

УДК 519.6, 004.421

DOI: 10.18799/24131830/2024/12/4731

Шифр специальности ВАК: 05.14.01

Прогнозирование электропотребления на основе метода главных компонент (РСА)

И.У. Рахмонов¹✉, Н.Н. Ниёзов¹, В.Я. Ушаков², Н.Н. Курбонов¹, А.М. Нажимова³

¹ Ташкентский государственный технический университет, Узбекистан, г. Ташкент

² Национальный исследовательский Томский политехнический университет, Россия, г. Томск

³ Каракалпакский государственный университет, Узбекистан, г. Нукус

✉ilider1987@yandex.ru

Аннотация. *Актуальность* исследования обусловлена необходимостью точного прогнозирования потребления электроэнергии для повышения эффективности и снижения затрат на промышленных предприятиях, что приводит к повышению конкурентоспособности товаров, производимых предприятием. Традиционные методы прогнозирования часто не учитывают сложные взаимодействия между различными факторами, влияющими на потребление энергии, и не обеспечивают необходимую точность прогнозов. Метод анализа главных компонент открывает многообещающую перспективу – уменьшение объема обрабатываемых данных (размерности) без значительной потери информации, что упрощает прогнозные модели при сохранении их точности. *Целью* исследования является разработка точной и эффективной модели прогнозирования электропотребления на промышленных предприятиях с использованием метода анализа главных компонент. Эта модель направлена на устранение ограничений, характерных для традиционных подходов к прогнозированию, путем уменьшения размерности данных и повышения точности предсказаний, что в конечном итоге позволяет повысить эффективность потребления электроэнергии и снизить финансовые затраты, в том числе из-за ошибок в прогнозировании. *Методы:* метод анализа главных компонент, позволивший уменьшить объем обрабатываемых данных (размерность) путем преобразования большого набора коррелированных переменных в меньший набор некоррелированных главных компонент. Исследование включает следующие этапы: импорт данных и факторный анализ, построение корреляционной матрицы, анализ выбранных и накопленных дисперсий для каждого фактора, построение матрицы факторных нагрузок, уменьшение размерности, разработка математической модели с использованием линейной регрессии и установка и валидация прогноза. *Результаты.* Применение метода главных компонент позволило создать модель прогнозирования электропотребления. Ее применение показало, что первая главная компонента объясняет 69,65 % общей дисперсии, вторая компонента – 17,28 %, т. е. в совокупности они объясняют почти 87 % дисперсии. Разработанная модель обеспечивает хорошее совпадение фактических и прогнозных значений электропотребления в нескольких временных интервалах со средним уровнем ошибки в пределах от +3 до –5 %. Это указывает на пригодность модели для прогнозирования электропотребления, хотя некоторые расхождения указывают на необходимость дальнейшего ее совершенствования.

Ключевые слова: прогнозирование электропотребления, метод главных компонент, уменьшение размерности данных, эффективность, снижение затрат, оптимизация энергопотребления, моделирование, корреляционная матрица, факторный анализ

Для цитирования: Прогнозирование электропотребления с использованием метода главных компонент (РСА) / И.У. Рахмонов, В.Я. Ушаков, Н.Н. Ниёзов, Н.Н. Курбонов // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2024. – Т. 335. – № 12. – С. 198–209. DOI: 10.18799/24131830/2024/12/4731

UDC 519.6, 004.421

DOI: 10.18799/24131830/2024/12/4731

Forecasting electricity consumption using the Principal Component Analysis method

I.U. Rakhmonov¹✉, N.N. Niyozov¹, V.Ya. Ushakov², N.N. Kurbonov¹, A.M. Najimova³

¹ Tashkent State Technical University, Tashkent, Uzbekistan

² National Research Tomsk Polytechnic University, Tomsk, Russian Federation

³ Karakalpak State University, Nukus, Uzbekistan

✉ ilider1987@asu.ru

Abstract. Relevance. The need for accurate forecasting of electricity consumption to improve efficiency and reduce costs at industrial enterprises, which leads to increased competitiveness of goods manufactured by the enterprise. Traditional forecasting methods often do not take into account complex interactions between various factors affecting energy consumption and do not provide the necessary forecast accuracy. The principal component analysis method offers a promising prospect – reducing the volume of processed data (dimensionality) without significant loss of information, which simplifies forecast models while maintaining their accuracy. **Aim.** To develop an accurate and efficient model for forecasting electricity consumption at industrial enterprises using the principal component analysis method. This model is aimed at eliminating the limitations of traditional forecasting approaches by reducing data dimensionality and increasing the accuracy of predictions, which ultimately improves the efficiency of electricity consumption and reduces financial costs, including those due to forecasting errors. **Methods.** The principal component analysis method, which allows us to reduce the volume of processed data (dimensionality) by transforming a large set of correlated variables into a smaller set of uncorrelated principal components. The study included the following stages: data import and factor analysis, correlation matrix construction, analysis of selected and accumulated variances for each factor, factor loading matrix construction, dimension reduction, development of a mathematical model using linear regression, and forecast installation and validation. **Results.** The application of the principal component analysis method allowed us to create a model for forecasting electricity consumption. Its application showed that the first principal component explains 69.65% of the total variance, the second component – 17.28%, i. e. together they explain almost 87% of the variance. The developed model provides good agreement between the actual and forecast values of electricity consumption in several time intervals with an average error level within the range of +3 to –5%. This indicates the suitability of the model for forecasting electricity consumption, although some discrepancies indicate the need for its further improvement.

Keywords: power consumption forecasting, principal component method, data dimensionality reduction, efficiency, cost reduction, energy consumption optimization, modeling, correlation matrix, factor analysis

For citation: Rakhmonov I.U., Ushakov V.Ya., Niyozov N.N., Kurbanov N.N. Forecasting electricity consumption using the principal component analysis method. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, 2024, vol. 334, no. 2, pp. 198–209. DOI: 10.18799/24131830/2024/12/4731

Введение

Рост потребности в точном прогнозировании электропотребления обусловлен необходимостью повышения его эффективности, сокращения платежей за электроэнергию и повышения конкурентоспособности производимых промышленным предприятием товаров. Традиционные методы прогнозирования, основанные на экспертных оценках или на оперировании с удельным расходом электроэнергии, часто не учитывают сложные взаимодействия между различными факторами, влияющими на потребление энергии, что приводит к недостаточно точным прогнозам. Метод главных компонент (Principal Component Analysis – PCA) является одним из методов многомерного статистического

анализа, который используется для уменьшения размерности (объема обрабатываемых данных) с наименьшей потерей полезной информации. В методе PCA большой набор коррелированных переменных преобразуют в меньший набор некоррелированных главных компонент, упрощая тем самым модель прогнозирования при сохранении ее предсказательной способности.

Уменьшение количества переменных в составе исходных данных за счёт исключения из рассмотрения второстепенных облегчает оперирование с ними, упрощает создание алгоритмов [1–5].

Обеспечиваемая методом PCA возможность уменьшения размерности данных упрощает визуализацию, обработку и анализ, при этом сохраняя

основную информацию, выраженную через максимальную дисперсию.

Потребность в уменьшении объема исходной информации объясняется следующим [6, 7]:

- необходимостью визуального отображения первоначально отобранных данных, что достигается на основе их проецирования в трехмерном пространстве, плоскости или в ряде чисел;
- стремлением упростить изучаемые модели в связи с необходимостью упростить расчёт и интерпретацию результатов;
- необходимостью уменьшить объём хранимых данных;
- низкой информативностью из-за малой изменчивости при переходе от одного объекта к другому;
- заменой информации в связи с корреляцией первоначальных признаков;
- необходимостью агрегирования первоначальных признаков.

Достижение конечной цели – повышение точности прогнозов методом PCA – осуществляется решением следующих задач:

1. Многомерные данные преобразуются в новое пространство с уменьшенной размерностью, где исходные переменные заменяются на главные компоненты с максимальной информативностью.
2. Выбираются главные компоненты, которые обеспечивают максимальную дисперсию данных, что позволяет сохранить наибольшее количество информации при минимальном числе измерений.
3. В результате преобразования для исходных данных создаются новые ортогональные оси (главные компоненты), между которыми отсутствует корреляция.
4. Снижается размерность данных с сохранением их основной структуры, что упрощает анализ и визуализацию.
5. Устраняется мультиколлинеарность, так как главные компоненты являются независимыми, и каждая из них вносит уникальный вклад в объяснение данных.

Кроме уменьшения размерности и удаления мультиколлинеарности, метод PCA имеет несколько преимуществ перед регрессионным анализом и анализом по Фурье.

Во-первых, PCA отличается простотой и универсальностью. Обеспечиваемая им возможность сжатия данных помогает экономить ресурсы хранения и ускоряет вычислительные процессы.

Это приводит к улучшению производительности алгоритмов машинного обучения за счет снижения числа признаков, с которыми они работают. PCA помогает устранить проблему мультиколлинеарно-

сти, позволяя преобразовывать исходные признаки в некоррелированные компоненты, что особенно полезно для регрессионных моделей и других алгоритмов, чувствительных к корреляции между признаками. Также метод способствует улучшению визуализации многомерных данных и тем самым позволяет преобразовать их в двух- или трёхмерные проекции и выявить скрытые структуры и аномалии в данных. PCA может улучшить качество данных, уменьшая влияние шума, так как при уменьшении размерности сохраняются только наиболее информативные компоненты, а слабые или шумовые компоненты отбрасываются.

Эти особенности делают его применимым для работы с большими наборами данных в различных областях, таких, например, как обработка изображений, биоинформатика и финансы, где требуется эффективный анализ высокоразмерных данных [8, 9].

В отличие от него регрессионный анализ требует более сложной настройки моделей (выбор зависимой переменной), а анализ по Фурье ограничен работой с данными, имеющими временную или периодическую природу.

Во-вторых, PCA позволяет анализировать широкий спектр данных, включая многомерные, неконтролируемые и неструктурированные наборы данных, а регрессионный анализ применим лишь в случаях, когда зависимости между переменными ясны и линейны. Анализ по Фурье, в свою очередь, наиболее применим для анализа периодических сигналов или данных с частотными характеристиками. Это ограничивает его использование в других областях.

В-третьих, PCA позволяет эффективно обрабатывать шум, фокусируясь на компонентах с высокой дисперсией, что особенно полезно при работе с большими и сложными наборами данных. В регрессионном анализе шум может влиять на оценку параметров модели, усложняя точные предсказания, а анализ по Фурье фильтрует шум только для периодических сигналов [10, 11].

Большие возможности, заложенные в методе PCA, подтверждены опытом применения многофакторного анализа, например, при прогнозировании электропотребления в быту [12–15].

В данной статье рассмотрено применение метода PCA для повышения точности прогнозирования электропотребления промышленными предприятиями, в частности металлургическими заводами. Авторами предложен укрупнённый алгоритм решения этой задачи с применением метода PCA.

Основные положения

Метод PCA впервые был применён Карлом Пирсоном в 1901 г. и заключается в следующем.

Пусть дана матрица X для разработки прогно-
 стической модели с использованием метода РСА:

$$X = \begin{pmatrix} \Phi 1_1 & \Phi 2_1 & \dots \\ \Phi 1_2 & \Phi 2_2 & \dots \\ \dots & \dots & \dots \\ \Phi 1_n & \Phi 2_n & \dots \end{pmatrix}.$$

Поскольку эти данные представлены в разных
 величинах, выполняется операция стандартизации
 данных. Из-за повышенной чувствительности РСА
 к масштабу данных факторы стандартизируются в
 диапазоне 0–1 относительно среднего значения μ_i :

$$F1' = \frac{F1 - \mu_{F1}}{\sigma_{F1}}, \dots, F2' = \frac{F2 - \mu_{F2}}{\sigma_{F2}}.$$

Здесь μ_{F1} и μ_{F2} – средние значения $F1$ и $F2$, а σ_{F1} и
 σ_{F2} – их дисперсии.

После стандартизации факторы X принимают
 вид X' :

$$X' = \begin{pmatrix} \Phi' 1_1 & \Phi' 2_1 & \dots \\ \Phi' 1_2 & \Phi' 2_2 & \dots \\ \dots & \dots & \dots \\ \Phi' 1_n & \Phi' 2_n & \dots \end{pmatrix}.$$

Для выявления связей между данными опреде-
 ляется ковариационная матрица:

$$C = \frac{1}{n-1} X'^T X'.$$

Значения и векторы ковариационной матрицы
 определяются путем решения следующего уравне-
 ния:

$$Cv = \lambda v$$

$$\det(C - \lambda I) = 0,$$

где v – направления главных компонент; λ – дис-
 персии главных компонент; I – единичная матрица.

Значения дисперсий ($\lambda_1, \lambda_2, \dots, \lambda_n$) показывают,
 как каждая главная компонента влияет на общую
 информацию. При выборе главных компонент це-
 лесообразно определить, какую роль главные ком-
 поненты играют в общей дисперсии данных
 [16–18]. Если кумулятивная дисперсия, объяснен-
 ная первыми k компонентами, обозначена как T , то
 она рассчитывается следующим образом:

$$T = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^n \lambda_i}.$$

В основном при выборе главных компонент вы-
 бираются те из них, которые имеют значения
 больше 0,7. С их помощью исходные данные про-
 ецируются в новое пространство размерности Z .
 Этот процесс сохраняет влияние сокращённых
 данных, при этом главные компоненты выделяются
 в виде:

$$Z = X' \cdot D,$$

где D – матрица, состоящая из выбранных k векто-
 ров.

Восстановление исходных параметров выполня-
 ется с помощью следующей формулы:

$$X' = Z \cdot D^T.$$

На следующем этапе на основе полученных
 данных формируется следующая модель:

$$W = \beta_0 + \beta_1 Z1 + \beta_2 Z2 + \dots + \beta_n ZN.$$

Параметры модели $\beta_0, \beta_1, \beta_2$ и β_n определяются
 с помощью метода наименьших квадратов следу-
 ющим образом:

$$\min_{\beta_0, \beta_1, \beta_2, \dots, \beta_n} \sum_{i=1}^n (W_i - (\beta_0 + \beta_1 Z1 + \beta_2 Z2 + \dots + \beta_n ZN))^2.$$

На основе приведенных выше выражений про-
 гнозирование потребления электроэнергии с ис-
 пользованием метода РСА осуществляется с ис-
 пользованием алгоритма, изображенного на рис. 1.
 Построение модели с использованием метода РСА
 начинается с получения исходных данных и их
 стандартизации для приведения всех переменных к
 одному масштабу. Затем создается ковариационная
 матрица C , которая рассчитывается для анализа
 взаимосвязей переменных. После этого путем
 нахождения собственных векторов и собственных
 значений ковариационной матрицы выделяют
 главные компоненты для уменьшения размерности
 данных. Далее выполняется расчет главных компо-
 нент. После этого анализируется накопленная дис-
 персия T для выбранных главных компонент, и ес-
 ли сумма объясненной дисперсии превышает или
 равна 70 %, то данные компоненты используются
 для дальнейшего построения модели. На следу-
 ющем этапе рассчитывается накопленная дисперсия.
 Затем строится регрессионная модель. В цикличе-
 ском процессе проверяется ошибка модели E , кото-
 рая рассчитывается на основе разности предсказан-
 ных и истинных значений. Если ошибка модели
 составляет менее 5 %, процесс считается завершен-
 ным и модель считается приемлемой для прогнози-
 рования. Если ошибка больше заданного порога,
 параметры модели корректируются и расчет по
 данной модели повторяется. При повторном расче-
 те используется Z -совокупность, содержащая фак-
 торы в виде матрицы размером $n \times m$. Принятые
 данные вводятся в разработанную модель РСА,
 после чего вычисляются m следующих прогнозных
 значений W , и результат выводится в файл в виде
 матрицы размером $1 \times m$.

Ниже достаточно подробно рассмотрено приме-
 нение метода РСА на основе многофакторного анали-
 за при прогнозировании электропотребления пред-
 приятием на примере металлургического завода.

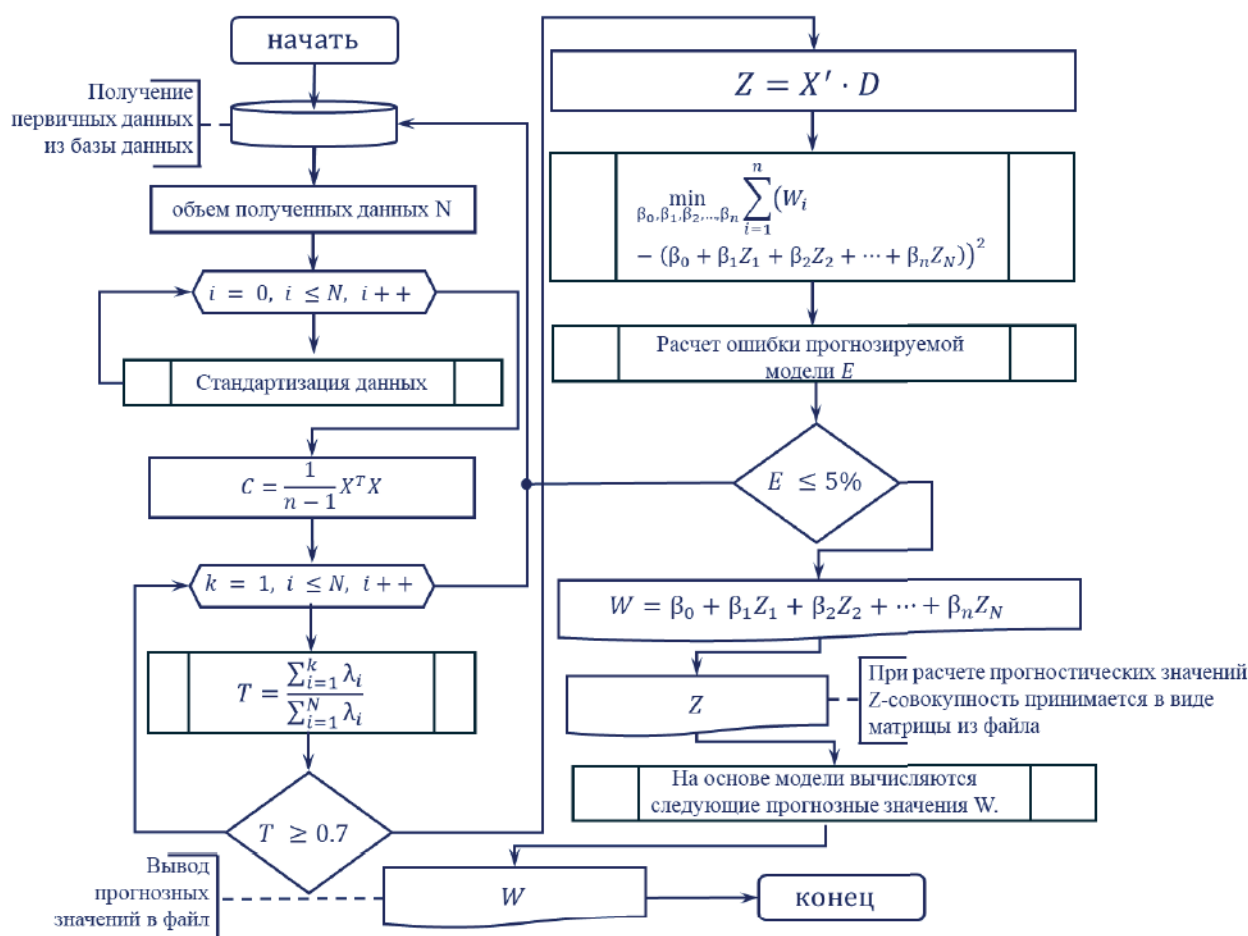


Рис. 1. Укрупнённая блок-схема алгоритма прогнозирования электропотребления методом главных компонент
 Fig. 1. Generalized block diagram of the electricity consumption forecasting algorithm using the PCA method

Расчет прогнозируемого электропотребления

На долю предприятий чёрной металлургии приходится около 10 % потребления электроэнергии, производимой в стране. Электросталеплавильные цеха этих предприятий являются крупнейшими потребителями электроэнергии. Потребление электроэнергии металлургическими предприятиями характеризуется большой нестабильностью, обусловленной проявлением множества факторов, таких как многозвенность технологической цепочки (плавка, прокатка, термообработка и др.), изменчивость производственных циклов, различие используемых технологий, изменение объемов производства, сезонные колебания. Например, в процессе плавки на электропотребление могут влиять такие параметры, как состав и количество сырья, температура, влажность, время обработки [19–21].

Именно данная особенность таких предприятий вынуждает разрабатывать и применять методы расчета прогнозных показателей энергопотребления при таких осложняющих обстоятельствах. Резкие изменения объёмов потребления электроэнергии вынуждают выполнять углубленный анализ состава

и количества исходных данных при прогнозировании потребления электроэнергии. Это, в свою очередь, требует разработки прогнозных моделей, обеспечивающих минимизацию погрешностей. Ниже приводятся данные по потреблению электрической энергии при выполнении плавки металла, показывающие величину ее нестабильности, табл. 1.

Известно, что металлургические предприятия характеризуются высокой энергоемкостью производственных процессов, значительными колебаниями электрической нагрузки, обусловленными внешними факторами, а также значительными масштабами использования вторичных энергоресурсов. Эти особенности усложняют прогнозирование электропотребления и требуют применения эффективных методов анализа данных, таких, например, как метод PCA. Металлургические предприятия характеризуются многокомпонентными и нелинейными процессами, такими как плавка, прокат, термообработка; для каждого из них характерен свой режим потребления энергии и высокая его корреляция с производственными показате-

лями (например, температурой печей, скоростью прокатки, качеством металла). Такой подход позволяет предприятиям более эффективно управлять электропотреблением, минимизировать издержки и оптимизировать использование энергетических ресурсов, что способствует повышению общей энергетической эффективности производства. Использование метода РСА помогает выделить ключевые факторы, влияющие на энергопотребление, путем декомпозиции исходного набора данных на ортогональные компоненты, которые представляют основную вариабельность процесса [22, 23].

Таблица 1. Потребление электрической энергии на проведение каждой из 100 плавов

Table 1. Electricity consumption for each of the 100 melts

№	W _{факт.} кВт·ч W _{ас} kW·h	№	W _{факт.} кВт·ч W _{ас} kW·h	№	W _{факт.} кВт·ч W _{ас} kW·h	№	W _{факт.} кВт·ч W _{ас} kW·h	№	W _{факт.} кВт·ч W _{ас} kW·h
1	32748	21	35976	41	37800	61	35858	81	36798
2	32758	22	35986	42	38304	62	35992	82	36786
3	32768	23	35996	43	38314	63	36002	83	36796
4	33442	24	35516	44	37998	64	36012	84	36806
5	33452	25	35526	45	38008	65	36176	85	37850
6	34282	26	35536	46	38018	66	36186	86	37860
7	34292	27	35708	47	37386	67	36196	87	37870
8	34302	28	35718	48	37396	68	36242	88	37956
9	35040	29	35728	49	37406	69	36252	89	37966
10	35050	30	35980	50	36928	70	36262	90	37946
11	35060	31	35990	51	36938	71	36320	91	37956
12	35104	32	36000	52	36948	72	36330	92	37966
13	35114	33	36132	53	36634	73	36340	93	39999
14	35124	34	36142	54	36644	74	36350	94	40019
15	35218	35	36152	55	36654	75	36788	95	40008
16	35228	36	37622	56	36442	76	36798	96	40018
17	35238	37	37632	57	36452	77	36808	97	40018
18	35472	38	37642	58	36462	78	36818	98	40028
19	35482	39	37780	59	35838	79	36828	99	40028
20	35492	40	37790	60	35848	80	36838	100	41388

Применение РСА в данном случае направлено на выделение главных факторов, которые наиболее убедительно объясняют колебания энергопотребления, что позволяет построить более точные модели для прогнозирования. Например, в процессе плавки на электропотребление могут влиять такие параметры, как состав и количество сырья, температура, влажность и время обработки. Учитывая эти особенности, применение метода РСА для прогнозирования электропотребления металлургическими предприятиями представляется наиболее перспективным в силу рассмотренных выше его возможностей и преимуществ.

Для решения поставленной задачи составлена корреляционная матрица R для определения взаимосвязей между факторами (табл. 2):

$$R = \begin{bmatrix} 1 & -0,31664 & -0,31602 & -0,31565 & -0,197091 & -0,21457 \\ -0,31664 & 1 & 0,999696 & 0,999277 & 0,519265 & 0,324651 \\ -0,31602 & 0,999696 & 1 & 0,999679 & 0,522692 & 0,264124 \\ -0,31565 & 0,999277 & 0,999679 & 1 & 0,521547 & 0,657814 \\ -0,19709 & 0,519265 & 0,522694 & 0,521547 & 1 & 0,321487 \\ -0,21457 & 0,324651 & 0,264124 & 0,657814 & 0,321487 & 1 \end{bmatrix}.$$

В качестве исходных данных выбраны следующие факторы: Φ_1 – загрузки печи, т; Φ_2 – количество отгруженного металла, т; Φ_3 – масса загружаемой металлической шихты, т; Φ_4 – продолжительность выплавки металла, мин.; Φ_5 – продолжительности работы под током, мин.

Таблица 2. Расчетно-экспериментальные значения факторов для 100 плавов стали

Table 2. Calculation of input data from the database for 100 steel melts

№ процесса плавов Melting process no.	W _{факт.} кВт·ч W _{ас} , kW·h	Φ_1 , т F ₁ , t	Φ_2 , т F ₂ , t	Φ_3 , т F ₃ , t	Φ_4 , мин F ₄ , min	Φ_5 , мин F ₅ , min
1	32748	132	130	128	57	53
5	33452	132	129	128	56	52
10	35050	132	131	129	55	51
15	35218	135	132	130	56	53
20	35492	134	131	130	55	51
25	35526	133	132	129	54	51
30	35980	134	132	129	59	53
35	36152	133	131	130	56	53
40	37790	133	130	128	55	53
45	38008	134	131	129	56	51
50	36928	134	132	130	54	53
55	36654	132	131	130	57	51
60	35848	131	129	128	55	53
65	36176	134	131	130	56	53
70	36262	133	132	130	59	53
75	36788	132	130	129	55	51
80	36838	131	129	128	59	53
85	37850	132	129	128	52	51
90	37946	133	131	130	58	53
95	40008	131	130	129	55	53
100	41388	131	130	129	58	53

Таблица 3. Вклады главных компонент в суммарную дисперсию исходных признаков

Table 3. Contributions of the main components to the total variance of the original characteristics

Факторы Factors	Общая дисперсия, % Total variance, %	Кумулятивная дисперсия, % Cumulative variance, %
1	69,65	69,65
2	17,28	86,94
3	13,04	99,98
4	0,014	99,99
5	0,003	100,0

Результаты анализа общей и накопленной дисперсий для 5 факторов приведены в табл. 3. Как видно, на долю первого фактора приходится 69,65 % от общей дисперсии, на долю второго – 17,28 %, и так далее.

Используя метод базовых компонент, получаем первичную матрицу весовых нагрузок факторов. Параллельно необходимо анализировать матрицу собственных значений корреляционной матрицы.

Матрица весовых нагрузок факторов имеет следующий вид:

$$U = \begin{bmatrix} -0,42577 & 0,90367 & -0,04563 & -0,29278 & -0,39021 \\ 0,97914 & 0,10209 & -0,17455 & -0,01866 & -0,00555 \\ 0,97982 & 0,10328 & -0,17077 & -0,00055 & 0,01065 \\ 0,97944 & 0,10346 & -0,17199 & 0,01924 & -0,00508 \\ 0,65063 & 0,12643 & 0,74879 & -0,88327 & -0,32493 \end{bmatrix}.$$

Если размер поля свойств меньше двух основных компонент, учитываются только первые два столбца указанной матрицы.

Зависимость главных компонент от централизованных нормированных начальных характеристик имеет следующий вид:

$$W_1 = -0,42577\Phi_1 + 0,97914\Phi_2 + 0,97982\Phi_3 + 0,97944\Phi_4 + 0,65063\Phi_5;$$

$$W_2 = 0,9036\Phi_1 + 0,10209\Phi_2 + 0,10209\Phi_3 + 0,10328\Phi_4 + 0,12634\Phi_5.$$

Анализируется следующая матрица факторных нагрузок для определения новых признаков:

$$R = \begin{bmatrix} -0,76494 & 0,983586 & -0,07316 & -0,60971 & -0,01433 \\ 0,96225 & -0,14342 & 0,03046 & -0,01863 & -0,00559 \\ 0,961577 & -0,14261 & -0,23434 & -0,00056 & 0,01065 \\ 0,261779 & -0,14229 & 0,03310 & 0,01926 & -0,00502 \\ 0,30014 & -0,07934 & -0,25058 & 0,59948 & 0,35618 \end{bmatrix}.$$

При этом исходные характеристики и коэффициенты корреляции основных компонент характеризуют корреляционную матрицу факторных нагрузок и составляют основу данных расчетов:

$$A = \begin{bmatrix} -0,76494 & 0,983586 \\ 0,96225 & -0,14342 \\ 0,961577 & -0,14261 \\ 0,261779 & 0,84229 \\ 0,30014 & -0,07934 \end{bmatrix}.$$

Исходные данные, для которых коэффициенты корреляции больше 0,7, непосредственно зависят от первого главного компонента: Φ_1 – загрузка печи, т; Φ_2 – масса отгруженного металла, т; Φ_3 – масса загружаемой металлической шихты, т. Вторая главная компонента непосредственно связана с исходными данными: Φ_1 – загрузка печи, т; Φ_4 – продолжительность выплавки металла, мин. По определенным таким способом главным факторам строится математическая модель с помощью линейной регрессии, имеющей следующий вид:

$$W = 36264 - 63,74\Phi_1 + 46,89\Phi_2. \quad (1)$$

По (1) определяются прогнозные значения исследуемого объекта и сравниваются с фактическими данными (рис. 1).

Уравнение (1) описывает аппроксимирующую зависимость прогнозируемого потребления электроэнергии W от двух основных компонент, полученных методом PCA и обозначенных как Φ_1 и Φ_2 . Применение PCA позволяет уменьшить количество (размерность) данных и упростить тем самым анализ и интерпретацию переменных. Коэффициенты 63,74 и 46,89 показывают величину вклада каждой компоненты в конечный результат. Они имеют размерность, обратную единицам соответствующих компонент, что обеспечивает размерность W в виде единиц потребленной электроэнергии (например, кВт·ч).

Область применения данной аппроксимирующей модели ограничена условиями и параметрами технологического процесса, для которого она разработана, а именно – металлургическими предприятиями. Данная модель применима только при сходных производственных параметрах и технологических характеристиках предприятия.

При оценке глубины прогнозирования важно учитывать временной диапазон или другие показатели, зависящие от специфики технологического процесса. В случае металлургического производства одним из таких показателей может служить количество плавов, на которые распространяется прогноз. Такое прогнозирование позволяет учитывать регулярные изменения в потреблении энергии, связанные с особенностями каждой плавки: используемые материалы, параметры процесса, режимы работы оборудования. Таким образом, глубина прогноза может быть определена как на временной основе (в часах или днях), так и на основе количества плавов, на протяжении которых сохраняется высокая точность прогнозируемых значений потребления энергии.

Как видно на рис. 2, фактические и прогнозные значения электропотребления совпадают в нескольких интервалах, например, 2–15, 18–20, 26–27, 33–34. В других интервалах наблюдаются значительные спады, которые негативно влияют на адекватность разработанной модели прогнозирования. Среднее значение погрешности расчетов по разработанной модели лежит в пределах от +3 до –5 % (рис. 3), что указывает на ее применимость на практике.

В табл. 4 приведены фактические и прогнозные значения расхода электроэнергии в процессах плавок стали для всех 100 точек прогнозирования. Применен шаг в 5 единиц, а ошибки прогнозирования определены по 100 точкам прогнозирования.

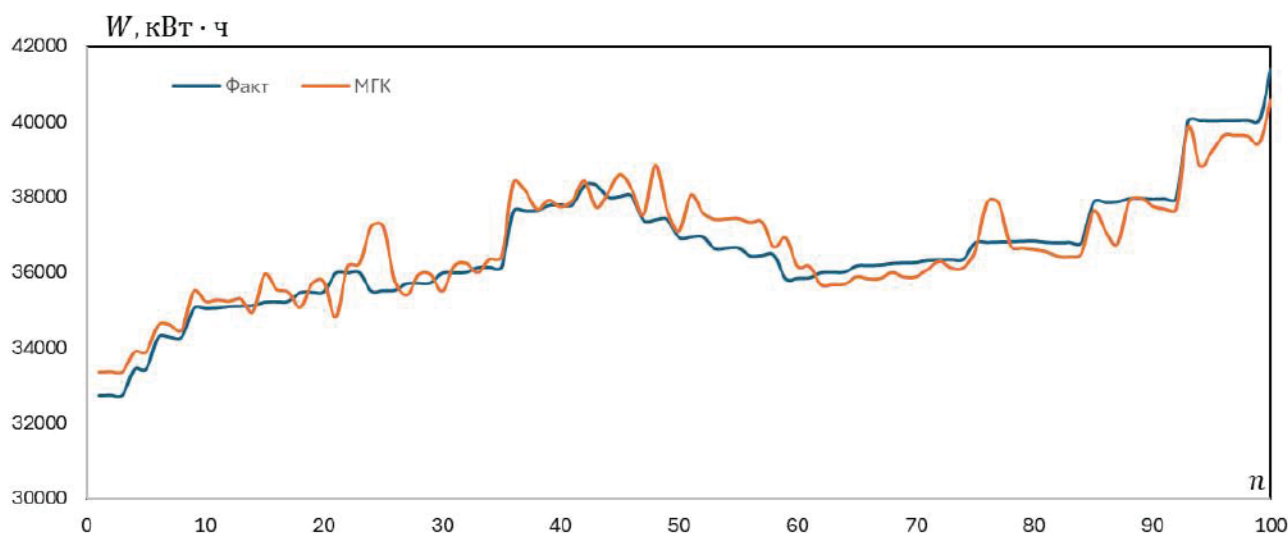


Рис. 2. Сравнение фактических и прогнозных величин электропотребления
Fig. 2. Comparison of actual and predicted values of power consumption

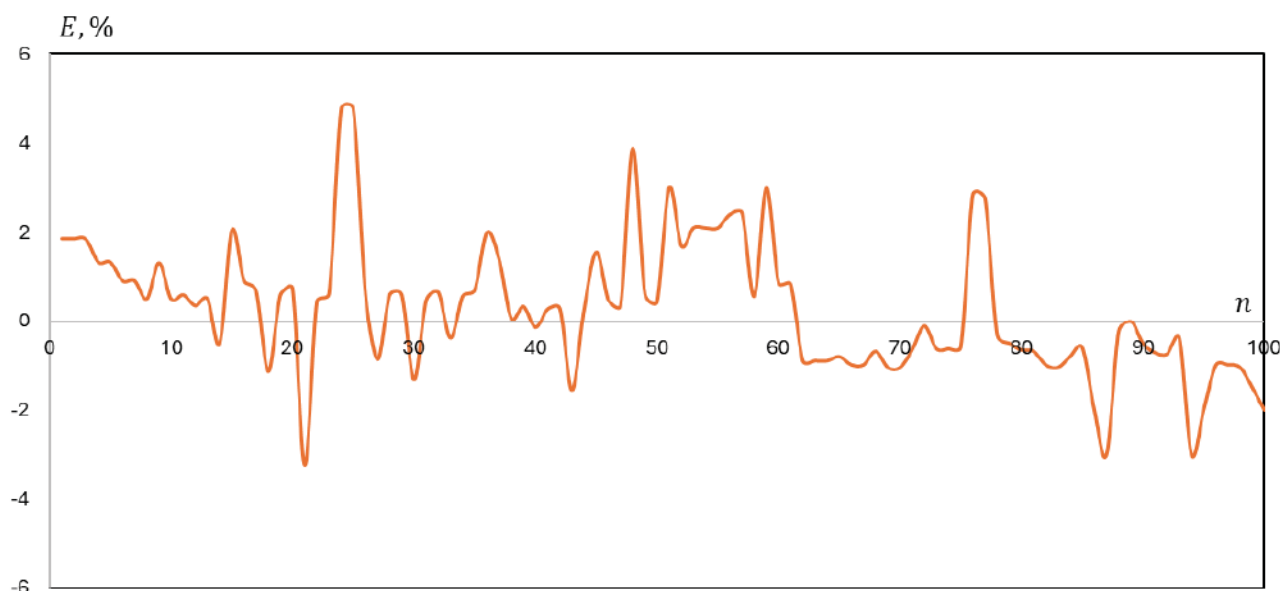


Рис. 3. Разница между фактическими и прогнозными величинами электропотребления
Fig. 3. Difference between actual and predicted electricity consumption values

Для более глубокого исследования адекватности разработанной модели прогнозирования электропотребления осуществлены расчеты погрешности моделей по ошибкам MAD, MAPE, SA, AMSE, BIAS, MAE, RMSE, MSE. Результаты приведены в табл. 4. Результаты приведенных ошибок показывают, что разработанная модель имеет минимальные значения.

Известно, что для получения модели прогнозирования с наименьшей погрешностью прогнозные и фактические значения сравниваются несколько раз, и на этом основании выбирается модель для прогнозирования. Исходя из этого, результат модели с наименьшей погрешностью приведен на рис. 3.

Заключение

Применение метода РСА позволяет уменьшить количество обрабатываемых данных и повысить точность прогнозов. Разработанная модель прогнозирования электропотребления, основанная на РСА, продемонстрировала высокую предсказательную способность, объясняя 69,65 % общей дисперсии первой главной компонентой и 17,28 % – второй, что в совокупности составляет почти 87 % общей дисперсии. Модель обеспечивает хорошее соответствие между фактическими и прогнозными значениями электропотребления (средняя величина ошибки находится в пределах от +3 до –5 %).

Таблица 4. Фактические и прогнозные величины электропотребления

Table 4. Actual and forecast values of electrical consumption

№ процесса плавок Melting process no.	W _ф /W _{ac}	W _{пор} /W _{for}	MAD	MAPE	SA	AMSE	BIAS	MAE	RMSE	MSE
	кВт·ч kW·h									
1	32748	33357,1	609,1128	1,86	1,17405	3,4596	0,249672	0,003208581	0,004141	2,29587525
5	33452	33898,6	446,5842	1,335		1,782225				
10	35050	35228,8	178,755	0,51		0,2601				
15	35218	35940	721,969	2,05		4,2025				
20	35492	35758,2	266,19	0,75		0,5625				
25	35526	37220,6	1694,5902	4,77		22,7529				
30	35980	35505,1	474,936	1,32		1,7424				
35	36152	36412,3	260,2944	0,72		0,5184				
40	37790	37733,3	56,685	0,15		0,0225				
45	38008	38600,9	592,9248	1,56		2,4336				
50	36928	37090,5	162,4832	0,44		0,1936				
55	36654	37423,7	769,734	2,1		4,41				
60	35848	36159,9	311,8776	0,87		0,7569				
65	36176	35883	293,0256	0,81		0,6561				
70	36262	35881,2	380,751	1,05		1,1025				
75	36788	36567,3	220,728	0,6		0,36				
80	36838	36594,9	243,1308	0,66		0,4356				
85	37850	37622,9	227,1	0,6		0,36				
90	37946	37752,5	193,5246	0,51		0,2601				
95	40008	39207,8	800,16	2		4				
100	41388	40560,2	827,76	2	4					

Тем не менее выявленные расхождения в некоторых временных интервалах указывают на необходимость дальнейшего совершенствования модели. Дальнейшие исследования целесообразно направить на учет дополнительных факторов и улучшение алгоритмов для повышения точности прогнозов. В целом метод PCA доказал свою эффективность и

перспективность для решения задач прогнозирования электропотребления, обеспечивая значительные преимущества в сравнении с традиционными методами. Применение этой методики позволяет промышленным предприятиям оптимизировать использование энергии, снизить затраты и повысить конкурентоспособность производимой ими продукции.

СПИСОК ЛИТЕРАТУРЫ

1. Кирпичникова И.М., Саплин Л.А., Саломохо К.Л. Прогнозирование объемов потребления электроэнергии // Вестник ЮУрГУ. Серия “Энергетика”. – 2014. – Т. 14. – № 2. – С. 16–21.
2. Improved estimation of electricity demand function by using artificial neural network, principal component analysis and data envelopment analysis / A. Kheirkhah, A. Azadeh, M. Saberi, A. Azaron, H. Shakouri // Computers & Industrial Engineering. – 2013. – Vol. 64. – № 1. – P. 425–441. DOI: <https://doi.org/10.1016/j.cie.2012.09.017>
3. A novel analysis and forecast method of electricity business expanding based on seasonal adjustment / Y. Zhang, X. Han, G. Yang, Y. Wang, L. Zhang, X. Miao // Asia-Pacific Power and Energy Engineering Conference (APPEEC). IEEE PES-2016. – 2016. – P. 707–711. DOI: <https://doi.org/10.1109/APPEEC.2016.7779608>
4. Taylor J.W., McSharry P.E. Short-term load forecasting methods: an evaluation based on European data // IEEE Transactions on Power Systems. – 2007. – Vol. 22. – № 4. – P. 2213–2219. DOI: <https://doi.org/10.1109/TPWRS.2007.907583>
5. Deep learning with long short-term memory neural networks combining wavelet transform and principal component analysis for daily urban water demand forecasting / B. Du, Q. Zhou, J. Guo, S. Guo, L. Wang // Expert Systems with Applications. – 2021. – Vol. 171. – № 114571. DOI: <https://doi.org/10.1016/j.eswa.2021.114571>
6. Ушаков В.Я., Рахмонов И.У., Жалилова Д.А. Прогнозирование электропотребления текстильными предприятиями на основе метода главных компонент // Проблемы науки. – 2022. – Т. 5. – № 73. – С. 22–26.
7. Analyzing and forecasting electrical load consumption in healthcare buildings / R. Gordillo-Orquera, L.M. Lopez-Ramos, S. Muñoz-Romero, P. Iglesias-Casarrubios, D. Arcos-Avilés, A.G. Marques, J.L. Rojo-Alvarez // Energies. – 2018. – Vol. 11. – № 493. DOI: <https://doi.org/10.3390/en11030493>
8. Analysis of user electricity consumption behavior based on density peak clustering with shared neighbors and attractiveness / Q. Li, G. Wang, Y. Zhang, Q. Yang // Concurrency and Computation: Practice & Experience. – 2023. – Vol. 35. – № 3. DOI: <https://doi.org/10.1002/cpe.7518>

9. Short term load forecasting in the industry for establishing consumption baselines: a French case / J. Blancarte, M. Batton-Hubert, X. Bay, M.A. Girard, A. Grau // Modeling and stochastic learning for forecasting in high dimensions / Eds. A. Antoniadis, J.M. Poggi, X. Brossat. – 2015. – Vol. 217 – P. 1–20. DOI: https://doi.org/10.1007/978-3-319-18732-7_1
10. Kesornsit W., Sirisathitkul Y. Hybrid machine learning model for electricity consumption prediction using random forest and artificial neural networks // Applied Computational Intelligence and Soft Computing. – 2022. – № 1562942. DOI: <https://doi.org/10.1155/2022/1562942>
11. Yuce B., Mourshed M., Rezgui Y. A smart forecasting approach to district energy management // Energies. – 2017. – Vol. 10. – № 1073. DOI: <https://doi.org/10.3390/en10081073>
12. Forecasting building energy consumption based on hybrid PSO-ANN prediction model / C. Hu, K. Li, G. Liu, L. Pan // 34th Chinese Control Conference (CCC). – 2015. – P. 8243–8247. DOI: <https://doi.org/10.1109/ChiCC.2015.7261053>
13. Golmohamadi H., Keypour R. A bi-level robust optimization model to determine retail electricity price in presence of a significant number of invisible solar sites // Sustainable Energy Grids & Networks. – 2018. – Vol. 13. – P. 93–111. DOI: <https://doi.org/10.1016/j.segan.2017.12.008>
14. Online clustering based fault data detection method for distributed PV sites / S. Wang, F. Gao, J. Wu, C. Zheng, X. Fu, F. Duan // Proceedings of the 39th Chinese Control Conference (CCC). – 2020. – P. 4341–4346. DOI: <https://doi.org/10.1109/ChiCC.2020.9292541>
15. Prediction of coal feeding during sintering in a rotary kiln based on statistical learning in the phase space / X. Zhang, L. Zhang, H. Chen, B. Dai // ISA Transactions. – 2018. – Vol. 83. – P. 248–260. <https://doi.org/10.1016/j.isatra.2018.09.015>
16. Прогнозирование электропотребления с помощью нейронных сетей с LSTM / И.У. Рахмонов, В.Я. Ушаков, Н.Н. Ниёзов, Н.Н. Курбонов // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2023. – Т. 334. – № 12. – С. 125–133. DOI: 10.18799/24131830/2023/12/4407
17. Математическое моделирование минимизации расходов электроэнергии промышленными предприятиями с непрерывным характером производства / И.У. Рахмонов, В.Я. Ушаков, А.М. Нажимова, К.К. Обидов, С.Р. Сулейманов // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2024. – Т. 335. – № 4. – С. 43–51. DOI: 10.18799/24131830/2024/4/4423
18. Rakhmonov I.U., Reymov K.M. Regularities of change of energy indicators of the basic technological equipment of the cotton cleaning industry // Journal of Physics: Conference Series. APITECH-2019. – 2019. DOI: 10.1088/1742-6596/1399/5/055038.
19. Analysis and calculation of optimum parameters of electric arc furnace / I.U. Rakhmonov, K.M. Reymov, A.M. Najimova, V.Ya. Ushakov, B.T. Seytmuratov // Journal of Physics: Conference Series. APITECH-2019. – 2019. DOI: 10.1088/1742-6596/1399/5/055048.
20. Taslimov A.D., Rakhmonov I.U. Optimization of complex parameters of urban distribution electric networks // Journal of Physics: Conference Series. APITECH-2019. – 2019. DOI: 10.1088/1742-6596/1399/5/055046.
21. Research on the ablation resistance of TiC particle-reinforced aluminium-based composite coatings on armature surface / C. Fan, L. Zhang, N.N. Kurbonov, I.U. Rakhmonov, G. Wang // Coatings. – 2024. – V. 14. – № 549. DOI: <https://doi.org/10.3390/coatings14050549>
22. Effect of heat treatment on structure of carbon shell-encapsulated pt nanoparticles for fuel cells / K. Davletbaev, S.S. Chougule, J. Min, K. Ko, Y. Kim, H. Choi, Y. Choi, A.A. Chavan, B. Pak, I.U. Rakhmonov et al. // Nanomaterials. – 2024. – Vol. 14. – № 924. DOI: <https://doi.org/10.3390/nano14110924>
23. Wang J. et al. Trap Distribution and Along-surface Discharge Characterization of Aromatic Compound-modified Silicone Gel // IEEE Transactions on Dielectrics and Electrical Insulation. – 2024. DOI: 10.1109/TDEI.2024.3456096

Информация об авторах

Икромжон Усмонович Рахмонов, доктор технических наук, профессор, заведующий кафедрой электро-снабжения Ташкентского государственного технического университета, Узбекистан, 100095, г. Ташкент, ул. Университетская, 2. ilider1987@yandex.ru, <http://orcid.org/0000-0003-2076-5919>

Нуъмон Низомиддинович Ниёзов, доктор философии по техническим наукам, доцент кафедры электро-снабжения Ташкентского государственного технического университета, Узбекистан, 100095, г. Ташкент, ул. Университетская, 2. nomon.niyozov_2422@mail.ru, <http://orcid.org/0000-0002-1031-3460>

Василий Яковлевич Ушаков, заслуженный деятель науки и техники РФ, доктор технических наук, профессор Инженерной школы энергетики Национального исследовательского Томского политехнического университета, Россия, 634050, г. Томск, пр. Ленина, 30. vyush@tpu.ru, <https://orcid.org/0000-0003-2931-2086>

Нурбек Нурулло угли Курбонов, доктор философии по техническим наукам, доцент кафедры электро-снабжения Ташкентского государственного технического университета, Узбекистан, 100095, г. Ташкент, ул. Университетская, 2. nurbek.kurbonov.96@gmail.com, <http://orcid.org/0009-0006-5400-3532>

Айсулу Махмудовна Нажимова, доктор философии по техническим наукам, доцент, доцент кафедры электроэнергетики Каракалпакского государственного университета, Узбекистан, 230100, г. Нукус, ул. Ч. Абдирова, 1. a_najimova@karsu.uz, <https://orcid.org/0009-0001-7336-8362>

Поступила в редакцию: 11.06.2024

Поступила после рецензирования: 30.10.2024

Принята к публикации: 28.11.2024

REFERENCES

1. Kirpichnikova I.M., Saplin L.A., Salomoho K.L. Forecasting volumes of electricity consumption. *Bulletin of SUSU. Series "Energy"*, 2014, vol. 14, no. 2, pp. 16–21. (In Russ.)
2. Kheirkhah A., Azadeh A., Saberi M., Azaron A., Shakouri H. Improved estimation of electricity demand function by using artificial neural network, principal component analysis and data envelopment analysis. *Computers & Industrial Engineering*, 2013, vol. 64, no. 1, pp. 425–441. DOI: <https://doi.org/10.1016/j.cie.2012.09.017>
3. Zhang Y., Han X., Yang G., Wang Y., Zhang L., Miao X. A novel analysis and forecast method of electricity business expanding based on seasonal adjustment. *Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, 2016, pp. 707–711. DOI: <https://doi.org/10.1109/APPEEC.2016.7779608>
4. Taylor J.W., McSharry P.E. Short-term load forecasting methods: an evaluation based on European data. *IEEE Transactions on Power Systems*, 2007, vol. 22, no. 4, pp. 2213–2219. DOI: <https://doi.org/10.1109/TPWRS.2007.907583>
5. Du B., Zhou Q., Guo J., Guo S., Wang L. Deep learning with long short-term memory neural networks combining wavelet transform and principal component analysis for daily urban water demand forecasting. *Expert Systems with Applications*, 2021, vol. 171, no. 114571. DOI: <https://doi.org/10.1016/j.eswa.2021.114571>
6. Ushakov V.Ya., Rakhmonov I.U., Jalilova D.A. Forecasting electricity consumption by textile enterprises based on the principal component method. *Problems of science*, 2022, vol. 5, no. 74, pp. 22–26. (In Russ.)
7. Gordillo-Orquera R., Lopez-Ramos L.M., Muñoz-Romero S., Iglesias-Casarrubios P., Arcos-Avilés D., Marques A.G., Rojo-Alvarez J.L. Analyzing and forecasting electrical load consumption in healthcare buildings. *Energies*, 2018, vol. 11, no. 493. DOI: <https://doi.org/10.3390/en11030493>
8. Li Q., Wang G., Zhang Y., Yang Q. Analysis of user electricity consumption behavior based on density peak clustering with shared neighbors and attractiveness. *Concurrency and Computation: Practice & Experience*, 2023, vol. 35, no. 3. DOI: <https://doi.org/10.1002/cpe.7518>
9. Blancarte J., Batton-Hubert M., Bay X., Girard M.A., Grau A. Short term load forecasting in the industry for establishing consumption baselines: a French case. *Modeling and stochastic learning for forecasting in high dimensions*. Eds. J.M. Antoniadis, X. Poggi, A. Brossat. 2015, vol. 217, pp. 1–20. DOI: https://doi.org/10.1007/978-3-319-18732-7_1
10. Kesomsit W., Sirisathitkul Y. Hybrid machine learning model for electricity consumption prediction using random forest and artificial neural networks. *Applied Computational Intelligence and Soft Computing*, 2022, no. 1562942. DOI: <https://doi.org/10.1155/2022/1562942>
11. Yuce B., Mourshed M., Rezgui Y. A smart forecasting approach to district energy management. *Energies*, 2017, vol. 10, no. 1073. DOI: <https://doi.org/10.3390/en10081073>
12. Hu C., Li K., Liu G., Pan L. Forecasting building energy consumption based on hybrid PSO-ANN prediction model. *34th Chinese Control Conference (CCC)*, 2015, pp. 8243–8247. DOI: <https://doi.org/10.1109/ChiCC.2015.7261053>
13. Golmohamadi H., Keypour R. A bi-level robust optimization model to determine retail electricity price in presence of a significant number of invisible solar sites. *Sustainable Energy Grids & Networks*, 2018, vol. 13, pp. 93–111. DOI: <https://doi.org/10.1016/j.segan.2017.12.008>
14. Wang S., Gao F., Wu J., Zheng C., Fu X., Duan F. Online clustering based fault data detection method for distributed PV sites. *Proceedings of the 39th Chinese Control Conference (CCC)*, 2020, pp. 4341–4346. DOI: <https://doi.org/10.1109/ChiCC.2020.9292541>
15. Zhang X., Zhang L., Chen H., Dai B. Prediction of coal feeding during sintering in a rotary kiln based on statistical learning in the phase space. *ISA Transactions*, 2018, vol. 83, pp. 248–260. DOI: <https://doi.org/10.1016/j.isatra.2018.09.015>
16. Rakhmonov I.U., Ushakov V.Ya., Niyozov N.N., Kurbonov N.N. Forecasting electricity consumption by LSTM neural network. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, 2023, vol. 334, no. 2, pp. 125–133. (In Russ.) Available at: <https://doi.org/10.18799/24131830/2023/12/4407> (accessed 15 September 2023).
17. Rakhmonov I.U., Ushakov V.Ya., Najimova A.M., Obidov K.K., Suleimanov S.R. Mathematical modeling of minimization of electricity consumption by industrial enterprises with continuous production. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, 2024, vol. 335, no. 4, pp. 43–51. (In Russ.) Available at: <https://doi.org/10.18799/24131830/2024/4/4423> (accessed 15 September 2023).
18. Rakhmonov I.U., Reymov K.M. Regularities of change of energy indicators of the basic technological equipment of the cotton cleaning industry. *Journal of Physics: Conference Series. APITECH-2019*. DOI: 10.1088/1742-6596/1399/5/055038.
19. Rakhmonov I.U., Reymov K.M., Najimova A.M., Ushakov V.Ya., Seytmuratov B.T. Analysis and calculation of optimum parameters of electric arc furnace. *Journal of Physics: Conference Series. APITECH-2019*. DOI: 10.1088/1742-6596/1399/5/055048.
20. Taslimov A.D., Rakhmonov I.U. Optimization of complex parameters of urban distribution electric networks. *Journal of Physics: Conference Series. APITECH-2019*. DOI: 10.1088/1742-6596/1399/5/055046.
21. Fan C., Zhang L., Kurbonov, N.N., Rakhmonov I.U., Wang G. Research on the ablation resistance of TiC particle-reinforced aluminium-based composite coatings on armature surface. *Coatings*, 2024, vol. 14, no. 549. DOI: <https://doi.org/10.3390/coatings14050549>
22. Davletbaev K., Chougule S.S., Min J., Ko K., Kim Y., Choi H., Choi Y., Chavan A.A., Pak B., Rakhmonov I.U. Effect of heat treatment on structure of carbon shell-encapsulated Pt nanoparticles for fuel cells. *Nanomaterials*, 2024, vol. 14, no. 924. DOI: <https://doi.org/10.3390/nano14110924>
23. Wang J. Trap distribution and along-surface discharge characterization of aromatic compound-modified silicone gel. *IEEE Transactions on Dielectrics and Electrical Insulation*, 2024. DOI: 10.1109/TDEI.2024.3456096

Information about the authors

Ikromjon U. Rakhmonov, Dr. Sc., Professor, Tashkent State Technical University, 2, Universitetskaya street, Tashkent, 100095, Uzbekistan. ilider1987@yandex.ru, <http://orcid.org/0000-0003-2076-5919>

Numon N. Niyozov, PhD, Associate Professor, Tashkent State Technical University, 2, Universitetskaya street, Tashkent, 100095, Uzbekistan. nomon.niyozov_2422@mail.ru, <http://orcid.org/0000-0002-1031-3460>

Vasily Ya. Ushakov, Dr. Sc., Professor, National Research Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russian Federation. vyush@tpu.ru, <https://orcid.org/0000-0003-2931-2086>

Nurbek N. Kurbonov, PhD, Associate Professor, Tashkent State Technical University, 2, Universitetskaya street, Tashkent, 100095, Uzbekistan. nurbek.kurbonov.96@gmail.com, <http://orcid.org/0009-0006-5400-3532>

Aysulu M. Najimova, PhD., Associate Professor, Karakalpak State University, 1, Ch. Abdirov street, Nukus, 230100, Uzbekistan. a_najimova@karsu.uz, <https://orcid.org/0009-0001-7336-8362>

Received: 11.06.2024

Revised: 30.10.2024

Accepted: 28.11.2024