



Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ
ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»**

Кафедра прикладной океанологии

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА (бакалаврская работа)

На тему Межгодовая изменчивость вылова Атлантическо-скандинавской сельди и возможность его прогноза статистическими методами

Исполнитель Мхеян Максим Александрович
(фамилия, имя, отчество)

Руководитель кандидат географических наук, доцент
(ученая степень, ученое звание)

Густоев Дмитрий Владимирович
(фамилия, имя, отчество)

«К защите допускаю»
Заведующий кафедрой _____
(подпись)

кандидат географических наук, доцент
(ученая степень, ученое звание)

Лукиянов Сергей Васильевич
(фамилия, имя, отчество)

« ____ » _____ 2022 г.

Санкт-Петербург
2022

Оглавление

Введение	3
1 Описание объекта исследования	5
1.1 Внешние признаки	5
1.2 Ареал обитания	6
1.3 Нерестилища, места нагулов и зимовки	7
1.4 Сельдь как хищник	8
1.5 Сельдь как добыча	9
1.6. Промысел сельди	10
1.7 Выводы по главе	11
2 Описание статистических методов исследования	12
2.1 Первичная статистика	12
2.2 Эмпирическая функция распределения и проверка её соответствие нормальному закону	12
2.3 Тренд временного ряда	14
2.4 Гармонический анализ	14
2.5 Вейвлет – анализ	16
2.6 Автокорреляционный анализ и авторегрессия 1 порядка	16
2.7 Выводы по главе	17
3 Прогноз и анализ изменчивости среднегодового вылова сельди в период с 1900 по 2013 год	18
3.1 Первичные статистики	18
3.2 Эмпирическая функция распределения	19
3.3 Детерминированная часть ряда	21
3.3.1 Тренд временного ряда	21
3.3.2 Гармонический анализ	25
3.3.3 Вейвлет-преобразование	29
3.4 Прогнозирование детерминированной и стохастической части ряда с заблаговременностью 1 год	30
3.4.1 Анализ ряда отклонений	30
3.4.2 Расчет и анализ автокорреляционной функции	31
3.4.3 Построение и оценка качества модели авторегрессии 1 порядка	32
3.4.4 Прогноз тренда с заблаговременностью 1 год	35
3.4.5 Прогноз гармоник с заблаговременностью 1 год	35
3.4.6 Сборка временного ряда с заблаговременностью 1 год	36
3.5 Выводы по главе	37
Заключение	38
Список использованной литературы	39
	2

Введение

Актуальность исследования. Рыбный промысел является одним из важнейших компонентов хозяйственной деятельности человека. Издревле рыба являлась одним из важнейших продуктов питания для человека, так как её относительно легко добыть, а также потому, что она является источником полезных жирных кислот, белка, фосфора, и других полезных веществ. Рыба водится в пресных и соленых водах, от озер до океанов, что позволяет большинству людей на планете самостоятельно вылавливать и употреблять в пищу рыбу. Львиную долю промысла составляет, конечно же, морская и океанологическая ловля.

В данной работе будет исследован вылов очень популярного в мировом промысле объекта – сельди, а конкретнее – один из её видов: Атлантическо-скандинавская сельдь.

Атлантическо-скандинавская сельдь пользуется огромным спросом не только в России, но и в мире. Вылов атлантическо-скандинавской сельди занимает 2 место по вылову в северном рыбохозяйственном бассейне, больше ловят только кильку.

Целью исследования является анализ среднегодового вылова Атлантическо-скандинавской сельди и попытка его прогнозирования статистическими методами.

Задачи, поставленные в работе:

1. Описать объект исследования;
2. Подготовить исходные данные;
3. Проанализировать исходные данные;
4. Предпринять попытку сделать прогноз.

Для осуществления поставленных задач было использовано ПО, а именно:

- «Пакет анализа» MS Excel;

- PAST4.

Исходные данные, а именно: статистический ряд вылова сельди в период с 1900 по 2013 год, дискретность – 1 год, были любезно предоставлены ПИНРО для учебных целей [1].

1 Описание объекта исследования

1.1 Внешние признаки

Исследуемый объект:

Атлантическо-скандинавская сельдь.



Рисунок 1.1. Атлантическо-скандинавская сельдь.

Отряд: Сельдеобразные

Семейство: Сельдевые

Род: Сельди

Вид: Атлантическо-скандинавская сельдь

Атлантическо-скандинавская сельдь (далее – сельдь) – стайная морская рыба, имеет темную верхнюю часть с оттенками черно-голубого или синезелёного цвета. Теша и боковые стороны имеют серебристую окраску. Плавники имеют серую окраску и небольшой размер. Самый большой плавник – хвостовой, имеет вильчатую форму. Голова маленькая, без чешуи. Челюсть чуть выдвинута вперед. Продолжительность жизни сельди составляет 25 лет.

Длина тела сельди может составлять до 45 см, а вес её, в среднем, 1,1 кг. Половозрелость наступает к 3-8 годам, при длине 24-30 см. В бедных поколениях сельди более 60% особей достигают половозрелости к 3-4 годам, в средних поколениях, порядка 70% особей достигают половозрелости к 4-5 годам, а в многочисленных поколениях порядка 70% особей достигают половозрелости к 5-

6 годам, а в самых многочисленных еще 30% особей достигают половозрелости к 7-8 годам [3].

1.2 Арел обитания

Обитает сельдь преимущественно в водах северной части Атлантического океана. Арел обитания сельди очень изменчив. Главной причиной этому является то, что сельдь является одновременно хищником и источником пищи для других животных [4].

В районах Баренцева/Белого морей имеются относительно небольшие запасы сельди, которые проводят там всю жизнь. Несмотря на то, что эти запасы относительно небольшие и имеют, по большей части, местное значение, они не являются объектами международного рыболовства.

Намного больший запас сельди находится в акватории Норвежского моря, где и расположены основные их нерестилища [2].

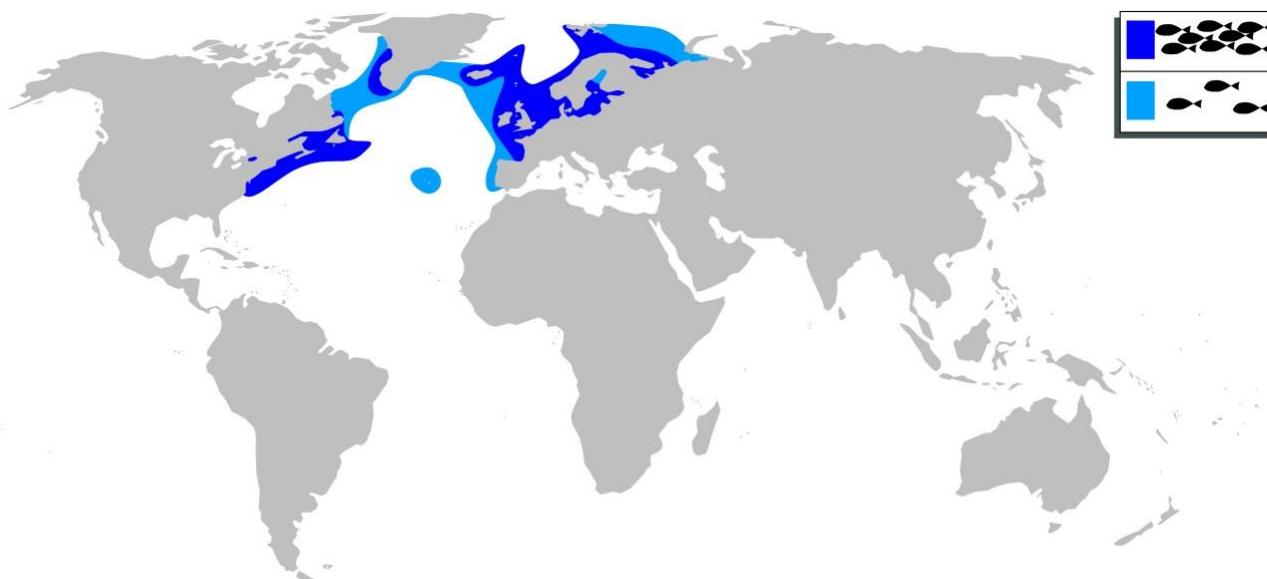


Рисунок 1.2. Арел обитания Атлантическо-скандинавской сельди.

1.3 Нерестилища, места нагулов и зимовки

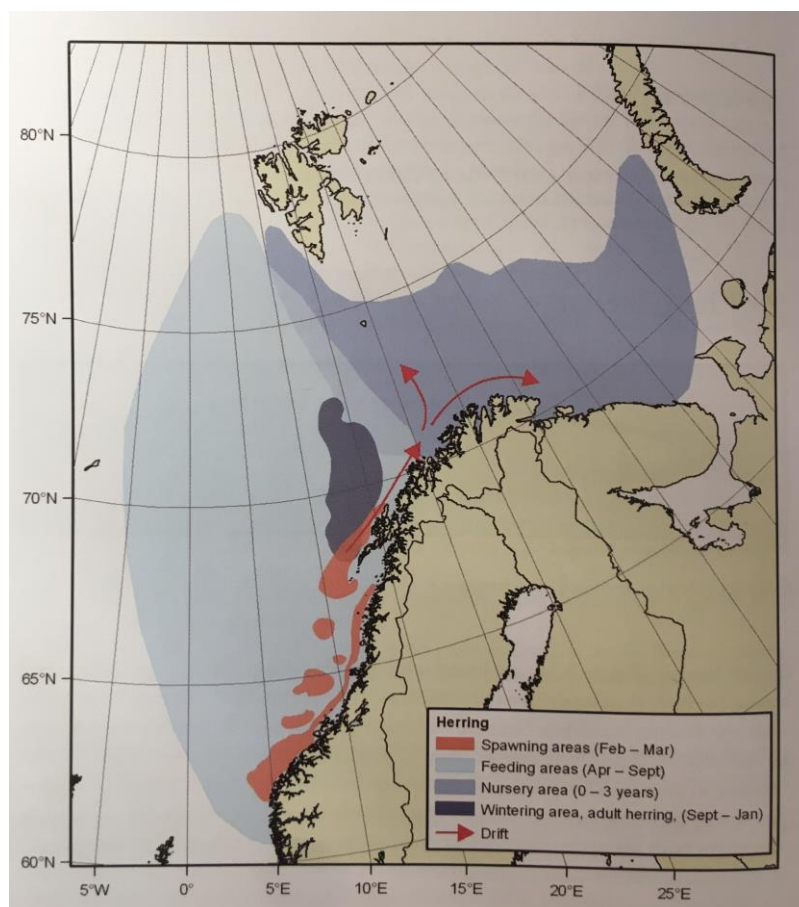


Рисунок 1.3. Основные нерестилища сельди, направление дрейфа личинок и районы нагула, зимовки и нагула в Норвежском и Баренцевом морях.

Как видно из рисунка, основным нерестилищем сельди является норвежская прибрежная зона.

Нерестится сельдь на глубинах от 30 до 100 м, севернее, в районе Лофотенских островов, от 150 до 200 м. Во время нереста в темное время суток сельдь держится практически у поверхности, а в светлое время суток пребывает на дне [2].

Температура в местах нереста варьируется от 5,0 до 8,5 °С, а соленость составляет 34,0 – 35,57‰.

По структуре икра сельди вязкая, диаметр икринки достигает от 1,6 до 2,1 мм. Откладывается икра на дно, в местах, где преобладают каменные, песчаные и ракушечные грунты [5].

По прошествии 10-15 дней после нереста вылупляются личинки, которые впоследствии появляются в верхних слоях воды и дрейфуют на север вдоль побережья Норвегии достигая Баренцева моря (показано стрелками на рисунке 1.3.), где дальше, ветвями Нордкапского течения, могут распространяться до островов Новой Земли [4].

Через два-три месяца они переживают метаморфоз и становятся мальками.

В период нагула сельдь передвигается стаями, в которых длинна и физиологическое состояние рыбы одинаково.

Осенью, после нагула, молодь сельди мигрирует в Баренцево море, в восточной части которого и зимует.

В возрасте 3-5 лет сельдь мигрирует с востока на запад в Норвежское море, где они объединяются со взрослыми особями. Они не возвращаются в Баренцево море во взрослом возрасте. Таким образом, Баренцево море является лишь рассадником для сельди [2].

1.4 Сельдь как хищник

Зоопланктон (ракообразные) является основным продуктом рациона сельди. Сезон кормления начинается в мае, когда ракообразные появляются в поверхностном слое, и продолжается до октября. Максимальная активность кормления приходится на июнь-июль. В феврале-марте происходит некоторое питание зимующим планктоном. 90% добычи состоит из 2 групп мелких ракообразных: euphausiids (криль) и *calanus finmarchicus* (калянус) [2].

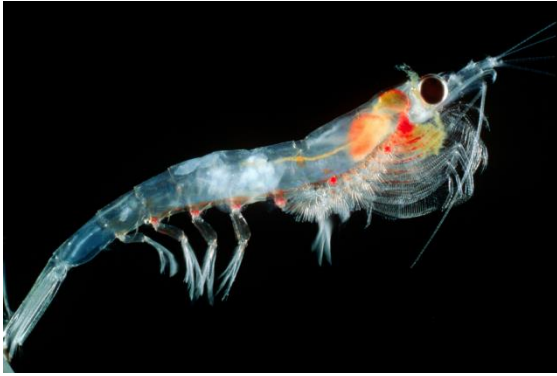


Рисунок 1.4. Euphausiids.

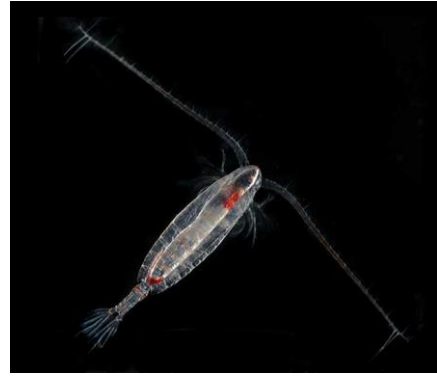


Рисунок 1.5. Calanus Finmarchicus.

Несмотря на то, что основным продуктом питания являются ракообразные, молодая сельдь также питается личиночными стадиями многих других организмов, включая мойву, двухстворчатых моллюсков, брюхоногих моллюсков, хетогнатов, иглокожих и других рыб [2].

1.5 Сельдь как добыча

1.5.1 Птицы

Мелкая сельдь в Баренцевом море и в северных норвежских фьордах, очевидно, в больших количествах поедается птицами.

1.5.2 Млекопитающие морские

Молодь сельди также является очень важным источником пищи для морских млекопитающих. В период с 1992-1995 годов киты-норки ежегодно потребляли примерно 630 000 тонн молодой сельди [Фолкоу (2000)], а популяцию гренландского тюленя в Баренцевом море ежегодно потребляла 210 000 тонн [Нильсен и др.].

1.5.3 Донная рыба

В Баренцевом море сельдь является ценной добычей северо-восточной арктической трески, особенно зимой, когда их ареалы обитания наиболее часто пересекаются. Летом сельдь и треска, как правило, находятся на разных глубинах: треска остается на дне, а сельдь у поверхности [2].

1.6. Промысел сельди

Атлантическо-скандинавская сельдь ранее являлась объектом двух различных промыслов - незрелой мелкой сельди (группы 0 и 1) и промыслом жирной сельди (незрелая сельдь 2-5 летнего возраста).

Российское рыболовство

До 1947 года СССР вылавливал, в основном, мелкую и жирную сельдь на мурманском побережье. Это являлось небольшим сетевым промыслом, производящийся летом в заливах Мурманского побережья. Ежегодный улов в этом промысле составил не более тысячи тонн. Позднее, до 1969 года, жирную сельдь также добывали в открытом море, с максимальным уловом в 93 000 тонн в 1967 году [2].

Норвежское рыболовство

В Норвегии, с начала 20-го века, постепенно развивался крупный кошельковый неводный промысел незрелой мелкой сельди в районах фьордов. Фьорд и прибрежные районы Финнмарка, граничащие с Баренцевым морем, стали важным районом для этого промысла. В северных районах также велся жаберный промысел жирной сельди [2].

С развитием гидролокатора, силового блока и кошельковых неводов из синтетического волокна промысел сельди полностью изменился. С начала 1960-х годов встало возможным ловить сельдь далеко в Баренцевом море. В 1960-х годах в Баренцевом море было выловлено около 100 000 тонн мелкой сельди и почти 350 000 тонн жирной сельди (1967г) [2].

Вклад промысла сельди в Баренцевом море в сокращение её запасов.

Промысел незрелой сельди в Баренцевом море и взрослой сельди в Норвежском море являются основными причинами сокращения запасов сельди во время весеннего нереста в начале 1970-х годов. Промысел незрелой сельди означал, что особи 1963-1968 годов выпуска были выловлены до того, как они смогли пополнить и без того сокращающийся нерестовый запас. В результате

промысел мелкой и жирной сельди в Баренцевом море внес значительный вклад в сокращение её запасов [2].

Нынешнее управление запасами сельди.

Промысел норвежской весенне-нерестовой сельди в настоящее время управляется региональной организацией по управлению прибрежными государствами, состоящей из пяти сторон, и было достигнуто международное соглашение о правилах ловли. Запасы восстановились, и промысел с нынешним среднегодовым уловом около миллиона тонн ведется за пределами Баренцева моря. В рамках режима управления, вылов незрелой сельди запрещен, и это обеспечивается минимальным размером в 25 см и закрытыми участками. Таким образом, Баренцево море является районом, где сельдь не вылавливается. Последствия промысла незрелой сельди в 1960-х годах сильно сказались на добыче и развитии сельди [2].

1.7 Выводы по главе

1. Атлантическо-скандинавская сельдь является очень интересным объектом для исследования. Эта рыба на протяжении всей жизни мигрирует на большие расстояния проходя различные этапы и борьбу за существование.

2. Данная рыба является одной из самых важных промысловых видов рыб. Она пользуется огромным спросом не только в России, но и в мире. Вылов сельди занимает 2 место по вылову в северном рыбохозяйственном бассейне, больше ловят только кильку.

3. Сельдь очень полезна для человека, так как она содержит в себе огромное количество питательных веществ, которые необходимы человеку, включая селен, витамин Д и омега-3, также она является низкокалорийной, что еще больше расширяет круг потребления данной рыбы.

2 Описание статистических методов исследования

2.1 Первичная статистика

Первичная статистика дает возможность охарактеризовать исследуемый процесс. Для прогноза и анализа межгодовой изменчивости вылова сельди понадобятся следующие характеристики:

1. Среднее арифметическое значение.

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (2.1)$$

* N – длина выборки.

2. Дисперсия и связанное с ней стандартное отклонение. Описывают средний разброс значений ряда от среднего арифметического значения.

$$D = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2, \quad \sigma = \sqrt{D}. \quad (2.2)$$

* D – дисперсия; σ – стандартное отклонение/

3. Размах вариации. Характеризует максимальный разброс значений ряда.

$$R = \max - \min \quad (2.3)$$

4. Коэффициент вариации. Показывает изменчивость характеристики в исследуемом статистическом ряде. Не имеет размерности.

$$C = \frac{\sigma}{\bar{x}} \cdot 100\%. \quad (2.4)$$

2.2 Эмпирическая функция распределения и проверка её соответствие нормальному закону

Функция распределения показывает взаимосвязь между возможными значениями случайной величины и вероятностью их появления.

Эмпирическая функция распределения называется функцией распределения, которая рассчитана по выборке. Для её изображения строится гистограмма [6].

1. Асимметрия. Определяет симметричность эмпирической функции распределения относительно среднего значения.

$$As = \frac{1}{N\sigma^3} \sum_{i=1}^N (x_i - \bar{x})^3 \quad (2.5)$$

2. Эксцесс. Определяет островершинность.

$$Ex = \left[\frac{1}{N\sigma^4} \sum_{i=1}^N (x_i - \bar{x})^4 \right] - 3 \quad (2.6)$$

3. Для построения гистограмм используются соотношения:

3.1. Размер интервала.

$$k_{max} \cong 5 \lg n \quad (2.7)$$

3.2. Размах интервала.

$$\Delta X_k = \frac{X_{max} - X_{min}}{k_{max}} \quad (2.8)$$

4. Функция плотности вероятности нормального закона распределения.

$$f(\bar{x}, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \quad (2.9)$$

5. Формула для перевода значений плотности вероятности нормального закона в значения соответствующих частот нормального закона, после перевода – результат округлить до целых (получится n_k).

$$n_k^* = f(x_k, \bar{x}, \sigma) \cdot \Delta X_k \cdot N \quad (2.10)$$

6. Статистический критерий χ^2 .

$$\chi^2 = \sum_{i=1}^{k_{max}} \frac{(m_k - n_k)^2}{n_k} \quad (2.11)$$

2.3 Тренд временного ряда

Тренд отражает наличие во временной изменчивости исследуемой характеристики длиннопериодного колебания с периодом, существенно превышающим длину выборки.

Тренд бывает линейный и нелинейным, в зависимости от того, на какую часть длиннопериодного колебания попала выборка.

Уравнение линейного тренда:

$$y = a_1 t + a_0 + \varepsilon \quad (2.12)$$

Уравнение нелинейного тренда:

$$y = a_2 t^2 + a_1 t + a_0 + \varepsilon \quad (2.13)$$

* t – время.

Основные характеристики тренда.

1. Коэффициент детерминации r^2 (для линейного тренда) или его нелинейный аналог – корреляционное отношение η^2 , характеризующий вклад тренда в общую дисперсию ряда. Вклад может быть значимым или незначимым, это проверяется путем проверки коэффициента корреляции r (далее – КК) или η на значимость. В случае незначимости – тренда нет [6].

2. Величина тренда – изменение характеристики по линейному тренду за определенный промежуток времени. Величина тренда равна коэффициенту a_1 линейного тренда и имеет размерность характеристики y за единицу дискретности [6].

2.4 Гармонический анализ

Гармонический анализ заключается в разложении любого ряда в ряд Фурье без остатка, то есть на конечное число гармоник.

Гармоника – тригонометрическая функция, имеющая период, кратный длине ряда, то есть каждая гармоника целое число раз «укладывается» в длину исходного ряда [6].

Формула гармоники:

$$G_k = A_k \cos(\omega_k \cdot t - \varphi_k), \quad \omega_k = 2\pi/T_k \quad (2.14)$$

* k – номер гармоники; A_k – амплитуда k -той гармоники; ω_k – частота k -той гармоники; T_k – период k -той гармоники; φ_k – фаза k -той гармоники; t – время.

Характеристики гармоники определяются на основании коэффициентов Фурье:

$$a_k = \frac{2}{N} \sum_{i=1}^N [x_i \sin(\omega_k \cdot t_i)] \quad (2.15)$$

$$b_k = \frac{2}{N} \sum_{i=1}^N [x_i \cos(\omega_k \cdot t_i)] \quad (2.16)$$

* x_i – исходный ряд; N – длина ряда; ω_k – частота k -той гармоники; t_i – ряд времени.

Характеристики гармоники определяются по формулам:

1. Амплитуда гармоники.

$$A_k = \sqrt{a_k^2 + b_k^2} \quad (2.17)$$

2. Фаза гармоники.

$$\varphi_k = \arctg \frac{a_k}{b_k} \pm \pi \quad (2.18)$$

3. Дисперсия гармоники.

$$D_k = \frac{A_k^2}{2} \quad (2.19)$$

4. Вклад гармоники в общую дисперсию ряда.

$$V_k = \frac{D_k}{D_y} \quad (2.20)$$

2.5 Вейвлет – анализ

Методология выполнения вейвлет-анализа подробно описана в электронном ресурсе [7].

2.6 Автокорреляционный анализ и авторегрессия 1 порядка

Анализ внутренней структуры исследуемого процесса проводится на основе автокорреляции.

Автокорреляция – корреляция статистического ряда самого с собой при разных сдвигах во времени.

Коэффициент автокорреляции $r(\tau)$ на каждом сдвиге:

$$r(\tau) = \frac{1}{\sigma_x^2(N-1-\tau)} \sum_{i=1}^{N-\tau} (x_i - \bar{x})(x_{i+\tau} - \bar{x}) \quad (2.21)$$

* N – длина реализации, τ – сдвиг, который меняется от 1 до максимума.

Относительно r рассчитывается критическое значение $r_{кр}$, соответствующее $t_{кр}$, при уровне значимости α и числе степеней свободы ν

$$r_{кр}(\tau) = \frac{-\sqrt{N-\tau-1} + \sqrt{N-\tau-1 + 4t_{кр}^2}}{2t_{кр}} \quad (2.22)$$

* N – длина реализации, τ – сдвиг АКФ, $t_{кр}(\alpha = 0,05; \nu = N - \tau - 1)$ – критерий Стьюдента.

Автопрогноз выполняется на основе модели парной линейной регрессии, которая называется авторегрессией 1 порядка. Уравнение АР1 формулируется в общем виде:

$$x_{i+k} = a_1 x_i + a_0 + \varepsilon \quad (2.23)$$

* k – заблаговременность прогноза.

2.7 Выводы по главе

1. Используя статистические методы можно дать оценку процессам, происходящим не только в гидрометеорологии, но и в других сферах жизни. Эти методы дают возможность четко и подробно узнать об изменениях конкретной характеристики, выявить цикличность этих изменений, а также сделать прогноз той или иной характеристики.

2. Стоит отметить, что природные процессы достаточно проблематично спрогнозировать статистическими методами, так как влияние внешних факторов, влияющих на вылов, большое множество, и точно учесть их не получится.

3 Прогноз и анализ изменчивости среднегодового вылова сельди в период с 1900 по 2013 год

3.1 Первичные статистики

Для ряда среднегодового вылова сельди рассчитаны первичные статистики.

Таблица 3.1. Описательные статистики среднегодового вылова сельди.

Вылов сельди		
Среднее	682,42	Тыс. Тонн
Стандартная ошибка	45,85	Тыс. Тонн
Медиана	553,50	Тыс. Тонн
Мода	302	
Стандартное отклонение	489,54	Тыс. Тонн
Дисперсия выборки	239646,02	Тыс. Тонн ²
Экцесс	1,70	
Асимметричность	1,23	
Размах вариации	2475	Тыс. Тонн
Минимум	103	Тыс. Тонн
Максимум	2578	Тыс. Тонн
Коэффициент вариации	6,72	%

Далее был построен график изменчивости среднегодового вылова сельди, на который были нанесены стандартное отклонение, среднее значение и размах вариации.

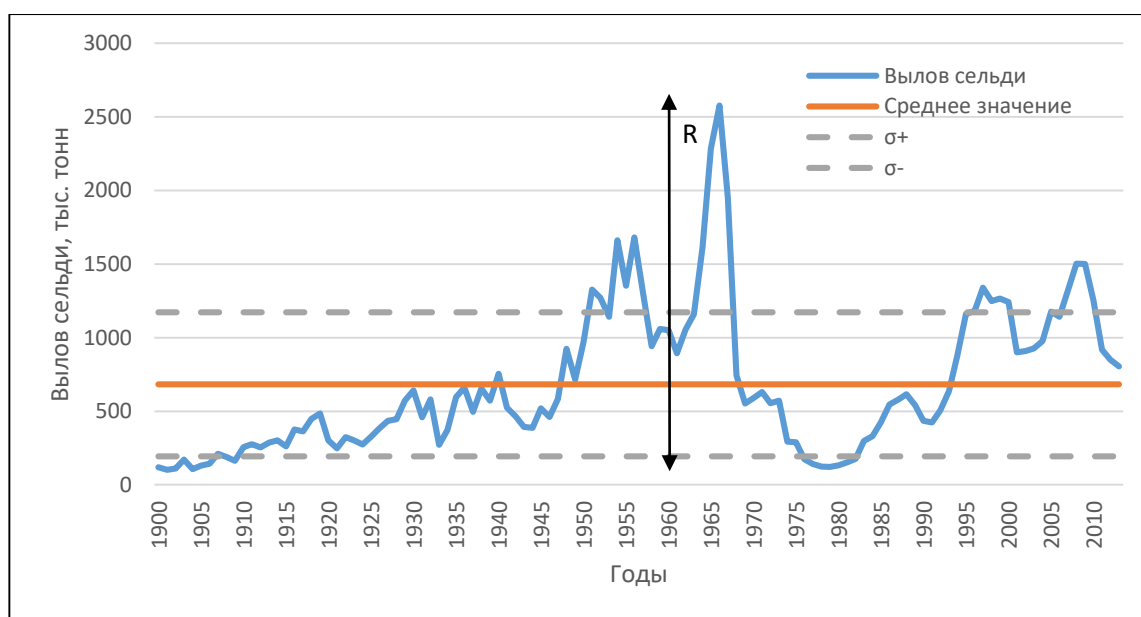


Рисунок 3.1. Среднегодовая изменчивость вылова сельди.

* $\sigma+$ и $\sigma-$ – стандартные отклонения от среднего значения.

Из таблицы 3.1 видно, что размах вариации составил 2475 тыс. тонн, а среднее значение 682,42 тыс. тонн. Коэффициент вариации, который показывает какую долю среднего значения составляет её средний разброс, составляет 6,72% от среднего значения, что меньше 33%, следовательно, сила влияющих факторов незначительна. Минимум вылова сельди наблюдался в 1901 и составлял 103 тыс. тонн, а максимум, в 1966 году – 2578 тыс. тонн.

3.2 Эмпирическая функция распределения

Для ряда среднегодового вылова сельди была рассчитана эмпирическая функция распределения (далее – ЭФР) и соответствующий ей нормальный закон распределения.

Таблица 3.2. Расчет ЭФР и расчет функции нормального закона распределения и критерия хи-квадрат (χ^2).

Номер интервала	Интервалы		Хср	Частота	Накопленная частота	Интегральная вероятность	f	nk*	nk	χ^2
	от	до								
1	103	350,50	226,75	34	34	0,30	0,0005284	149,09	149,00	88,76
2	350,50	598,00	474,25	32	66	0,58	0,0007445	210,06	210	150,88
3	598,00	845,50	721,75	10	76	0,67	0,0008123	229,19	229	209,44
4	845,50	1093,00	969,25	14	90	0,79	0,0006864	193,67	194	167,01
5	1093,00	1340,50	1216,75	15	105	0,92	0,0004492	126,74	127	98,77
6	1340,50	1588,00	1464,25	3	108	0,95	0,0002276	64,23	64	58,14
7	1588,00	1835,50	1711,75	3	111	0,97	0,0000893	25,21	25	19,36
8	1835,50	2083,00	1959,25	1	112	0,98	0,0000272	7,66	8	6,13
9	2083,00	2330,50	2206,75	1	113	0,99	0,0000064	1,80	1	0,00
10	2330,50	2578,00	2454,25	1	114	1,00	0,0000012	0,33	1	0,00
									χ^2_*	798,48
									$\chi^2_{кр}$	14,07

Далее была построена гистограмма и кривая интегральной ЭФР. На гистограмму были нанесены определенные по рисунку 3.2 моды и медиана, а также среднее значение и стандартные отклонения.

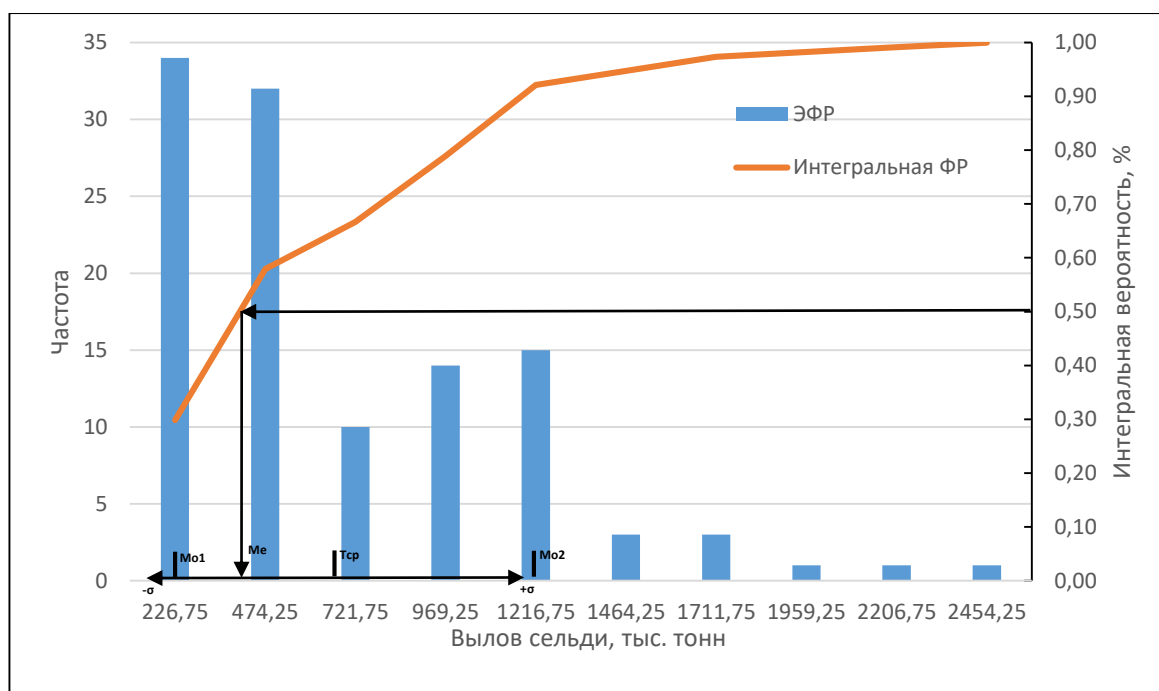


Рисунок 3.2. ЭФР значений вылова сельди.

По рисунку 3.2 видно, что выделяется двухмодальное распределение. Первая мода ($Mo1=226,75$ тыс. тонн) характеризует относительно малоинтенсивный вылов сельди, а вторая мода ($Mo2=1216,65$ тыс. тонн) характеризует относительно интенсивный вылов сельди. Медиана для этого ряда, определенная по графику, составляет 412 тонн, а вычисленная математически – 302 тыс. тонн (таблица 3.1.). Асимметричность составила 1,27, что значительно больше нуля, следовательно, ЭФР обладает положительным «хвостом» и основная масса наблюдений, а также медиана, будет меньше среднего значения. Эксцесс составляет 1,7, что больше нуля, значит ЭФР является относительно островершинной.

Асимметрия и эксцесс больше нуля. Исходя из того, что нормальному закону присуще равенство среднего значения, моды и медианы, а также нулевые значения асимметричности и эксцесса и симметричность относительно среднего [6], то можно сделать вывод о том, что ЭФР в большей степени отличается от нормального закона.

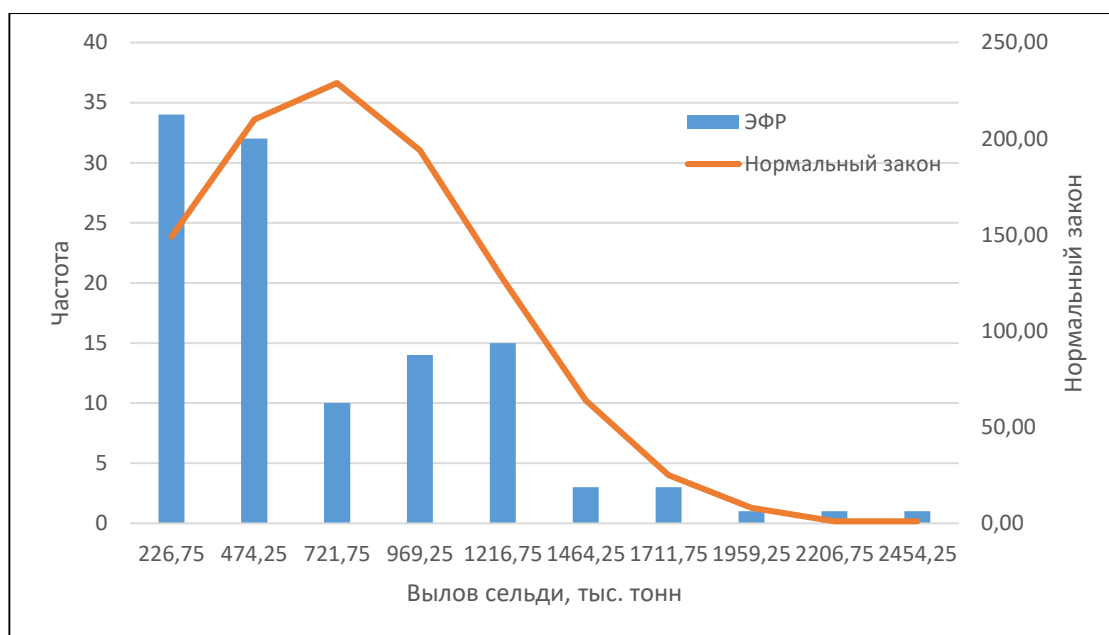


Рисунок 3.3. ЭФР и соответствующий ей нормальный закон распределения значений вылова сельди.

Для проверки соответствия ЭФР нормальному закону, нужно сравнить χ^2 и $\chi^2_{кр}$

$$\chi^2 = 798,48$$

$$\chi^2_{кр} = 2,18$$

$\chi^2 > \chi^2_{кр}$ различия между ЭФР и нормальным законом статистически значимы, ЭФР не соответствует нормальному закону. Среди факторов, влияющих на характеристику, есть один или несколько преобладающих. Налицо вмешательство человека (перевылов сельди в конце 1960-х).

3.3 Детерминированная часть ряда

3.3.1 Тренд временного ряда

Таблица 3.3. Рассчитанные характеристики линейного тренда как линейной регрессии.

Регрессионная статистика	
Множественный R	0,48
R-квадрат	0,23
Нормированный R-квадрат	0,22
Стандартная ошибка	431,90
Наблюдения	114

Дисперсионный анализ				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Регрессия	1	6187922,13	6187922,13	33,17
Остаток	112	20892077,66	186536,41	
Итого	113	27079999,79		

	Коэффициенты	Стандартная ошибка	t-статистика
Вылов сельди	275,33	81,44	3,38
<i>t</i> ₁ (<i>a</i> ₁)	7,08	1,23	5,76

Чтобы оценить значимость линейного тренда, нужно проверить его множественный R, он же коэффициент корреляции *r*. Для того, чтобы его проверить формулируется нулевая гипотеза $H_0: r = 0$ и альтернативная $H_1: r \neq 0$. Для проверки гипотез используется статистический критерий Стьюдента, рассчитаем его критическое значение $t_{кр}$, а также критерий Стьюдента t^* .

$$t^* = 6,56$$

$$t_{кр} = 1,98$$

Сравниваем t^* с $t_{кр}$

$6,56 > 1,98$, нулевая гипотеза отвергается, коэффициент корреляции $r = 0,23$ статистически значим, тренд влияет в формирование изменчивости исходного ряда.

Также для оценки линейного тренда нужно проверить на значимость линейный коэффициент $a_1 = 7,08$. Для того, чтобы его проверить формулируется нулевая гипотеза $H_0: r = 0$ и альтернативная $H_1: r \neq 0$. Для проверки гипотез используется статистический критерий Стьюдента, рассчитаем его критическое значение $t_{кр}$, а также критерий Стьюдента t^* .

$$t^*_{a_1} = 5,76$$

$$t_{кр} = 1,98$$

Сравниваем t^* с $t_{кр}$

$5,76 > 1,98$, нулевая гипотеза отвергается, линейный тренд значим.

Таблица 3.4. Рассчитанные характеристики нелинейного тренда как множественной регрессии.

Регрессионная статистика	
Множественный R	0,51
η -квадрат	0,26
Нормированный R-квадрат	0,25
Стандартная ошибка	424,81
Наблюдения	114,00

Дисперсионный анализ				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Регрессия	2	7048761,32	3524380,66	19,53
Остаток	111	20031238,47	180461,61	
Итого	113	27079999,79		

	Коэффициенты	Стандартная ошибка	t-статистика
Вылов сельди (a_0)	75,84	121,49	0,62
t_1 (a_1)	17,40	4,88	3,57
t_1^2 (a_2)	-0,09	0,04	-2,18

Чтобы оценить значимость нелинейного тренда, нужно проверить его множественный R, он же корреляционное отношение η . Для того, чтобы его проверить формулируется нулевая гипотеза $H_0: \eta = 0$ и альтернативная $H_1: \eta \neq 0$. Для проверки гипотез используется статистический критерий Стьюдента, рассчитаем его критическое значение $t_{кр}$, а также критерий Стьюдента t^* .

$$t^* = 7,30$$

$$t_{кр} = 1,98$$

Сравниваем t^* с $t_{кр}$

$7,30 > 1,98$, нулевая гипотеза отвергается, корреляционное отношение $\eta = 0,26$ значим, тренд влияет в формировании изменчивости исходного ряда.

Также для оценки линейного тренда нужно проверить на значимость нелинейные коэффициенты $a_1 = 17,40$ и $a_2 = -0,09$. Для того, чтобы его проверить формулируется нулевая гипотеза $H_0: \eta = 0$ и альтернативная $H_1: \eta \neq 0$. Для проверки гипотез используется статистический критерий Стьюдента, рассчитаем его критическое значение $t_{кр}$, а также критерий Стьюдента t^* .

$$t^*_{a_1} = 3,57$$

$$t^*_{a_2} = 2,18$$

$$t_{кр} = 1,98$$

Сравниваем t^* с $t_{кр}$

$3,57 > 1,98$, нулевая гипотеза отвергается, нелинейный тренд значим.

$2,18 > 1,98$, нулевая гипотеза отвергается, нелинейный тренд значим.

Проверка на значимость коэффициентов регрессии показала, что оба тренда статистически значимы, и вносят вклад в формирование изменчивости исходного ряда, но более предпочтительным является нелинейный тренд, так как он вносит больший вклад в дисперсию выборки, чем линейный (26% против 23%). Стоит отметить, что тренды значимы из-за человека, а не из-за природы.

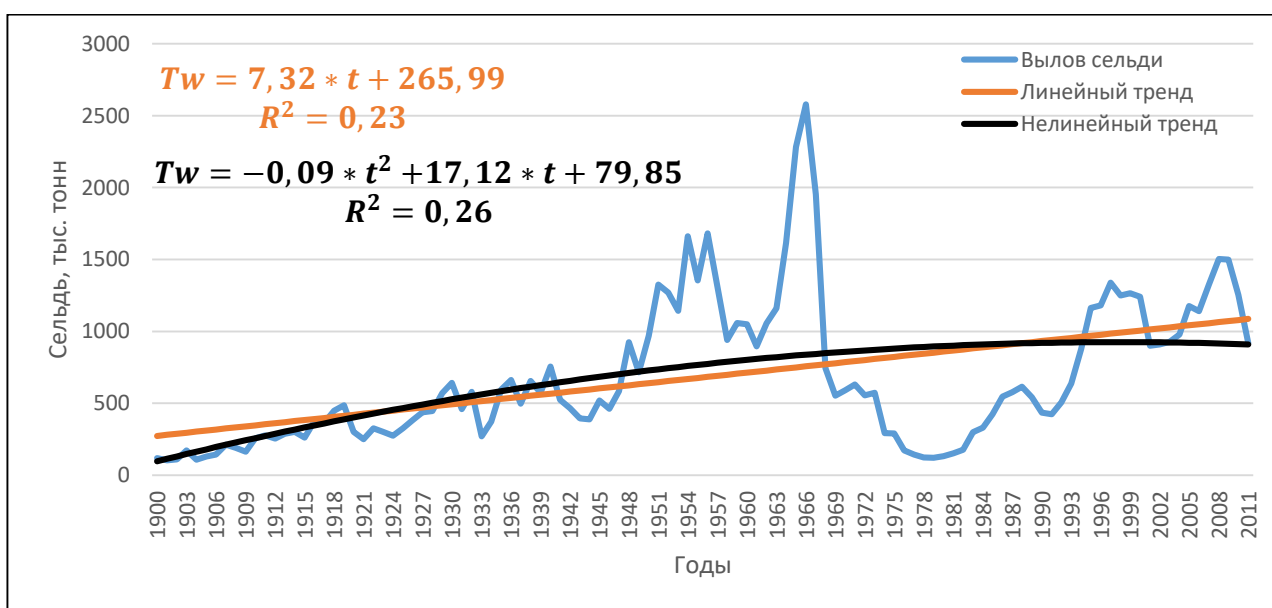


Рисунок 3.4. Среднегодовая изменчивость вылова сельди и ее линейный и нелинейный тренды.

Уравнение линейного тренда: $T_w = 7,32 * t + 265,99$

Уравнение нелинейного тренда: $T_w = -0,09 * t^2 + 17,12 * t + 79,85$

Характер тренда: рост до 1996, затем падение. Нелинейный тренд вносит значительный вклад в дисперсию исходного ряда. Величина тренда равна

коэффициенту a_1 линейного тренда и имеет размерность характеристики y за единицу дискретности. $a_1 = 7,08$ тыс. тонн/год. Тогда величина тренда $Tr = 70,8$ тыс. тонн/10 лет.

3.3.2 Гармонический анализ

Из ряда среднегодового вылова сельди вычитается статистически значимый тренд – нелинейный. Получается ряд отклонений. По значениям отклонений строится периодограмма.

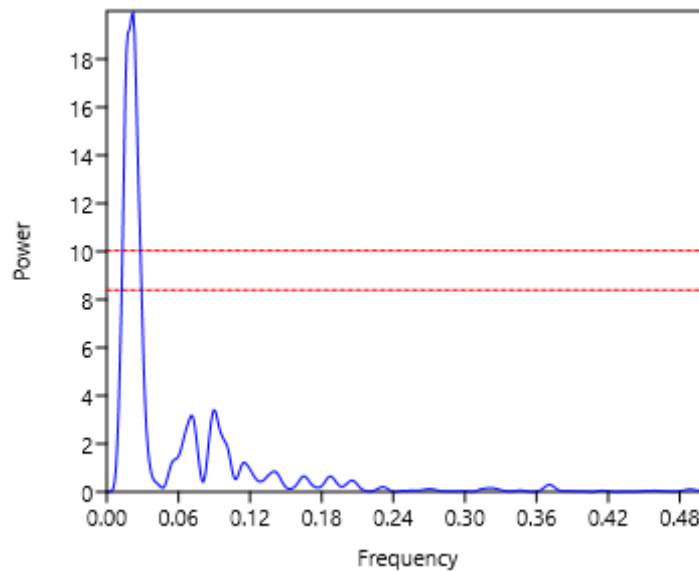


Рисунок 3.5. Периодограмма для ряда среднегодовых отклонений для вылова сельди.

По периодограмме (рисунок 3.5) определяем частоты гармоник (ω), которые соответствуют пикам периодограммы. В данном случае их 5:

$$\begin{aligned}\omega &= 0,22 \text{ (год}^{-1}\text{)} \\ \omega &= 0,071 \text{ (год}^{-1}\text{)} \\ \omega &= 0,091 \text{ (год}^{-1}\text{)} \\ \omega &= 0,115 \text{ (год}^{-1}\text{)} \\ \omega &= 0,5 \text{ (год}^{-1}\text{)}\end{aligned}$$

Таблица 3.5. Расчет характеристик гармоник.

	G ₁	G ₂	G ₃	G ₄	G ₅
ω год ⁽⁻¹⁾	0,022	0,071	0,091	0,115	0,5
T	45,2	14,1	11,0	8,7	2,0
ω рад/с	0,14	0,44	0,57	0,72	3,14
a	309,77	-130,40	-59,34	-27,73	0,76
b	-166,83	48,99	129,70	-80,56	-15,79
A	351,84	139,30	142,63	85,20	15,81
ϕ рад	2,06	-1,21	-0,43	3,47	3,09
ϕ год	14,9	-2,7	-0,8	4,8	1,0
D	61894,81	9701,73	10171,80	3629,15	125,01
V	0,35	0,05	0,06	0,02	0,001

Проверка гармоник на значимость.

Вклад гармоник в общую дисперсию ряда (V из расчета характеристик гармоник) – аналог коэффициента детерминации (r^2), следовательно, вклад может быть аналогично проверен на значимость как коэффициент детерминации. Для того, чтобы его проверить формулируется нулевая гипотеза $H_0: r = 0$ и альтернативная $H_1: r \neq 0$. Для проверки гипотез используется статистический критерий Стьюдента, рассчитаем его критическое значение $t_{кр}$, а также критерий Стьюдента t^* .

Гармоника 1 (G₁).

$$t^* = 9,61$$

$$t_{кр} = 1,98$$

Сравниваем t^* с $t_{кр}$

9,61 > 1,98, нулевая гипотеза отвергается, вклад гармоник в общую дисперсию ряда V=0,35 (35%) значим. Гармоника вносит значительный вклад в дисперсию ряда отклонений.

Гармоника 2 (G₂).

$$t^* = 2,62$$

$$t_{кр} = 1,98$$

Сравниваем t^* с $t_{кр}$

$2,62 > 1,98$, нулевая гипотеза отвергается, вклад гармоник в общую дисперсию ряда $V=0,05$ (5%) значим. Гармоника вносит значительный вклад в дисперсию ряда отклонений.

Гармоника 3 (G_3).

$$\begin{aligned}t^* &= 2,69 \\t_{кр} &= 1,98\end{aligned}$$

Сравниваем t^* с $t_{кр}$

$2,69 > 1,98$, нулевая гипотеза отвергается, вклад гармоник в общую дисперсию ряда $V=0,06$ (6%) значим. Гармоника вносит значительный вклад в дисперсию ряда отклонений.

Гармоника 4 (G_4).

$$\begin{aligned}t^* &= 1,55 \\t_{кр} &= 1,98\end{aligned}$$

Сравниваем t^* с $t_{кр}$

$1,55 < 1,98$, нулевая гипотеза принимается, вклад гармоник в общую дисперсию ряда $V=0,02$ (2%) незначим. Гармоника не вносит значительный вклад в дисперсию ряда отклонений.

Гармоника 5 (G_5).

$$\begin{aligned}t^* &= 0,28 \\t_{кр} &= 1,98\end{aligned}$$

Сравниваем t^* с $t_{кр}$

$0,28 < 1,98$, нулевая гипотеза принимается, вклад гармоник в общую дисперсию ряда $V=0,001$ (0,1%) незначим. Гармоника не вносит значительный вклад в дисперсию ряда отклонений.

Проверка гармоник на значимость показала, что значительный вклад в дисперсию ряда отклонений вносят только 3 гармоники из 5 выявленных.

Восстанавливаем ряд отклонений по уравнениям значимых гармоник:

$$G_1 = 351,84 * \cos(0,14 * t_i + 2,06)$$

$$G_2 = 139,77 * \cos(0,44 * t_i - 1,21)$$

$$G_3 = 142,63 * \cos(0,45 * t_i - 0,43)$$

Далее находим сумму гармоник, и строим по полученным значениям совмещенный график отклонений и суммы гармоник.

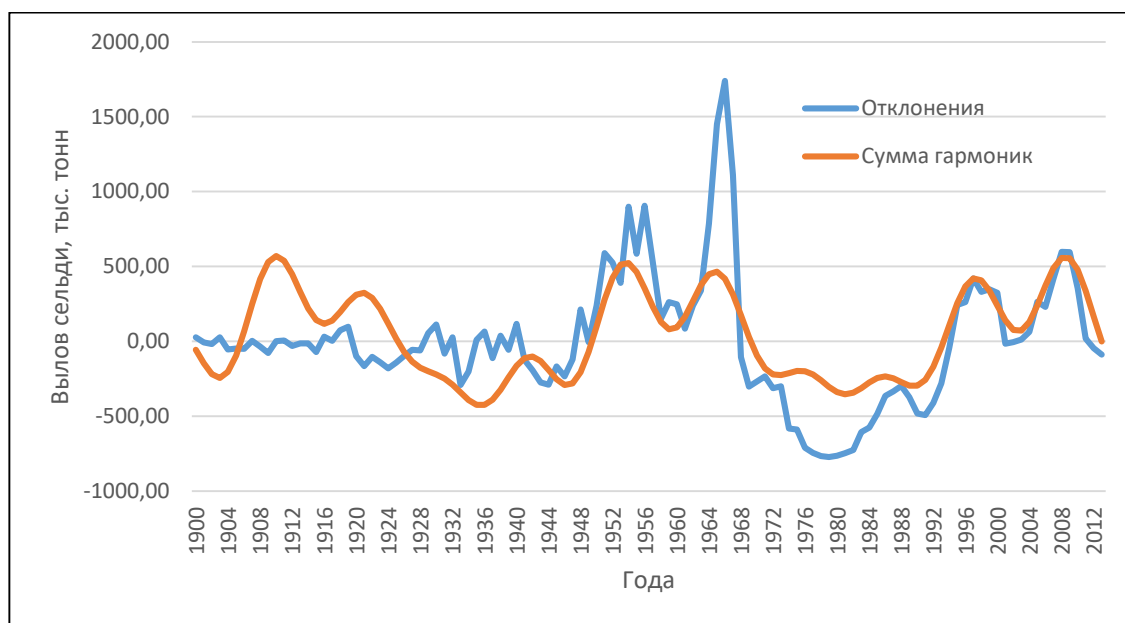


Рисунок 3.6. Совмещенный график отклонений и суммы гармоник.

Из выявленных периодов гармонических колебаний в 45,2; 14,1; 11,0; лет, все вносят значительный вклад в дисперсию исходного ряда.

Гармоника с периодом в 45,2 года связана с процессом 35-летнего колебания климата, выражающееся в переходе от холодных и влажных лет к теплым и сухим (Брикнеров цикл). Фаза составляет 14,9 лет, что означает то, что каждые 45,2 года (542 месяца), первый максимум наступит через 14,9 лет (14,9 месяцев) от начала периода. В период с 1 января 1900 года по 1 марта 1945 года первый максимум придется на 1 ноября 1914 года.

Гармоника с периодом в 11 лет связана с 11-летним циклом солнечной активности. Фаза составляет 0,8 года, что означает то, что каждые 11 лет (132 месяцев), первый максимум наступит через 0,8 года (9 месяцев) от начала

периода. В период с 1 января 1900 года по 1 января 1911 года первый максимум придется на 1 октября 1900 года.

Гармоника с периодом в 14,1 года связана с отголосками 11-летнего цикла солнечной активности. Фаза составляет 2,7 года, что означает то, что каждые 14,1 лет (169 месяцев), первый максимум наступит через 2,7 лет (32 месяца) от начала периода. В период с 1 января 1900 года по 1 февраля 1914 года первый максимум придется на 1 сентября 1902 года

Вклад G_1 в дисперсию ряда отклонений – 25%, G_2 – 5%, G_3 -5%. Суммарный вклад в дисперсию ряда всех статистически значимых гармоник – 35%, Наибольший вклад у 1 гармоники (G_1).

Из рисунка 3.6, видно, что сумма гармоник значения хорошо описывают ход ряда отклонений исходного вылова сельди.

3.3.3 Вейвлет-преобразование

Для ряда отклонений было построено вейвлет-преобразование.

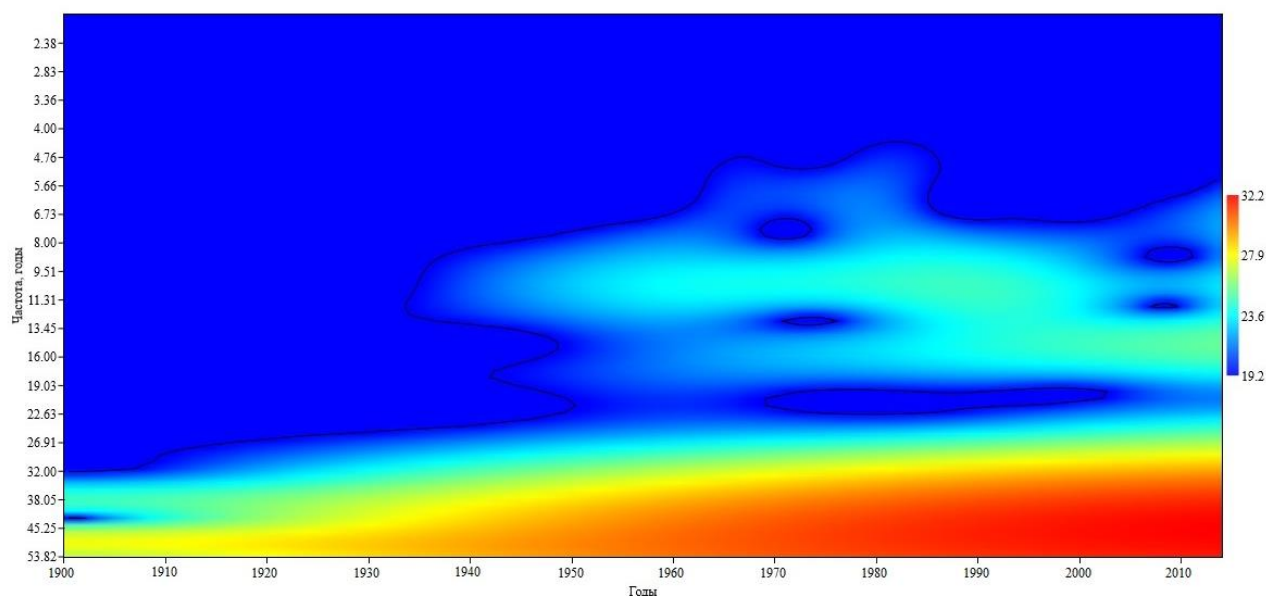


Рисунок 3.7. Вейвлет-преобразование.

Из рисунка 3.7 видно, что вклад в дисперсию происходит на 38-45-ти летней частоте и составляет, в среднем, 30%, на протяжении всего периода, но

максимальный вклад наблюдается с 1940 до 2013 года, и составляет 33%, чуть меньший вклад в дисперсию ряда происходит на 11-ти летней частоте, в период с 1940 по 2000 год, и составляет 24%, на 14-ти летней частоте вклад аналогичен 11-ти летней, но в период с 1980 по 2013 год.

Таблица 3.6. Вклад в дисперсию детерминированной части ряда.

	Дисперсия, тыс. тонн ²	Часть от общей дисперсии, %
Исходный ряд	239646	100
Тренд	62378,42	26
Гармоники	84526,58	35

3.4 Прогнозирование детерминированной и стохастической части ряда с заблаговременностью 1 год

3.4.1 Анализ ряда отклонений

Из исходного ряда вычитаются неслучайные составляющие – тренд и сумма значимых гармоник, получается ряд случайных отклонений, по которым будет предпринята попытка сделать автокорреляционный прогноз.

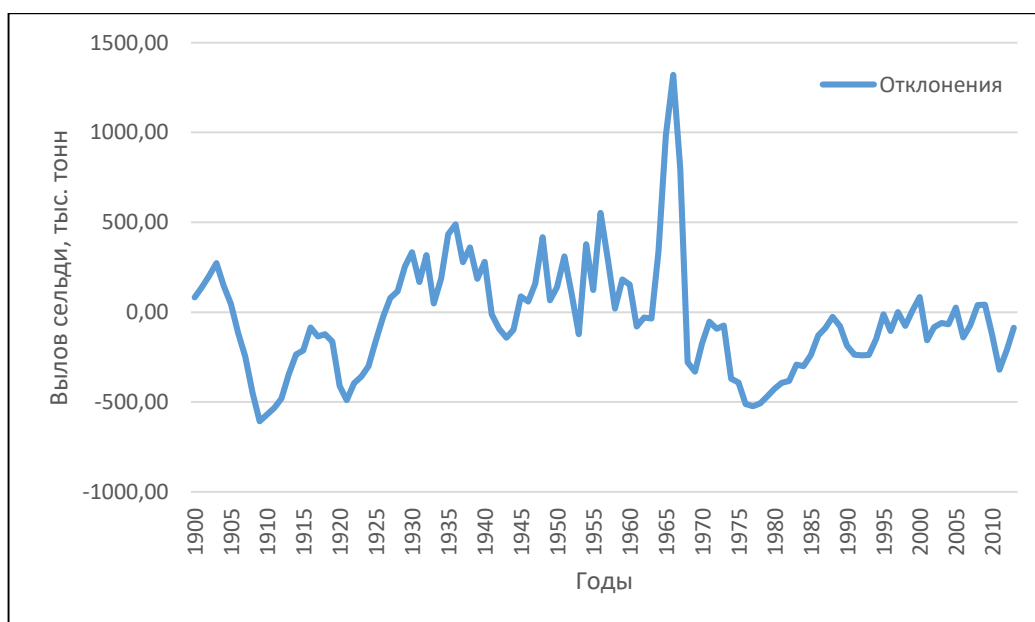


Рисунок 3.8. График хода стохастической части ряда.

Из рисунка 3.8, видно, что стохастическая часть ряда значительна, и вклад в дисперсию достаточно велик.

3.4.2 Расчет и анализ автокорреляционной функции

По ряду отклонений была рассчитана автокорреляционная функция (далее – АКФ) и уровни значимости, а также построен график.

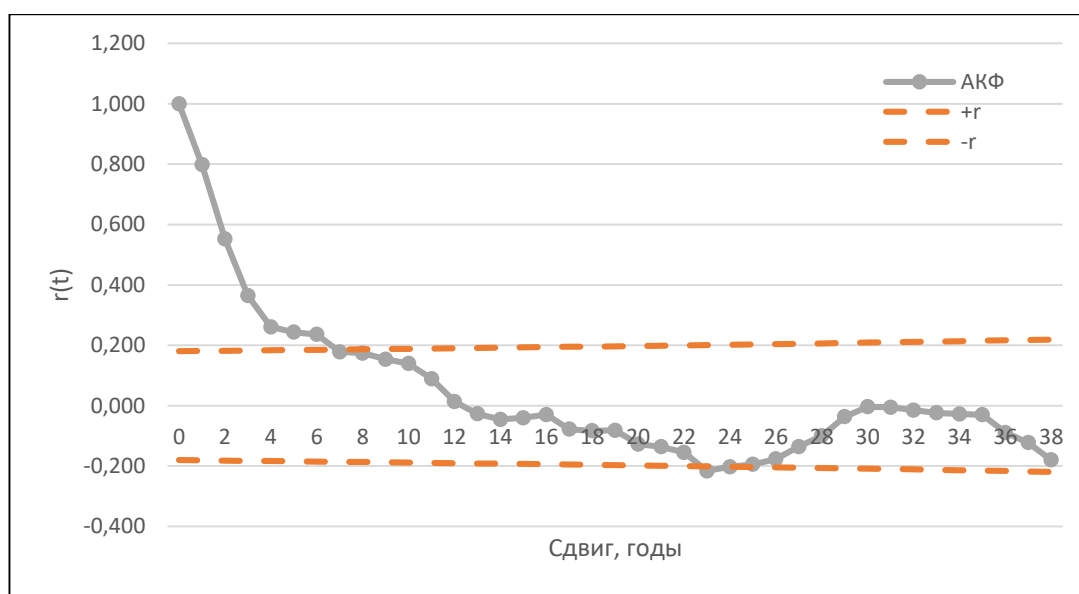


Рисунок 3.9. АКФ среднегодового вылова сельди.

Анализ АКФ

1. Инерционность процесса.

Инерционность процесса определяется по радиусу корреляции. Радиус корреляции – это сдвиг при первом пересечении функцией нуля. Инерционность процесса показывает, насколько долго характеристика сохраняет свое предыдущее состояние. Данный процесс имеет инерционность 13 лет, и является инерционным.

2. Периодичность процесса.

Если в исследуемом процессе присутствуют циклические колебания, то они отражаются на АКФ в виде локальных максимумов. Тогда период циклического колебания определяется по сдвигу АКФ, соответствующему

локальному максимуму. В данном процессе локальных максимумов не выявлено, следовательно, периодических составляющих в процессе нет.

3. Тип процесса.

По рисунку 3.9 видно, что тип процесса – красный шум.

4. Возможность автопрогноза.

Наблюдается несколько значимых коэффициентов автокорреляции, а значит можно дать автопрогноз с разной заблаговременностью, в данном случае от 0 до 4 месяцев. Так как прогноз зависит от величины коэффициента автокорреляции (чем больше, тем лучше прогноз), то в данном случае самым оптимальным будет автопрогноз с заблаговременностью 1 месяц.

3.4.3 Построение и оценка качества модели авторегрессии 1 порядка

Атопрогноз выполняется на основе модели парной линейной регрессии, которая называется авторегрессией 1 порядка (далее – АР1). Уравнение АР1 формулируется в общем виде:

$$x_{i+k} = a_1 x_i + a_0 \pm \varepsilon$$

k – заблаговременность прогноза.

С заблаговременностью 1 год уравнение будет иметь вид:

$$BC_{i+1} = a_1 BC_i + a_0 \pm \varepsilon$$

*BC – вылов сельди

Таблица 3.7. Рассчитанные характеристики АР1 для среднегодового вылова сельди.

Регрессионная статистика	
Множественный R	0,80
R-квадрат	0,64
Нормированный R-квадрат	0,63
Стандартная ошибка	189,18
Наблюдения	113,00

Дисперсионный анализ				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Регрессия	1	6936991,68	6936991,68	193,84
Остаток	111	3972449,07	35787,83	
Итого	112	10909440,75		

	Коэффициенты	Стандартная ошибка	t-статистика
a(0)	-6,94	17,96	-0,39
a(1)	0,80	0,06	13,92

Проверка качества модели.

1. Оценка значимости коэффициентов регрессии.

Для оценки значимости коэффициентов регрессии $a_0 = -6,94$ и $a_2 = 0,8$ формулируем 2 гипотезы – нулевую $H_0: r = 0$ и альтернативную $H_1: r \neq 0$. Для проверки гипотез используется статистический критерий Стьюдента, рассчитаем его критическое значение $t_{кр}$, а также критерий Стьюдента t^* .

$$t^*_{a_0} = 0,39$$

$$t^*_{a_1} = 13,92$$

$$t_{кр} = 1,98$$

Сравниваем t^* с $t_{кр}$

$0,39 < 1,98$, нулевая гипотеза принимается, коэффициент a_0 незначим.

$13,92 > 1,98$, нулевая гипотеза отвергается, коэффициент a_1 значим.

2. Оценка адекватности модели по критерию Фишера.

Для оценки значимости коэффициентов регрессии формулируем 2 гипотезы – нулевую $H_0: D_{\bar{y}} = D_{\sigma}$ и альтернативную $H_1: D_{\bar{y}} \neq D_{\sigma}$. Для проверки гипотез используется критерий Фишера (F^*), который сравнивается с $F_{кр}$.

$$F^* = 193,84$$

$$F_{кр} = 3,93$$

Сравниваем F^* с $F_{кр}$

$193,84 > 3,93$ нулевая гипотеза о равенстве дисперсий отвергается. Проверка адекватности регрессионной модели по критерию Фишера при уровне значимости 0.05 показала, что модель адекватна.

3. Оценка коэффициента детерминации.

Коэффициент детерминации $r^2 = 0,64$, что меньше 0,7, это означает, что данной моделью описывается 64% дисперсии исходного ряда, а также свидетельствует нам о том, что относительная часть зависимой переменной, которая определяется независимой переменной, совсем невелика – данных независимых переменных недостаточно для описания дисперсии исходного ряда.

4. Оценка стандартной ошибки модели.

Стандартная ошибка модели $\sigma_\varepsilon = 62,81$ тыс. тонн, что меньше стандартного отклонения ряда зависимой выборки $\sigma_y = 311,91$ тыс. тонн.

Так как стандартная ошибка модели меньше стандартного отклонения ряда зависимой выборки, то модель можно использовать для прогноза.

Данная модель не будет обладать хорошим качеством, так как при оценке значимости коэффициентов регрессии коэффициент a_0 оказался незначим и коэффициент детерминации меньше 0,7.

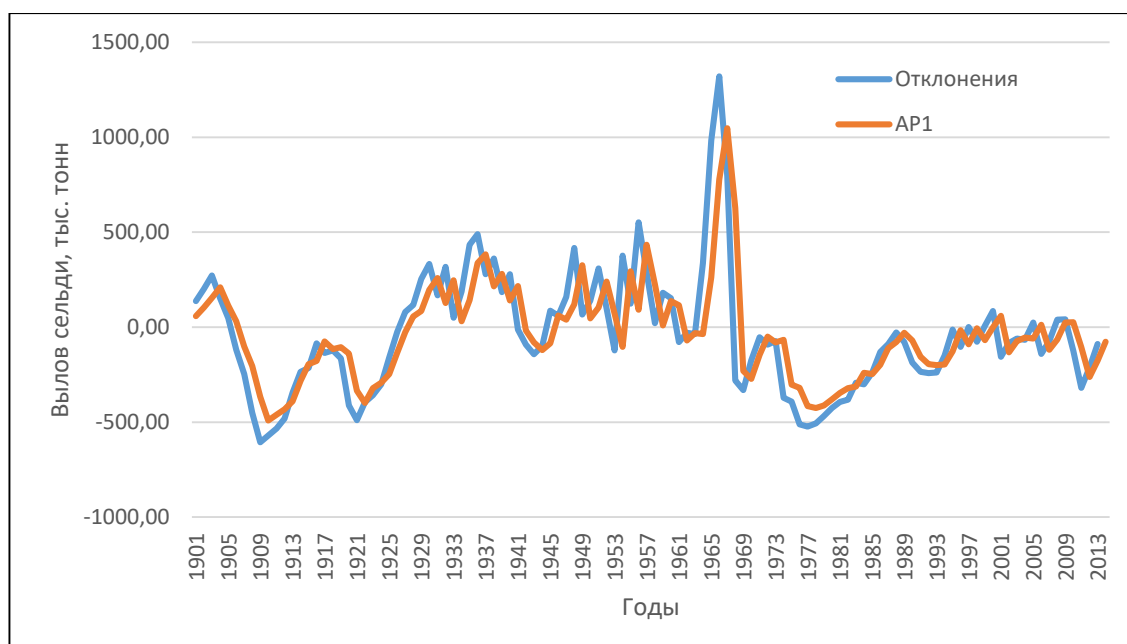


Рисунок 3.10. Совмещенный график отклонений и модели AR1 с заблаговременностью 1 год.

Модель AR1 хорошо описывает ход отклонений. Вклад в дисперсию исходного ряда составляет 26%.

3.4.4 Прогноз тренда с заблаговременностью 1 год

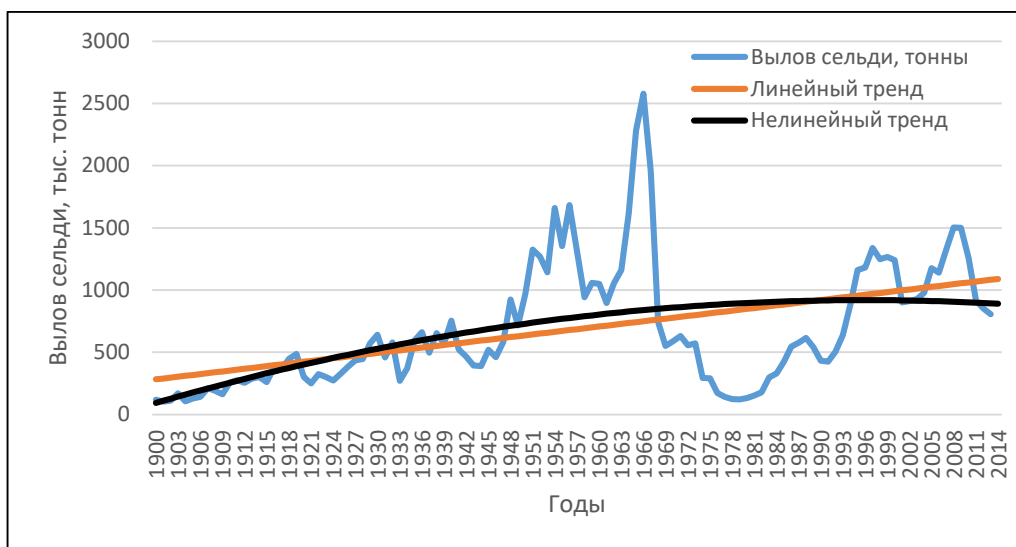


Рисунок 3.11. Ряд вылова сельди и его линейный и нелинейный тренды с заблаговременностью 1 год.

Из рисунка 3.11 видно, что тенденция роста линейного тренда продолжилась, а нелинейный продолжил уменьшаться. Вклад в дисперсию исходного ряда составляет 25%.

3.4.5 Прогноз гармоник с заблаговременностью 1 год

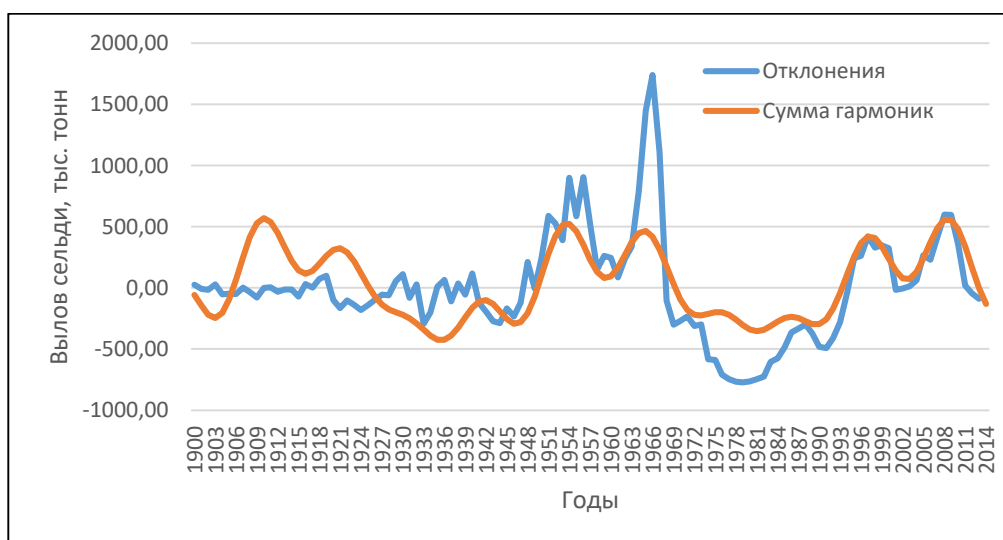


Рисунок 3.12. Совмещенный график отклонений и суммы гармоник с заблаговременностью 1 год.

Сумма гармоник хорошо описывает ряд отклонений и вносит наибольший вклад в дисперсию исходного ряда – 35%.

3.4.6 Сборка временного ряда с заблаговременностью 1 год

Уравнение:

$$BC = Tr + C + K \pm \varepsilon$$

* Tr – значимый тренд; C – сумма гармоник; K – Модель AP1.

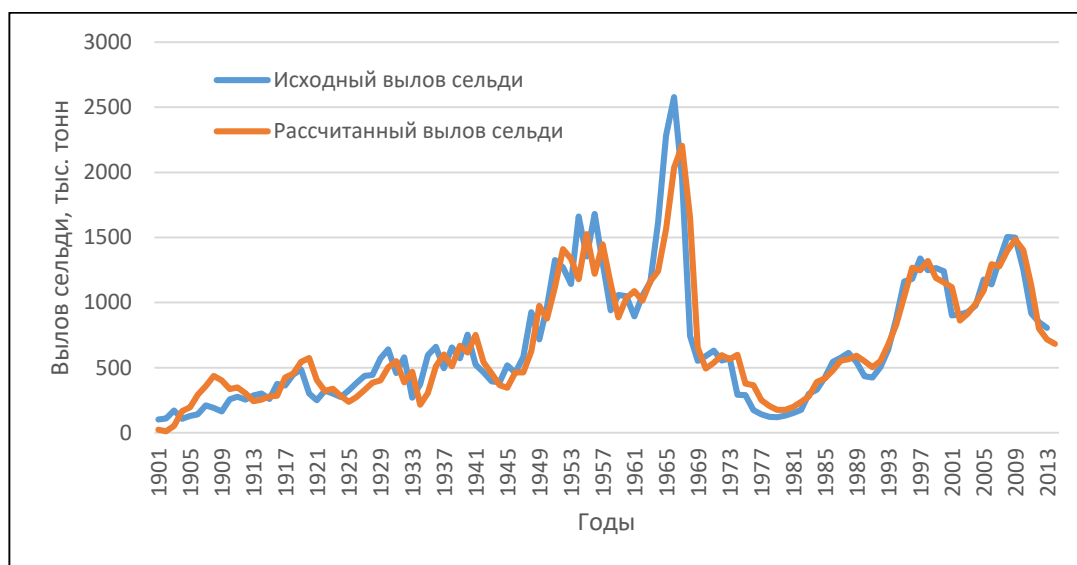


Рисунок 3.13. Исходный и рассчитанный ряд среднегодового вылова сельди с заблаговременностью 1 год.

Из рисунка 3.13 видно, что рассчитанные значения вылова сельди с заблаговременностью 1 год неплохо описывают фактическую изменчивость среднегодового вылова сельди.

Таблица 3.8. Вклад в дисперсию каждой составляющей прогностической модели.

	Дисперсия, тыс. тонн ²	Часть от общей дисперсии, %
Исходный ряд	239646	100
Тренд	59638	25
Гармоники	84435	35
AP1	61476	26
Шум	34098	14

3.5 Выводы по главе

Выполнив статистический анализ ряда вылова сельди, можно сказать, что тенденция нелинейного тренда к уменьшению продолжится, а линейного – к росту, наиболее сильно на характеристику влияет Брикнеровский цикл. ЭФР не соответствует нормальному закону. Среди факторов, влияющих на характеристику, есть один или несколько преобладающих. Налицо вмешательство человека (перевылов сельди в конце 1960-х). Исходя из всех проведенных расчетов, можно сделать соответствующие выводы:

1. Статистическими методами можно пользоваться при прогнозировании вылова сельди.

2. Ряд можно разделить на детерминированную и стохастическую части.

3. Детерминированной частью ряда описывается 60% дисперсии исходного ряда (25% вклад нелинейного тренда, 35% вклад суммы гармоник), 40% описывает стохастическая часть ряда.

4. По стохастической части ряда был сделан прогноз с заблаговременностью 1 год. Прогноз нельзя назвать точным, так как при оценке значимости коэффициентов регрессии коэффициент a_0 оказался незначим и коэффициент детерминации меньше 0,7. Модель AP1 описывает 26% дисперсии исходного ряда, 25% вклад нелинейного тренда, 35% вклад суммы гармоник. Вклад шума составил 14%.

Заключение

В данной работе был исследован статистический ряд вылова сельди в период с 1900 по 2013 год.

Данная рыба является очень важным объектом промысла не только в России, в которой вылов сельди занимает 2 место по вылову в северном рыбохозяйственном бассейне, но во всем мире.

На основе статистических данных были произведены расчеты, и их анализ. Исходя из полученных результатов можно сказать о том, что статистические методы можно использовать при анализе и прогнозе вылова сельди. На вылов сельди влияет, в первую очередь, человек, и только во вторую – природные факторы. Тенденция вылова с 1900 года росла, а с развитием гидролокатора, силового блока и кошельковых неводов из синтетического волокна промысел сельди достиг своего пика в конце 1960-х годов, что привело к её перевылову. Так вылов в 1966 году составил 2578 тыс. тонн, а в 1969 году этот показатель составлял всего 552 тыс. тонн, а к 1980 году и вовсе 132 тыс. тонн. Только к концу 1980-х годов вылов восстановился, и продолжил тенденцию роста. Наиболее сильно из природных факторов на вылов влияет Брикнеровский цикл.

Прогноз вылова показал, что вылов сельди продолжит уменьшение, которое началось в 2009 году, и в 2014 году он должен составить 681 тыс. тонн. Так как данных позднее 2013 года в открытом доступе нет, то проверить этот прогноз на соответствие не предоставляется возможным.

Ситуация, которая произошла в конце 1960-х годов, должна послужить уроком для будущих поколений.

Список использованной литературы

1. Сайт ПИНРО. [Электронный ресурс] URL: <http://pinro.vniro.ru/ru/> (дата обращения 09.10.2021).
2. Владимир Ожигин «Баренцево море». (дата обращения 03.04.2022).
3. Рыба Баренцева моря [Электронный ресурс] URL: <http://barenzevo.arktifiksh.com/index.php/ryba-barentseva-morya/257-seld-atlanticheskaya> (дата обращения 04.04.2022).
4. Биология Атлантическо-скандинавской сельди [Электронный ресурс] URL: <https://pandia.ru/text/78/002/68274.php> (Дата обращения 06.04.2022).
5. Жизнь Атлантическо-скандинавской сельди [Электронный ресурс] URL: <https://moreprodukty.info/seld-atlanticheskaya/>
6. Гордеева С.М. Практикум по курсу Статистические методы обработки и анализа гидрометеорологической информации. / С.М. Гордеева. - СПб: РГГМУ, 2017.
7. Методика выполнения Вейвлет-анализа. [Электронный ресурс] URL: <https://habr.com/ru/post/449646/>

