

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	5
1 ТЕОРЕТИЧЕСКАЯ ЧАСТЬ.....	6
1.1 Понятие временного ряда	6
1.2 Прогнозирование временных рядов на основе линейных авторегрессионных моделей	9
2 ПРАКТИЧЕСКАЯ ЧАСТЬ	11
2.1 Анализ временных рядов на примере данных о производстве электроэнергии	11
2.2 Прогнозирование временных рядов и оценка качества	16
ЗАКЛЮЧЕНИЕ	18
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	19
ПРИЛОЖЕНИЯ.....	21
Приложение А	22

ВВЕДЕНИЕ

Энергетический рынок тесно взаимодействует с другими секторами экономики, эффективное управление производственными ресурсами требует точных прогнозов. Разработка программы прогнозирования временных рядов позволит предпринимателям и управленцам в энергетической сфере адаптироваться к изменениям спроса и принимать обоснованные решения о распределении ресурсов.

Цель данной работы заключается в создании инструмента, способного анализировать данные о производстве электроэнергии и прогнозировать их будущее изменение. Путем применения линейных авторегрессионных моделей мы сможем оценить тенденции и паттерны в производстве электроэнергии, что поможет оптимизировать производственные процессы и минимизировать риски нехватки или избытка энергии.

Этот подход к разработке программы прогнозирования временных рядов предоставит энергетическим компаниям и регуляторам инструмент для более точного планирования и управления производством электроэнергии. Точные прогнозы спроса на энергию позволят эффективно использовать ресурсы и обеспечивать стабильное функционирование энергетической системы.

Задачи, решаемые в данной курсовой работе:

- использовать знания математической статистики с использованием современных средств обработки данных: платформы RStudio и языка программирования R;
- описать необходимые формулы и понятия;
- выполнить анализ временных рядов на основе собранных данных;
- оценить качество решения на основе исходного набора данных.

1 ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

1.1 Понятие временного ряда

Временные ряды — это актуальный инструмент, применимый во множестве решений, от предсказания цен на акции, прогнозов погоды, планирования бизнеса, до распределения ресурсов. Несмотря на то, что прогнозирование может быть сведено к построению контролируемой регрессии, существуют особенности, связанные с временным характером наблюдений, которые необходимо учитывать, используя специальные инструменты. [1.1]

Временной ряд — собранный в разные моменты времени статистический материал о значении каких-либо параметров исследуемого процесса. Каждая единица статистического материала называется измерением или отсчётом. [1.2]

Большинство методов исследования временных рядов включает различные способы фильтрации шума, позволяющие увидеть регулярную составляющую более отчетливо. Большинство регулярных составляющих временных рядов принадлежит к двум классам: они являются либо трендом, либо сезонной составляющей.

Тренд представляет собой общую систематическую линейную или нелинейную компоненту, которая может изменяться во времени. Сезонная составляющая — это периодически повторяющаяся компонента. Оба эти вида регулярных компонент часто присутствуют в ряде одновременно. [1.3]

Для определения наличия сезонности можно использовать тест Льюнг–Бокса. Данный тест предназначен для нахождения автокорреляции временных рядов. Вместо тестирования на случайность каждого отдельного коэффициента, он проверяет на отличие от нуля сразу несколько коэффициентов автокорреляции. [1.4]

Тест Льюнг–Бокса может быть определён следующим образом.

Выдвигаются две конкурирующие гипотезы:

- данные являются случайными (нулевая гипотеза);
- данные не являются случайными (альтернативная гипотеза).

Проводится статистическое испытание:

$$\tilde{Q} = n(n + 2) \sum_{k=1}^m \frac{\hat{p}_k^2}{n-k},$$

где n — число наблюдений;

\hat{p}_k^2 — автокорреляция k -го порядка;

m — число проверяемых лагов.

Нулевая гипотеза отвергается и признаётся наличие автокорреляции, если выполняется условие:

$$\tilde{Q} > X_{1-\alpha, m}^2,$$

где $X_{1-\alpha, m}^2$ — квантили распределения хи-квадрат с m степенями свободы.

Временные ряды классифицируются по следующим признакам:

- по количеству показателей, для которых определяются уровни в каждый момент времени: одномерные и многомерные временные ряды;
- по характеру временного параметра: моментные и интервальные временные ряды. В моментных временных рядах уровни характеризуют значения показателя по состоянию на определенные моменты времени. В интервальных рядах уровни характеризуют значение показателя за определенные периоды времени;
- по наличию пропущенных значений: полные и неполные временные ряды;

- по расстоянию между датами и интервалами времени выделяют равноотстоящие, когда даты регистрации или окончания периодов следуют друг за другом с равными интервалами и неполные, когда принцип равных интервалов не соблюдается;
- временные ряды бывают детерминированными и случайными: первые получают на основе значений некоторой неслучайной функции (ряд последовательных данных о количестве дней в месяцах), вторые есть результат реализации некоторой случайной величины;
- в зависимости от наличия основной тенденции выделяют стационарные ряды и нестационарные.

В стационарных временных рядах статистические свойства не зависят от времени, поэтому результат легко предсказать. Большинство статистических методов предполагают, что все временные ряды должны быть стационарными.

В нестационарных временных рядах статистические свойства меняются со временем. Они показывают сезонные эффекты, тренды и другие структуры, которые зависят от временного показателя.

Тест Дики-Фуллера (Augmented Dickey-Fuller) используется для проверки стационарности временного ряда. Стационарность означает, что статистические свойства временного ряда, такие как среднее и дисперсия, остаются постоянными во времени. Этот тест помогает определить, присутствует ли единичный корень в временном ряде, что указывает на нестационарность. [1.5]

В Тесте Дики-Фуллера рассматривается нулевая гипотеза о том, что временной ряд имеет единичный корень (то есть, он нестационарен), против альтернативной гипотезы, что ряд стационарен. Если p -значение теста меньше уровня значимости, нулевая гипотеза отвергается, и можно сделать вывод, что ряд стационарен. Если p -значение больше, то нет достаточных доказательств для того, чтобы считать ряд стационарным, и может потребоваться преобразование данных.

Таким образом, временные ряды являются важным инструментом для множества приложений, от финансового прогнозирования до планирования

бизнеса и распределения ресурсов. Для определения сезонности и автокорреляции используется тест Льюнг-Бокса, который помогает выявить наличие сезонности. Стационарность временного ряда проверяется с помощью теста Дики-Фуллера, который определяет, остаются ли статистические свойства ряда постоянными во времени.

1.2 Прогнозирование временных рядов на основе линейных авторегрессионных моделей

В анализе временных рядов выделяются две основные задачи: задача идентификации и задача прогноза.

Модель ARIMA (AutoRegressive Integrated Moving Average) представляет собой инструмент для анализа временных рядов, который объединяет в себе несколько ключевых компонентов, таких как интегрированная авторегрессия, скользящее среднее, и возможность учета дополнительных внешних факторов. Эта математическая модель широко применяется в различных областях, где необходимо строить прогнозы на основе имеющихся данных. [1.6]

Основным назначением моделей ARIMA в целом является решение задач, связанных с прогнозированием. Они позволяют вычислять будущие значения временного ряда, используя информацию о его предыдущих состояниях.

Модель ARIMA представляет собой интегрированный и гибкий подход к анализу временных рядов, обеспечивая надежные инструменты для прогнозирования и понимания динамики данных в различных областях применения. Процесс авторегрессии — последовательная зависимость элементов временного ряда, выражается следующим уравнением:

$$x(t) = \alpha + \varphi_1 \cdot x(t-1) + \varphi_2 \cdot x(t-2) + \varphi_3 \cdot x(t-3) + \dots + \epsilon,$$

где α — свободный член (константа);

$\varphi_1, \varphi_2, \varphi_3, \dots$ — параметры авторегрессии.

Скользящее среднее представляет собой широкий класс функций, значения которых в каждой точке определения равны некоторому усредненному значению исходной функции за предыдущий период времени. Этот метод анализа данных является мощным инструментом для выявления трендов, устранения шума и выявления паттернов в последовательных значениях временных рядов. [1.7]

Процесс скользящего среднего основывается на том, что каждый элемент временного ряда подвергается суммарному воздействию предыдущих ошибок или изменений.

$$x_t = \mu + \epsilon_t - \theta_1 \cdot \epsilon_{t-1} - \theta_2 \cdot \epsilon_{t-2} - \dots,$$

где μ — константа;

$\theta_1, \theta_2, \theta_3, \dots$ — параметры скользящего среднего;

$\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \dots$ — шумовые компоненты.

Таким образом, модель ARIMA используется для анализа и прогнозирования временных рядов, объединяя в себе компоненты интегрированной авторегрессии и скользящего среднего. Ее основное назначение заключается в вычислении будущих значений временных рядов. Компонент авторегрессии модели описывает зависимость текущих значений ряда от предыдущих, а скользящее среднее помогает выявлять тренды и устранять шум.

Рассмотрим применение анализа временных рядов на примере данных о производстве электроэнергии.

2 ПРАКТИЧЕСКАЯ ЧАСТЬ

2.1 Анализ временных рядов на примере данных о производстве электроэнергии

В практической работе проведём анализ временных рядов на примере данных о производстве электроэнергии.

В качестве исследуемых данных на вход подаются данные о производстве электроэнергии за 2010-2022 года.

Набор данных «Производство электроэнергии за 2010-2022 года» включает в себя следующие столбцы данных:

1. COUNTRY — Название страны.
2. CODE_TIME — Месяц и год в кратком формате.
3. TIME — Месяц и год в полном формате.
4. YEAR — Год.
5. MONTH — Месяц, представленный в виде числа.
6. MONTH_NAME — Название месяца.
7. PRODUCT — Тип энергетического продукта.
8. VALUE — Количество произведенной электроэнергии в гигаватт-часах.
9. DISPLAY_ORDER — Порядок, в котором должны отображаться продукты.
10. yearToDate — Количество выработанной электроэнергии за текущий год до текущего месяца в ГВтч.
11. previousYearToDate — Количество выработанной электроэнергии за предыдущий год до текущего месяца в ГВтч.
12. share — Доля продукта в общем объеме производства электроэнергии в стране в десятичном формате.

Первоначально данные о производстве электроэнергии загружаются из файла energyProduction.csv [2.1]. Данные содержат информацию о месячном производстве электроэнергии в разных странах. После загрузки данных необходимо переименовать столбцы и преобразовать значения производства электроэнергии в числовой формат для дальнейшего анализа.

Далее данные агрегируются по годам и странам, чтобы определить годовое производство электроэнергии для каждой страны [2.2]. Это позволяет увидеть общие тенденции и сравнить производство электроэнергии между разными странами.

После агрегирования данных рассчитывается общее производство электроэнергии для каждой страны за весь период наблюдений. На основе этих данных определяется топ стран с наибольшим производством электроэнергии. Для наглядности строится столбчатая диаграмма, на которой показаны страны, занимающие первые места по производству электроэнергии. На Рисунке 2.1 можно увидеть первые страны по производству электроэнергии.

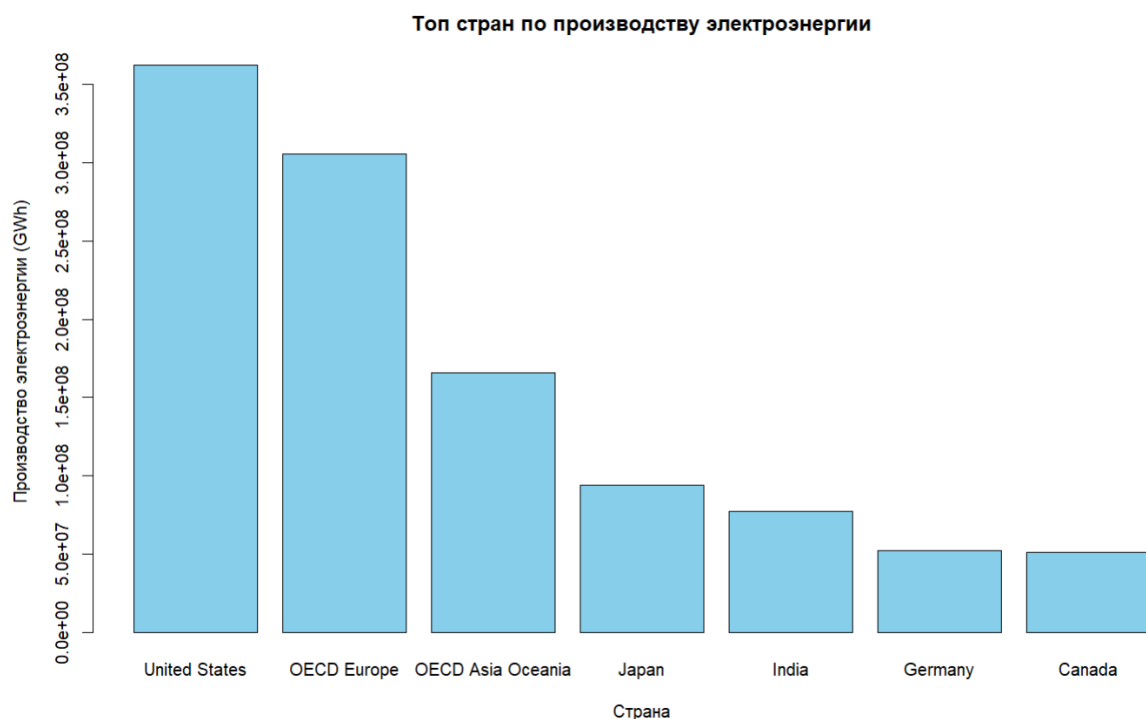


Рисунок 2.1 — Диаграмма первых стран по производству электроэнергии

Как можно отметить, среди ведущих стран по производству электроэнергии, помимо межгосударственных объединений, находятся такие

крупные державы, как США, Япония и Индия. Эти страны демонстрируют значительное производство электроэнергии, что отражает их важную роль на мировом энергетическом рынке. В дальнейшем, проводится анализ временного ряда на примере Японии.

Для анализа временных рядов рассматриваются данные производства электроэнергии в Японии. Строится график ежемесячного производства электроэнергии. На Рисунке 2.2 видно, как менялось производство электроэнергии в Японии с течением времени.



Рисунок 2.2 —Ежемесячное производство электроэнергии в Японии

Преобразуем значения из столбца с данными в формат временного ряда, начиная с самой ранней даты, указанной в соответствующем столбце времени, и устанавливаем месячную частоту. В результате создается временной ряд, представляющий данные по месяцам, начиная с первой доступной даты. [2.3]

Для более детального анализа проводится декомпозиция временного ряда, которая позволяет выделить тренд, сезонность и остаточные компоненты. На Рисунке 2.3 декомпозиции видно, как сезонные колебания и общий тренд влияют на производство электроэнергии в Японии.

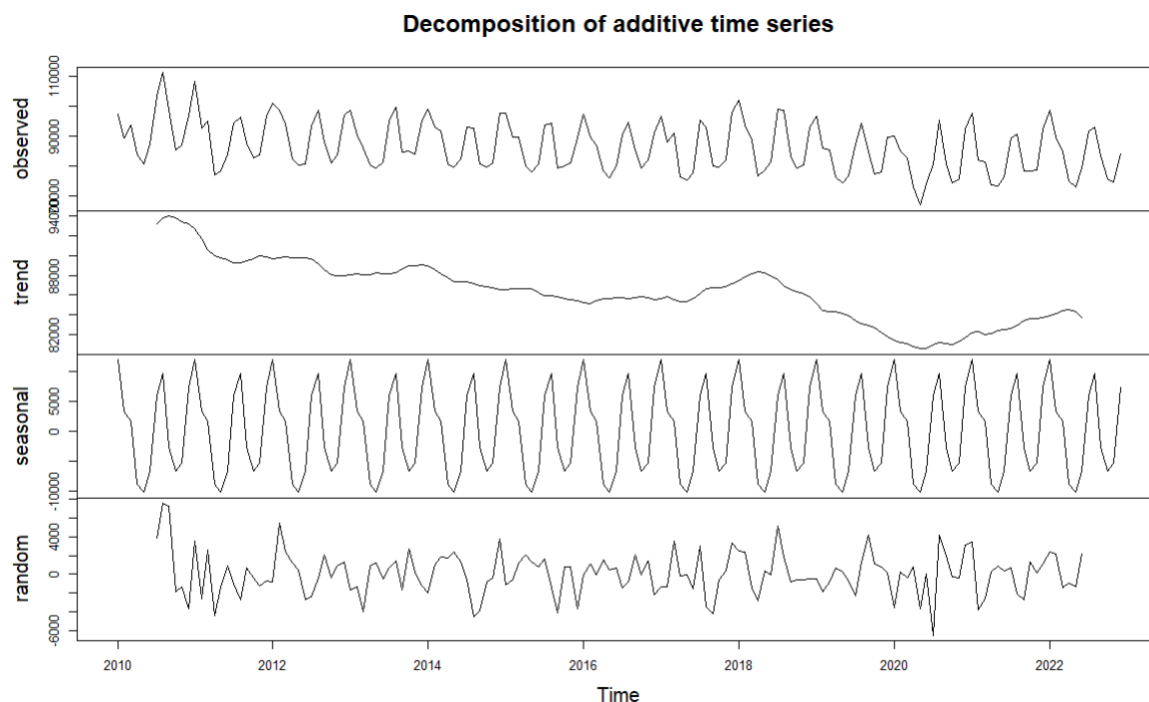


Рисунок 2.3 —Декомпозиция временного ряда

На первом графике представлен исходный временной ряд данных, отражающий общее производство электроэнергии за весь период наблюдений. Второй график иллюстрирует выявленный тренд. Тренд демонстрирует ярко выраженное снижение. Это свидетельствует о падении производства электроэнергии в Японии. Третий график демонстрирует сезонные колебания временного ряда. Сезонность выражена довольно четко, с повторяющимися циклическими изменениями примерно с периодичностью в полгода, что указывает на наличие регулярных сезонных факторов, влияющих на производство электроэнергии.

Проводится тест Дики-Фуллера для проверки стационарности временного ряда и тест Льюнга-Бокса для выявления сезонности. На рисунке 2.4 можно заметить результат тестирования.

В обоих тестах p -значение оказалось малым, что указывает на стационарность ряда и наличие ярко выраженной сезонности.

```

> # Тест Дики-Фуллера на стационарность
> adf_test_jp_result <- adf.test(ts_jp_data)
Предупреждение:
В adf.test(ts_jp_data) : p-value smaller than printed p-value
> adf_test_jp_result

Augmented Dickey-Fuller Test

data: ts_jp_data
Dickey-Fuller = -5.4289, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary

> # Тест Льюнга-Бокса на сезонность
> ljung_box_test <- Box.test(ts_jp_data, lag = 12, type = "Ljung-Box")
> ljung_box_test

```

Box-Ljung test

```

data: ts_jp_data
X-squared = 333, df = 12, p-value < 2.2e-16

```

Рисунок 2.4 —Тест Дики-Фуллера и Льюнга-Бокса

Для анализа тренда временного ряда рассчитывается уравнение тренда с использованием модели `tslm`. Это позволяет определить общий тренд в данных и оценить скорость изменений производства электроэнергии с течением времени. На Рисунке 2.5 показано уравнение тренда `виден`, который накладывается на исходные данные, показывая направление и скорость изменений.

График производства электроэнергии в Японии с линией тренда

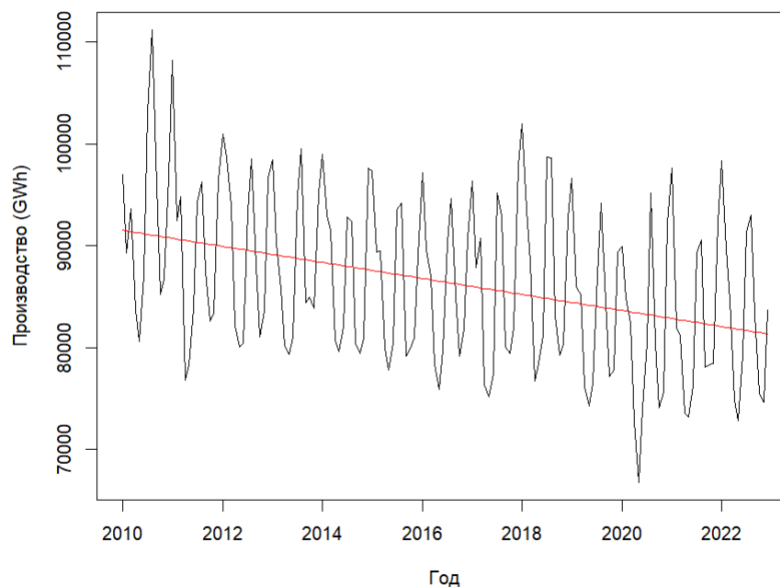


Рисунок 2.5 —Уравнение тренда

Отметим, что тренд является нисходящим.

Таким образом, проведён анализ временных рядов производства электроэнергии в Японии, включающий построение графика ежемесячного производства, декомпозицию временного ряда для выделения тренда, сезонности и остаточных компонентов, а также проверку стационарности и сезонности с использованием тестов Дики-Фуллера и Льюнга-Бокса. Результаты показали, что временной ряд обладает стационарностью и выраженной сезонностью. Построенное уравнения тренда с использованием модели `tslm` помогли определить и оценить скорость изменений.

Выполним прогнозирование временного ряда.

2.2 Прогнозирование временных рядов и оценка качества

Для прогнозирования будущих значений используется модель `auto.arima` [2.4]. Прогноз строится на 5 лет вперёд. На графике прогноза показаны будущие значения производства электроэнергии вместе с доверительными интервалами, которые позволяют оценить неопределённость прогноза. Прогнозируемые значения показывают ожидаемое развитие производства электроэнергии в Японии с учётом текущих тенденций, их можно увидеть на Рисунке 2.6.

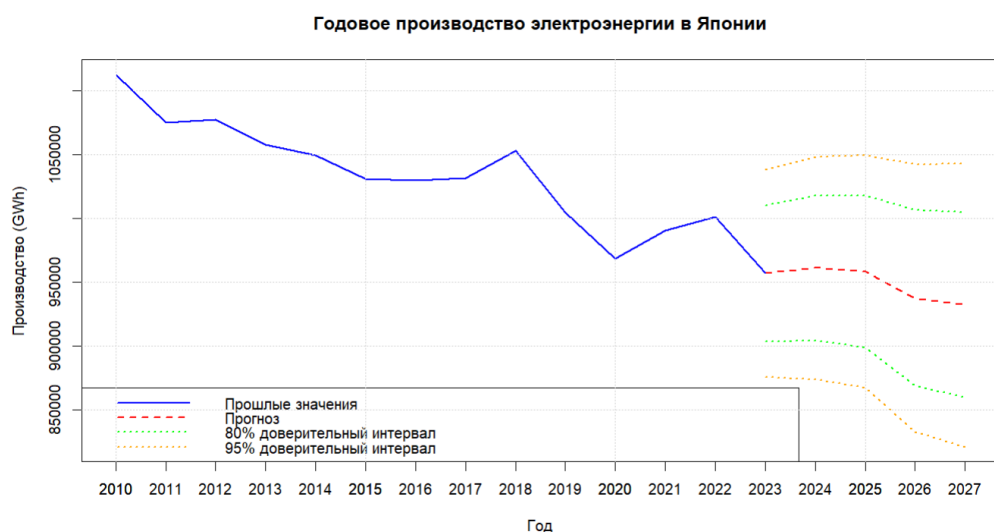


Рисунок 2.6 — Прогнозируемые значения производства электроэнергии в Японии

Как можно заметить из прогноза, производство электроэнергии в Японии будет следовать нисходящему тренду. Перейдём к рассмотрению основных метрик данной модели на Рисунке 2.7

```
> summary(jp_forecast_model)
Series: ts_jp_data
ARIMA(2,0,0)(2,1,0)[12] with drift

Coefficients:
      ar1      ar2      sar1      sar2      drift
    0.3962  0.1321 -0.6656 -0.4608 -66.9491
s.e.  0.0844  0.0859  0.0851  0.0891  22.3434

sigma^2 = 9639490: log likelihood = -1364.03
AIC=2740.06  AICc=2740.67  BIC=2757.88

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -41.95989 2930.705 2180.435 -0.1132378 2.511176 0.6758349 -0.006933581
```

Рисунок 2.7 — Метрики модели ARIMA для производства электроэнергии в Японии

Метрики качества модели, такие как $AIC = 2740.06$ и $BIC = 2757.88$, свидетельствуют о хорошей подгонке модели. Ошибки на тренировочном наборе включают $RMSE = 2930.705$, $MAE = 2180.435$ и $MAPE = 2.511176$ %, что указывает на приемлемую точность прогноза.

Таким образом, проведенный анализ временных рядов производства электроэнергии за 2010-2022 годы позволил выявить ключевые тенденции и сезонные колебания в данных. Агрегирование данных по годам и странам показало ведущие страны по объему производства электроэнергии, а детальный анализ временного ряда для Японии выявил наличие стационарности и значительной сезонности.

Прогнозирование с использованием модели `auto.arima` предоставило точные предсказания на ближайшие пять лет, демонстрируя спад в производстве электроэнергии. Эти результаты помогают лучше понять динамику производства электроэнергии и могут быть использованы для дальнейшего планирования и оптимизации энергетических ресурсов.

ЗАКЛЮЧЕНИЕ

Цель данной курсовой работы — создание инструмента, способного анализировать данные о производстве электроэнергии и прогнозировать их будущее изменение — достигнута.

В ходе практической работы проведён глубокий анализ временных рядов данных о производстве электроэнергии в Японии. Применение метода декомпозиции временных рядов, сопряженный с использованием модели `auto.arima`, позволил сформировать прогнозы на 5 лет вперёд. Полученные результаты наглядно отображены на графиках, что обеспечило четкое представление о текущих тенденциях и предполагаемых изменениях в производстве электроэнергии.

Получены результаты тестов Дики-Фуллера и Льюнга-Бокса, которые подтвердили стационарность временного ряда и наличие сезонных колебаний. Эти результаты укрепляют обоснованность прогнозов и повышают уверенность в их применении при принятии решений в энергетической отрасли.

Задачи, выполненные в данной курсовой работе:

- использовать знания математической статистики с использованием современных средств обработки данных: платформы RStudio и языка программирования R;
- описать необходимые формулы и понятия;
- выполнить анализ временных рядов на основе собранных данных;
- оценить качество решения на основе исходного набора данных.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

1.1. Подкорытова О.А. Анализ временных рядов / О.А. Подкорытова, М.В. Соколов — Москва: Издательство Юрайт, 2020. — 143 с.

1.2. Нильсен Э. А. Практический анализ временных рядов: учебное пособие — Москва: Издательство Диалектика Вильямс, 2021 — 544 с.

1.3. Прогнозирование и временные ряды: учебное пособие / К. О. Кизбикенов. — Барнаул: АлтГПУ, 201 [Электронный ресурс]. / URL: <https://library.altspu.ru/dc/pdf/kizbikenov.pdf>

1.4. Кильдишев Г.С., Френкель А.А. Анализ временных рядов и прогнозирование. — Москва.: Ленанд, 2023. — 104 с.

1.5. Рассел Д., Кон Р. Тест Дики-Фуллера — Москва: Книга по Требованию, 2013 — 96 с.

1.6. Прогнозирование временных рядов методом ARIMA: учебное пособие / П. Ю. Трофимов, В. Ю. Носков. — Екатеринбург: УрФУ, [Электронный ресурс]. / URL: https://elar.urfu.ru/bitstream/10995/52487/1/978-5-9908685-0-2_2017_72.pdf

1.7. Введение в анализ временных рядов: учебное пособие для вузов/Н. В. Артамонов, Е. А. Ивин, А. Н. Курбацкий, Д. Фантацини. — Вологда: ВолНИЦ РАН, [Электронный ресурс]. / URL: mse.msu.ru/wp-content/uploads/2021/03/Введение-в-анализ-временных-рядов-1.pdf

ПРАКТИЧЕСКАЯ ЧАСТЬ

2.1. R: Documentation [Электронный ресурс]. / URL: <https://www.r-project.org/other-docs.html>

2.2. R. Documentation — tidyverse [Электронный ресурс] / URL: <https://www.rdocumentation.org/packages/tidyverse/versions/2.0.0>

2.3. R. Documentation — tseries [Электронный ресурс] / URL:
<https://cran.r-project.org/web/packages/tseries/tseries.pdf>

2.4. R. Documentation — forecast [Электронный ресурс] / URL:
<https://www.rdocumentation.org/packages/forecast/versions/8.22.0>

ПРИЛОЖЕНИЯ

Приложение А — Код, написанный на языке программирования R

Приложение А

Ниже приведен код, написанный на языке R с использованием среды разработки RStudio.

Листинг 1 — *kursovaya.R*

```
if (!require("tidyverse")) install.packages("tidyverse")
if (!require("forecast")) install.packages("forecast")
if (!require("tseries")) install.packages("tseries")
library(tidyverse)
library(forecast)
library(tseries)
# Загрузка данных
data <- read.csv("energyProduction.csv", sep=";", fileEncoding =
"Latin1", stringsAsFactors = FALSE)
colnames(data) <- c("COUNTRY", "CODE_TIME", "TIME", "YEAR", "MONTH",
"MONTH_NAME", "PRODUCT", "VALUE", "DISPLAY_ORDER", "yearToDate",
"previousYearToDate", "share")
data$VALUE <- as.numeric(data$VALUE)
# Расчёт годового производства
annual_data <- data %>%
  group_by(COUNTRY, PRODUCT, YEAR) %>%
  summarize(ANNUAL_VALUE = sum(VALUE, na.rm = TRUE)) %>%
  ungroup()
total_production_by_country <- annual_data %>%
  group_by(COUNTRY) %>%
  summarize(TOTAL_PRODUCTION = sum(ANNUAL_VALUE))
# Топ стран по производству электроэнергии
top_countries <- total_production_by_country %>%
  arrange(desc(TOTAL_PRODUCTION))
top_countries_10 <- top_countries[4:10, ]
# Построение столбчатой диаграммы для первых 10 стран
barplot(top_countries_10$TOTAL_PRODUCTION,
        names.arg = top_countries_10$COUNTRY,
        main = "Топ стран по производству электроэнергии",
        xlab = "Страна",
        ylab = "Производство электроэнергии (GWh)",
        col = "skyblue")
# Для Японии
jp_data <- data %>%
  filter(COUNTRY == "Japan" & PRODUCT == "Electricity supplied") %>%
  select(TIME, VALUE)
ts_jp_data <- ts(jp_data$VALUE, start = c(year(min(jp_data$TIME)),
month(min(jp_data$TIME))), frequency = 12)
# Построение графика временного ряда
plot(ts_jp_data, main = "Ежемесячное производство электроэнергии в
Японии",
      ylab = "Производство (GWh)", xlab = "Год")
jp_annual_data <- annual_data %>%
  filter(COUNTRY == "Japan" & PRODUCT == "Electricity supplied") %>%
  select(YEAR, ANNUAL_VALUE)
# Декомпозиция временного ряда
decomposed_jp_ts <- decompose(ts_jp_data)
plot(decomposed_jp_ts)
# Вычисление уравнения тренда
jp_trend_model <- tslm(ts_jp_data ~ trend)
```

```

summary(jp_trend_model)
intercept <- coef(jp_trend_model) ["(Intercept)"]
slope <- coef(jp_trend_model) ["trend"]
plot(ts_jp_data, main = "График производства электроэнергии в Японии
с линией тренда", xlab = "Год", ylab = "Производство (GWh)")
lines(fitted(jp_trend_model), col = "red")
# Тест Дики-Фуллера на стационарность
adf_test_jp_result <- adf.test(ts_jp_data)
adf_test_jp_result
# Тест Льюнга-Бокса на сезонность
ljung_box_test <- Box.test(ts_jp_data, lag = 12, type = "Ljung-Box")
ljung_box_test
# Прогноз
jp_forecast_model <- auto.arima(ts_jp_data, d = 0)
jp_forecast_values <- forecast(jp_forecast_model, h = 12*5) # Прогноз
на 5 лет вперед
summary(jp_forecast_model)
jp_forecast_values
jp_forecast_data <- as.data.frame(jp_forecast_values)
jp_forecast_data$Year <- substr(rownames(jp_forecast_data), 5, 8)
sum_jp_by_year <- jp_forecast_data %>%
  group_by(Year) %>%
  summarise(
    Total_Point_Forecast = sum(`Point Forecast`),
    Total_Lo_80 = sum(`Lo 80`),
    Total_Hi_80 = sum(`Hi 80`),
    Total_Lo_95 = sum(`Lo 95`),
    Total_Hi_95 = sum(`Hi 95`)
  )
sum_jp_by_year
plot(jp_annual_data$YEAR, jp_annual_data$ANNUAL_VALUE, type = "l",
c(us_annual_data$ANNUAL_VALUE[nrow(us_annual_data)],
  xlim = c(min(jp_annual_data$YEAR),
max(as.integer(sum_jp_by_year$Year))),
  ylim = c(min(sum_jp_by_year$Total_Lo_95),
max(jp_annual_data$ANNUAL_VALUE)),
  xlab = "Год", ylab = "Производство (GWh)", main = "Годовое
производство электроэнергии в Японии")
lines(as.integer(sum_jp_by_year$Year),
sum_jp_by_year$Total_Point_Forecast, col = "red", lty = "dashed", lwd = 2)
lines(as.integer(sum_jp_by_year$Year), sum_jp_by_year$Total_Lo_80,
col = "green", lty = "dotted", lwd = 2)
lines(as.integer(sum_jp_by_year$Year), sum_jp_by_year$Total_Hi_80,
col = "green", lty = "dotted", lwd = 2)
lines(as.integer(sum_jp_by_year$Year), sum_jp_by_year$Total_Lo_95,
col = "orange", lty = "dotted", lwd = 2)
lines(as.integer(sum_jp_by_year$Year), sum_jp_by_year$Total_Hi_95,
col = "orange", lty = "dotted", lwd = 2)
lines(c(max(jp_annual_data$YEAR),
as.integer(sum_jp_by_year$Year[1])),
c(jp_annual_data$ANNUAL_VALUE[nrow(jp_annual_data)],
sum_jp_by_year$Total_Point_Forecast[1]), col = "blue", lwd = 2)
legend("bottomleft", legend = c("Прошлые значения", "Прогноз", "80%
доверительный интервал", "95% доверительный интервал"),
  col = c("blue", "red", "green", "orange"), lty = c(1, 2, 3,
3), lwd = 2)
axis(1, at = seq(min(jp_annual_data$YEAR),

```