2 от глубины деревьев.

Fig. 3. Dependence of the reliability of the model R^2 on the depth of the trees.

*Источник: Составлено авторами Source: compiled by the author

Из рисунка 3 видно, что оптимальное значение глубины деревьев совпадает с точкой, в которой оценка обучения достигла своего максимального значения и где оценка тестирования стабилизируется. Количество деревьев было принято равным 100, поскольку увеличение количества деревьев свыше этого значения оказывает слабое влияние на качество модели. Таким образом, оптимальными значениями гиперпараметров являются 20 для глубины деревьев и 100 для их количества. Настройка параметров глубины и количества деревьев реализована с помощью функции GridSearchCV (библиотеки sklearn), которая подбирает наилучшие параметры модели. Оценки моделей приведены в таблице 3.

Таблица 3
Table 3

Достоверности моделей
Model validities

	Параметры		ВЭС-1		ВЭС-2		ВЭС-3	
	Глубина деревьев	Количество деревьев	MAE, м/с	nMAE, %	MAE, м/с	nMAE, %	MAE, м/с	nMAE, %
RF	20	100	1,33	15,81	2,19	17,25	1,44	17,71
XGB			0,68	8,16	0,70	8,66	0,74	9,14
ADA			0,63	7,49	0,67	8,46	0,68	8,40

*Источник: Составлено авторами Source: compiled by the author

Алгоритм обучается на временных рядах скорости ветра, полученных по почасовым данным за прошедшую неделю (wind-i). В дополнение к данным о скорости ветра рассчитывались данные о сезонности, включая названия месяцев (month), дней (day) и часов (hour). Это позволит определить уровень изменения скорости ветра по месяцам, то есть при прогнозировании скорости ветра будет учитываться особенность сезона. На рисунке 4 показана роль этих переменных при построении модели на основе вышеперечисленных алгоритмов и показан отрезок процесса тестирования модели на рисунке 5.

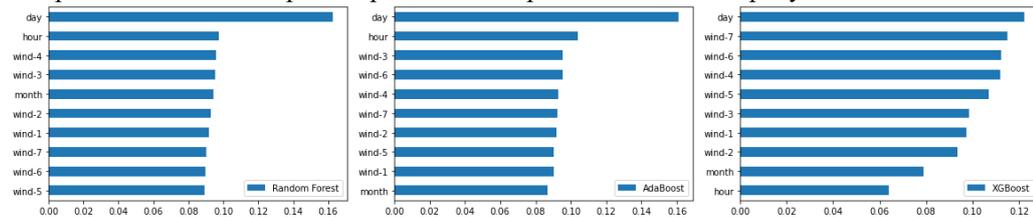


Рис. 4. Роль переменных при построении модели.

Fig. 4. The role of variables in building models.

*Источник: Составлено авторами Source: compiled by the author

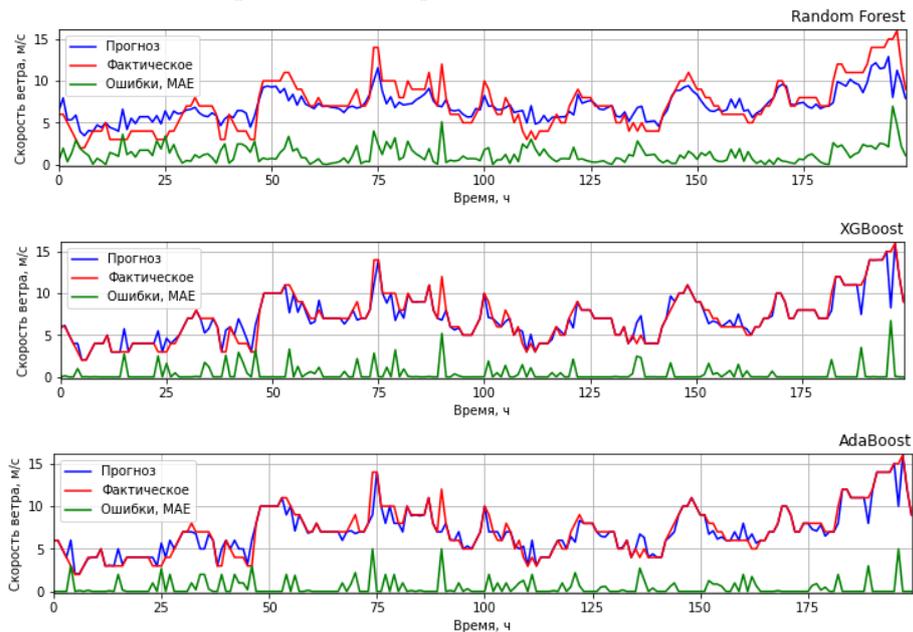


Рис. 5. Отрезок процесса тестирования модели.

Fig. 5. Samples of the testing process.

*Источник: Составлено авторами Source: compiled by the author

Из таблицы 3 и рисунка 5 видно, что модели, построенные по алгоритмам экстремального бустинга (XGBoost) и адаптивного бустинга (AdaBoost) имеют меньшую погрешность по сравнению с моделью случайного леса (Random Forest).

На основе данных о скорости ветра, спрогнозированных моделью адаптивного бустинга, был построен временной ряд выработки ветровых электростанций с помощью графика зависимости генерации от скорости ветра. На рисунке 6 показан график зависимости генерации от скорости ветра для ветровой электростанции «Сайншанд», оснащенной 25 турбинами типа Vestas V110-2.0.

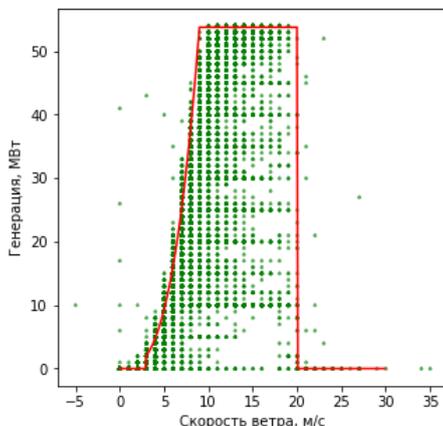


Рис. 6. Зависимость генерации от скорости ветра.

Fig. 6. Power curve of wind turbine.

*Источник: Составлено авторами Source: compiled by the author

В таблице 4 приведены результаты прогнозирования суточных графиков выработки ветровых электростанций, которые выбраны случайным образом.

Таблица 4

Результаты прогнозирования выработки ВЭС.

	ВЭС-1		ВЭС-2		ВЭС-3	
	MAE, МВт	nMAE, %	MAE, МВт	nMAE, %	MAE, МВт	nMAE, %
1	0,78	2,18	3,52	7,30	4,35	17,09
2	0,58	1,63	4,15	8,45	2,52	5,05
3	7,96	20,54	0,63	1,26	1,07	2,35
4	1,79	14,50	6,85	31,17	5,97	29,10
5	3,01	11,45	1,97	5,08	6,82	19,77
6	0,82	22,29	3,12	7,99	0,85	19,30
7	1,22	6,80	3,66	25,25	1,73	7,17
8	1,49	4,48	2,13	4,17	6,42	15,34
9	2,17	16,65	2,75	13,32	1,74	10,91
10	4,43	22,56	6,46	19,91	3,35	6,73
Средняя ошибка	2,42	12,30	3,52	12,39	3,48	13,28

*Источник: Составлено авторами Source: compiled by the author

Заключение (Conclusions)

В данной работе было рассмотрено применение метода машинного обучения для прогнозирования почасового графика выработки ветровых электростанций. Из-за стохастических и резко переменных характеристик изменения скорости ветра возникает необходимость применения более сложных подходов к решению задачи прогнозирования скорости ветра, а именно искусственного интеллекта и машинного обучения. В данном случае были применены ансамблевые методы машинного обучения, преимущество которых заключается в возможности определения нелинейных и нестационарных зависимостей во временных рядах. В работе были построены модели скорости ветра с применением алгоритмов случайного леса (Random Forest), адаптивного бустинга (AdaBoost) и экстремального бустинга (XGBoost). В результате экспериментов наилучшее качество показал алгоритм адаптивного бустинга: средняя ошибка по модулю данного алгоритма не превышает 1,0 м/с.

С помощью графика зависимости выработки от скорости ветра и полученных данных о скорости ветра были спрогнозированы суточные почасовые графики генерации для трех ветровых электростанций, которые расположены в различных регионах Монголии. Средняя

ошибка по модулю прогнозирования (MAE) составила от 2,42 до 3,48 МВт; нормализованная средняя ошибка по модулю в процентах (nMAE) составила от 12,3 до 13,3 %. Таким образом, ошибка прогнозирования составляет от 5,0 до 7,0 % установленной мощности соответствующих ветровых электростанций.

Можно сделать вывод, что данный метод машинного обучения может быть применен в решении задачи прогнозирования суточного графика выработки ветроэлектростанций на практике. Повышение точности прогнозирования выработки ветроустановок имеет высокую значимость для эффективного функционирования и планирования режимов энергосистемы.

Литература

1. Olaofe Z. O. A 5-day wind speed and power forecasts using a layer recurrent neural network (LRNN) // Sustainable Energy Technologies and Assessments. 2014. Т. 6. С. 1-24. doi: [10.1016/j.seta.2013.12.001](https://doi.org/10.1016/j.seta.2013.12.001)
2. Santos M., González M. Factors that influence the performance of wind farms // Renewable Energy. 2019. Т. 135. С. 643-651. doi: [10.1016/j.renene.2018.12.033](https://doi.org/10.1016/j.renene.2018.12.033)
3. Ouyang T., Kusiak A., He Y. Predictive model of yaw error in a wind turbine // Energy. 2017. Т. 123. С. 119-130. doi: [10.1016/j.energy.2017.01.150](https://doi.org/10.1016/j.energy.2017.01.150)
4. Matrenin P. V., Osgonbaatar T., Sergeev N. N. Overview of Renewable Energy Sources in Mongolia // 2022 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON). IEEE. 2022. С. 700-703. doi: [10.1109/SIBIRCON56155.2022.10016986](https://doi.org/10.1109/SIBIRCON56155.2022.10016986)
5. Zhao L, Muhammad S.N., Hafiz M.J., Ahmed N.A. A review on proliferation of artificial intelligence in wind energy forecasting and instrumentation management // Environmental Science and Pollution Research. 2022. №. 29. Т. 29. С. 43690-43709. doi: [10.1007/s11356-022-19902-8](https://doi.org/10.1007/s11356-022-19902-8)
6. Manusov V., Matrenin P., Nazarov M., Beryozkina S. Short-Term Prediction of the Wind Speed Based on a Learning Process Control Algorithm in Isolated Power Systems // Sustainability. 2023. №. 2. Т. 15. С. 1730. doi: [10.3390/su15021730](https://doi.org/10.3390/su15021730)
7. Tiwari S. Wind speed forecasting methods for wind energy generation // 2022 1st International Conference on Informatics (ICI). IEEE, 2022. С. 143-147. doi: [10.1109/ICI53355.2022.9786880](https://doi.org/10.1109/ICI53355.2022.9786880)
8. Weather forecasting for renewable energy system: a review / Meenal R. и др // Archives of Computational Methods in Engineering. 2022. №. 5. Т. 29. С. 2875-2891. doi: [10.1007/s11831-021-09695-3](https://doi.org/10.1007/s11831-021-09695-3)
9. Li X., Sabas J. F., Mendéz V. D. Wind energy forecasting using multiple ARIMA models // 2022 IEEE 18th International Conference on Automation Science and Engineering (CASE). IEEE, 2022. С. 2034-2039. doi: [10.1109/CASE49997.2022.9926516](https://doi.org/10.1109/CASE49997.2022.9926516)
10. A survey of artificial neural network in wind energy systems / Marugán A. P. и др // Applied energy. 2018. Т. 228. С. 1822-1836. doi: [10.1016/j.apenergy.2018.07.084](https://doi.org/10.1016/j.apenergy.2018.07.084)
11. Zhang Y., Li R. Short term wind energy prediction model based on data decomposition and optimized LSSVM // Sustainable Energy Technologies and Assessments. 2022. Т. 52. С. 102025. doi: [10.1016/j.seta.2022.102025](https://doi.org/10.1016/j.seta.2022.102025)
12. Regression model for predicting the speed of wind flows for energy needs based on fuzzy logic / Khasanzoda N. и др // Renewable Energy. 2022. Т. 191. С. 723-731. doi: [10.1016/j.renene.2022.04.017](https://doi.org/10.1016/j.renene.2022.04.017)
13. A new wind power prediction method based on chaotic theory and Bernstein Neural Network / Wang C. и др // Energy. 2016. Т. 117. С. 259-271. doi: [10.1016/j.energy.2016.10.041](https://doi.org/10.1016/j.energy.2016.10.041)
14. Buhan S., Özkazanç Y., Çadırcı I. Wind pattern recognition and reference wind mast data correlations with NWP for improved wind-electric power forecasts // IEEE Transactions on Industrial Informatics. 2016. №. 3. Т. 12. С. 991-1004. doi: [10.1109/TII.2016.2543004](https://doi.org/10.1109/TII.2016.2543004)
15. Negative correlation learning-based RELM ensemble model integrated with OVMD for multi-step ahead wind speed forecasting / Peng T. и др. // Renewable Energy. 2020. Т. 156. С. 804-819. doi: [10.1016/j.renene.2020.03.168](https://doi.org/10.1016/j.renene.2020.03.168)
16. Smart wind speed forecasting approach using various boosting algorithms, big multi-step forecasting strategy / Li Y. и др. // Renewable energy. 2019. Т. 135. С. 540-553. doi: [10.1016/j.renene.2018.12.035](https://doi.org/10.1016/j.renene.2018.12.035)
17. Wang L., Guo Y., Li X. Wind speed prediction using measurements from neighboring locations and combining the extreme learning machine and the AdaBoost algorithm // Energy Reports. 2022. Т. 8. С. 1508-1518. doi: [10.1016/j.egyvr.2021.12.062](https://doi.org/10.1016/j.egyvr.2021.12.062)

18. Dosdoğru A. T., İpek A. B. Hybrid boosting algorithms and artificial neural network for wind speed prediction // *International Journal of Hydrogen Energy*. 2022. №. 3. Т. 47. С. 1449-1460. doi: [10.1016/j.ijepes.2021.107365](https://doi.org/10.1016/j.ijepes.2021.107365)
19. A multi-level model for hybrid short term wind forecasting based on SVM, wavelet transform and feature selection / Abedinia O. и др.// 2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe). IEEE, 2022. С. 1-6. doi: [10.1109/EEEIC/ICPSEurope54979.2022.9854519](https://doi.org/10.1109/EEEIC/ICPSEurope54979.2022.9854519)
20. Lahouar A., Slama J. B. H. Hour-ahead wind power forecast based on random forests // *Renewable energy*. 2017. Т. 109. С. 529-541. doi: [10.1016/j.renene.2017.03.064](https://doi.org/10.1016/j.renene.2017.03.064)
21. Breiman L. Random forests // *Machine learning*. 2001. Т. 45. С. 5-32. doi: [10.1023/a:1010933404324](https://doi.org/10.1023/a:1010933404324)
22. Jorgensen K. L., Shaker H. R. Wind power forecasting using machine learning: State of the art, trends and challenges // 2020 IEEE 8th International Conference on Smart Energy Grid Engineering (SEGE). IEEE, 2020. С. 44-50. doi: [10.1109/SEGE49949.2020.9181870](https://doi.org/10.1109/SEGE49949.2020.9181870)
23. Freund Y., Schapire R. E. A decision-theoretic generalization of on-line learning and an application to boosting // *Journal of computer and system sciences*. 1997. №. 1. Т. 55. С. 119-139. doi: [10.1006/jcss.1997.1504](https://doi.org/10.1006/jcss.1997.1504)
24. Phan Q. T., Wu Y. K., Phan Q. D. A hybrid wind power forecasting model with XGBoost, data preprocessing considering different NWP // *Applied Sciences*. 2021. №. 3. Т. 11. С. 1100. doi: [10.3390/app11031100](https://doi.org/10.3390/app11031100)
25. Ouyang T., Kusiak A., He Y. Modeling wind-turbine power curve: A data partitioning and mining approach // *Renewable Energy*. 2017. Т. 102. С. 1-8. doi: [10.1016/j.renene.2016.10.032](https://doi.org/10.1016/j.renene.2016.10.032)

Авторы публикации

Осгонбаатар Тувшин – аспирант, Новосибирский государственный технический университет. Email: o.tuvshin.21@gmail.com.

Русина Анастасия Георгиевна – д-р техн. наук, декан факультета энергетики, Новосибирский государственный технический университет. Email: anastasiarusina@gmail.com.

Матренин Павел Викторович – канд. техн. наук, доцент кафедры систем электроснабжения предприятий, Новосибирский государственный технический университет. Email: pavel.matrenin@gmail.com.

Сергеев Никита Николаевич – студент, младший научный сотрудник Межкафедральной научно-исследовательской лаборатории обработки, анализа и представления данных в электроэнергетических системах, Новосибирский государственный технический университет. Email: nikita.n.sergeev@gmail.com

References

1. Olaofe Z. O. A 5-day wind speed and power forecasts using a layer recurrent neural network (LRNN). *Sustainable Energy Technologies and Assessments*, 2014. vol. 6, pp.1-24. doi: [10.1016/j.seta.2013.12.001](https://doi.org/10.1016/j.seta.2013.12.001)
2. Santos M., González M. Factors that influence the performance of wind farms, *Renewable Energy*, 2019, vol.135, pp.643-651. doi: [10.1016/j.renene.2018.12.033](https://doi.org/10.1016/j.renene.2018.12.033)
3. Ouyang T., Kusiak A., He Y. Predictive model of yaw error in a wind turbine. *Energy*. 2017. vol.123. pp. 119-130. doi: [10.1016/j.energy.2017.01.150](https://doi.org/10.1016/j.energy.2017.01.150)
4. Matrenin P. V., Osgonbaatar T., Sergeev N. N. Overview of Renewable Energy Sources in Mongolia. 2022 *IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*. IEEE, 2022. pp. 700-703. doi: [10.1109/SIBIRCON56155.2022.10016986](https://doi.org/10.1109/SIBIRCON56155.2022.10016986)
5. Zhao L, Muhammad S.N., Hafiz M.J., Ahmed N.A. A review on proliferation of artificial intelligence in wind energy forecasting and instrumentation management. *Environmental Science and Pollution Research*. 2022. vol. 29. pp. 43690-43709. doi: [10.1007/s11356-022-19902-8](https://doi.org/10.1007/s11356-022-19902-8)

6. Manusov V., Matrenin P., Nazarov M., Beryozkina S. Short-Term Prediction of the Wind Speed Based on a Learning Process Control Algorithm in Isolated Power Systems. *Sustainability*. 2023. vol. 2(15). pp. 1730. doi: [10.3390/su15021730](https://doi.org/10.3390/su15021730)
7. Tiwari, Seemant. Wind speed forecasting methods for wind energy generation. *2022 1st International Conference on Informatics (ICI)*. IEEE. 2022. pp. 143-147. doi: [10.1109/ICI53355.2022.9786880](https://doi.org/10.1109/ICI53355.2022.9786880)
8. Meenal, R., et al. Weather forecasting for renewable energy system: a review. *Archives of Computational Methods in Engineering*. 2022. vol. 5(29). pp. 2875-2891. doi: [10.1007/s11831-021-09695-3](https://doi.org/10.1007/s11831-021-09695-3)
9. Li, Xiaou, Juan Francisco Sabas, and Vicente Duarte Mendéz. Wind energy forecasting using multiple ARIMA models. *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)*. IEEE, 2022. pp. 2034-2039. doi: [10.1109/CASE49997.2022.9926516](https://doi.org/10.1109/CASE49997.2022.9926516)
10. Marugán A. P. et al. A survey of artificial neural network in wind energy systems. *Applied energy* 228. 2018. pp.1822-1836. Doi: [10.1016/j.apenergy.2018.07.084](https://doi.org/10.1016/j.apenergy.2018.07.084)
11. Zhang Y., and Li R. Short term wind energy prediction model based on data decomposition and optimized LSSVM. *Sustainable Energy Technologies and Assessments*. 2022. vol.52. pp. 102025. doi: [10.1016/j.seta.2022.102025](https://doi.org/10.1016/j.seta.2022.102025)
12. Khasanzoda, Nasrullo, et al. Regression model for predicting the speed of wind flows for energy needs based on fuzzy logic. *Renewable Energy*. 2022. vol.191. pp. 723-731. doi: [10.1016/j.renene.2022.04.017](https://doi.org/10.1016/j.renene.2022.04.017)
13. Wang C, et.al. A new wind power prediction method based on chaotic theory and Bernstein Neural Network. *Energy*. 2016. vol. 117. pp. 259-271. doi: [10.1016/j.energy.2016.10.041](https://doi.org/10.1016/j.energy.2016.10.041)
14. Buhan S., Özkazanç Y., Çadircı I. Wind pattern recognition and reference wind mast data correlations with NWP for improved wind-electric power forecasts. *IEEE Transactions on Industrial Informatics*. 2016. vol. 3(12). pp. 991-1004. doi: [10.1109/TII.2016.2543004](https://doi.org/10.1109/TII.2016.2543004)
15. Peng T., et.al. Negative correlation learning-based RELM ensemble model integrated with OVMD for multi-step ahead wind speed forecasting. *Renewable Energy*. 2020. vol.156. pp. 804-819. doi: [10.1016/j.renene.2020.03.168](https://doi.org/10.1016/j.renene.2020.03.168)
16. Li Y., et.al. Smart wind speed forecasting approach using various boosting algorithms, big multi-step forecasting strategy. *Renewable energy*. 2019. vol.135. pp.540-553. doi: [10.1016/j.renene.2018.12.035](https://doi.org/10.1016/j.renene.2018.12.035)
17. Wang L., Guo Y., Li X. Wind speed prediction using measurements from neighboring locations and combining the extreme learning machine and the AdaBoost algorithm. *Energy Reports*. 2022. vol. 8. pp. 1508-1518. doi: [10.1016/j.egy.2021.12.062](https://doi.org/10.1016/j.egy.2021.12.062)
18. Dosdoğru A. T., İpek A. B. Hybrid boosting algorithms and artificial neural network for wind speed prediction. *International Journal of Hydrogen Energy*. 2022. no. 3. vol. 47. pp. 1449-1460. doi: [10.1016/j.ijepes.2021.107365](https://doi.org/10.1016/j.ijepes.2021.107365)
19. Abedinia O., et.al. A multi-level model for hybrid short term wind forecasting based on SVM, wavelet transform and feature selection. *2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*. IEEE, 2022. pp. 1-6. doi: [10.1109/EEEIC/ICPSEurope54979.2022.9854519](https://doi.org/10.1109/EEEIC/ICPSEurope54979.2022.9854519)
20. Lahouar A., Slama J. B. H. Hour-ahead wind power forecast based on random forests. *Renewable energy*. 2017. vol. 109. pp. 529-541. doi: [10.1016/j.renene.2017.03.064](https://doi.org/10.1016/j.renene.2017.03.064)
21. Breiman, Leo. Random forests. *Machine learning*. 2001. vol.45. pp. 5-32. doi: [10.1023/a:1010933404324](https://doi.org/10.1023/a:1010933404324)
22. Jorgensen K. L., Shaker H. R. Wind power forecasting using machine learning: State of the art, trends and challenges. *2020 IEEE 8th International Conference on Smart Energy Grid Engineering (SEGE)*. IEEE, 2020. pp. 44-50. doi: [10.1109/SEGE49949.2020.9181870](https://doi.org/10.1109/SEGE49949.2020.9181870)
23. Freund Y., Schapire R. E. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*. 1997. no. 1. vol. 55. pp. 119-139. doi: [10.1006/jcss.1997.1504](https://doi.org/10.1006/jcss.1997.1504)
24. Phan Q. T., Wu Y. K., Phan Q. D. A hybrid wind power forecasting model with XGBoost, data preprocessing considering different NWPs. *Applied Sciences*. 2021. no. 3. vol. 11. pp. 1100. doi: [10.3390/app11031100](https://doi.org/10.3390/app11031100)
25. Ouyang T., Kusiak A., He Y. Modeling wind-turbine power curve: A data partitioning and mining approach. *Renewable Energy*. 2017. vol. 102. pp. 1-8. doi: [10.1016/j.renene.2016.10.032](https://doi.org/10.1016/j.renene.2016.10.032)

Authors of the publication

Osgonbaatar Tuvshin – Novosibirsk State Technical University, Novosibirsk, Russia. Email: o.tuvshin.21@gmail.com.

Anastasia G. Rusina – Novosibirsk State Technical University, Novosibirsk, Russia. Email: anastasiarusina@gmail.com.

Pavel V. Matrenin – Novosibirsk State Technical University, Novosibirsk, Russia. Email: pavel.matrenin@gmail.com.

Nikita N. Sergeev – Novosibirsk State Technical University, Novosibirsk, Russia. Email: nikita.n.sergeev@gmail.com

Шифр научной специальности: 2.4.3 Электроэнергетика.

Получено ***20.10.2023 г.***

Отредактировано ***09.11.2023 г.***

Принято ***07.02.2024 г.***