УДК 621.311.243 DOI: 10.18324/2077-5415-2026-3-66-73

Современные подходы к прогнозированию солнечной радиации и выбору мест размещения фотоэлектрических станций: от регрессионных моделей к многокритериальному анализу

И.Н. Погодаева[a], Н.В. Савина[b]

Братский государственный университет, ул. Макаренко, 40, Братск, Россия
[a] m.inessa13@yandex.ru, [b] nickolai.savina@yandex.ru.
[a] https://orcid.org/0000-0003-3176-2422, [b] https://orcid.org/ 0009-0005-6316-8446.
Статья поступила 04.09.2026, принята 23.09.2026.

В статье проанализированы современные подходы к прогнозированию солнечной радиации и выбору мест размещения фотоэлектрических станций с учётом климатических особенностей и нормативной базы Российской Федерации. Выявлено, что регрессионные модели машинного обучения (линейная регрессия, метод опорных векторов) обеспечивают повышение точности прогнозов на 15–20 % по сравнению с физическими моделями, что подтверждается зарубежными исследованиями 2024–2025 гг. Определены ключевые климатические предикторы (температура, влажность, облачность, давление) на основе суточных данных метеостанции г. Иркутска за период 2007–2025 гг., разделённых на обучающую (2007–2015) и тестовую (2016–2025) выборки. Сформированы и валидированы 17 регрессионных моделей – 10 известных и 7 новых, включая 3 модели на основе SVR; последние показали минимальную среднюю абсолютную ошибку (463,2 Вт · ч/м²·сут) против 627,52 у известных моделей. Разработано приложение на Python, автоматизирующее полный цикл регрессионного анализа: загрузку данных, интерактивный ввод уравнения, обучение, расчёт метрик (R^2, RMSE, MAE, MBE), визуализацию и экспорт результатов в Excel с детализированными отчётами. Обоснована необходимость интеграции полученных прогностических моделей с методами многокритериального принятия решений (MCDM) в единой среде Python, что позволяет лицу, принимающему решение, проводить анализ чувствительности и получать воспроизводимые результаты без переключения между разными программами.

Ключевые слова: солнечная радиация; регрессионные модели; метод опорных векторов; машинное обучение; Python; многокритериальный анализ; валидация, прогнозирование.

Modern approaches to forecasting solar radiation and selecting sites for photovoltaic plants: from regression models to multi-criteria analysis

I.N. Pogodaeva[a], N.V. Savina[b]

Bratsk State University; 40, Makarenko St., Bratsk, Russia
[a] m.inessa13@yandex.ru, [b] nickolai.savina@yandex.ru.
[a] https://orcid.org/0000-0003-3176-2422, [b] https://orcid.org/ 0009-0005-6316-8446.
Received 04.09.2026, accepted 23.09.2026.

The paper analyzes modern approaches to forecasting solar radiation and siting photovoltaic plants, accounting for the climatic conditions and regulatory framework of the Russian Federation. It is revealed that machine learning regression models (linear regression, support vector regression) improve forecast accuracy by 15–20% compared to physical models, as confirmed by foreign studies of 2024–2025. Key climatic predictors (temperature, humidity, cloud cover, pressure) are determined from daily data from the Irkutsk weather station for 2007–2025, divided into training (2007–2015) and test (2016–2025) sets. Seventeen regression models were built and validated – 10 known and 7 new, including 3 SVR-based models; the latter showed the minimum mean absolute error (463.2 Wh/m²/day) versus 627.52 for the known models. A Python application was developed that automates the full cycle of regression analysis: data loading, interactive equation input, training, calculation of metrics (R^2, RMSE, MAE, MBE), visualization, and export of results to Excel with detailed reports. The need to integrate the obtained predictive models with multi-criteria decision-making (MCDM) methods in a unified Python environment is substantiated, which allows the decision-maker to conduct sensitivity analysis and obtain reproducible results without switching between different programs.

Keywords: solar radiation; regression models; support vector regression; machine learning; Python; multi-criteria analysis; validation; forecasting.

Введение. Развитие солнечной генерации в Российской Федерации является одним из стратегических направлений энергетической политики. Согласно прогнозам, совокупная установленная мощность возобновляемых источников энергии (ВИЭ) в стране возрастёт с 9,8 ГВт в 2025 году до 18,4 ГВт к 2035 году, причём доля солнечных электростанций (СЭС) оценивается в 5,3 ГВт [1]. Достижение этих показателей поддерживается системой государственных мер: отбор проектов ВИЭ, проводимый Минэнерго, а также программы поддержки генерации на основе ВИЭ на период 2025–2035 гг.,

которые задают чёткие ориентиры для инвесторов и разработчиков.

Важным этапом нормативно-правового регулирования является утверждение ГОСТ Р 70787-2023 «Единая энергетическая система и изолированно работающие энергосистемы. Возобновляемые источники энергии. Технические требования к фотоэлектрическим солнечным станциям». Данный стандарт устанавливает технические требования к СЭС как для работы в Единой энергосистеме, так и в изолированных территориальных энергосистемах, что особенно значимо для удалённых и труднодоступных регионов.

Одной из ключевых особенностей России является её разнородность по климатическим условиям. Особый интерес для солнечной энергетики представляют районы Крайнего Севера – здесь, несмотря на суровый климат, в летние месяцы наблюдается полярный день, обеспечивающий круглосуточное поступление солнечной радиации (СР) в течение нескольких недель. Это создаёт уникальный потенциал для сезонной генерации, особенно для гибридных энергокомплексов (ГЭК), которые позволяют снизить зависимость от дорогостоящего завозного топлива и повысить надёжность энергоснабжения арктических и северных территорий.

Однако реализация этого потенциала невозможна без достоверной оценки энергетического ресурса. Ключевым критерием при выборе мест размещения СЭС остаётся уровень суммарной СР, который в условиях Крайнего Севера сильно варьирует в зависимости от сезона, облачности, снежного покрова и угла падения лучей. Поэтому первоочередной задачей выступает точное прогнозирование характеристик для конкретной местности.

Прежде чем применять методы многокритериального анализа для выбора оптимальных площадок, необходимо количественно оценить доступный солнечный ресурс. В международных исследованиях последних лет именно регрессионные модели машинного обучения выступают базовым инструментом для такого прогнозирования.

Так в исследовании [2] показано, что алгоритмы из библиотеки Scikit-Learn (Random Forest, XGBoost) существенно повышают точность прогнозов по сравнению с математической моделью CMA-WSP v2.0. Машинное обучение эффективно корректирует систематические ошибки базовых моделей, что подтверждается снижением среднеквадратичной ошибки RMSE на 15–20 %.

В исследовании [3] при оценке СР в Томске применялась LASSO-регрессия. Авторы сделали акцент на интерпретируемости модели и учёте климатической изменчивости; было показано, что LASSO позволяет выделять наиболее значимые предикторы (температуру, влажность, облачность) даже при ограниченных выборках, а значение коэффициента детерминации R^2 достигло 0,89.

В работе [4] сопоставлена многомерная гребневая регрессия (MRR) и многомерная LASSO-регрессия (MLR) для почасового прогнозирования. Оба метода показали сопоставимую точность (MAE ~ 45 Вт/м²) и были признаны пригодными для оперативного использования в условиях неполных метеоданных.

В исследовании [5] авторы из Египта провели всесторонний анализ на основе 12-летнего набора данных с 56 станций, применив автоматизированные алгоритмы машинного обучения (Random Forest, XGBoost). Результаты подтвердили, что комплексный подход к предобработке данных и выбору гиперпараметров позволяет строить высокоточные прогнозы для различных климатических зон ($R^2 > 0,92$).

Исследование [6] в Омане использовало фреймворк SARIMAX для долгосрочного прогнозирования, продемонстрировав, что временные ряды с учётом сезонности дают надёжные оценки для планирования мощностей СЭС (средняя абсолютная процентная ошибка MAPE менее 8 %).

Таким образом, регрессионные модели являются неотъемлемым начальным этапом анализа: они превращают метеорологические данные в количественные оценки потенциала СР, которые затем поступают на вход многокритериальных процедур.

Методы исследования. Установление количественной зависимости между потоком суммарной СР и совокупностью метеорологических, географических и геометрических факторов в исследуемом районе требует последовательного выполнения нескольких этапов

Первый этап посвящается анализу природно-климатических условий территории. Прежде всего необходимо определить тип климата и выявить его характерные закономерности – сезонные колебания температуры и режим выпадения осадков. Важно учитывать, что локальный климатический режим может модифицироваться под влиянием таких факторов, как близость крупных водоёмов, горный рельеф, антропогенные объекты и пр. Понимание этих особенностей позволяет обоснованно решить, требуется ли построение отдельных регрессионных моделей для разных сезонов года.

На этом же этапе выполняется сбор метеорологических данных из открытых источников. Основным хранилищем информации о СР служит Мировой центр радиационных данных (МЦРД) [7], который предоставляет суточные данные с 1964 г. по настоящее время по следующим показателям: суммарная и рассеянная радиация, радиационный баланс подстилающей поверхности, продолжительность солнечного сияния. Однако в ряде пунктов наблюдения встречаются пропуски – от одного дня до нескольких лет.

Многолетние архивы по температуре, влажности, облачности и атмосферному давлению доступны на сайте «Расписание погоды» [8], где представлены фактические данные наземных метеостанций в форматах SYNOP и METAR. В табл. 1 перечислены параметры, наиболее востребованные при построении регрессионных моделей оценки СР.

Таблица 1. Параметры, используемые для построения регрессионных моделей

Параметр	Единица измерения	Обозначение
Общая облачность	балл	C_t
Нижняя облачность	балл	C_l
Температура	°C	T
Относительная влажность	%	RH
Давление	мм рт. ст	P

Следует отметить, что не все метеостанции оснащены оборудованием для регистрации полного набора параметров. Например, продолжительность солнечного сияния, являющаяся одним из ключевых предикторов, измеряется далеко не везде. Это обстоятельство повышает актуальность поиска альтернативных комбинаций доступных параметров, способных обеспечить приемлемую точность модели.

Второй этап – формирование набора регрессионных моделей и их калибровка. Модели могут описывать как линейные, так и нелинейные связи между предикторами и потоком СР. На практике достаточно высокую точность часто даёт линейная модель вида [9]:

$$\frac{H}{H_0} = \sum_{i=1}^{n} k_i \cdot x_i, \quad (1)$$

где H – поток суммарной СР; H_0 – поток внеатмосферной СР; k_i – параметры модели; n – количество параметров; x_i – коэффициенты линейной регрессии.

Поток внеатмосферной СР определяется по формуле (2):

$$H_0 = \frac{12 \cdot 360 \cdot G_{SC}}{\pi} \cdot \left[1 + 0,033 \cos\left(\frac{360 \cdot N}{365}\right)\right] \cdot \left\{\cos(L) \cdot \cos(\delta) \cdot \left[\sin(h_2) - \sin(h_1)\right] + \left[\frac{\pi \cdot (h_2 - h_1)}{180}\right] \cdot \frac{1}{2} \sin(L) \cdot \sin(\delta)\right\}, \quad (2)$$

где G_{sc} – солнечная постоянная, равная 1367 Вт/м^2; N – порядковый номер дня от 1 до 365; L – широта местности, рад, δ – угол склонения солнца, рад; h_2 и h_1 – часовые углы солнца, соответствующие интервалу оценки, рад.

Набор моделей может быть составлен на основе ранее опубликованных работ [10], а также дополнен новыми комбинациями известных зависимостей или включением дополнительных параметров. Калибровка коэффициентов линейной регрессии выполняется по большей части данных (до 80 %) наземных измерений.

Для этого могут использоваться средства автоматизации Microsoft Excel (надстройка «Поиск решений») или специализированные пакеты.

Третий этап – оценка точности полученных моделей на данных, зарезервированных для валидации (оставшиеся 20 %). Качество моделей оценивается с помощью статистических показателей, среди которых наиболее распространены средняя абсолютная ошибка (MAE) и средняя ошибка смещения (MBE) [11]:

$$MAE = \frac{1}{t} \cdot \sum_{i=1}^{t} \left| H_i - H_m \right|, \quad (4)$$

где t – количество измерений СР, по которым проводится сравнение; H_i – расчётное значение потока СР, H_m – измеренное значение потока СР.

Чем ниже значение MAE, тем выше точность модели [12].

MBE является индикатором среднего отклонения расчётных значений от измеренных и описывает систематическую ошибку или погрешность:

$$MBE = \frac{1}{t} \cdot \sum_{i=1}^{t} \left(H_i - H_m \right) \cdot \quad (5)$$

Положительные значения MBE указывают на завышение моделью оценок СР, а отрицательные – на их занижение оценок [12].

Метод опорных векторов (SVR) зарекомендовал себя как эффективный инструмент для оценки СР, часто превосходящий другие алгоритмы машинного обучения [13, 14]. Рассмотрим его ключевые особенности применительно к нашей задаче.

SVR относится к моделям обучения с учителем и используется как для классификации, так и для регрессии. В процессе обучения каждой точке данных присваивается вес, отражающий её значимость для определения решающей границы. Точки, лежащие на границе и оказывающие наибольшее влияние, называются опорными векторами. Прогноз для новой точки строится на основе взвешенного расстояния до этих опорных векторов [15].

Важным преимуществом SVR является использование ядер – функций, позволяющих проецировать исходные данные в пространство более высокой размерности, где линейные зависимости становятся применимы даже для нелинейных исходных соотношений [16]. Применяются полиномиальные, гауссовы (RBF), сигмоидные и линейные ядра. Благодаря «kernel trick» обучение выполняется без явного построения полного N-мерного представления, что существенно экономит вычислительные ресурсы [16].

В данном исследовании построение SVR-моделей выполнено на Python с использованием библиотеки scikit-learn. Выбрано гауссово ядро с радиальной базовой функцией (RBF), для которой расстояние между точками x_i и x_j определяется как:

$$k_{rbf}(x_i, x_j) = e^{-\gamma \left\| x_i - x_j \right\|^2}. \quad (6)$$

где x_i, x_j – точки данных; $\left\| x_i - x_j \right\|$ – евклидово расстояние между точками; γ – параметр, регулирующий ширину гауссового ядра.

Настройка регрессионных моделей осуществлялась выбором двух параметров – C и γ.

Настройка моделей сводится к подбору двух гиперпараметров: C (регуляризация).и γ. Параметр γ определяет радиус влияния отдельной точки: малые значения γ соответствуют «гладким» моделям низкой сложности, большие – более изогнутым границам. Параметр C устанавливает баланс между допустимой ошибкой и сложностью модели: увеличение C повышает влияние каждой точки на форму решающей поверхности. Правильный выбор C и γ критически важен для достижения высокой точности регрессии.

Разработка регрессионных моделей. В качестве объекта исследования для построения регрессионных моделей выбран г. Иркутск.

Выбор обусловлен наличием репрезентативных метеорологических данных, а также актуальностью развития солнечной энергетики в восточной части России, где точная оценка радиационного потенциала является ключевым условием для обоснованного размещения СЭС.

Для построения и верификации моделей используются метеорологические данные наземных метеостанций

г. Иркутска, включающие климатические параметры, представленные в табл. 1.

Выбор именно этих параметров обусловлен их доступностью в открытых базах данных наземных измерений, а также доказанной значимостью для оценки потока суммарной СР. В исследовании участвуют как известные регрессионные модели, включающие перечисленные параметры, так и новые модели, дополнительно учитывающие высоту солнца для повышения точности оценок.

Временной горизонт исследования разделяется на два периода:

– период обучения (калибровки) моделей: 2007–2015 гг.;

– период прогнозирования (валидации): 2016–2025 гг. (для доступных данных – валидация, для последующих лет – экстраполяция).

Валидация на данных 2016–2025 гг. позволяет оценить качество моделей на независимом временном интервале. Для верификации используются открытые базы данных наземных измерений метеостанций.

Программная реализация регрессионных моделей на Python. Разработка моделей выполняется в среде Python. Выбор данного языка обусловлен наличием развитой экосистемы библиотек для научных вычислений и машинного обучения. Весь вычислительный конвейер включает следующие этапы:

– загрузка и предобработка данных с использованием pandas и numpy;

– реализация регрессионных моделей с применением scikit-learn;

– обучение и валидация моделей с разделением на обучающую (2007–2015) и тестовую (2016–2025) выборки;

– оценка точности с расчётом метрик качества (MAE, RMSE, MBE, R^2).

В работе калибруются и валидируются 17 моделей. Формирование набора моделей обусловлено исследованием, проводимым авторами ранее (табл. 2) [10].

Таблица 2. Выбранные регрессионные модели оценки уровня суммарной СР

Модель	Обозначение	Формула	Переменные
Модель 1	M1	$H/H_0=x_1+x_2\cdot RH$	H_0, RH
Модель 2	M2	$H/H_0=x_1+x_2\cdot T_{max}+x_3\cdot T_{min}+x_4\cdot (T_{max}-T_{min})+x_5\cdot RH$	H_0, T_{max}, T_{min}, RH
Модель 3	M3	$H/H_0=x_1+x_2\cdot(T_{max}-T_{min})$	H_0, T_{max}, T_{min}
Модель 4	M4	$H/H_0=x_1+x_2\cdot(T_{max}-T_{min})^{0,5}$	H_0, T_{max}, T_{min}
Модель 5	M5	$H/H_0=x_1+x_2\cdot(T_{max}/T_{min})+x_3\cdot(T_{max}/T_{min})^2$	H_0, T_{max}, T_{min}
Модель 6	M6	$H/H_0=x_1+x_2\cdot T_{max}+x_3\cdot T_{min}$	H_0, T_{max}, T_{min}
Модель 7	M7	$H/H_0=x_1+x_2\cdot T_{max}+x_3\cdot T_{min}+x_4\cdot P$	H_0, T_{max}, T_{min},P
Модель 8	M8	$H/H_0=x_1+x_2\cdot C_t$	H_0, C_t
Модель 9	M9	$H/H_0=x_1+ x_2\cdot C_t+ x_3\cdot C_t^2$	H_0, C_t
Модель 10	M10	$H/H_0=x_1+ x_2\cdot(T_{max}-T_{min})^{0,5}+ x_3\cdot C_t$	H_0, T_{max}, T_{min}, C_t
Модель 11	M11	$H/H_0=x_1+x_2\cdot P+x_3\cdot RH+x_4\cdot T+x_5\cdot C_l+$ $+x_6\cdot C_t+x_7\cdot Sin\alpha$	H_0, P, RH, T, C_l, C, $sin\alpha$
Модель 12	M12	$H/H_0=x_1+x_2\cdot(T_{max}T_{min})+x_3\cdot C_t+x_4\cdot C_l$	H_0, T_{max}, T_{min}, C_l, C_t
Модель 13	M13	$H/H_0=x_1+x_2\cdot(T_{max}/T_{min})+x_3\cdot P+x_4\cdot C_t+x_5\cdot C_l$	H_0, T_{max}, T_{min}, C_l, C_t, P
Модель 14	M14	$H/H_0=x_1+x_2\cdot(T_{max}+T_{min})+x_3\cdot RH+x_4\cdot C_t+x_5\cdot C_l$	H_0, T_{max}, T_{min}, C_l, C_t, RH
Модель 15	M15	C=5000 $gamma$=1	H_0, C_l, C_t, $sin\alpha$
Модель 16	M16	C=5000; $gamma$=1	H_0, C_l, C_t, $sin\alpha$, RH,
Модель 17	M17	C=5000; $gamma$=1	H_0, C_l, C_t, $sin\alpha$, RH, T_{max}, T_{min}

Для автоматизации процессов построения, обучения и оценки регрессионных моделей разработано специализированное приложение на языке Python. Программа предоставляет удобный интерфейс для работы с климатическими данными и позволяет лицу, принимающему решение (ЛПР), оперативно получать количественные оценки СР на основе вводимых параметров.

Для сравнения точности для каждой модели рассчитываются следующие показатели:

– средняя абсолютная ошибка (MAE) – характеризует величину ошибки в единицах измеряемой величины;

– среднеквадратичная ошибка (RMSE) – чувствительна к крупным выбросам;

– средняя ошибка смещения (MBE) – позволяет выявить систематическое завышение или занижение прогнозов.

На основе анализа средней ошибки смещения определены модели, показавшие наибольшую точность при оценке месячных и годовых сумм суммарной СР: известная регрессионная модель, использующая данные о влажности воздуха, а также регрессионная модель на основе метода опорных векторов.

На первом этапе программа осуществляет загрузку исходных данных из внешнего файла. Поддерживаются форматы CSV (с различными разделителями) и Excel. Пользователь указывает путь к файлу, после чего программа автоматически определяет структуру таблицы, выводит список доступных столбцов и предлагает выбрать:

– целевую переменную (зависимую) – суммарную СР (Вт · ч/м² · сут);

– набор предикторов (независимых переменных) – из предложенного перечня (влажность, температура, давление, облачность и др.).

Программа выполняет первичную проверку данных: удаляет строки с пропусками (или предлагает методы интерполяции), проверяет выбросы с использованием

межквартильного размаха и выводит пользователю сводную статистику. По желанию пользователь может задать временной диапазон для разделения на обучающую и тестовую выборки – в нашем исследовании это 2007–2015 для обучения и 2016–2025 для прогнозирования.

Пользователь может выбрать одну или несколько регрессионных моделей из реализованного набора. Программа включает:

– известные классические модели (линейная регрессия, гребневая регрессия, LASSO, ElasticNet, регрессия на основе деревьев решений);

– модели машинного обучения из scikit-learn (Random Forest, Gradient Boosting, SVR);

– новые предложенные модели, в том числе с использованием метода опорных векторов.

Для каждой модели предусмотрена возможность настройки гиперпараметров через графический интерфейс или конфигурационный файл. По умолчанию используются значения, оптимально подобранные на основе кросс-валидации по обучающей выборке.

После выбора модели и настройки параметров программа:

– разделяет данные на обучающий и тестовый наборы в соответствии с заданными годами;

– обучает модель на периоде 2007–2015;

– выполняет прогнозирование для периода 2016–2025 (для доступных годов – валидация, для последующих – экстраполяция);

– рассчитывает метрики качества: MAE, RMSE, MBE, а также коэффициент детерминации R^2.

Все метрики выводятся в виде таблицы, позволяя сравнить эффективность различных моделей. Для моделей, поддерживающих оценку важности признаков, программа строит столбчатую диаграмму вклада каждого климатического параметра.

При работе с программой в среде Python результаты валидации на данных 2016–2025 гг. показали, что модели M11, M12, M13, M14 обеспечивают более высокую точность:

– минимальная MAE для моделей M1-M10 – 637,33 Вт · ч/м² · сут;

– для регрессионных моделей (M11, M12, M13, M14) – 506,4 Вт · ч/м² · сут;

– для моделей на основе метода опорных векторов (M15, M16, M17) – 460,4 Вт · ч/м² · сут.

Наилучший результат достигнут для SVR-моделей.

Для визуализации полученных данных программа предоставляет интерактивные графики:

– сопоставление фактических и предсказанных значений СР на тестовом периоде (точечный график с линией идеального совпадения);

– временные ряды фактических и прогнозных значений по месяцам/годам;

– график остатков – позволяет визуально оценить наличие систематических ошибок и гетероскедастичность;

– при наличии географической привязки – пространственное распределение ошибок.

Результат работы программы представлен на рис. 1.

Графики строятся с использованием matplotlib, seaborn и plotly (интерактивные элементы).

Результат работы программы представлен на рис. 1

По окончании расчётов программа генерирует структурированный отчёт в формате PDF или HTML, который включает:

– описание использованных данных и выбранных моделей;

– численные значения метрик качества для каждой модели;

– все построенные графики;

– выводы о наиболее точной модели и её прогнозных способностях.

Отчёт может быть сохранён локально и использован для дальнейшего обоснования выбора площадок под СЭС. Таким образом, разработанное программное обеспечение полностью покрывает этап от загрузки данных до получения интерпретируемых результатов, снижая трудоёмкость расчётов и повышая их воспроизводимость.

Многокритериальный анализ при выборе мест размещения СЭС. Получив достоверный прогноз СР для каждого потенциального участка, исследователи сталкиваются со следующей задачей: как среди множества конкурирующих критериев (экономических, экологических, социальных, технических) выбрать наилучшую площадку. Именно здесь применимы методы многокритериального принятия решений (MCDM). Современные исследования в этой области выделяют несколько ключевых сложностей: многокритериальность (до 47 критериев, как показано в работе [17] и неоднозначность результатов разных MCDM-методов (средняя корреляция между TOPSIS, VIKOR, SAW, COPRAS, PROMETHEE составляет всего 0,11).

В связи с этим в исследовании [18] предпочтение отдаётся гибридным системам (AHP+TOPSIS, AHP+CRITIC, CRITIC+EDAS), а также объективным методам взвешивания без привлечения экспертов (CRITIC).

Основные критерии выбора СЭС варьируются от 4 до 15 и включают СР, стабильность энергоснабжения, погодоустойчивость, экологические и социальные факторы.

Именно из-за высокой сложности, неопределённости и больших объёмов данных традиционные ручные расчёты в Excel или специализированных MCDM-программах становятся чрезмерно трудоёмкими. Поэтому в современных исследованиях всё чаще применяют Python для автоматизации всех этапов: от сбора и предобработки растровых и векторных данных (библиотеки rasterio, geopandas) до расчёта объективных весов (CRITIC) и ранжирования альтернатив (TOPSIS, EDAS, CoCoSo), а также визуализации результатов в виде интерактивных карт (folium, plotly). Например, в работе [Ramos-Diez et al., 2026] реализована модульная структура, позволяющая ЛПР получить карту пригодности местности одной командой.

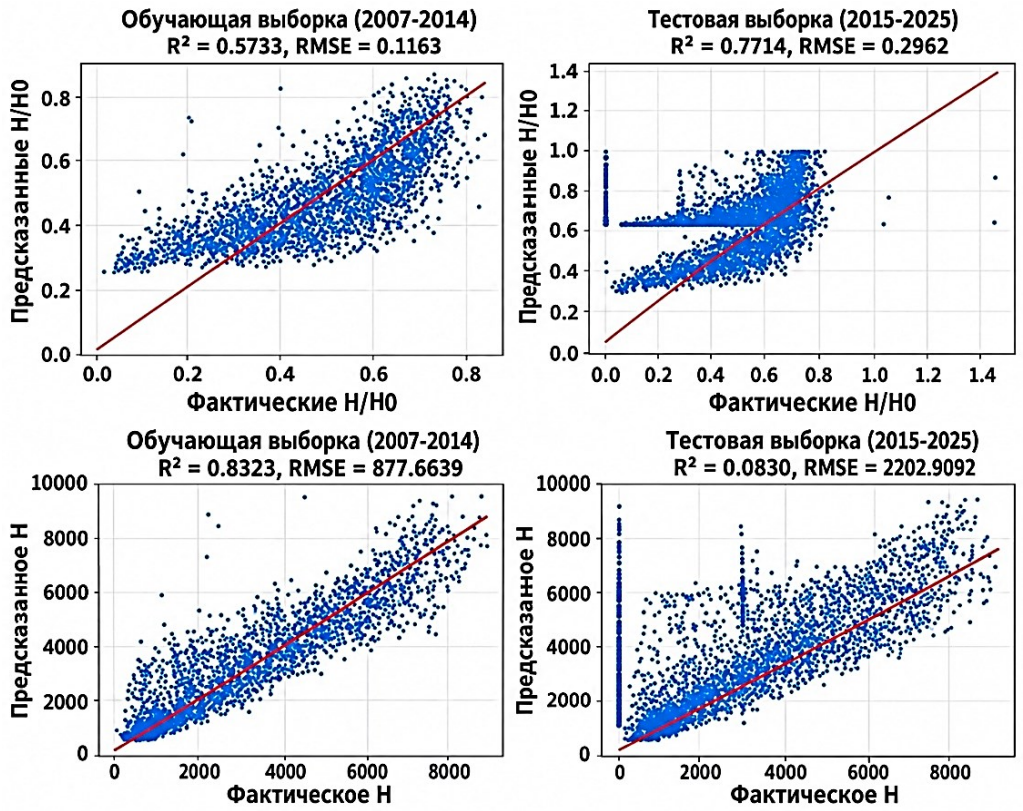

Рис. 1. Результат работы программы

Заключение. В ходе выполнения данного исследования были решены задачи, связанные с разработкой и валидацией регрессионных моделей для прогнозирования суммарной СР, а также обоснован переход к многокритериальному анализу при выборе мест размещения фотоэлектрических станций. На основе анализа отечественной и зарубежной научной литературы последних лет подтверждена эффективность методов машинного обучения (линейная регрессия, LASSO, гребневая регрессия, Random Forest, XGBoost, SARIMAX, SVR) для повышения точности прогнозов по сравнению с физическими моделями, что демонстрирует снижение RMSE на 15–20 % и достижение $R^2 > 0{,}92$ в различных климатических условиях.

В качестве объекта исследования выбран г. Иркутск, характеризующийся континентальным климатом и перспективностью для развития солнечной энергетики в восточной части России. На основе суточных метеорологических данных за 2007–2025 гг. (температура, влажность, давление, облачность) сформированы обучающая (2007–2015) и тестовая (2016–2025) выборки. Проведена калибровка и валидация 17 регрессионных моделей. Наилучший результат достигнут для SVR-моделей (M15–M17) с MAE = 460,4 Вт · ч/м² · сут, что подтверждает целесообразность применения ядерных методов для аппроксимации нелинейных зависимостей в условиях ограниченных метеоданных.

Разработанное программное приложение на Python автоматизирует полный цикл регрессионного анализа: от загрузки данных из Excel/CSV, интерактивного ввода уравнения пользователем, обучения и расчёта метрик (MAE, RMSE, MBE, R²) до визуализации и генерации структурированного отчёта. Блок-схема программы и примеры графиков наглядно иллюстрируют логику работы и возможности анализа чувствительности. Использование библиотек pandas, scikit-learn, matplotlib и openpyxl обеспечивает воспроизводимость, гибкость и низкий порог входа для лиц, принимающих решения.

Обоснована необходимость интеграции прогностических моделей с методами многокритериального принятия решений (MCDM) – такими как AHP, TOPSIS, CRITIC, EDAS, – позволяющими учитывать экономические, экологические, социальные и технические критерии при выборе площадок под СЭС. В единой среде Python, дополненной библиотеками rasterio, geopandas, folium и plotly, возможна комплексная обработка пространственных данных и построение интерактивных карт пригодности территорий, что снижает трудоёмкость и повышает прозрачность принимаемых решений.

Полученные результаты имеют практическую значимость для регионов России со схожими климатическими условиями, особенно для удалённых и арктических территорий, где достоверная оценка солнечного потенциала и обоснованный выбор мест размещения СЭС позволяют снизить зависимость от дорогостоящего завозного топлива и повысить энергетическую безопасность. Перспективы дальнейших исследований включают расширение набора предикторов (например, учёт аэрозольной оптической толщины, данных реанализа), применение глубокого обучения (LSTM, сверточных сетей) для учёта пространственно-временных корреляций, а также разработку веб-интерфейса для удалённого доступа к инструменту.

Таким образом, предложенный подход, объединяющий регрессионное прогнозирование и многокритериальный анализ в рамках Python-экосистемы, представляет собой эффективный инструментарий для научного обоснования и практической реализации проектов солнечной генерации, способствуя переходу к устойчивой и низкоуглеродной энергетике.

Литература

1. Russia’s renewable capacity to reach 18.4GW by 2035, forecasts GlobalData, [Электронный ресурс] URL: https://www.power-technology.com/analyst-comment/russia-renewable-capacity-2035-forecasts-globaldata/?cf-view&Top-Article-Share-Button-clicked:%20Copy%20Link&Top-Article-Share-Button-clicked:%20Copy%20Link&Top-Article-Share-Button-clicked:%20Copy%20Link, March 16, 2026 (дата обращения: 01.06.2026)
2. Wang Dan & Shen Yanbo & Ye Dong & Yang Yanchao & Da Xuanfang & Mo Jingyue. (2024). Evaluation of Scikit-Learn Machine Learning Algorithms for Improving CMA-WSP v2.0 Solar Radiation Prediction. Atmosphere. 15. 994. DOI: 10.3390/ atmos15080994.
3. Akpuluma David & Yurchenko Alexey. (2024). Advancing Solar Irradiation Prediction in Extreme Climates: A LASSO Regression Analysis in Tomsk. Proceedings of the International Conference on Applied Innovations in IT. 12.
4. B. Arseven and S. Çınar «Short-Term Solar Radiation Forecasting with Discrete Wavelet Transform Based Boosted Machine Learning Methods», *C.R. Acad. Bulg. Sci.*, vol. 78, no. 4, pp. 543–551, Apr. 2025.
5. Hassan Basma & Samman Ahmad & Morsy Mostafa & Abdaldym Abdallah & Alsubhi Yazeed & Almazroui, Mansour & Elhag, Mohamed & Aboelkhair, Hassan. (2025). Advanced Regression Approaches for High-Fidelity Solar Radiation Prediction: Analysis of a 12-Year Meteorological Dataset from Egypt. Earth Systems and Environment. DOI: 10.1007/s41748-025-00892-9.
6. Baloch Mazhar & Honnurvali Mohamed & Kabbani, Adnan & Jumani, Touqeer & Tahir, Sohaib. (2024). An Intelligent SARIMAX-Based Machine Learning Framework for Long-Term Solar Irradiance Forecasting at Muscat, Oman. Energies. 17. DOI: 6118. 10.3390/en17236118.
7. Мировой центр радиационных данных. [Электронный ресурс]. URL: http://wrdc.mgo.rssi.ru/ (дата обращения: 01.06.2026).
8. Расписание погоды. [Электронный ресурс]. URL: https://rp5.ru/ (дата обращения: 01.06.2026).
9. Naserpour S., Zolfaghari H., Zeaiean Firouzabadi, P., 2020. Calibration and evaluation of sunshine-based empirical models for estimating daily solar radiation in Iran. Sustain. Energy Technol. Assessment, DOI: 10.1016/j.seta.2020.100855
10. Маленкова И.Н., Шакиров В.А. Оценка суммарной солнечной радиации с суточной детализацией на основе регрессионных моделей // Вестник ИрГТУ. 2023. № 1. URL: https://cyberleninka.ru/article/n/otsenka-summarnoy-solnechnoy-radiatsii-s-sutochnoy-detalizatsiey-na-osnove-regressionnyh-modeley (дата обращения: 16.06.2026).
11. Besharat F., Dehghan A.A., Faghih A.R. Empirical models for estimating global solar radiation: A review and case study. Renew Sustain Energy Rev 2013; 21:798–821. Empirical models for estimating global solar radiation: A review and case study, DOI: 10.1016/j.rser.2012.12.043
12. Olanrewaju M. Oyewola, Tchilabalo E. Patchali, Olusegun O. Ajide, Satyanand Singh, Olaniran J. Matthew, Global solar radiation predictions in Fiji Islands based on empirical models, Alexandria Engineering Journal Volume 61, Issue 11, November 2022, Pages 8555–8571, DOI: 10.1016/j.aej.2022.01.065
13. Dongyu Jia, Liwei Yang, Tao Lv, Weiping Liu, Xiaoqing Gao, Jiaxin Zhou, 2022, "Evaluation of machine learning models for predicting daily global and diffuse solar radiation under different weather/pollution conditions", Renewable Energy 187 (2022), pages 896–906. DOI:10.1016/j.renene.2022.02.002
14. Nematchoua M.K., Orosa J.A., Afaifia M. Prediction of daily global solar radiation and air temperature using six machine learning algorithms; a case of 27 European countries. Ecological Informatics. 69:101643. 2022. DOI: 10.1016/j.ecoinf.2022.101643
15. Мюллер А. Гвидо С. Введение в машинное обучение с помощью Python. Руководство для специалистов по работе с данными.:Пер. с англ. СПб.: ООО «Альфа-книга», 2017. 480 с.
16. Елагина Е.А., Маргун А.А. Исследование методов машинного обучения в задаче идентификации клеток крови // Научно-технический вестник информационных технологий, механики и оптики. 2021. Т. 21, № 6. С. 903–911. DOI:10.17586/2226-1494-2021-21-6-903-91
17. Jong F.C., Ahmed M.M. (2024). Multi-Criteria Decision-Making Solutions for Optimal Solar Energy Sites Identification: A Systematic Review and Analysis. IEEE Access, DOI: 10.1109/ACCESS.2024.3461948.
18. Nguyen Viet Tinh & Chaysiri Rujira. (2025). CRITIC-CoCoSo Model Application in Hybrid Solar-Wind Energy Plant Location Selection Problem: A Case Study in Vietnam. Energy Engineering. 122. 515-536. DOI:10.32604/ee.2024.057786.
19. Ramos-Diez I., Barilari S., Ljunggren J., Hellsten S., Ferreras-Alonso N. (2026). GIS-based MCDM methodology for identifying suitable areas for land-use-based adaptation and mitigation solutions focused on solar energy. ISPRS IJGI, DOI: 10.3390/ijgi15040157.

References

1. Russia’s renewable capacity to reach 18.4GW by 2035, forecasts GlobalData, [electronic resource] URL: https://www.power-technology.com/analyst-comment/russia-renewable-capacity-2035-forecasts-globaldata/?cf-view&Top-Article-Share-Button-clicked:%20Copy%20Link&Top-Article-Share-Button-clicked:%20Copy%20Link&Top-Article-Share-Button-clicked:%20Copy%20Link, March 16, 2026 (date of request: 01/06/2026)
2. Wang Dan & Shen Yanbo & Ye Dong & Yang Yanchao & Da Xuanfang & Mo Jingyue. (2024). Evaluation of Scikit-Learn Machine Learning Algorithms for Improving CMA-WSP v2.0 Solar Radiation Prediction. Atmosphere. 15. 994. DOI: 10.3390/atmos15080994.
3. Akpuluma David & Yurchenko Alexey. (2024). Advancing Solar Irradiation Prediction in Extreme Climates: A LASSO Regression Analysis in Tomsk. Proceedings of the International Conference on Applied Innovations in IT. 12.
4. B. Arseven and S. Çınar «Short-Term Solar Radiation Forecasting with Discrete Wavelet Transform Based Boosted Machine Learning Methods» C.R. Acad. Bulg. Sci., vol. 78, no. 4, pp. 543–551, Apr. 2025.
5. Hassan Basma & Samman Ahmad & Morsy Mostafa & Abdaldym Abdallah & Alsubhi, Yazeed & Almazroui, Mansour & Elhag, Mohamed & Aboelkhair, Hassan. (2025). Advanced Regression Approaches for High-Fidelity Solar Radiation Prediction: Analysis of a 12-Year Meteorological Dataset from Egypt. Earth Systems and Environment. DOI: 10.1007/s41748-025-00892-9.

6. Baloch Mazhar & Honnurvali Mohamed & Kabbani Adnan & Jumani Touqeer & Tahir, Sohaib. (2024). An Intelligent SARIMAX-Based Machine Learning Framework for Long-Term Solar Irradiance Forecasting at Muscat, Oman. Energies. 17. DOI: 6118. 10.3390/en17236118.
7. World Radiation Data Center. [electronic resource]. URL: http://wrdc.mgo.rssi.ru/ (date of request: 01/06/2026).
8. Weather forecast. [electronic resource]. URL: https://rp5.ru/ (date of request: 01/06/2026).
9. Naserpour S., Zolfaghari, H. Zeaiean Firouzabadi, P. 2020. Calibration and evaluation of sunshine-based empirical models for estimating daily solar radiation in Iran. Sustain. Energy Technol. Assessment, DOI: 10.1016/j.seta.2020.100855
10. Malenkova I.N., Shakirov V.A. ASSESSMENT OF TOTAL SOLAR RADIATION WITH DAILY DETAIL BASED ON REGRESSION MODELS // Bulletin of IrSTU. 2023. №1. URL: https://cyberleninka.ru/article/n/otsenka-summarnoy-solnechnoy-radiatsii-s-sutochnoy-detalizatsiey-na-osnove-regressionnyh-modeley (date of request: 16/06/2026).
11. Besharat F., Dehghan A.A., Faghih A.R. Empirical models for estimating global solar radiation: A review and case study. Renew Sustain Energy Rev 2013; 21:798–821. Empirical models for estimating global solar radiation: A review and case study, DOI: 10.1016/j.rser.2012.12.043
12. Olanrewaju M. Oyewola, Tchilabalo E. Patchali, Olusegun O. Ajide, Satyanand Singh, Olaniran J. Matthew, Global solar radiation predictions in Fiji Islands based on empirical models, Alexandria Engineering Journal Volume 61, Issue 11, November 2022, Pages 8555–8571, DOI: 10.1016/j.aej.2022.01.065
13. Dongyu Jia, Liwei Yang, Tao Lv, Weiping Liu, Xiaoqing Gao, Jiaxin Zhou, 2022, «Evaluation of machine learning models for predicting daily global and diffuse solar radiation under different weather/pollution conditions», Renewable Energy 187 (2022), pages 896–906. DOI:10.1016/j.renene.2022.02.002
14. Nematchoua MK, Orosa JA, Afaifia M. Prediction of daily global solar radiation and air temperature using six machine learning algorithms; a case of 27 European countries. Ecological Informatics. 69:101643. 2022. DOI: 10.1016/j.ecoinf.2022.101643
15. Müller A., Guido S. Introduction to Machine Learning with Python. A Guide for Data Professionals: Translated from English – St. Petersburg: LLC "Alpha-Book", 2017. 480 p.
16. Elagina E.A., Margun A.A. Research into Machine Learning Methods for Blood Cell Identification // Scientific and Technical Journal of Information Technologies, Mechanics and Optics. 2021. T21, №6. P. 903-911. DOI:10.17586/2226-1494-2021-21-6-903-91
17. Jong F.C., Ahmed M.M. (2024). Multi-Criteria Decision-Making Solutions for Optimal Solar Energy Sites Identification: A Systematic Review and Analysis. IEEE Access, DOI: 10.1109/ACCESS.2024.3461948.
18. Nguyen, Viet Tinh & Chaysiri, Rujira. (2025). CRITIC-Co-CoSo Model Application in Hybrid Solar-Wind Energy Plant Location Selection Problem: A Case Study in Vietnam. Energy Engineering. 122. 515–536. DOI:10.32604/ee.2024.057786.
19. Ramos-Diez I., Barilari S., Ljunggren J., Hellsten S., Ferreras-Alonso N. (2026). GIS-based MCDM methodology for identifying suitable areas for land-use-based adaptation and mitigation solutions focused on solar energy. ISPRS IJGI, DOI: 10.3390/ijgi15040157.