



# Elektr energiyasi iste'molini uzoq muddatli prognozlash uchun dastlabki ma'lumotlarni tasdiqlash

Kaxraman R. Allayev<sup>1</sup>, Toxir F. Maxmudov<sup>1,a)</sup>, Denis Y. Losev<sup>2</sup>

<sup>1</sup> Akademik, DSc, prof., Toshkent davlat texnika universiteti, Toshkent, 100095, O'zbekiston; [kahramon-allaev@mail.ru](mailto:kahramon-allaev@mail.ru) <https://orcid.org/0000-0003-0056-8854>

<sup>1,a)</sup> PhD, dots., Toshkent davlat texnika universiteti, Toshkent, 100095, O'zbekiston; [t.maxmudov@tdtu.uz](mailto:t.maxmudov@tdtu.uz) <https://orcid.org/0000-0002-0357-0303>

<sup>3</sup> ERHB bosh mutaxassisi, Energetika vazirligi huzuridagi "Milliy dispecherlik markazi" DUK, Toshkent, 100000, O'zbekiston; [d.losev@gubkin.uz](mailto:d.losev@gubkin.uz) <https://orcid.org/0009-0008-5712-9522>

**Dolzarbli:** elektr energetikasining hozirgi kundagi rivojlanishi iste'molchilarning talabdagi noaniqlikning oshishi, taqsimlangan energiya manbalarining ulushining ko'payishi hamda yangi raqamli va aqlli texnologiyalarning joriy etilishi bilan kechmoqda. Bunday sharoitda elektr energiyasiga bo'lgan uzoq muddatli talabni prognozlash strategik rejalashtirish, investitsiya qarorlarini asoslash, yangi energiya tizimlarini loyihalash hamda barqaror energiya rivojlanish strategiyasini shakllantirish uchun muhim vositaga aylanadi. Prognoz aniqligi tahlil va modellashtirish jarayonlarida foydalaniladigan dastlabki ma'lumotlarning sifatiga to'g'ridan-to'g'ri bog'liqdir. Avtomatlashtirilgan hisobga olish tizimlaridan olingan haqiqiy iste'mol ma'lumotlari ko'pincha anomal qiymatda uchrashi, ma'lumotlarning yetishmasligi va texnik hamda tashkiliy sabablar bilan bog'liq tizimli buzilishlarni o'z ichiga oladi. Ma'lumotlarni tasdiqlash bosqichini e'tiborsiz qoldirish prognoz modellari buzilishiga olib kelishi mumkin. Bu esa quvvat ehtiyojlari, energiya ishlab chiqarish va yetkazib berish hajmlari, shuningdek, tarmoq infratuzilmasi hisoblarida xatoliklarga sabab bo'ladi. Shu sababli, dastlabki ma'lumotlarni tasdiqlash uzoq muddatli talabni prognozlashning eng muhim bosqichi hisoblanadi va hisob-kitoblarning ishonchiligi hamda axborot to'liqligini ta'minlaydi. O'zgarishlarga chidamli statistik usullar, xususan, interkvartil oralig'i tahlilidan foydalanish prognoz ssenariylarining ishonchiligi sezilarli darajada oshirishga va energetika sohasida noto'g'ri strategik qarorlar qabul qilish xavfini minimallashtirishga imkon beradi.

**Maqsad:** elektr energiyasini uzoq muddatli prognozlashga mo'ljallangan matematik modellarni ishlab chiqishda foydalanish uchun dastlabki ma'lumotlarni tasdiqlashdir.

**Usullari:** interkvartil oralig'i statistik usuli qo'llanildi.

**Natijalar:** interkvartil oralig'i (IQR) usulining qo'llanilishi elektr energiyasi, harorat, aholi soni va YaIM bo'yicha dastlabki ma'lumotlarda hech qanday chetlanmalar (anomalialar) yo'qligini ko'rsatdi. Bu ma'lumotlarning yuqori sifatli va barqaror ekanligini tasdiqlaydi hamda energiya iste'molini uzoq muddatli prognozlash bo'yicha ishonchli modellarni yaratish uchun mustahkam asos bo'lib xizmat qiladi.

**Kalit so'zlar:** uzoq muddatli prognozlash, elektr energiyasi iste'moli, interkvartil tahlil, elektr energetika tizimi, demografik omil, iqtisodiy omil, iqlimiy omil.

## Верификация исходных данных для долгосрочного прогнозирования потребления электроэнергии

Кахраман Р. Аллаев<sup>1</sup>, Тохир Ф. Махмудов<sup>1,a)</sup>, Денис Ю. Лосев<sup>2</sup>

<sup>1</sup> Академик, DSc, проф., Ташкентский государственный технический университет, Ташкент, 100095, Узбекистан; [kahramon-allaev@mail.ru](mailto:kahramon-allaev@mail.ru) <https://orcid.org/0000-0003-0056-8854>

<sup>1,a)</sup> PhD, доц., Ташкентский государственный технический университет, Ташкент, 100095, Узбекистан; [t.maxmudov@tdtu.uz](mailto:t.maxmudov@tdtu.uz) <https://orcid.org/0000-0002-0357-0303>

<sup>2</sup> Главный специалист ОРЭР ГУП «Национальный диспетчерский центр» при Министерстве энергетики, Ташкент, 100000, Узбекистан; [d.losev@gubkin.uz](mailto:d.losev@gubkin.uz) <https://orcid.org/0009-0008-5712-9522>

**Актуальность:** современное развитие электроэнергетики сопровождается повышением неопределённости в поведении потребителей, ростом доли распределённых источников энергии, а также внедрением новых цифровых и интеллектуальных технологий. В этих условиях долгосрочное прогнозирование спроса на электроэнергию становится важнейшим инструментом стратегического планирования, обоснования инвестиционных решений, проектирования новых энергосистем и формирования устойчивой энергетической стратегии развития. Точность прогноза напрямую зависит от качества исходных данных, используемых в аналитических и моделирующих процедурах. Реальные данные о потреблении, поступающие из автоматизированных систем учета, нередко содержат выбросы, пропуски, систематические искажения, вызванные как техническими, так и организационными причинами. Пренебрежение этапом верификации данных способно привести к искажению прогностических моделей, что, в свою очередь, порождает ошибки в оценке потребностей в мощности, в объемах производства и поставок энергии, а также в расчётах сетевой инфраструктуры. В этой связи верификация исходных данных рассматривается как самый важный этап долгосрочного прогнозирования

**For citation:** K.R. Allaev, T.F. Makhmudov, D.Y. Losev. Verification of initial data for long-term electricity consumption forecasting. Scientific and technical journal of Problems of Energy and Sources Saving, 2025, no. 4, pp. 1-8.

<https://doi.org/10.5281/zenodo.18409623>

Received: 02.04.2025

Revised: 16.04.2025

Accepted: 08.07.2025

Published: 27.12.2025

**Copyright:** © Kakhraman R. Allaev, Tokhir F. Makhmudov, Denis Y. Losev, 2025. Submitted to Problems of Energy and Sources Saving for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).



спроса, обеспечивающий надёжность и информационную состоятельность расчетов. Применение устойчивых статистических методов, таких как межквартильный анализ, позволяет значительно повысить достоверность прогнозных сценариев и минимизировать риски ошибочных стратегических решений в электроэнергетике.

**Цель:** целью исследования является верификация исходных данных для использования в разработке математических моделей долгосрочного прогнозирования потребления электроэнергии.

**Методы:** использован статистический метод межквартильного размаха.

**Результаты:** применение метода межквартильного размаха (IQR) показало отсутствие выбросов в исходных данных по электроэнергии, температуре, населению и ВВП. Это свидетельствует о высоком качестве и стабильности данных, что обеспечивает надёжную основу для построения долгосрочных прогнозных моделей энергопотребления.

**Ключевые слова:** долгосрочное прогнозирование, потребление электроэнергии, межквартильный анализ, электроэнергетическая система, демографический фактор, экономический фактор, климатический фактор.

## Verification of initial data for long-term electricity consumption forecasting

Kakhraman R. Allaev<sup>1</sup>, Tokhir F. Makhmudov<sup>1,a)</sup>, Denis Y. Losev<sup>2</sup>

<sup>1</sup> Academician, DSc, prof., Tashkent State Technical University, Tashkent, 100095, Uzbekistan; [kahramon-allaev@mail.ru](mailto:kahramon-allaev@mail.ru)  
<https://orcid.org/0000-0003-0056-8854>

<sup>1,a)</sup> PhD., assoc. prof, Tashkent State Technical University, Tashkent, 100095, Uzbekistan; [t.maxmudov@tdtu.uz](mailto:t.maxmudov@tdtu.uz)  
<https://orcid.org/0000-0002-0357-0303>

<sup>2</sup> Lead Specialist of the Department of Power Flow Calculations, SUE "National Dispatch Centre" under the Ministry of Energy, Tashkent, 100000, Uzbekistan; [d.losev@gubkin.uz](mailto:d.losev@gubkin.uz) <https://orcid.org/0009-0008-5712-9522>

**Relevance:** modern developments in the electric power industry are accompanied by increasing uncertainty in consumer behavior, a growing share of distributed energy sources, and the implementation of new digital and intelligent technologies. Under these conditions, long-term forecasting of electricity demand becomes a critical tool for strategic planning, substantiation of investment decisions, design of new power systems, and formation of a sustainable energy development strategy. The accuracy of the forecast directly depends on the quality of the initial data used in analytical and modeling procedures. Actual consumption data obtained from automated metering systems often contain outliers, gaps, and systematic distortions caused by both technical and organizational factors. Neglecting the data verification stage can lead to distortions in predictive models, which, in turn, results in errors in assessing capacity needs, energy production and supply volumes, as well as in grid infrastructure calculations. Therefore, verification of initial data is considered the most important stage of long-term demand forecasting, ensuring the reliability and informational completeness of calculations. The application of robust statistical methods, such as interquartile range analysis, allows for a significant improvement in the credibility of forecast scenarios and minimizes the risks of erroneous strategic decisions in the electric power sector.

**Aim:** to verify the initial data for use in the development of mathematical models for long-term electricity consumption forecasting.

**Methods:** there was used the interquartile range statistical method.

**Results:** the application of the interquartile range (IQR) method revealed no outliers in the initial data on electricity consumption, temperature, population, and GDP. This indicates high quality and stability of the data, providing a reliable foundation for developing long-term energy consumption forecasting models.

**Keywords:** long-term forecasting, electricity consumption, interquartile analysis, electric power system, demographic factor, economic factor, climatic factor.

### 1. Введение (Introduction)

Эффективное функционирование электроэнергетических систем в долгосрочной перспективе требует разработки обоснованных прогнозов спроса на электрическую энергию. Такие прогнозы необходимы для планирования генерирующих мощностей, проектирования электросетевой инфраструктуры, формирования инвестиционных программ и выработки государственной энергетической стратегии развития. Одним из важнейших факторов, определяющих точность прогноза, является качество исходных данных, формирующих основу прогнозных модели.

В международной литературе подчёркнута важность предварительной очистки данных с применением робастных методов — межквартильного размаха (interquartile range, IQR), z-оценки, регрессионного контроля выбросов и методов восстановления пропусков с помощью машинного обучения. Так, в [1] демонстрируется эффективность гибридных моделей ARIMA + LSTM для прогнозирования энергопотребления на национальном уровне. Аналогично, в [2], при сравнительном анализе моделей временных рядов, показана высокая точность методов регрессии, ARIMA и SVR для прогнозирования бытового энергопотребления.

Как результат исследований отечественных учёных, в [3] описывается разработанный в среде MATLAB программный инструмент для верификации энергетических данных,



включающий фильтрацию на наличие нулевых, абсурдных, пропущенных или повторяющихся значений.

Данные о фактическом потреблении электроэнергии, получаемые из автоматизированных систем коммерческого учета, как правило, содержат аномальные и некорректные значения, вызванные сбоями оборудования, ошибками передачи, а также человеческим фактором. Без предварительной обработки такие данные способны исказить результаты статистического анализа и машинного обучения, лежащих в основе современных методов прогнозирования.

В связи с этим важнейшим этапом становится верификация и очистка исходных данных, направленная на выявление и исключение выбросов, устранение пропусков и проверку логической согласованности информации. В настоящем исследовании рассматривается применение робастных статистических методов, в частности, метода межквартильного размаха (IQR), для эффективной верификации данных, используемых в задачах долгосрочного прогнозирования спроса.

## 2. Методы и материалы (Methods and materials)

В данном исследовании использованы исходные данные [4-8] по четырем ключевым факторам, влияющим на потребление электроэнергии в долгосрочной перспективе: потребление электроэнергии за текущий год ( $x_1$ ), среднегодовая температура ( $x_2$  – климатический фактор), численность населения ( $x_3$  – демографический фактор) и валовый внутренний продукт (ВВП) ( $x_4$  – экономический фактор). Целью исследования является верификация исходных данных для использования при разработке математических моделей долгосрочного прогнозирования потребления электроэнергии.

Анализ проводился на основе статистического массива данных, включающего информацию за 32 года наблюдений, который показан в таблице 1 [9].

**Таблица 1.** Исходные информация, предназначенные для верификации  
**Table 1.** Initial data for verification purposes

годы	1991	1992	1993	1994	1995	1996	1997	1998
Уп, ТВт·ч	50,851	50,705	49,521	48,337	47,153	47,057	46,961	46,864
t, °C	14,4	14,4	13,4	14,3	15,0	14,0	15,1	14,7
N, млн.чел	20,857	21,354	21,847	22,277	22,684	23,128	23,56	23,954
W, млрд. \$	13,8	13,0	13,1	12,9	13,4	13,9	14,7	15,0
годы	1999	2000	2001	2002	2003	2004	2005	2006
Уп, ТВт·ч	46,768	46,672	47,178	47,683	48,189	48,694	49,200	50,920
t, °C	15,1	15,4	15,7	15,2	14,8	15,9	15,3	15,5
N, млн.чел	24,312	24,65	24,965	25,272	25,568	25,864	26,167	26,488
W, млрд. \$	17,1	13,8	11,4	9,69	10,1	12,0	14,3	17,3
годы	2007	2008	2009	2010	2011	2012	2013	2014
Уп, ТВт·ч	48,950	49,400	49,950	51,700	52,400	52,500	54,200	55,400
t, °C	15,3	15,1	15,1	15,8	14,9	14,5	15,8	14,4
N, млн.чел	26,868	27,303	27,767	28,001	29,123	29,555	29,993	30,493
W, млрд. \$	22,3	29,5	33,7	49,8	60,2	67,5	73,2	80,8
годы	2015	2016	2017	2018	2019	2020	2021	2022
Уп, ТВт·ч	57,280	58,319	54,634	59,975	60,843	64,228	70,535	74,096
t, °C	15,9	16,5	15,6	15,7	16,4	15,2	16,7	16,5
N, млн.чел	31,023	31,575	31,121	32,567	33,256	33,905	34,559	35,271
W, млрд. \$	86,2	86,1	69,7	58,7	67,3	66,4	77,3	90,1

Структура исходных данных включает ежегодные значения с 1991 по 2022 год для каждой



из переменных:

x1 (потребление электроэнергии) – входной параметр  $Y_n$ , характеризующий годовое потребление электроэнергии энергосистемы, измеренное в Тераватт-часах (ТВт·ч);

x2 (среднегодовая температура) – входной параметр  $t$ , характеризующий среднегодовую температуру воздуха страны в градусах Цельсия;

x3 (численность населения) – входной параметр  $N$ , характеризующий данные о численности населения страны, измеренные в миллионах человек;

x4 (ВВП) – входной параметр  $W$ , характеризующий данные о валовом внутреннем продукте страны, измеренные в миллиардах долларов США.

Исследование включает следующие процедуры:

**1-шаг. Исключение нулевых значений (нулевых данных).**

Данная процедура направлена на удаление записей, в которых значения ключевых параметров равны нулю, что в ряде случаев может свидетельствовать о сбоях измерительного оборудования, ошибках ввода или передачи данных.

Нулевые значения, не обусловленные реальными процессами (например, физически невозможное нулевое потребление в рабочие часы), исключаются из анализа с целью повышения достоверности статистических характеристик и прогнозных моделей.

**2-шаг. Выявление и удаление выбросов (ошибочных данных).**

Выбросы представляют собой аномально высокие или низкие значения, выходящие за пределы типичного диапазона данных. Они могут быть результатом ошибок измерения, сбоев системы или единичных нетипичных событий.

Для их идентификации применяются статистические методы, такие как межквартильный размах (IQR), z-оценка или междисперсионный анализ. Обнаруженные выбросы либо удаляются, либо подлежат коррекции в зависимости от контекста задачи.

В данной работе для целей верификации данных по потреблению электрической энергии применяется метод межквартильного размаха - IQR, относящийся к числу устойчивых (робастных) статистических методов.

Метод основан на вычислении первого ( $Q_1$ ) и третьего ( $Q_3$ ) квартилей, которые отражают пороговые значения, ниже которых расположены соответственно 25 % и 75 % элементов выборки. Межквартильный размах определяется как:

$$IQR = Q_3 - Q_1, \quad (1)$$

Данный интервал охватывает центральные 50 % наблюдаемых значений и характеризует область типичного (нормального) потребления электроэнергии без учета экстремальных отклонений. Верификация осуществляется посредством применения критерия Тьюки: все значения, выходящие за пределы интервала:

$$[Q_1 - 1,5 \cdot IQR, Q_3 + 1,5 \cdot IQR], \quad (2)$$

рассматриваются как аномальные, и подлежат дополнительной проверке или исключению из анализа.

Преимуществом данного подхода является его устойчивость к выбросам и отсутствие необходимости в предположении о нормальности распределения данных, что особенно актуально для реальных массивов данных с сезонной и поведенческой составляющей, характерных для энергопотребления.

Применение метода IQR позволяет существенно повысить достоверность результатов анализа, обеспечить корректную агрегацию потребления и формирование обоснованных энергетических профилей. Методология может быть эффективно использована как в локальных, так и в распределённых энергетических системах, включая интеллектуальные платформы мониторинга и учета электроэнергии.

**3-шаг. Удаление абсолютно одинаковых записей (статических по значению).**

Наличие множественных повторяющихся записей с идентичными значениями параметров на протяжении значительного времени может указывать на неработающее оборудование, зависшие сенсоры или ошибки при регистрации данных.

Такие «статические» участки искажают динамическую структуру данных, занижают дисперсию и влияют на адекватность прогнозных моделей. Повторяющиеся записи исключаются для сохранения информативности и вариативности выборки.

**4-шаг. Восстановление утраченных данных (при наличии пропусков).**

Пропущенные значения (отсутствующие значения или пустые поля) могут возникать в результате сбоев при передаче данных, отключения оборудования или ошибок в базах данных.

В зависимости от объема и характера пропусков применяются методы их восстановления: линейная интерполяция, скользящее среднее, регрессионное восстановление или использование медианных значений. Это позволяет сохранить полноту выборки и избежать искажения

результатов анализа.

### 3. Результаты и обсуждение (Results and discussion)

#### 1-шаг. Исключение нулевых значений (нулевых данных).

Проверка массива на наличие нулевых значений показала, что ни в одной из переменных отсутствующих или равных нулю значений не обнаружено. Следовательно, этап восстановления отсутствующих данных через интерполяцию и экстраполяцию не потребовался.

#### 2-шаг. Выявление и удаление выбросов (ошибочных данных).

Для выявления выбросов в исследуемых данных используем метод межквартильного размаха (IQR), который широко применяется при очистке числовых временных рядов. Выбросом считается значение, выходящее за пределы:

$$\begin{aligned} \text{нижняя граница} &= Q_1 - 1,5 \cdot IQR, \\ \text{верхняя граница} &= Q_3 + 1,5 \cdot IQR; \end{aligned}$$

где:

$Q_1$  — первый квартиль (25%),

$Q_3$  — третий квартиль (75%),

$IQR = Q_3 - Q_1$ .

а). Проверим на выявление выбросов параметра  $x_1$ :

$Q_1 = 47,153$ ;

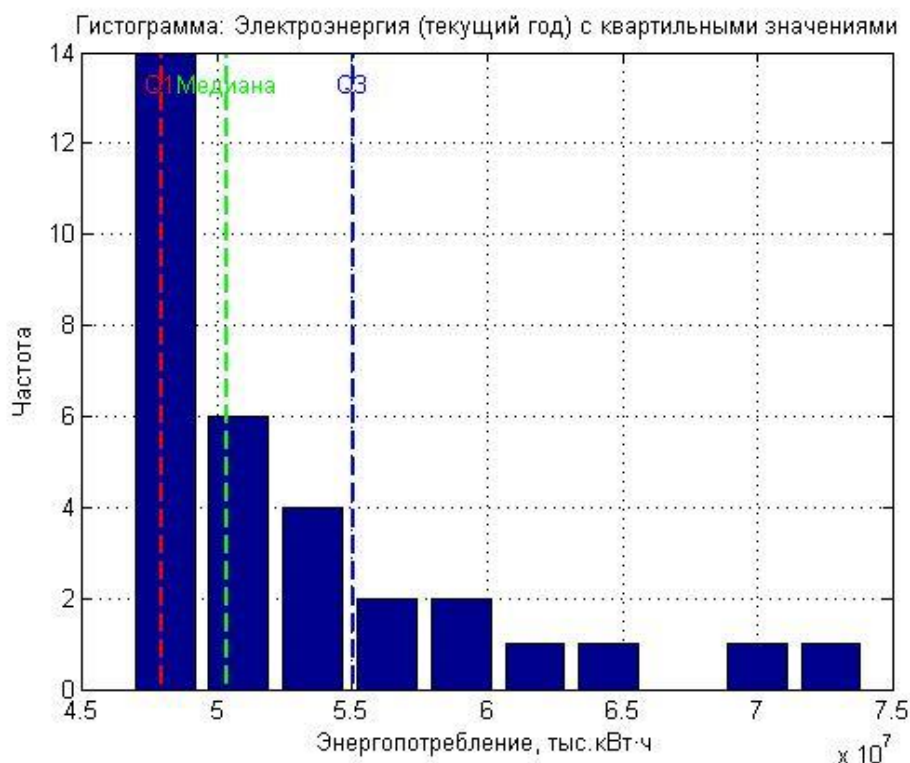
$Q_3 = 57,280$ ;

$IQR = 57,280 - 47,153 = 10,127$ ;

нижняя граница  $= 47,153 - 1,5 \cdot 10,127 = 32,0725$ ;

верхняя граница  $= 57,280 + 1,5 \cdot 10,127 = 72,3605$ .

**Вывод:** все значения  $x_1$  находятся в пределах  $[32,0725; 72,3605]$ . **Выбросов нет.**



**Рис. 1.** Гистограмма потребления электроэнергии (текущий год) с квартильным анализом  
**Fig. 1.** Histogram of Electricity Consumption with Quartile Analysis

б). Проверим на выявление выбросов параметра  $x_2$ :

$Q_1 = 14,4$ ;

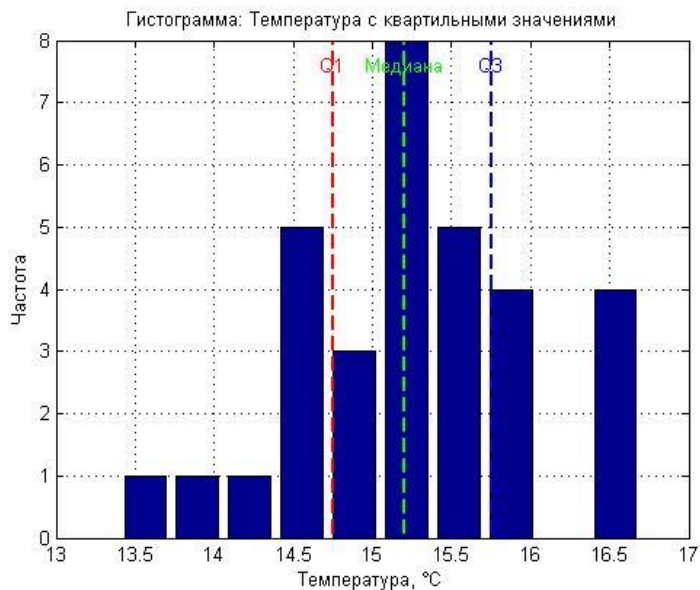
$Q_3 = 15,9$ ;

$IQR = 15,9 - 14,4 = 1,5$ ;

нижняя граница  $= 14,4 - 1,5 \cdot 1,5 = 12,15$ ;

верхняя граница  $= 15,9 + 1,5 \cdot 1,5 = 18,15$ .

**Вывод:** все значения  $x_2$  находятся в пределах [12,15; 18,15]. **Выбросов нет.**



**Рис. 2.** Гистограмма среднегодовой температуры с квартильным анализом  
**Fig. 2.** Histogram of Average Annual Temperature with Quartile Analysis

в). Проверим на выявление выбросов параметра  $x_3$ :

$$Q_1=24,312;$$

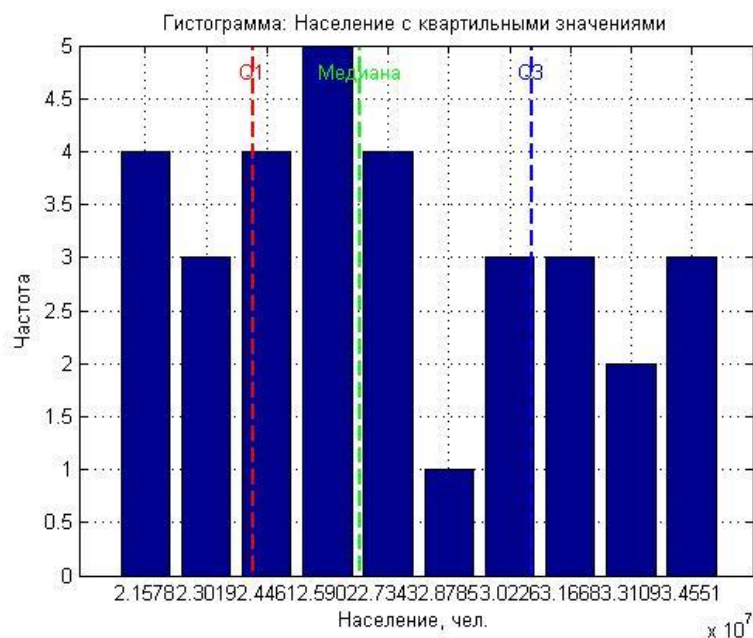
$$Q_3=31,5753;$$

$$IQR = 31,5753-24,312=7,2633;$$

$$\text{нижняя граница}=24,312-1,5 \cdot 7,2633=13,422;$$

$$\text{верхняя граница}=31,5753+1,5 \cdot 7,2633=42,6453.$$

**Вывод:** все значения  $x_3$  находятся в пределах [13,422; 42,6453]. **Выбросов нет.**



**Рис. 3.** Гистограмма населения страны с квартильным анализом  
**Fig. 3.** Histogram of the Country's Population with Quartile Analysis

г). Проверим на выявление выбросов параметра  $x_4$ :

$$Q_1=13,4;$$

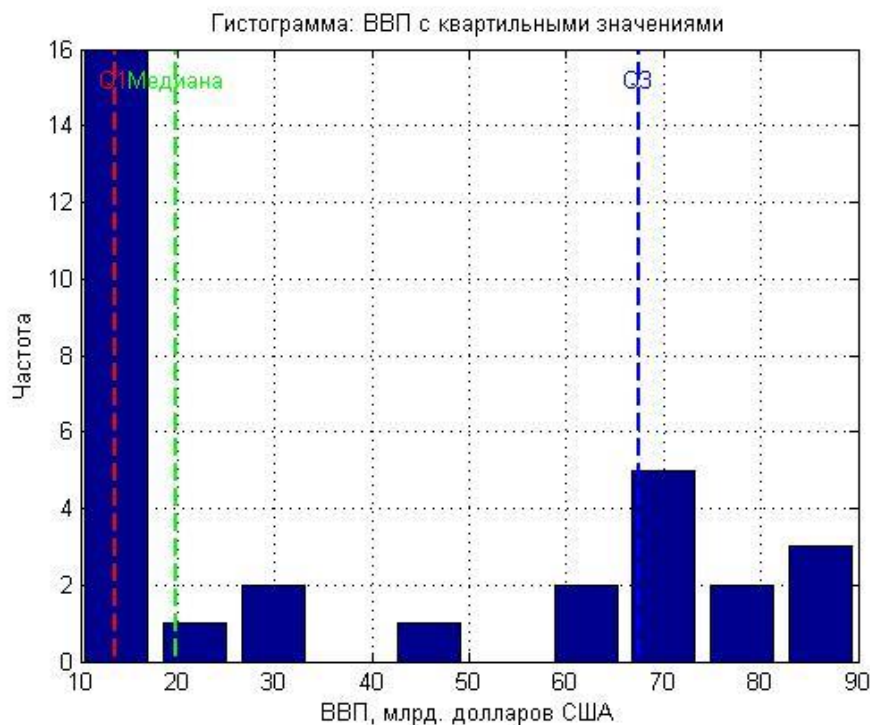
$$Q_3=67,5;$$

$$IQR = 67,5-13,4=54,1;$$

$$\text{нижняя граница}=13,4-1,5 \cdot 54,1=-67,25;$$

верхняя граница= $67,5+1,5\cdot 54,1=148,15$ .

**Вывод:** все значения  $x_i$  находятся в пределах  $[-67,25; 148,15]$ . **Выбросов нет.**



**Рис. 4.** Гистограмма ВВП страны с квартильным анализом

**Fig. 4.** Histogram of the Country's GDP with Quartile Analysis

### 3-шаг. Удаление абсолютно одинаковых записей (статических по значению).

Полный проход по базе показал отсутствие строк, полностью дублирующих друг друга. Все строки уникальны по совокупности признаков. Таким образом, этап удаления статических повторений оказался избыточным.

### 4-шаг. Восстановление утраченных данных (при наличии пропусков).

Отсутствующих значений в анализируемой базе не выявлено. Следовательно, этап восстановления утраченных данных, включая методы линейной, полиномиальной и нейросетевой интерполяции, не проводился.

## 4. Заключение (Conclusion)

В настоящем исследовании рассмотрен процесс верификации исходных данных, используемых для долгосрочного прогнозирования спроса на электроэнергию. Проведён комплексный анализ статистической информации за период 1991–2022 гг., включающей ключевые макроэкономические, демографические и климатические показатели, оказывающие существенное влияние на энергопотребление.

Результаты анализа подтвердили высокое качество исходных данных: в массиве не выявлено нулевых или пропущенных значений, отсутствуют дублирующие строки. Применение метода межквартильного размаха (IQR) для обнаружения выбросов показало отсутствие аномальных значений во всех исследуемых параметрах, что свидетельствует о стабильности и достоверности статистической базы. Робастный характер метода IQR обеспечил устойчивость оценки в условиях возможных сезонных и поведенческих колебаний, присущих временным рядам энергопотребления.

Выборка в целом демонстрирует высокую степень согласованности и логической непротиворечивости.

Статистическая верификация данных подтвердила их пригодность для последующего использования в построении регрессионных и методов машинного обучения, ориентированных на решение задач долгосрочного прогнозирования потребления электрической энергии.



## ЛИТЕРАТУРА

1. Tatiana Gonzalez Grandon, Johannes Schwenzer, Thomas Steens, Julia Breuing. Electricity Demand Forecasting with Hybrid Statistical and Machine Learning Algorithms: Case Study of Ukraine. arXiv preprint arXiv:2304.05174; <https://doi.org/10.48550/arXiv.2304.05174>.
2. Bilal, M., Kim, H., Fayaz, M. и Pawar, P. (2022). Comparative Analysis of Time Series Forecasting Approaches for Household Electricity Consumption Prediction. arXiv preprint arXiv:2207.01019; <https://doi.org/10.48550/arXiv.2207.01019>.
3. Muzaffar Khudayarov, Sarvar Qurbonov. Data verification for forecasting of building energy consumption. ICECAE 2023. E3S Web of Conferences 434, 01003 (2023), [doi.org/10.1051/e3sconf/202343401003](https://doi.org/10.1051/e3sconf/202343401003).
4. Marwen Elkamel, Lily Schleider, Eduardo L. Pasilliao, Ali Diabat and Qipeng P. Zheng, Long-Term Electricity Demand Prediction via Socioeconomic Factors—A Machine Learning Approach with Florida as a Case Study. *Energies* 2020, 13(15), 3996; <https://doi.org/10.3390/en13153996>.
5. Marwen Elkamel, Lily Schleider, Eduardo L. Pasilliao, Ali Diabat and Qipeng P. Zheng, Long-Term Electricity Demand Prediction via Socioeconomic Factors—A Machine Learning Approach with Florida as a Case Study. *Energies* 2020, 13(15), 3996; <https://doi.org/10.3390/en13153996>.
6. Yun BAI, Simon Camal and Andrea Michiorri. Investigation of the Impact of Economic and Social Factors on Energy Demand through Natural Language Processing. June 2024 DOI: 10.48550/arXiv.2406.06641.
7. K.R. Allaev. Prospects for energy development in Uzbekistan until 2035 and beyond. *Problems of energy and sources saving. Tashkent-2023*, pp. 25-35.
8. Sarychev A. Modeling of complex systems in conditions of structural uncertainty: Regression and autoregressive models. 2016 - 284 p.
9. Аллаев К.Р. Электроэнергетика Узбекистана и мира. –Ташкент: Фан ва технология. 2009. - 463 с.

## REFERENCES

1. Demand Forecasting with Hybrid Statistical and Machine Learning Algorithms: Case Study of Ukraine. arXiv preprint arXiv:2304.05174; <https://doi.org/10.48550/arXiv.2304.05174>.
2. Bilal, M., Kim, H., Fayaz, M. и Pawar, P. (2022). Comparative Analysis of Time Series Forecasting Approaches for Household Electricity Consumption Prediction. arXiv preprint arXiv:2207.01019; <https://doi.org/10.48550/arXiv.2207.01019>.
3. Muzaffar Khudayarov, Sarvar Qurbonov. Data verification for forecasting of building energy consumption. ICECAE 2023. E3S Web of Conferences 434, 01003 (2023), [doi.org/10.1051/e3sconf/202343401003](https://doi.org/10.1051/e3sconf/202343401003).
4. Marwen Elkamel, Lily Schleider, Eduardo L. Pasilliao, Ali Diabat and Qipeng P. Zheng, Long-Term Electricity Demand Prediction via Socioeconomic Factors—A Machine Learning Approach with Florida as a Case Study. *Energies* 2020, 13(15), 3996; <https://doi.org/10.3390/en13153996>.
5. Marwen Elkamel, Lily Schleider, Eduardo L. Pasilliao, Ali Diabat and Qipeng P. Zheng, Long-Term Electricity Demand Prediction via Socioeconomic Factors—A Machine Learning Approach with Florida as a Case Study. *Energies* 2020, 13(15), 3996; <https://doi.org/10.3390/en13153996>.
6. Yun BAI, Simon Camal and Andrea Michiorri. Investigation of the Impact of Economic and Social Factors on Energy Demand through Natural Language Processing. June 2024 DOI: 10.48550/arXiv.2406.06641.
7. K.R. Allaev. Prospects for energy development in Uzbekistan until 2035 and beyond. *Problems of energy and sources saving. Tashkent-2023*, pp. 25-35.
8. Sarychev A. Modeling of complex systems in conditions of structural uncertainty: Regression and autoregressive models. 2016 - 284 p.
9. Allaev K.R. Electric power industry of Uzbekistan and the world. - Tashkent: Fan va technology. 2009. - 463 p.