

Вестник Череповецкого государственного университета, 2026, № 1 (130), с. 36–46.  
Cherepovets State University Bulletin, 2026, no. 1 (130), pp. 36–46.

Научная статья  
УДК 007.51:007.3  
<https://doi.org/10.23859/1994-0637-2026-1-130-3>  
<https://elibrary.ru/ahxavy>

### Прогнозирование электропотребления с помощью выделения функций из временного ряда

Константин Сергеевич Серегичев<sup>1</sup>✉, Алексей Александрович Суконщиков<sup>2</sup>,  
Вячеслав Алексеевич Горбунов<sup>3</sup>

<sup>1,2,3</sup>Вологодский государственный университет,  
Вологда, Россия

<sup>1</sup>[seregichevks@yandex.ru](mailto:seregichevks@yandex.ru), <https://orcid.org/0009-0006-3804-8706>

<sup>2</sup>[avt@vogu35.ru](mailto:avt@vogu35.ru), <https://orcid.org/0000-0002-8673-3314>

<sup>3</sup>[gorbunov1945@inbox.ru](mailto:gorbunov1945@inbox.ru), <https://orcid.org/0009-0000-0623-9977>

**Аннотация.** В статье рассматриваются методы прогнозирования электропотребления на основе анализа временных рядов. В рамках исследования предложен метод автоматизированного выделения характеристик из временных последовательностей, направленный на их дальнейшее использование в задачах машинного обучения. Данный подход позволяет эффективно преобразовывать исходные данные в информативные признаки, пригодные для построения и обучения прогностических моделей. Исследование проводится на данных о потреблении электроэнергии в коммерческих объектах, что позволяет повысить точность прогнозирования и оптимизировать управление энергохозяйством. В работе представлены результаты сравнения различных комбинаций оценочных показателей и функций потерь в рамках одной модели машинного обучения. Особое внимание уделено оценке неопределенности прогнозов и анализу влияния временных факторов на точность предсказаний. Результаты исследования демонстрируют, что использование современных методов машинного обучения позволяет достичь высокой точности прогнозирования электропотребления на основе исторических данных.

**Ключевые слова:** прогнозирование, временные ряды, машинное обучение, электропотребление, tsfresh, квантильная регрессия

**Для цитирования:** Серегичев К. С., Суконщиков А. А., Горбунов В. А. Прогнозирование электропотребления с помощью выделения функций из временного ряда. *Вестник Череповецкого государственного университета*, 2026, № 1 (130), с. 36–46. <https://doi.org/10.23859/1994-0637-2026-1-130-3>; EDN: AHXAVY

---

© Серегичев К. С., Суконщиков А. А., Горбунов В. А., 2026

## Forecasting electricity consumption using time series feature extraction

Konstantin S. Seregichev<sup>1✉</sup>, Aleksey A. Sukonshchikov<sup>2</sup>, Vyacheslav A. Gorbunov<sup>3</sup>

<sup>1,2,3</sup>Vologda State University,  
Vologda, Russia

<sup>1✉</sup>seregichevks@yandex.ru, <https://orcid.org/0009-0006-3804-8706>

<sup>2</sup>avt@vogu35.ru, <https://orcid.org/0000-0001-8673-3314>

<sup>3</sup>gorbunov1945@inbox.ru, <https://orcid.org/0009-0000-0623-9977>

**Abstract.** The article discusses methods for predicting power consumption based on time series analysis. As part of the study, a method for the automated extraction of characteristics from time sequences is proposed, aimed at their further use in machine learning tasks. This approach makes it possible to effectively transform the initial data into informative features suitable for building and training predictive models. The study considers data on electricity consumption in commercial facilities, which makes it possible to increase forecasting accuracy and optimize energy management. The paper presents the results of comparing various combinations of estimated indicators and loss functions within the framework of a single machine learning model. Special attention is paid to estimating the uncertainty of forecasts and analyzing the influence of time factors on the accuracy of predictions. The results of the study demonstrate that the use of modern machine learning methods makes it possible to achieve high accuracy in predicting power consumption based on historical data.

**Keywords:** forecasting, time series, machine learning, electricity consumption, tsfresh, quantile regression

**For citation:** Seregichev K. S., Sukonshchikov A. A., Gorbunov V. A. Forecasting electricity consumption using time series feature extraction. *Cherepovets State University Bulletin*, 2026, no. 1 (130), pp. 36–46. (In Russ.) <https://doi.org/10.23859/1994-0637-2026-1-130-3>; EDN: AHXAVY

### Введение

Направление большинства современных исследований при прогнозировании электропотребления закономерно направлено на определение и использование различных косвенных параметров, имеющих технологическое или природное происхождение. Предпочтение преимущественно отдается тем параметрам, измерение которых в автоматическом режиме было уже организовано и не требует дополнительных затрат. Основными среди них являются температура окружающей среды и естественная освещенность<sup>1</sup>. Последнюю, в случае отсутствия прямых измерений, можно заменить данными из открытых источников метеослужбы<sup>2</sup> или функцией биодинамической освещенности.

<sup>1</sup> Бугаец В. А. Краткосрочное прогнозирование электропотребления энергорайонов и региона с учетом метеофакторов: специальность 05.14.02 "Электрические станции и электроэнергетические системы": автореф. дис. ... канд. техн. наук. Новочеркасск, 2015. 22 с. EDN: ZPRVLZ

<sup>2</sup> Надтока И. И., Звозникова И. А. Сравнительный анализ расчетных и фактических данных естественной освещенности, используемых при прогнозировании электропотребления // Студенческая научная весна-2011: материалы региональной научно-технической конференции студентов, аспирантов и молодых ученых вузов Ростовской области, Новочеркасск,

### Основная часть

На предыдущем этапе исследования<sup>1</sup> было проведено сравнение статистической модели и различных моделей нейронных сетей для оценки их возможностей при прогнозировании электропотребления. В качестве исходных данных использовались временные ряды полной электрической мощности. Использование дополнительных параметров, косвенным образом влияющих на потребление электроэнергии предприятием, было исключено по условиям поставленной задачи.

Прогнозирование электропотребления с использованием моделей машинного обучения и нейронных сетей только на основе предыдущих значений оказалось невозможным. При этом, было выявлено, что исследуемые временные ряды имеют внутреннюю взаимную зависимость между соседними значениями и в рамках определенных интервалов. Это показало значительно более успешное применение алгоритмов: сезонные авторегрессионные интегрированные скользящие средние с экзогенными регрессорами (SARIMAX). Результаты прогнозирования с их применением приведены на рис. 1.

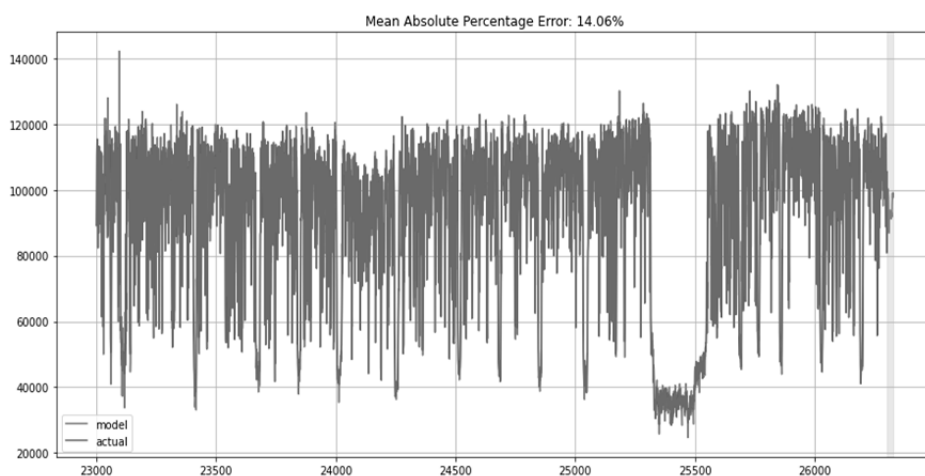


Рис. 1. Графики действительного и прогнозируемого электропотребления с использованием SARIMAX

Отсюда можно сделать предположение, что данные об электропотреблении предприятия уже несут в себе информацию о производственных циклах и других косвенных параметрах, которые возможно использовать при построении моделей прогнозирования.

---

01 января – 31 декабря 2011 года. Новочеркасск: Южно-Российский государственный политехнический университет (НПИ) имени М. И. Платова, 2011. С. 335–336. EDN: JTQDYF

<sup>1</sup> Серегичев К. С., Суконщиков А. А. Кратковременное прогнозирование электропотребления современных энергоемких предприятий // Вестник Вологодского государственного университета. Серия: Технические науки. 2024. № 2 (24). С. 36–39. EDN: NTQAME

Вот перечень некоторых основных функций, которые необходимо использовать для выделения вышеуказанной информации из временного ряда:

- анализ частотных компонент – для выявления периодичности как явно задаваемой производственным циклом, так и не явно – внутренней корпоративной культурой производственной деятельности;
- автокорреляция – для выявления повторяющихся шаблонов и паттернов;
- количество пиков данных – для выявления циклов активности;
- коэффициент вариации – для выявления относительной изменчивости данных и ее величины;
- разброс данных – для выявления смены режимов работы;
- мера регулярности – для определения степени энтропии, низкая степень которой может указывать на цикличность данных;
- выявление гармоник – что также позволяет выявлять циклы различного происхождения в данных.

Восстанавливать эту информацию в человеко-читаемый вид нет необходимости. Для работы моделей прогнозирования требуются только значимые для данного исследования математические признаки, функции и зависимости.

Данные об электропотреблении предприятий являются коммерческой тайной. Получение их для экспериментов с целью построения моделей прогнозирования затруднительно и требует как подписания соответствующих договоров NDA, так и регулярного отвлечения эксплуатационного персонала.

Для решения этой проблемы было проведено сравнительное исследование графиков электропотребления на промышленном предприятии и в полуэтаже, который занимает компания АО «НПП Энергопромсервис». Помимо офисных, там имеются производственные помещения, помещения для тестирования и наладки, собственный мини-ЦОД. Система отопления и кондиционирования выполнена на основе приточно-вытяжной вентиляции, работающей ступенчато. Количество сотрудников – более 100. Сравнение электропотребления показало схожесть форм при различии в амплитудах. Аналогично наблюдаются цикличность и сезонность временных рядов. Имеют место быть значительные перепады по потребляемой мощности из-за коммутации мощного оборудования.

Для получения исходных данных использовано ПО SEDMAX. Данное программное обеспечение осуществляет автоматический учет и диспетчеризацию потребления электроэнергии на вводно-распределительном устройстве. Временные ряды для построения и обучения моделей прогнозирования представляют собой почасовые значения средней полной мощности. Прогнозирование в ходе данного исследования основано только на предыдущих значениях этого временного ряда. Пример выгрузки исходных данных приведен в табл. 1.

Таблица 1

## Временной ряд средней полной мощности (первые пять значений)

	time	P, кВт
0	2021-07-20 09:00:00	26.210
1	2021-07-20 10:00:00	25.596
2	2021-07-20 11:00:00	23.652
3	2021-07-20 12:00:00	25.469
4	2021-07-20 13:00:00	24.318

Общим для этих данных является то, что они упорядочены по независимой переменной. Наиболее распространенной независимой переменной является время (временной ряд)<sup>1</sup>. Далее необходимо вычислить различные характеристики, такие как максимальное или минимальное значение полной мощности, его среднее значение или количество временных температурных пиков<sup>2</sup>. Графическое представление принципа выделения функций на основе анализа временного ряда приведено на рис. 2. Признаки, полученные таким образом, применяются для характеристики временного ряда, то есть они позволяют по-новому взглянуть на его структуру и изменения во времени. Для того, чтобы не рассчитывать все эти характеристики вручную, была использована библиотека *tsfresh*, которая автоматизирует этот процесс, рассчитывая и возвращая все эти характеристики автоматически<sup>3</sup>.

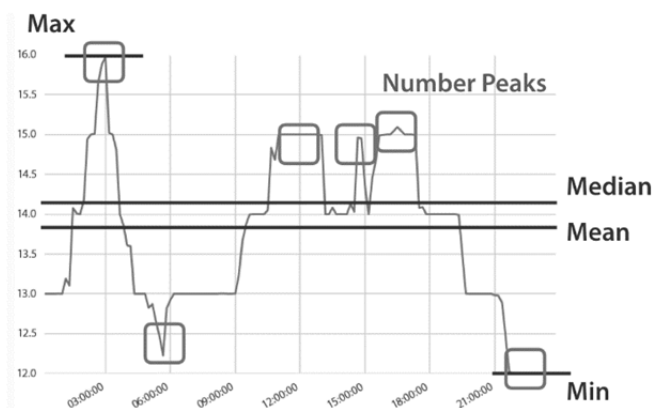


Рис. 2. Представление способа разделения функций на основе анализа временного ряда

<sup>1</sup> Хайкин С. Нейронные сети: полный курс. Москва: Вильямс, 2016. 1104 с.

<sup>2</sup> Brownlee J. Deep Learning for Time Series Forecasting. Machine Learning Mastery, 2021. 320 p.

<sup>3</sup> Feature extraction settings *tsfresh*. URL: [https://tsfresh.readthedocs.io/en/latest/text/feature\\_extraction\\_settings.html](https://tsfresh.readthedocs.io/en/latest/text/feature_extraction_settings.html). (дата обращения: 24.03.2025)

В научных кругах по всему миру регулярно проводятся исследования с применением этой библиотеки<sup>1</sup>. Благодаря этому инструменту возможно получить фрейм данных `extracted_features` с более чем 1200 различными извлеченными признаками. Стоит уточнить, что в каждом конкретном случае количество значимых признаков обычно существенно меньше. Оно зависит от сложности исследуемых процессов и величины «временного окна», используемого для прогнозирования.

Полученные характеристики можно использовать для описания временных рядов, часто они дают новое представление о временных рядах и их динамике. Их также можно использовать для кластеризации временных рядов и развития моделей машинного обучения, которые выполняют задачи классификации или регрессии на основе временных рядов. В табл. 2 показаны некоторые из этих функций.

Таблица 2

**Список некоторых функций, поддерживаемых в текущей версии `tsfresh`**

№	Наименование	Описание функции
1	<code>autocorrelation</code> (x, задержка)	Вычисляет автокорреляцию заданного запаздывания по формуле
2	<code>c3</code> (x, задержка)	Использует статистику <code>c3</code> для измерения нелинейности во временных рядах
3	<code>change_quantiles</code> (x, ql, qh, isabs, f_agg)	Сначала определяется коридор, заданный квантилями ql и qh распределения x
4	<code>cid_ce</code> (x, нормализация)	Этот функциональный калькулятор позволяет оценить сложность временного ряда (более сложный временной ряд имеет больше пиков, впадин и т. д.)
5	<code>large_standard_deviation</code> (x, t)	Имеют ли временные ряды большое стандартное отклонение?
6	<code>last_location_of_maximum</code> (x)	Возвращает относительное последнее местоположение максимального значения x
7	<code>count_below</code> (x, t)	Возвращает процент значений в x, которые меньше, чем t

Для построения модели прогнозирования воспользуемся ансамблевым методом в машинном обучении. Ансамблевые методы объединяют прогнозы нескольких базовых моделей, построенных с помощью заданного алгоритма обучения, чтобы повысить обобщаемость и надежность по сравнению с одной моделью. В ходе данного исследования была применена модель дерева регрессии с повышением градиента (GBRT).

При обучении модели будут использованы такие статистические показатели, как квантили. Они удобны, когда необходимо получить интервал, в который с высокой

<sup>1</sup> Kennedy, A. Modeling predicted microlensing event detections using systematic time series feature analysis / A. Kennedy, N. Gemma, N. Rattenbury, A. V. Kempa-Liehr // *Astronomy and Computing*. 2021. P. 1–14. <https://doi.org/10.1016/j.ascom.2021.100460>

вероятностью попадают все интересующие вас объекты исследования, или когда необходимо отфильтровать экстремальные значения.

Воспользуемся правильной стратегией разделения данных для оценки, которая учитывает временную структуру набора данных, чтобы оценить способность нашей модели прогнозировать точки данных в будущем (чтобы избежать «предвзятости» модели с использованием значений из запаздывающих признаков в обучающем наборе данных). Обучив модель и оценив ее производительность на основе средней абсолютной ошибки в процентах (MAPE), мы получили значение: 0,146 %.

Ошибка обобщения, измеренная с помощью перемешанного обучающего и тестового наборов данных, слишком оптимистична. Обобщение с помощью разделения на обучающий и тестовый наборы данных по времени, скорее всего, более точно отражает реальную эффективность регрессионной модели. Оценка разброса точности работы модели и стандартного отклонения (CV MAPE):  $0,212 \pm 0,06$  %. Вычислим несколько комбинаций оценочных показателей и функций потерь, которые приведены в табл. 3.

Таблица 3

Потери, минимизирующие каждую метрику

Loss Metric	fit_time	MAPE	MSE	MAE	pinball loss_05	pinball loss_50	pinball loss_95
"squared_error"	"22.83 ±3.46 s"	"0.08 ±0.02"	"4.8 ±1.4"	"1.1 ±0.2"	"0.6 ±0.1"	"0.6 ±0.1"	"0.5 ±0.1"
"poisson"	"23.16 ±5.08 s"	"0.08 ±0.02"	"4.9 ±1.5"	"1.1 ±0.2"	"0.6 ±0.1"	"0.6 ±0.1"	"0.5 ±0.1"
"absolute_error"	"25.98 ±5.27 s"	"0.07 ±0.02"	"5.4 ±1.3"	"1.1 ±0.2"	"0.5 ±0.1"	"0.6 ±0.1"	"0.6 ±0.1"
"quantile 5"	"22.20 ±4.63 s"	"0.13 ±0.02"	"24.0 ±4.7"	"2.4 ±0.4"	"0.2 ±0.1"	"1.2 ±0.2"	"2.2 ±0.4"
"quantile 50"	"25.94 ±4.95 s"	"0.07 ±0.02"	"5.4 ±1.3"	"1.1 ±0.2"	"0.5 ±0.1"	"0.6 ±0.1"	"0.6 ±0.1"
"quantile 95"	"21.40 ±4.73 s"	"0.21 ±0.05"	"15.6 ±3.5"	"2.6 ±0.4"	"2.4 ±0.4"	"1.3 ±0.2"	"0.3 ±0.0"

Вместо моделирования ожидаемого значения распределения  $Y | X$ , подобно методу наименьших квадратов и потерям Пуассона, выполним оценку квантили условного распределения.  $Y | X = x_i$ . Ожидается, что это случайная величина для указанной точки данных, поскольку мы ожидаем, что количество потребленной электроэнергии не может быть на 100 % точно предсказано на основе функций. На это могут влиять другие переменные, которые должным образом не учитываются существующими функциями с задержкой.

Квантильная регрессия позволяет дать более точное описание этого распределения, не делая сильных предположений о его форме.

Визуализация производительности модели в отношении 5-го перцентиля, среднего, медианы и 95-го перцентиля приведена на рис. 3.

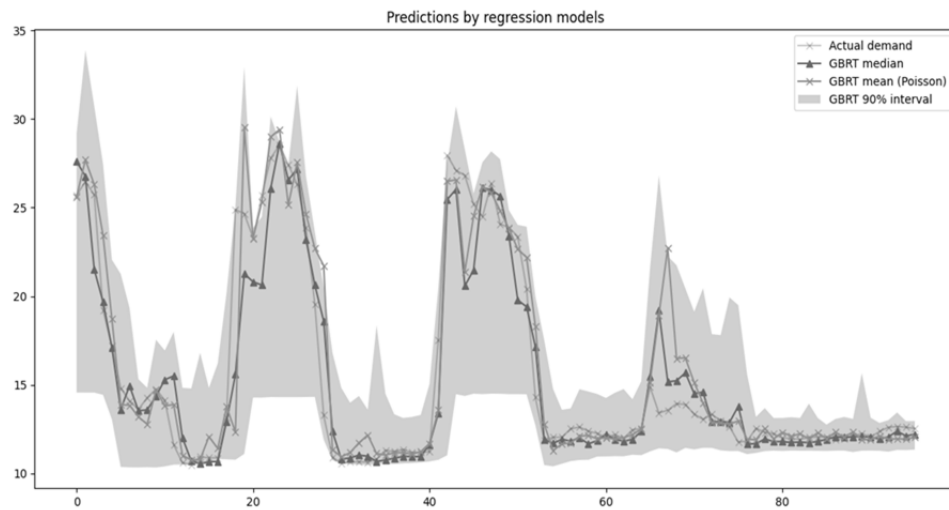


Рис. 3. Визуализация прогнозов, сделанных с помощью регрессионных моделей

Здесь интересно заметить, что область на рис. 3, обозначенная как 90% интервал, между оценками перцентилей 5 и 95 % имеет ширину, которая меняется в зависимости от времени суток. Ночью данная область намного уже, что является правильным в том смысле, что фактическое потребление остается в ее пределах. В течение дня область намного шире: неопределенность растет, что объяснимо трудовым процессом в офисе. Мы также можем видеть, что в выходные дни реальное потребление так же попадает в оценки 5 и 95 %. Наконец, стоит отметить, что данная модель является регрессионной, что говорит о том, что она проводит усреднение предсказаний. Например, для более точного предсказания пиков потребления потребовалась отдельная настройка, при этом видно ухудшение качества предсказаний минимального потребления. Сравнение производительности моделей на основе различных метрик приведено на рис. 4.

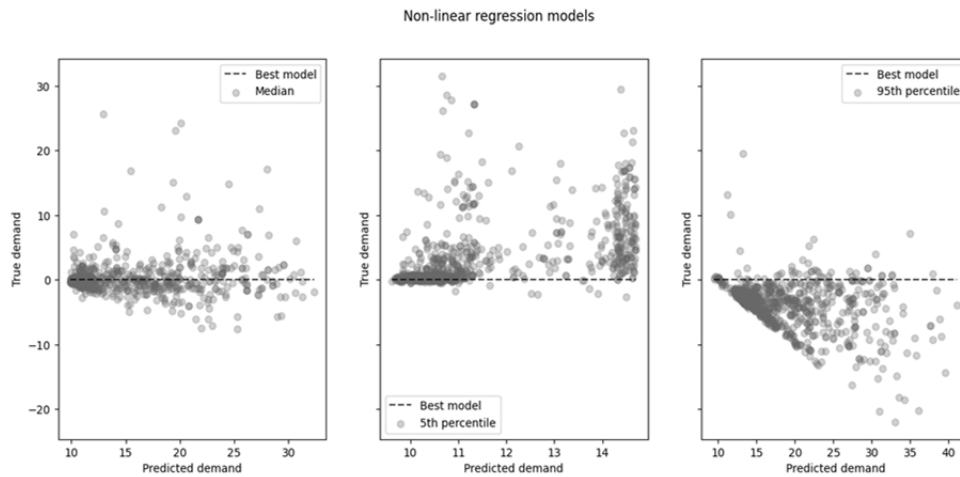


Рис. 4. Сравнение производительности моделей нелинейной регрессии с лучшими моделями

### Выводы

В данном исследовании прогнозирования временных рядов с использованием запаздывающих функций было проведено сравнение наивной регрессии (с использованием стандартизированной `train_test_split`) с надлежащей стратегией оценки временных рядов с использованием `TimeSeriesSplit`. Было отмечено, что модель, обученная с использованием `train_test_split` значения по умолчанию, `shuffle` установленного равным `True`, выдавала чуть более оптимистичную среднюю процентную ошибку (MAPE). Данные, полученные в результате разделения по времени, лучше отражают производительность нашей модели регрессии временных рядов. Мы также проанализировали прогнозную неопределенность нашей модели с помощью квантильной регрессии. Прогнозы, основанные на использовании 5-го и 95-го перцентилей, `loss = "quantile"` предоставляют нам количественную оценку неопределенности прогнозов, сделанных нашей моделью регрессии временных рядов. Оценка неопределенности также может быть выполнена с помощью MAPE, который предоставляет реализацию, основанную на недавних работах по методам конформного прогнозирования, и оценивает как алеаторическую, так и эпистемологическую неопределенность одновременно. Кроме того, функциональные возможности, предоставляемые `sktime`, могут быть использованы для расширения оценочных средств `scikit-learn` за счет использования рекурсивного прогнозирования временных рядов, которое позволяет динамично прогнозировать будущие значения.

Результаты исследования показывают, что современные возможности вычислительной техники, языков программирования и разработок в области машинного обучения позволяют с высокой точностью прогнозировать электрическую мощность, потребляемую коммерческим объектом, только на основе исторических данных о ее потреблении.

### Список литературы / References

Бугаец В. А. *Краткосрочное прогнозирование электропотребления энергорайонов и региона с учетом метеофакторов*: автореф. дис. ... канд. техн. наук. Новочеркасск: [б. и.], 2015. 22 с. EDN: ZPRVLZ

Bugaets V. A. *Short-term forecasting of electricity consumption in energy districts and regions considering meteorological factors*: Abstract Cand. thesis in Technical Sciences. Novocherkassk: [s. n.], 2015. 22 p. (In Russ.) EDN: ZPRVLZ

Надтока И. И., Звозникова И. А. Сравнительный анализ расчетных и фактических данных естественной освещенности, используемых при прогнозировании электропотребления. *Студенческая научная весна-2011: Материалы региональной научно-технической конференции*. Новочеркасск: ЮРГПУ, 2011. С. 335–336. EDN: JTQDYF

Nadtoka I. I., Zvoznikova I. A. Comparative analysis of calculated and actual natural illumination data used in electricity consumption forecasting. *Student Scientific Spring-2011: Proceedings of the Regional Scientific and Technical Conference*. Novocherkassk: IURGPU, 2011, pp. 335–336. (In Russ.) EDN: JTQDYF

Серегичев К. С., Суконщиков А. А. Кратковременное прогнозирование электропотребления современных энергоемких предприятий. *Вестник Вологодского государственного университета. Серия: Технические науки*, 2024, № 2 (24), с. 36–39. EDN: NTQAME

Seregichev K. S., Sukonshchikov A. A. Short-term prediction of electricity consumption for modern energy-intensive enterprises. *Bulletin of Vologda State University. Series: Technical Sciences*, 2024, no. 2 (24), pp. 36–39. (In Russ.) EDN: NTQAME

Хайкин С. *Нейронные сети: полный курс*. Москва: Вильямс, 2016. 1104 с.

Khaikin S. *Neural Networks: A Comprehensive Foundation*. Moscow: Vil'iams, 2016. 1104 p. (In Russ.)

Brownlee J. *Deep Learning for Time Series Forecasting*. Machine Learning Mastery, 2021. 320 p.

*Feature extraction settings tsfresh*. Available at: [https://tsfresh.readthedocs.io/en/latest/text/feature\\_extraction\\_settings.html](https://tsfresh.readthedocs.io/en/latest/text/feature_extraction_settings.html) (accessed: 24.03.2025)

Kennedy A., Gemma N., Rattenbury N., Kempa-Liehr A. V. Modeling predicted microlensing event detections using systematic time series feature analysis. *Astronomy and Computing*, 2021, pp. 1–14. <https://doi.org/10.1016/j.ascom.2021.100460>

### Сведения об авторах

**Константин Сергеевич Серегичев** – аспирант; [seregichevks@yandex.ru](mailto:seregichevks@yandex.ru), <https://orcid.org/0009-0006-3804-8706>, Вологодский государственный университет (д. 1, ул. Галкинская, 160000 Вологда, Россия); **Konstantin S. Seregichev** – Postgraduate Student, [seregichevks@yandex.ru](mailto:seregichevks@yandex.ru), <https://orcid.org/0009-0006-3804-8706>, Vologda State University (1, Galinskaya ul., 160000 Vologda, Russia).

**Алексей Александрович Суконщиков** – кандидат технических наук, доцент; [avt@vogu35.ru](mailto:avt@vogu35.ru), <https://orcid.org/0000-0002-8673-3314>, Вологодский государственный универ-

ситет (д. 1, ул. Галкинская, 160000 Вологда, Россия); **Aleksey A. Sukonshchikov** – Candidate of Technical Sciences, Associate Professor, avt@vogu35.ru, <https://orcid.org/0000-0002-8673-3314>, Vologda State University (1, Galkinskaya ul., 160000 Vologda, Russia).

**Вячеслав Алексеевич Горбунов** – доктор физико-математических наук, профессор; gorbunov1945@inbox.ru, <https://orcid.org/0009-0000-0623-9977>, Вологодский государственный университет (д. 1, ул. Галкинская, 160000 Вологда, Россия); **Vyacheslav A. Gorbunov** – Doctor of Physical and Mathematical Sciences, Professor, gorbunov1945@inbox.ru, <https://orcid.org/0009-0000-0623-9977>, Vologda State University (1, Galkinskaya ul., 160000 Vologda, Russia).

**Заявленный вклад авторов:** К. С. Серегичев – 70 %, А. А. Суконщиков – 20 %, В. А. Горбунов – 10 %. Авторы заявляют об отсутствии конфликта интересов.

**Contribution of the authors:** K. S. Seregichev – 70 %, A. A. Sukonshchikov – 20 %, V. A. Gorbunov – 10 %. The authors declare no conflicts of interests.

---

Статья поступила в редакцию 21.04.2025; одобрена после рецензирования 11.12.2025; принята к публикации 14.01.2026.

The article was submitted 21.04.2025; Approved after reviewing 11.12.2025; Accepted for publication 14.01.2026.