

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»**

Кафедра дифференциальных уравнений и
математической экономики

**МОДЕЛИ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ ЦЕН НА
РОССИЙСКОМ ОПТОВОМ РЫНКЕ ЭЛЕКТРОЭНЕРГИИ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 441 группы

направления 09.03.03 Прикладная информатика

механико-математического факультета

Горшуковой Регины Васильевны

Научный руководитель
Профессор, д.э.н., профессор _____

В.А. Балаш

Заведующий кафедрой
Зав. кафедрой, д.ф.-м.н., профессор _____

С.И.Дудов

Саратов 2021

Введение. В настоящее время в России действует оптовый рынок электроэнергии и мощности. Данный рынок необходим в современном обществе, потому что электроэнергия является ресурсом, который потребляется практически всеми жителями нашей планеты. Цена данного ресурса оказывает большое влияние на экономику. Все участники оптового рынка электроэнергии и мощности функционируют в условиях конкуренции, в этих условиях им необходимо выработать рыночную стратегию и планировать будущие финансовые потоки. В данных условиях им необходимо прогнозировать цены на электроэнергию на так называемом «рынке на сутки вперед».

Актуальность данной бакалаврской работы связана с тем, что в условиях высокой неопределённости на российском рынке, нужно прогнозировать модели, независимые от прогноза экзогенных переменных, поскольку возможная ошибка предсказания данных параметров приведет к нежелательной ошибке прогноза искомого показателя - цены на рынке электроэнергии.

Целью бакалаврской работы является нахождение метода, который окажется более точным на обучающей и тестовой выборке.

Для достижения поставленной цели в работе необходимо решить следующие **задачи**: в теоретической части исследовать оптовый рынок электроэнергии и мощности, основные принципы его функционирования, основные методы краткосрочного прогнозирования цен в электроэнергетике; в практической части производим построение моделей: naïve, snaive, ETS, ARIMA, TBATS, NNAR, Prophet.

Выявим более точный метод с помощью реализации их программой R. Построение моделей основывается на данных с сайта АО «АТС» с 31.12.2018 по 31.12.2020.

Основное содержание работы. Работа состоит из введения, четырёх разделов, заключения, списка использованных источников, содержащего 21 наименований. Общий объем работы составляет 54 страниц, включая 3 таблицы.

Во **введении** обосновывается актуальность темы работы, формулируется цель работы и решаемые задачи, отмечается практическая значимость полученных результатов.

В **первом** разделе приводятся основные понятия оптового рынка электроэнергии и мощности и его основные принципы функционирования.

ОРЭиМ действует в так называемых ценовых и неценовых зонах. Ценовая зона делится ещё на две отдельные зоны, на первую и вторую.

В первую входят Приволжский округ и Уральский округ, Северо – Западный, Южный, Центральный и Северо – Кавказский округ. Во вторую входит Сибирский федеральный округ.

В неценовую входят те регионы, в которых по техническим причинам невозможна организация рыночных отношений. Регионы, входящие в данную зону: регионы Дальнего Востока, Калининградская область, Архангельская область, Республика Коми.

ОРЭиМ функционирует по правилам, которые определило Правительство Российской Федерации - это Постановление Правительства Российской Федерации №1172 от 27 декабря 2010 года.

Во **втором** разделе приводятся основные методы краткосрочного прогнозирования цен в электроэнергетике.

Исследователи изучили потребности данного рынка и предложили ряд подходов, по которым можно создавать модели прогнозирования. Данный ряд делится на три группы: 1. модели, которые строятся с помощью теории игр; 2. модели, которые называются фундаментальные или имитационные; 3. модели, которые строятся с помощью анализа временных рядов, они подразделяются на два направления: традиционные статистические модели и модели искусственного интеллекта.

В **третьем** разделе производится построение моделей прогноза цен на электроэнергию на рынке «на сутки вперед» и анализ данных.

Naïve метод иначе называют «наивный метод». Он так называется из-за своей простоты, потому что он предполагает сохранение тенденций в будущем и строится на основании последнего наблюдения.

Snaiive метод иначе называют «сезонный метод». Данное название основывается на том, что он в отличие от предыдущего метода naïve строится на основании данных за последний сезон.

ETS метод иначе называют «Метод экспоненциального сглаживания». Каждый метод помечается парой букв (T,S), которые определяют тип "трендовых" и "сезонных" компонентов, комбинации представлены в таблице 1.

Таблица 1

| Тренд | Сезонная составляющая | | |
|-------------------------|-----------------------|-------------|------------------|
| | N | A | M |
| | (None) | (Additive) | (Multiplicative) |
| N(None) | (N,N) | (N,A) | (N,M) |
| A(Additive) | (A,N) | (A,A) | (A,M) |
| A_d (Additive damped) | (A_d ,N) | (A_d ,A) | (A_d ,M) |

Для каждого метода существует две модели: с аддитивными ошибками (A) и с мультипликативными ошибками (M). Для отличия этих моделей, а также для отличия моделей от методов, добавляем третью букву ETS(*,*,*).

ARIMA метод расшифровывается как авторегрессионная интегрированная скользящая средняя. Многие из моделей являются частными случаями модели ARIMA, как показано в таблице 2.

Таблица 2

| | |
|-------------------------------|---------------|
| Белый шум | ARIMA (0,0,0) |
| Случайное блуждание | ARIMA(0,1,0) |
| Случайное блуждание с дрейфом | ARIMA(0,1,0) |
| Авторегрессия | ARIMA(p,0,0) |
| Скользящее среднее значение | ARIMA(0,0,q) |

Модель выглядит так: ARIMA(p,d,q), где p - порядок следования авторегрессионной части; d - степень первой вовлеченной дифференцировки; q - порядок расположения части скользящей средней.

TBATS метод основывается на комбинациях из рядов Фурье с экспоненциальной сглаживающей моделью и преобразованиях Бокса – Кокса.

NNAR метод иначе называется авторегрессией нейронной сети, обозначается NNAR(p,k), где p - запаздывающие входы и k - число скрытых узлов.

Этот метод основывается на данных временных рядов с запаздывающими значениями. Они используются в качестве входных данных для нейросети.

Prophet метод состоит в подгонке аддитивных регрессионных моделей со следующими четырьмя основными компонентами: тренд, годовая сезонность, недельная сезонность, праздники.

Произведем анализ моделей. Чтобы реализовать наивный метод прогнозирования используется функция $\text{naive}(y, h)$, где $h=21$. Результаты представлены на рисунке 1 для Приволжского округа (ПО) и рисунке 2 для Сибирского федерального округа (СФО).

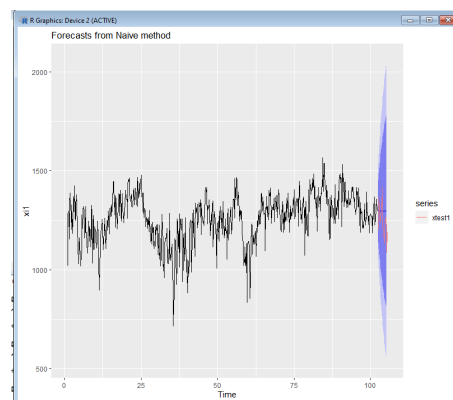
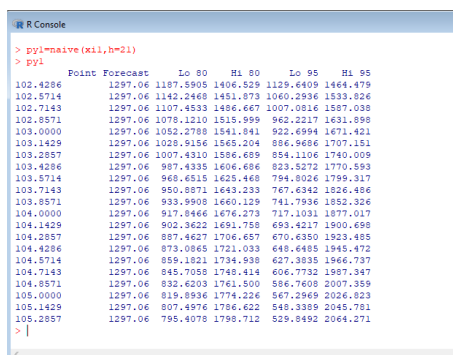


Рисунок 1 — Приволжский округ

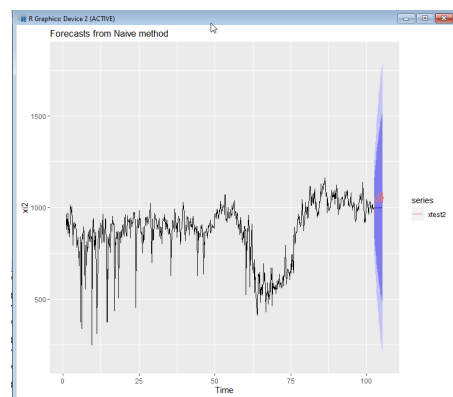
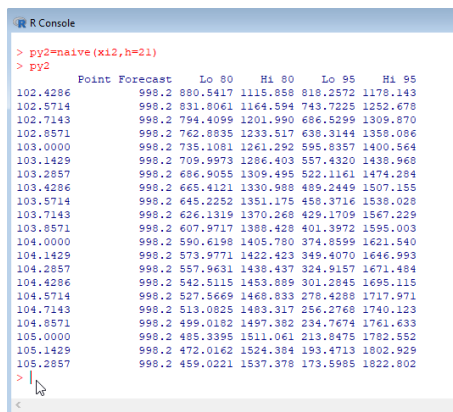


Рисунок 2 — Сибирский федеральный округ

Из графиков видно в качестве прогноза прямую линию, которая параллельна оси абсцисс, она показывает, что цена будет стабильна.

Функция $\text{naive}(y, h)$ реализуется для сезонного наивного метода, $h=21$ и сезонность семидневная. Результаты для рядов модели naive представлены на рисунке 3 для ПО и 4 для СФО.

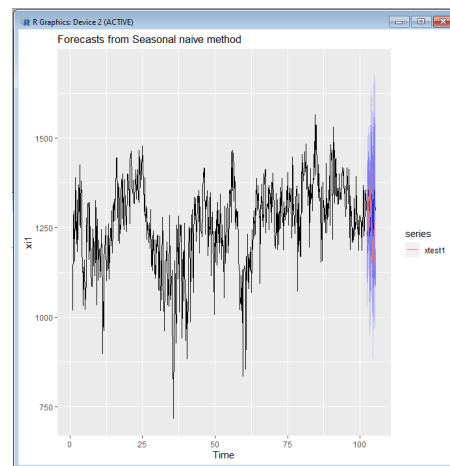
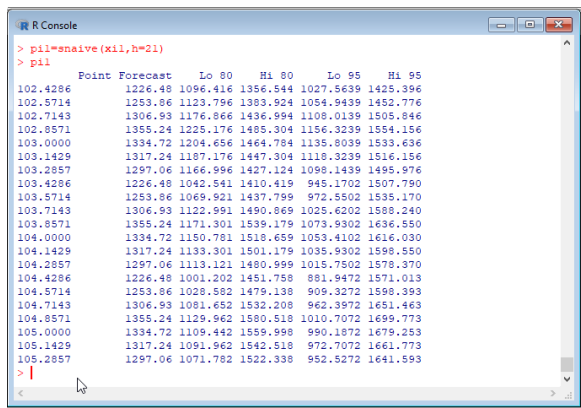


Рисунок 3 — Приволжский округ

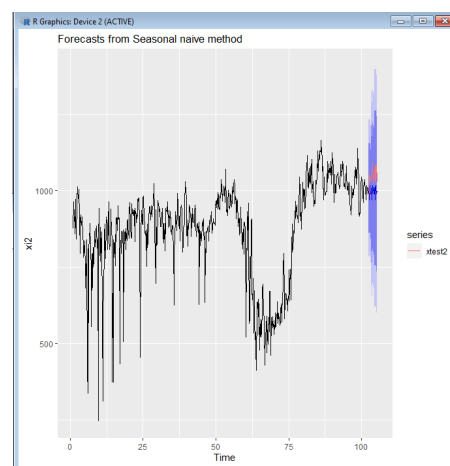
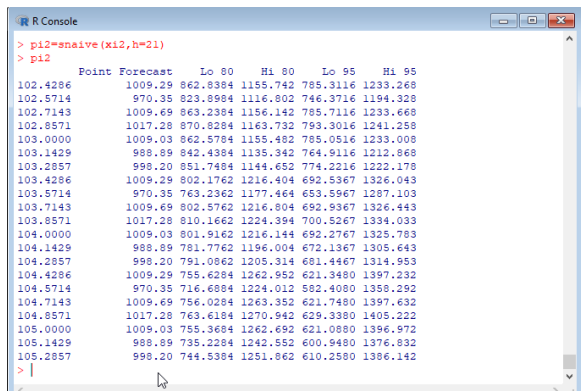


Рисунок 4 — Сибирский федеральный округ

На основе представленного графика для ПО видно, что метод `snaiive` является более точным, чем предыдущий метод `naive`. Из графика, что он повторяет динамику исходного ряда и позволяет прогнозировать сезонные всплески. Подобное замечено и на графике для СФО.

Метод экспоненциального сглаживания оценивает временной ряд и автоматически выбирает модель с помощью реализации функции `ets(y, h)`. Результаты представлены на рисунке 5 для ПО и рисунке 6 для СФО.

В результате видим, что в данном случае для ПО и СФО выбрана модель $ETS(A, N, A)$: модель с отсутствующим трендом и аддитивной сезонностью с аддитивной ошибкой.

Выбор подходящих значений для параметров p , d и q модель авторегрессии и скользящего среднего реализуется с помощью функции

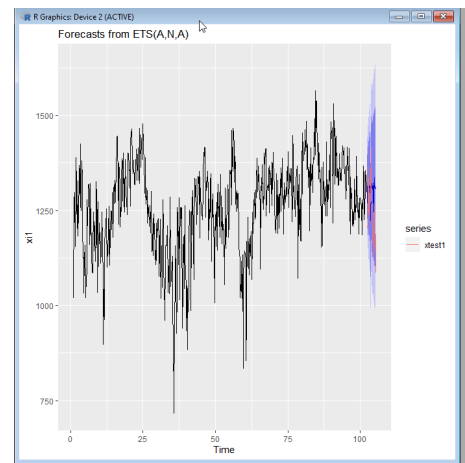
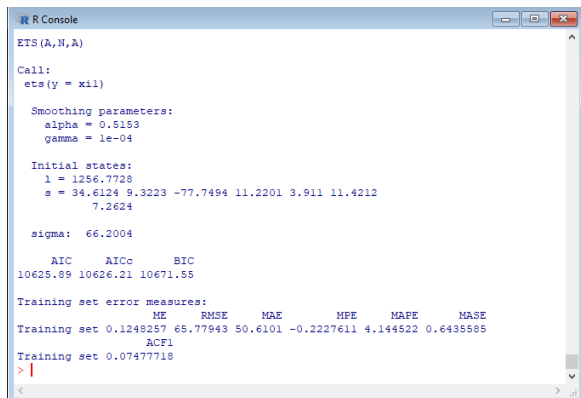


Рисунок 5 — Приволжский округ

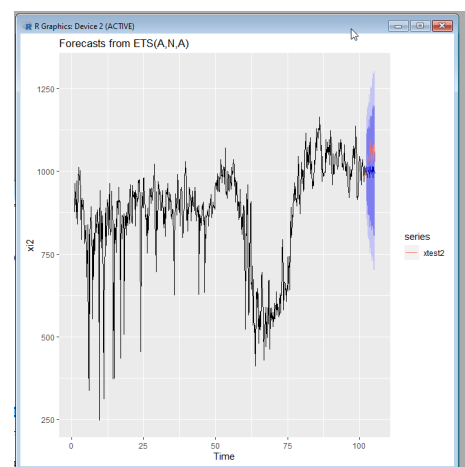
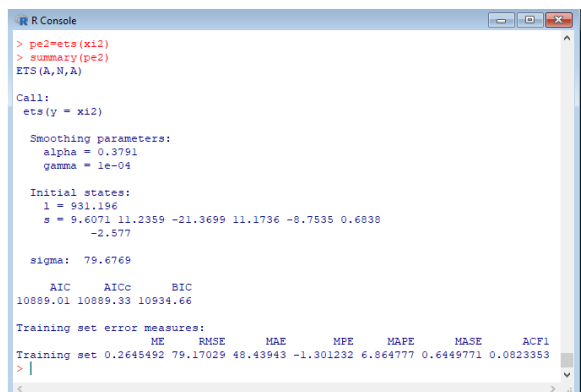


Рисунок 6 — Сибирский федеральный округ

auto.arima(), которая автоматически выбирает параметры. Результаты для рядов модели ARIMA представлены на рисунке 7 для ПО и рисунке 8 для СФО.

По полученным результатам можно сделать вывод, что ряд по ПО соответствует модели ARIMA(0,1,2): $y_t = c - 0.4373\varepsilon_{t-1} - 0.1537\varepsilon_{t-2} + \varepsilon_t$. Стандартное отклонение: $\sqrt{5094} = 71.37$. Ряд по СФО соответствует модели ARIMA(3,1,1): $y_t = c + 0.2813y_{t-1} + 0.1977y_{t-2} - 0.1150y_{t-3} - 0.8677\varepsilon_{t-1} + \varepsilon_t$. Стандартное отклонение: $\sqrt{3667} = 77.81$.

Функция tbats(y) реализуется для модель TBATS. Значение показателей функции представлены в таблице 3. Результаты для рядов модели TBARS представлены на рисунке 9 для ПО и рисунке 10 для СФО.

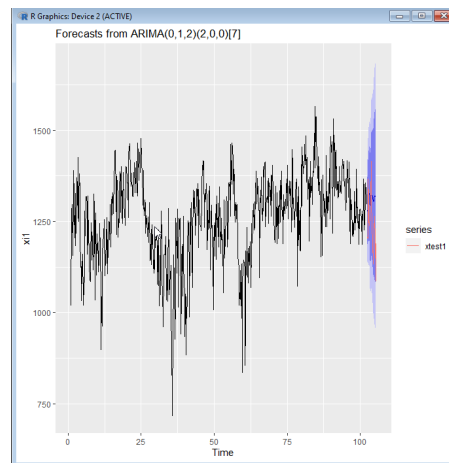
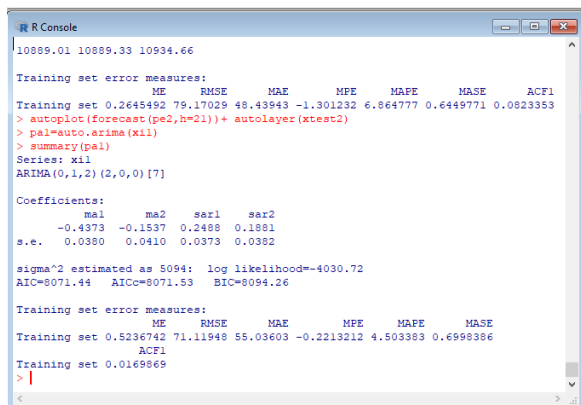


Рисунок 7 — Приволжский округ

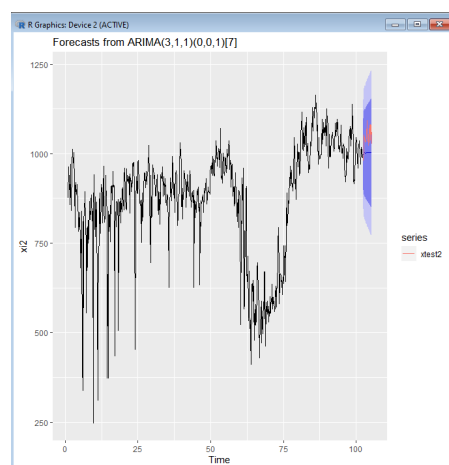
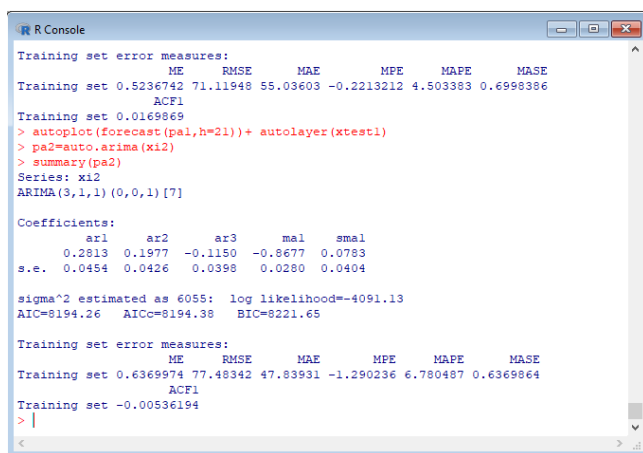


Рисунок 8 — Сибирский федеральный округ

Таблица 3

| Компонент | Значение |
|-----------|-------------------------------------|
| 1 | Параметр преобразования Бокса-Кокса |
| {0,0} | Ошибка ARMA |
| - | Параметр демпфирования |
| {<0,0>} | Сезонный период, слагаемые Фурье |

Ряд по ПО соответствует модели $TBATS(1, \{2,1\}, -, \langle 7,3 \rangle)$, где 1 означает, что преобразования Бокса-Кокса не было, $\{2,1\}$ - ARMA(2,1), - - влияние внутри колебательной системы отсутствует, $\langle 7,3 \rangle$ - семидневная сезонность, три члена ряда Фурье. По СФО наблюдаем, что результат $BATS(1, \{3,1\}, -, -)$, где 1 означает, что преобразования Бокса-Кокса не было, $\{3,1\}$ - ARMA(3,1), -


```

R R Console
> pt1=tbats(x11)
> pt1
TBATS(1, (2,1), -, (<7,3>))

Call: tbats(y = x11)

Parameters
Alpha: -0.003478109
Gamma-1 Values: 0.0002154546
Gamma-2 Values: -9.840115e-05
AR coefficients: 1.260799 -0.289311
MA coefficients: -0.699866

Seed States:
      [,1]
[1,] 1247.308179
[2,]  28.538465
[3,] -13.153196
[4,] -1.390427
[5,]  3.681324
[6,] -24.697375
[7,]  24.272449
[8,]  0.000000
[9,]  0.000000
[10,] 0.000000

Sigma: 63.93601
AIC: 10597.53

```

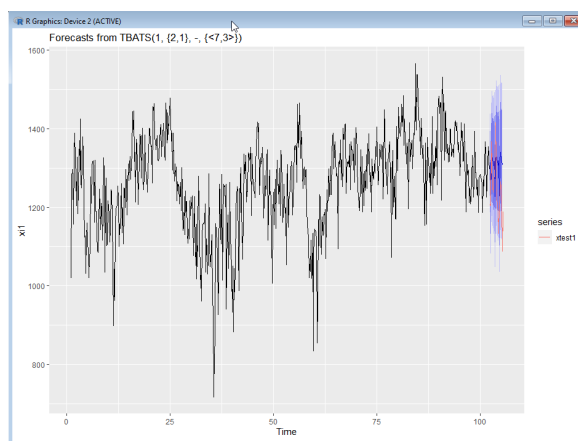


Рисунок 9 — Приволжский округ

```

R R Console
> pt2=tbats(x12)
> pt2
BATS(1, (3,1), -, -)

Call: tbats(y = x12)

Parameters
Alpha: 0.1612782
AR coefficients: 0.41386 0.139708 -0.157094
MA coefficients: -0.161171

Seed States:
      [,1]
[1,] 903.2952
[2,]  0.0000
[3,]  0.0000
[4,]  0.0000
[5,]  0.0000

Sigma: 77.62022
AIC: 10860.93
> |

```

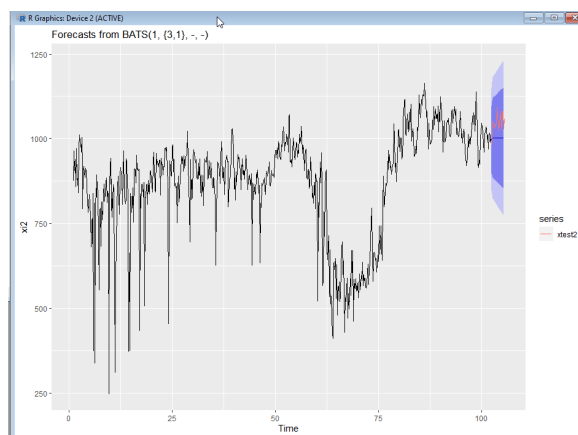


Рисунок 10 — Сибирский федеральный округ

- влияние внутри колебательной системы отсутствует, - - сезонность отсутствует. BATS отличается от TBATS только тем, как моделирует сезонные эффекты. В BATS используется более традиционный подход, при котором может моделировать только целочисленную длину периода.

Для построения нейросетевой модели ряда реализуется функции $nnetar(y, lambda)$, где трансформацию Бокса-Кокса $lambda=0$, чтобы гарантировать, что прогнозы будут положительными. Результаты представлены на рисунке 11 для ПО и рисунке 12 для СФ.

Ряд по ПО соответствует модели $NNAR(22, 1, 12)_7$ модель имеет $Y_{t-1}, Y_{t-2}, Y_{t-3}$ и Y_{t-7} , и двенадцать нейронных скрытых узлов. По СФО мы наблюдаем, что модель выдала результат $NNAR(28, 1, 14)_7$ модель имеет $Y_{t-1}, Y_{t-2}, Y_{t-3}$ и Y_{t-7} , и четырнадцать нейронных скрытых узлов.

```

> pn1=nnetar(xil, lambda=0)
> pn1
Series: xil
Model: NNAR(22,1,12) [7]
Call: nnetar(y = xil, lambda = 0)

Average of 20 networks, each of which is
a 22-12-1 network with 289 weights
options were - linear output units

sigma^2 estimated as 0.0009992
> |

```

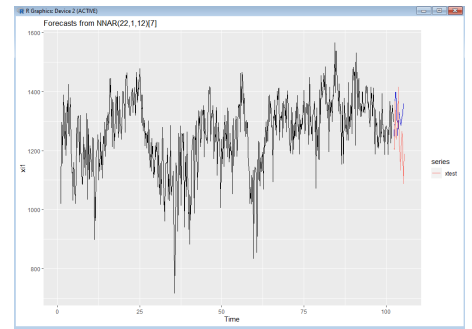


Рисунок 11 — Приволжский округ

```

> pn2=nnetar(xi2, lambda=0)
> pn2
Series: xi2
Model: NNAR(28,1,14) [7]
Call: nnetar(y = xi2, lambda = 0)

Average of 20 networks, each of which is
a 28-14-1 network with 421 weights
options were - linear output units

sigma^2 estimated as 0.001858
> |

```

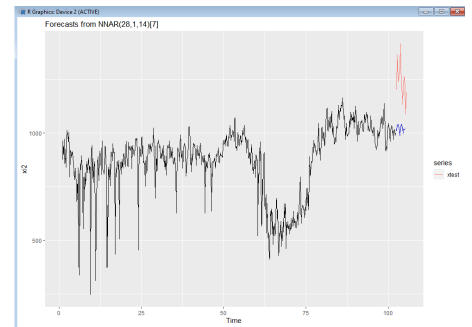


Рисунок 12 — Сибирский федеральный округ

Результаты для модели prophet первых нескольких предсказанных значений отклика, их (принятые по умолчанию) 80%-ные доверительные границы и график представлены на рисунке 13 для ПО и рисунке 14 для СФО.

```

> future_df1=make_future_dataframe(M1, periods = 90)
> forecast_M1=predict(M1, future_df1)
> head(forecast_M1[, c("yhat", "yhat_lower", "yhat_upper")])
      yhat yhat_lower yhat_upper
1 1247.130  1142.090  1356.774
2 1249.211  1147.404  1353.853
3 1243.122  1141.027  1348.781
4 1259.842  1149.734  1362.980
5 1181.369  1082.100  1283.200
6 1275.839  1175.419  1375.084

```

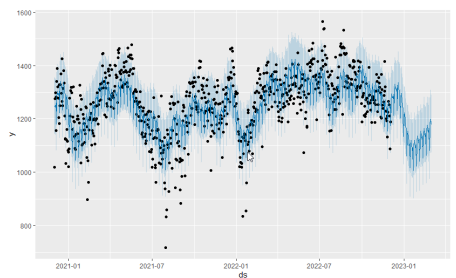


Рисунок 13 — Приволжский округ

Черные точки на рисунке 14 и 15 соответствуют значениям отклика из обучающей выборки. Сплошная голубая линия - это предсказанные моделью значения отклика, а огибающая эту линию светло-голубая "лента" это 80% доверительные интервалы соответствующих предсказанных значений. Прогнозные значения y на следующие 90 дней видны в правой части графика.

С помощью функции `prophet_plot_components()` можно изобразить отдельные компоненты модели на рисунке 15 для ПО и СФО.

```
> head(forecast_M2[, c("yhat", "yhat_lower", "yhat_upper")])
      yhat yhat_lower yhat_upper
1 1123.965 1012.2048 1236.375
2 1102.460  982.9227 1209.297
3 1108.243  996.5791 1210.429
4 1103.108  984.4449 1216.769
5 1109.197  999.2110 1222.248
6 1106.611 1001.2560 1217.841
```

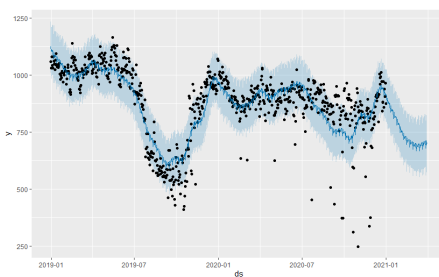


Рисунок 14 — Сибирский федеральный округ

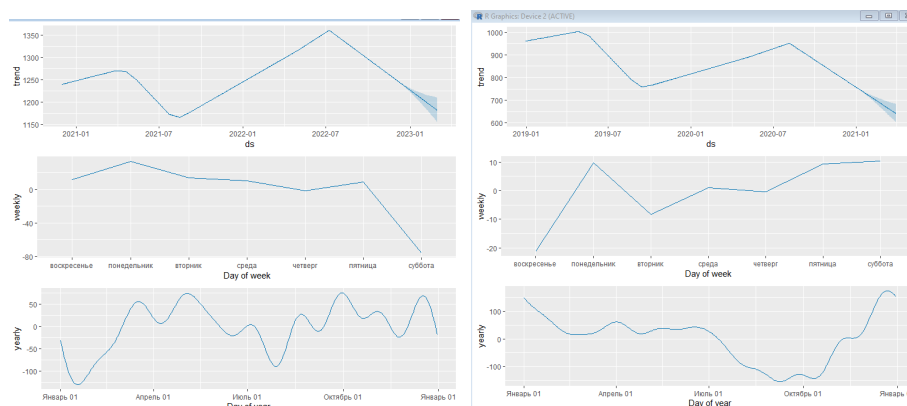


Рисунок 15 — Приволжский округ и Сибирский федеральный округ

На рисунке 15 в верхней части изображен тренд, посередине изображена недельная сезонность, а внизу можно увидеть годовую сезонность.

В ПО наблюдаем, что тренд сложный, в этом временном ряду максимальная цена по понедельникам, а минимальная цена в субботу. Если рассматривать внутри-годовые колебания, можно заметить закономерность, что высокая цена обычно в весенние месяцы, в октябре и декабре, а минимальная в феврале. В СФО наблюдаем, что тренд сложный, в этом временном ряду максимальная цена по понедельникам, пятницам и субботам, а минимальная цена в воскресенье. Если рассматривать внутри-годовые колебания, можно заметить закономерность, что высокая цена обычно в зимние месяцы, а минимальная в осенний период.

В **четвёртом** разделе рассмотрим оценку точности результатов прогнозирования и выберем итоговую модель.

В качестве характеристики точности моделей могут быть использованы следующие показатели: Рассмотрим для моделей следующие показатели точности: средняя ошибка (ME), среднеквадратичная ошибка корня (RMSE),

средняя абсолютная ошибка (MAE), средняя процентная ошибка прогнозирования (MPE), средняя абсолютная процентная ошибка (MAPE)

Все показатели можно найти с помощью функции `accuracy()`. Если точность прогноза равна приближена к 0, то выбранная модель описывает фактические значения на 100%, т. е. очень точно, но надо сразу обратить внимание, что такого точного показателя никогда не будет. На рисунке 16 и 17 представлены ошибки для моделей.

| > test_acc1 | | | | | > test_acc2 | | | | |
|------------------|-----------|-----------|--------------|-----------|------------------|------------|------------|------------|------------|
| | RMSE | MAE | MPE | MAPE | | RMSE | MAE | MPE | MAPE |
| Training Naive | 85.41949 | 65.02702 | -0.226166140 | 5.3321607 | Training Naive | 91.80924 | 52.52381 | -1.1580274 | 7.4186554 |
| Test Naive | 87.84790 | 72.64429 | -4.405293673 | 6.0016185 | Test Naive | 54.06156 | 49.21667 | 4.4861418 | 4.6586591 |
| Training Snaive | 101.48967 | 78.64102 | -0.324604208 | 6.4221521 | Training Snaive | 114.27680 | 75.10256 | -1.4998415 | 10.1638642 |
| Test Snaive | 88.47031 | 75.92143 | -4.490046638 | 6.4492722 | Test Snaive | 53.72693 | 48.08286 | 4.2792750 | 4.5595184 |
| Training ETS | 65.77943 | 50.61010 | -0.222761137 | 4.1445222 | Training ETS | 79.17029 | 48.43943 | -1.3012322 | 6.8647767 |
| Test ETS | 64.90326 | 57.41469 | -1.591359152 | 4.4912785 | Test ETS | 45.63811 | 42.59594 | 3.5066615 | 4.0656656 |
| Training ARIMA | 71.11948 | 55.03603 | -0.221321234 | 4.5033833 | Training ARIMA | 77.48342 | 47.83531 | -1.2802361 | 6.7804873 |
| Test ARIMA | 63.82293 | 52.67811 | -2.123191681 | 4.1425536 | Test ARIMA | 45.92014 | 41.07898 | 3.5971947 | 3.9051783 |
| Training TBATS | 63.93601 | 48.67271 | -0.490124543 | 4.0040983 | Training TBATS | 77.62022 | 48.11062 | -1.2954602 | 6.8117235 |
| Test TBATS | 66.21942 | 58.97428 | -1.826497347 | 4.6108413 | Test TBATS | 45.88885 | 40.80009 | 3.5807581 | 3.8772473 |
| Training NNAR | 24.44892 | 18.62365 | -0.001262405 | 1.4974164 | Training NNAR | 22.78907 | 14.90076 | -0.0284911 | 1.7535443 |
| Test NNAR | 80.74936 | 65.58627 | -2.795414859 | 5.1334634 | Test NNAR | 32.22142 | 24.11411 | 1.5032646 | 2.2990398 |
| Training Prophet | 236.30674 | 185.48211 | 0.138246653 | 0.1407537 | Training Prophet | 365.25306 | 278.61378 | 0.3237311 | 0.3562858 |
| Test Prophet | 172.59355 | 122.49002 | 0.095039796 | 0.1049404 | Test Prophet | 2761.86537 | 1998.72860 | 0.7579347 | 1.9358667 |

Рисунок 16 — Приволжский округ и Сибирский федеральный округ

Из рисунков видно, что для ПО на обучающей выборке наиболее точной моделью является TBATS, а на тестовой выборке ARIMA. Для СФО на обучающей и тестовой выборке наиболее точной моделью является NNAR.

В **заключении** приведены результаты бакалаврской работы.

Основные результаты

1. Определены основные понятия оптового рынка электроэнергии и мощности и его принципы функционирования;
2. Изучены основные методы краткосрочного прогнозирования цен в электроэнергетике;
3. Построены модели и графики на так называемом «рынке на сутки вперед»;
4. Проанализированы результаты Приволжского округа и Сибирского федерального округа;
5. Выявлены точные модели для Приволжского округа и Сибирского федерального округа.