

Санкт-Петербургский государственный университет
Прикладная математика - процессы управления
Кафедра Моделирования экономических систем

Проколова Виктория Викторовна

Выпускная квалификационная работа магистра

**Статистическое исследование последствий
чрезвычайных ситуаций в России**

Направление 01.04.02 «Прикладная математика и информатика»
Основная образовательная программа ВМ.5759 «Цифровая экономика»

Научный руководитель,
кандидат физ.-мат. наук,
доцент кафедры МТИиСР
Панкратова Я. Б.

Санкт-Петербург

2021

Содержание

Введение	3
Постановка задачи	5
Обзор литературы	6
Глава 1. Описательная статистика чрезвычайных ситуаций в России	8
Глава 2. Анализ временного ряда численности погибших	13
2.1. Предобработка временного ряда	13
2.2. Моделирование временного ряда	17
2.3. Прогнозирование временного ряда	23
Глава 3. Визуализация количества погибших на карте России	29
Заключение	31
Список литературы	32
Приложение	34

Введение

На мир повлияло большое количество техногенных аварий, серьезных стихийных бедствий, таких как землетрясения, цунами и ураганы уносящие тысячи человеческих жизней и разрушающих экономику стран. Защита населения и территорий от катастрофических угроз является одной из основных проблем устойчивого развития общества как в международном, так и в государственном масштабах.

Чрезвычайные ситуации (в дальнейшем будем использовать сокращение ЧС) приводят к физическим, экономическим и социальным потерям для живых и неживых существ и влияют на повседневную жизнь и деятельность человека. Хотя, например, стихийные катастрофы возникают в результате природных процессов Земли, их последствия зависят от человеческого фактора. Размеры повреждений формируются в зависимости от плотности населения и эффективности аварийно-спасательных служб. Человеческий фактор может оказывать положительное или отрицательное влияние на результаты любой катастрофы.

Чрезвычайные ситуации приводят к экономическим потерям, таким как ущерб инфраструктуре, рост уровня безработицы, потеря сырья, производственные потери, увеличение государственных расходов и реконструкции.

Основной процент жертв от ЧС в России составляют техногенные катастрофы. В 2019 году данный процент был равен 75,94 от общего числа ЧС, тем самым представляя серьезную опасность социально-экономическому развитию страны [2]. В мире Россия занимает 3 место по количеству пожаров в среднем в год. В 2017 году в России число жертв от пожаров составило 7816 человек, при этом в США - 3400 че-

ловек [10]. Принимая во внимание тот факт, что численность населения США больше России практически в 2,5 раза, а количество погибших от пожаров меньше в 3 раза, показывает актуальность проблемы возникновения чрезвычайных ситуаций.

Состояние защиты от чрезвычайных ситуаций в Российской Федерации оценивается национальным стандартом ГОСТ Р 22.10.02-2016 [6]. По итогам 2019 года требования стандарта были выполнены всего в 6 субъектах Российской Федерации из 85: в Республиках Дагестан, Северной Осетии – Алании, Чеченской Республике, Камчатском крае, Сахалинской области и в Чукотском автономном округе. В 1 субъекте (Воронежской области) фактический индивидуальный риск ЧС и пожаров превысил установленный допустимый уровень более чем в 10 раз, т. е. считается неприемлемым [2].

Работа имеет следующую структуру. В первой главе проводится описательная статистика чрезвычайных ситуаций в России. Вводятся понятия, производится сбор данных и их анализ. Во второй главе производится прогноз ежемесячной динамики последствий техногенных чрезвычайных ситуаций (численности погибших и финансовых потерь) с применением моделей ARIMA/SARIMA и Хольта-Уинтерса. В третьей главе проводится анализ субъектов Российской Федерации: статистика количества пожаров и количества погибших.

Постановка задачи

Целью работы является изучение теоретических основ чрезвычайных ситуаций в России и анализирование данных, характеризующих число погибших и принесенный материальный ущерб от чрезвычайных ситуаций. Для выполнения поставленных целей необходимо решить ряд задач:

- Изучить научные публикации и законы, связанные с чрезвычайными ситуациями.
- Собрать данные по показателям ЧС.
- Применить модели для прогноза на 2020 год числа пострадавших в техногенных пожарах.
- Проанализировать субъекты РФ.

Обзор литературы

В ходе написания работы и проведения данного исследования была использована научная, учебно-методическая литература, публикации из научных изданий, а также правовые и нормативно-технические документы.

Понятие чрезвычайной ситуации было взято в Федеральном законе от 21.12.1994 № 68-ФЗ (ред. от 01.04.2020) «О защите населения и территорий от чрезвычайных ситуаций природного и техногенного характера» [5].

Классификация ЧС утверждена в следующих документах:

- Постановление Правительства РФ от 21.05.2007 № 304 «О классификации чрезвычайных ситуаций природного и техногенного характера» [7].
- Приказ МЧС РФ от 21.11.2008 №714 «Об утверждении Порядка учета пожаров и их последствий» [11].
- Приказ от 24.02.2009 №92 О внесении изменений в приказ МЧС России от 08.07.2004 №329 «Об утверждении критериев информации о чрезвычайных ситуациях» [12].

Причем в последнем представленном приказе [12] было принято решение статистику по пожарам формировать отдельно от общих данных по чрезвычайным ситуациям.

Для сбора данных по ЧС был использован государственный доклад «О состоянии защиты населения и территорий от ЧС природного и техногенного характера», который ежегодно публикуют МЧС России и Министерство Российской Федерации по делам гражданской обороны, чрез-

вычайным ситуациям и ликвидации последствий стихийных бедствий. В данном источнике содержится статистика о произошедших чрезвычайных ситуациях за год, как в целом по стране, так и отдельно по регионам РФ [2].

Ежегодно НИИ противопожарной обороны министерства РФ по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий» публикует сборник по статистике ежемесячных пожаров и основных последствий, а также показателям пожарной деятельности как в целом по России и отдельно по каждому региону РФ [1].

Для изучения регрессионного анализа была использована книга Буре В.М., Парилина Е.М., Седаков А.А. Методы прикладной статистики в R и Excel [14]. При проведении анализа временных рядов была рассмотрена книга Буре В.М., Евсеева Е.А. Основы эконометрики [13].

Глава 1. Описательная статистика чрезвычайных ситуаций в России

Для того, чтобы анализировать данные по чрезвычайным катастрофам, необходимо ввести основные определения и понятия ЧС.

Определение 1. [5] *Чрезвычайная ситуация* - это обстановка на определенной территории, сложившаяся в результате аварии, опасного природного явления, катастрофы, распространения заболевания, представляющего опасность для окружающих, стихийного или иного бедствия, которые могут повлечь или повлекли за собой человеческие жертвы, ущерб здоровью людей или окружающей среде, значительные материальные потери и нарушение условий жизнедеятельности людей.

На Рис.1 представлена классификация чрезвычайных ситуаций в России.



Рис. 1: Классификация ЧС в России

Рассмотрим подробнее, какие ситуации относятся к тем или иным

чрезвычайным катастрофам в зависимости от характера источника возникновения.

Техногенные ЧС	Взрывы, пожары, аварии на химически опасных объектах, выбросы радиационных веществ на радиационно опасных объектах, аварии с выбросом экологически опасных веществ, обрушение зданий, аварии на системах жизнеобеспечения, транспортные катастрофы
Биолого-социальные ЧС	Особо опасная или широко распространенная инфекционная болезнь людей, сельскохозяйственных животных и растений
Природные	Землетрясения, цунами, наводнения, извержения вулканов, оползни, сели, ураганы, смерчи, бури, природные пожары
Экологические	Аномальные изменения состояния природной среды: загрязнения биосферы, разрушение озонового слоя, опустынивание, кислотные дожди

Основными показателями ЧС являются [8]:

- Количество чрезвычайных ситуаций;
- Число погибших и пострадавших (понесших материальный или физический ущерб);
- Материальные убытки.

Проанализируем ситуацию в России, связанную с техногенными, природными и биолого-социальными ЧС.

В 2019 году на территории РФ произошло 266 чрезвычайных ситуаций, из них:

- 202 техногенных ЧС (75,94%). Всего погибло 498 человека, пострадало - 2532, принесенный материальный ущерб равен 4752,476 млн руб.
- 49 природных ЧС (18,42%). Всего погибло 34 человека, пострадало - 118374 человека, материальные убытки равны 15019,962 млн руб.

- 15 биолого-социальных ЧС (5,64%). Число погибших равно 0, число пострадавших равно 5 людям, и материальный ущерб равен 736,45 млн руб.
- Крупных террористических актов зафиксировано не было.

Было замечено, что огромную опасность для жизни людей представляют техногенные аварии (около 93,61% от общего числа погибших), где крупную долю занимают ДТП с тяжкими последствиями (число погибших от 5 человек и более/число пострадавших от 10 человек и более) - 58,91%, авиационные катастрофы и гидродинамические аварии. В то время как стихийные бедствия являются главной причиной большого количества пострадавших людей (97,90% от общего числа пострадавших), где главными факторами возникновения природного ЧС являются сильные дождь и снегопад, крупный град, опасные гидрологические явления (99,04% от общего числа пострадавших от природных ЧС) и крупные природные пожары.

На Рис.2 и Рис.3 представлены диаграммы числа погибших и пострадавших от разных типов чрезвычайных ситуаций за 2013 - 2019 года. Можно заметить, что пик погибших приходился на 2016 год, после же в целом замечается тенденция снижения.



Рис. 2: Количество погибших от чрезвычайных ситуаций

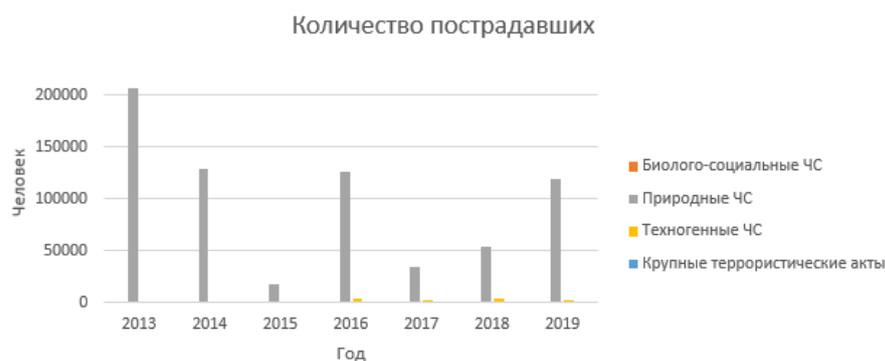


Рис. 3: Количество пострадавших от чрезвычайных ситуаций

Теперь рассмотрим обстановку, связанную с пожарами в Российской Федерации. В целом число погибших и пострадавших снижается с каждым годом, но в 2019 году произошел резкий скачок количества пожаров по сравнению с 2018 годом (на 257,6%), главной причиной которых было неосторожное обращение с огнем (72%). На Рис.4 представлена общая статистика количества погибших и пострадавших от пожаров за 2013-2019 года.



Рис. 4: Статистика погибших и пострадавших от пожаров

На Рис. 5 представлен полученный материальный ущерб от чрезвычайных ситуаций и пожаров за 2013 - 2019 года. Как можно заметить, в 2013 году максимальные финансовые убытки были в результате природных ЧС.



Рис. 5: Принесенный материальный ущерб

После проведения общей описательной статистики всех последствий чрезвычайных ситуаций и пожаров, было замечено, что крупную долю погибших и материальный ущерб обуславливают техногенные ЧС и пожары. Именно по этой причине дальнейший анализ в следующих главах будет посвящен данным типам катастроф.

Глава 2. Анализ временного ряда численности погибших

2.1. Предобработка временного ряда

Определение 2. [13] Совокупность значений изучаемого показателя y за несколько последовательных моментов (периодов) времени: $y = 1, 2, 3, \dots, n$ называют временным рядом (в дальнейшем будем использовать аббревиатуру ВР).

На Рис.6 представлен временной ряд, состоящий из количества погибших в техногенных пожарах за период с 2013 по 2019 годы.



Рис. 6: Количество погибших в техногенных пожарах с 2013 по 2019 годы

Значения временного ряда складываются из следующих компонент [13]:

- Тренд: систематическая составляющая долговременного действия.
- Сезонная компонента (период колебания не превышает 1 года): регулярные изменения, обусловленные влиянием на наблюдаемое значение показателя некоторого из экзогенных факторов, действующих с заранее известной периодичностью.

- Циклическая компонента (период колебания больше 1 года): регулярная компонента, не связанная с сезонностью.
- Случайная компонента: накладывается на системные компоненты – обусловлена действием большого числа побочных причин.

Произведем декомпозицию временного ряда, то есть выделим основные системные компоненты.

Декомпозиция аддитивной модели временного ряда:

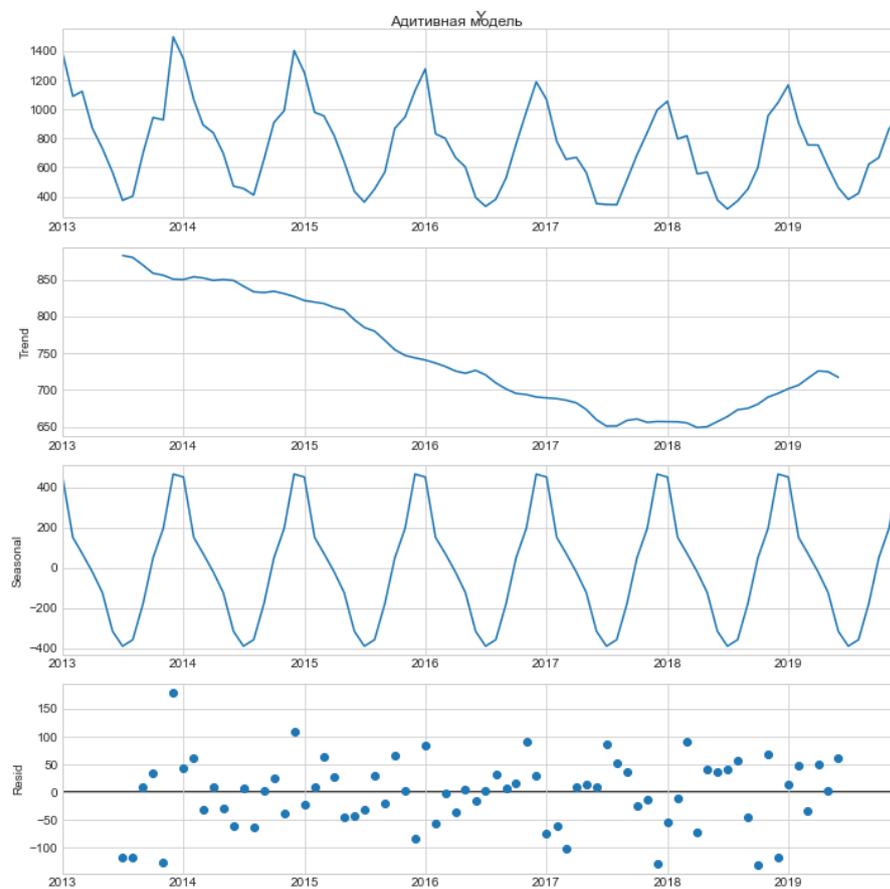


Рис. 7: Декомпозиция аддитивной модели

Декомпозиция мультипликативной модели временного ряда:

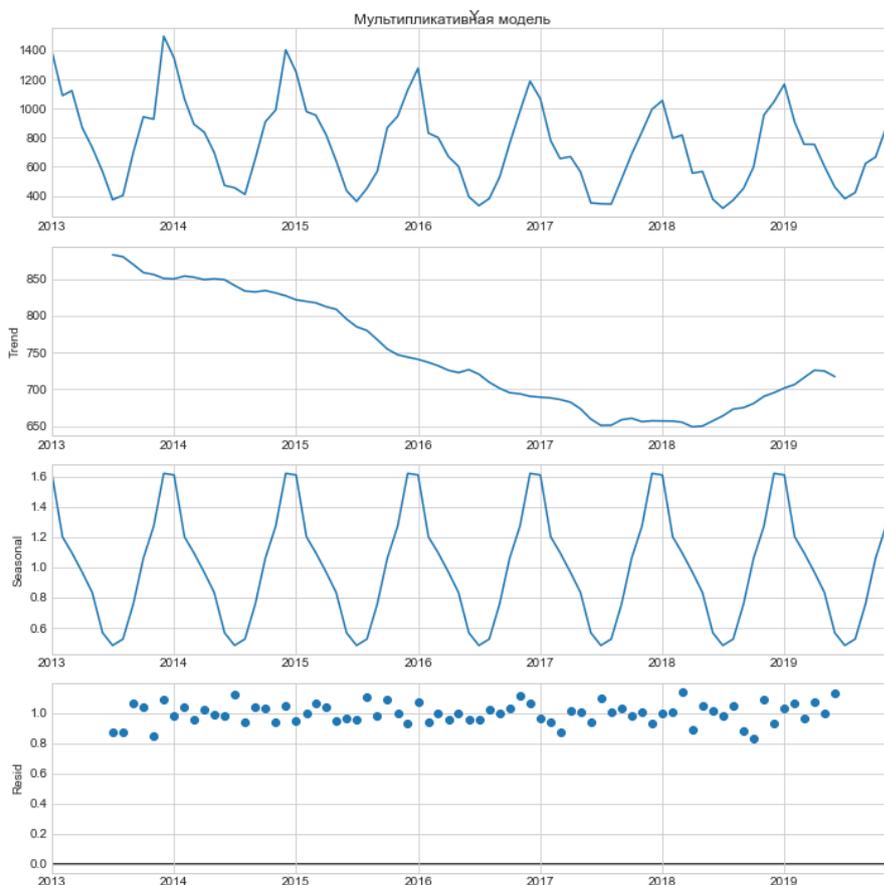


Рис. 8: Декомпозиция мультипликативной модели

Визуально можно заметить, что имеем дело с нисходящим линейным трендом, а также годовой сезонностью, где наблюдается спад числа погибших в теплое время года и подъем в холодное время года., что может быть обусловлено перегрузкой электросетей.

После выделения системных компонент (декомпозиции) ВР может быть представлен в виде суперпозиции основных компонент:

- Аддитивная модель - временной ряд состоит из суммы соответствующих компонент:

$$y(t) = u(t) + s(t) + \epsilon(t), \quad (1)$$

- Мультипликативная модель - временной ряд представляется произ-

ведением компонент:

$$y(t) = u(t)s(t)\epsilon(t), \quad (2)$$

где $y(t)$ – наблюдения ВР, $u(t)$ – тренд, $s(t)$ – сезонная компонента, $\epsilon(t)$ – случайная компонента.

2.2. Моделирование временного ряда

1. Определение наличия тренда

Осуществим проверку на случайность временного ряда для подтверждения наличия тренда. [13]

Основные критерии для проверки гипотезы случайности:

- 1) критерий серий, основанный на медиане выборки;
- 2) критерий “восходящих и нисходящих” серий.

Критерий серий, основанный на медиане выборки

- а) Ранжируем ряд и посчитаем медиану (медиана = 743).
- б) Вычислим последовательность d_i :

	Дата	Количество погибших, чел	d_i
0	2013-01	1388	1
1	2013-02	1089	1
2	2013-03	1123	1
3	2013-04	869	1
4	2013-05	732	-1
...
79	2019-08	423	-1
80	2019-09	623	-1
81	2019-10	667	-1
82	2019-11	865	1
83	2019-12	954	1

Рис. 9: Последовательность d_i

- в) Определим максимальную длину серии и количество серий:

Максимальная длина серии: $t_{max} = 8$;

Количество серий: $\nu(n) = 15$.

- г) Проверим гипотезу о случайности:

$$\nu(n) = 15 < \nu_{cr} = 33.572 \text{ и } t_{max} = 8 > t_{cr} = 6.367$$

Следовательно, гипотеза на 5% уровне значимости о случайности ряда отклоняется.

-Критерий “восходящих и нисходящих” серий

а) Вычислим последовательность d_i :

	Дата	Количество погибших, чел	d_i
0	2013-01	1388	-1
1	2013-02	1089	1
2	2013-03	1123	-1
3	2013-04	869	-1
4	2013-05	732	-1
...
79	2019-08	423	1
80	2019-09	623	1
81	2019-10	667	1
82	2019-11	865	1
83	2019-12	954	Nan

Рис. 10: Последовательность d_i

б) Определим максимальную длину серии и количество серий:

Максимальная длина серии: $t_{max} = 8$;

Количество серий: $\nu(n) = 24$.

в) Проверим гипотезу о случайности:

$t_0 = 6$, так как $n = 84$ и $26 < n \leq 153$.

$\nu(n) = 24 < \nu_{cr} = 48.175$ и $t_{max} = 8 > t_0 = 6$.

Следовательно, гипотеза на 5% уровне значимости о случайности ряда отклоняется.

2.Определение наличия сезонности

Для выявления сезонной компоненты будем использовать автокорре-

ляционную функцию.

$$acf_k = \frac{\sum_{i=k+1}^n (x_i - x^k)(x_{i-k} - x^{k+1})}{\sqrt{\left(\sum_{i=k+1}^n (x_i - x^k)^2\right) \left(\sum_{i=k+1}^n (x_{i-k} - x^{k+1})^2\right)}}$$

$$x^k = \frac{1}{n-k} \sum_{i=k+1}^n x_i,$$

$$x^{k+1} = \frac{1}{n-k} \sum_{i=k+1}^n x_{i-k}.$$

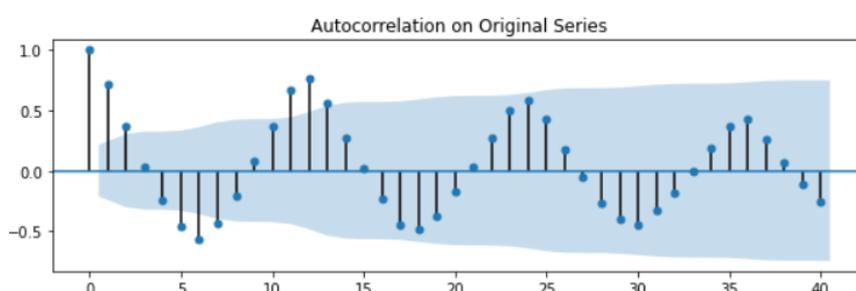


Рис. 11: Определение сезонности с помощью автокорреляции

На рисунке 11 мы видим, что на 12, 24, 36 лаге значение автокорреляционной функции стремится к 1, а значит, имеем дело с годовой сезонностью.

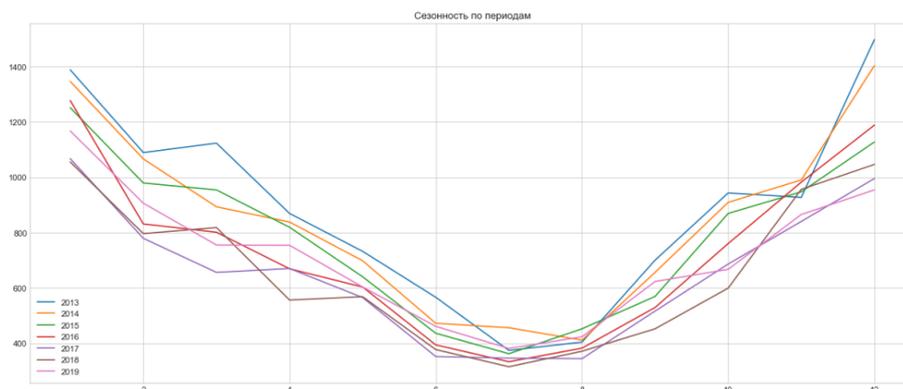


Рис. 12: Сезонные периоды (2013 - 2019)

На рис. 12 наблюдаются небольшие различия в форме сезонной волны, что, возможно, говорит о непостоянстве сезонных периодов за 2013

– 2019 года. Вследствие этого можно предположить, что имеет место мультипликативная модель, а не аддитивная.

3. Моделирование временного ряда

Самым известным способом определения тенденции развития является сглаживание временного ряда. Суть данного алгоритма проста: процесс сглаживания сводится к замене исходных наблюдений ВР вычисленными значениями, которые меньше подвержены колебательному процессу, также более четко просматривается тенденция развития.

Метод скользящей средней [15]:

$$\hat{y}_t = \frac{\sum_{i=t-p}^{t+p} y_t}{2p + 1},$$

где y_t – наблюдение исходного ВР в момент времени t ; \hat{y}_t – значение скользящей средней в момент t ; $2p + 1$ – длина интервала сглаживания.

Затем нужно рассчитать отклонения вычисленных значений сглаженного ряда от реальных.

$$e = y_t - \hat{y}_t$$

Оценки сезонной составляющей вычисляются следующим образом:

$$S_l = y_t - \bar{e},$$

где $\bar{e} = \frac{1}{m} \sum_{l=1}^m e_l$, $m = 12$

После выполненных вычислений и расчетов значений тренда была получена мультипликативная модель временного ряда погибших от техногенных пожаров.

$$\hat{y} = (784.79 - 0.702t) * S_l,$$

где $l = \overline{1, 12}$, а значения S_l представлены в таблице ниже.

Сезонная компонента	Вычисленные значения
S_1	1.608
S_2	1.2
S_3	1.095
S_4	0.966
S_5	0.832
S_6	0.566
S_7	0.483
S_8	0.525
S_9	0.757
S_{10}	1.059
S_{11}	1.271
S_{12}	1.619

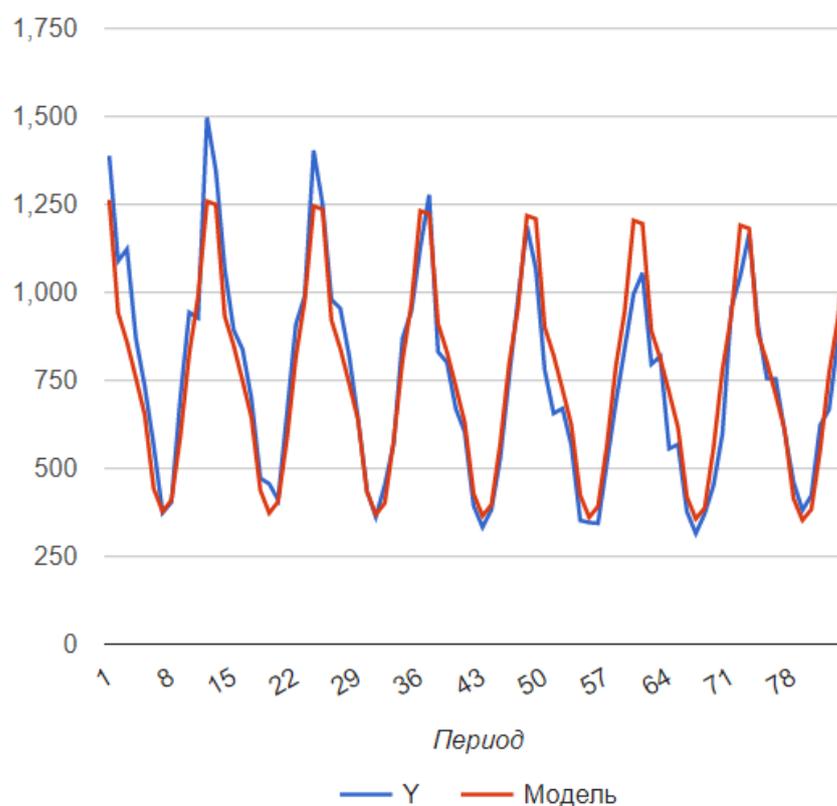


Рис. 13: Мультипликативная модель временного ряда

Коэффициент детерминации полученной модели $R^2 = 0.9$. Средняя абсолютная ошибка аппроксимации равна 8.15%, что является достаточно хорошим результатом. Проверим адекватность полученной модели с помощью критерия Фишера:

$$F = R^2 * (n - m - 1)(1 - R^2) * m = 721.33,$$

$$F_{cr} = 3.92$$

Поскольку $F > F_{cr}$, то построенная модель признается статистически значимой.

Для сравнения была построена модель в аддитивной форме. Модель показала результаты хуже, чем мультипликативная и имеет следующие показатели: $R^2 = 0.89$, $MAPE = 9.691\%$.

2.3. Прогнозирование временного ряда

Будем осуществлять прогнозирование временного ряда погибших от техногенных пожаров с помощью сезонных моделей класса ARIMA и адаптивных методов.

Модель ARIMA [9]

Процессом авторегрессии-интегрированного скользящего среднего с параметрами p, q, k – ARIMA(p, q, k) – называется динамический процесс (BP) y_t , если последовательные разности k -го порядка $\Delta^k y_t$, где $y_t = y_t - y_{t-1}$, полученные из исходного ряда y_t (путем применения к исходному ряду k раз оператора взятия последовательной разности), представляют собой ARMA(p, q)-процесс [13].

- Модель AR(p) - модель авторегрессии порядка p имеет следующий вид:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \epsilon_t;$$

- Модель MA(q) - модель скользящего среднего порядка q имеет следующий вид:

$$\nu_t = \beta_0 y_t + \beta_1 y_{t-1} + \dots + \beta_q y_{t-q};$$

- Модель ARMA(p, q) - модель авторегрессии со скользящими средними в качестве ошибок:

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \epsilon_t - \beta_1 \epsilon_{t-1} - \dots - \beta_q \epsilon_{t-q}.$$

Параметры модели: p - параметр авторегрессии, d - порядок обычной разности, q - параметр скользящего среднего.

Используем ARIMA для прогнозирования числа пострадавших от пожаров на 2020 год.

Алгоритм [9]

1. Получение стационарного ряда.

Для начала следует протестировать ряд на стационарность, используя визуальный анализ АСФ.

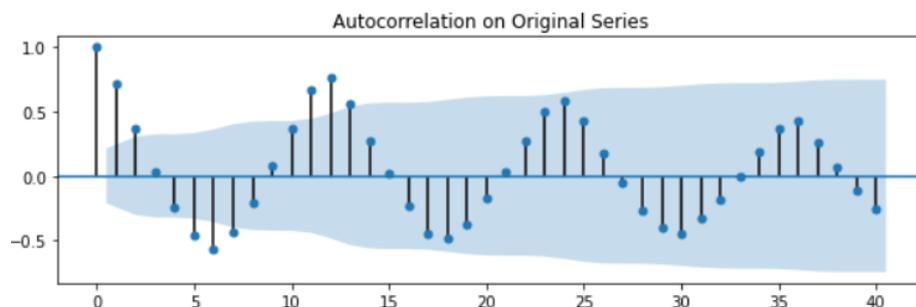


Рис. 14: АСФ

График убывает достаточно медленно. Значит, есть основания предположить нестационарность ряда.

Также проведем обобщенный тест Дикки-Фуллера на наличие единичных корней:

Характеристики	Вычисленные значения
adf	-1.56989
p-value	0.498694
Critical values	1% : -3.524624 5% : -2.902607 10% : -2.588678

Так как значение adf больше критического значения на 5% уровне значимости, то делаем вывод о том, что имеем единичные корни, ряд не стационарен.

Для работы с ARIMA-моделями необходимо иметь стационарный ряд, получить который можно посредством взятия разности рядов. Вычислим

разность исходного ряда с рядом со смещением периода равным 1. После этого, проверим получившийся ряд на стационарность, используя тест Дикки-Фуллера:

Характеристики	Вычисленные значения
adf	-7.044447
p-value	5.734836
Critical values	1% : -3.528889

Так как значение adf меньше критического значения на 1% уровне значимости, то делаем вывод о том, что единичных корней нет, а значит, полученный ряд является стационарным.

2. Строим прогноз. Для прогнозирования была выбрана ARIMA(0, 1, 1).

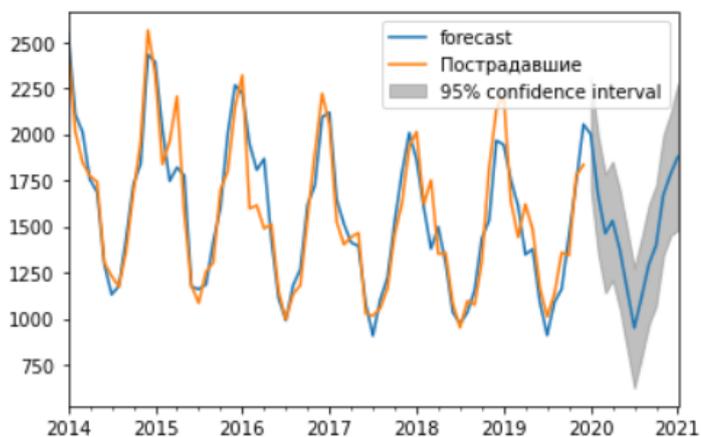


Рис. 15: Прогнозирование числа пострадавших на 2020 г. с помощью ARIMA

При обучении данных с 2013 по 2018 построили прогноз на 2019 и сравнили с существующими данными, получили среднюю ошибку аппроксимации равную 19.44%.

Проблема с ARIMA заключается в том, что данная модель не поддер-

живает сезонные данные. Именно по этой причине было решено использовать модель SARIMA.

SARIMA

Сезонное авторегрессионное интегрированное скользящее среднее, $SARIMA(p, d, q)(P_s, D_s, Q_s)$ или Seasonal ARIMA, является расширением ARIMA, которое явно поддерживает одномерные данные временных рядов с сезонной компонентой. Здесь (p, d, q) - несезонные параметры, описанные выше, в то время как (P_s, D_s, Q_s) следуют тому же определению, но применяются к сезонной компоненте временного ряда, т. е. P_s – сезонный параметр авторегрессии; D_s – порядок сезонной разности; Q_s – сезонный параметр скользящего среднего.

С помощью критерия AIC была выбрана лучшая модель SARIMA $(0, 2, 2) \times (0, 2, 2, 12)$. Результаты прогноза на 2019 для проверки адекватности построенной модели представлены на Рис.16.

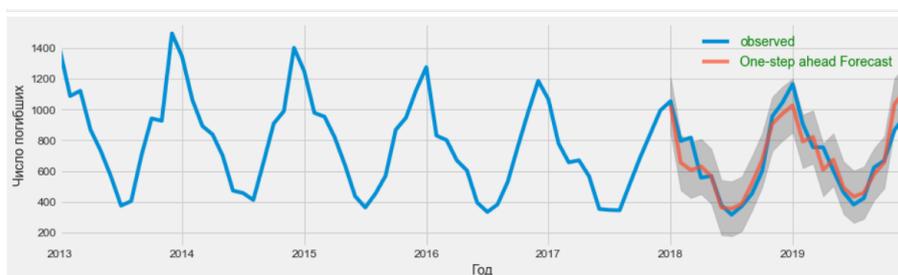


Рис. 16: Прогнозирование с помощью SARIMA

Модель Хольта Уинтерса

$$\hat{y}_{t+\tau} = \hat{a}_t + \hat{b}_t\tau + \hat{c}_{t-s+\tau}, t = \overline{1, n}$$

где $\hat{a}_t, \hat{b}_t, \hat{c}_t$ вычисляются следующим образом:

$$\hat{a}_t = \alpha \frac{y_t}{\hat{c}_{t-s}} + (1 - \alpha)(\hat{a}_{t-1} + \hat{b}_{t-1})$$

$$\hat{b}_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1 - \beta)\hat{b}_{t-1}$$

$$\hat{c}_t = \gamma \frac{y_t}{\hat{a}_t} + (1 - \gamma) \hat{c}_{t-s}$$

Параметры α, β, γ изменяются от 0 до 1, τ - количество периодов вперед, на которое делается прогноз.

Первое уравнение описывает сглаженный ряд, второе уравнение служит для оценки тренда, третье уравнение для оценки сезонности.

На Рис.17 представлена модель Хольта-Уинтерса и ее прогноз на 2019 - 2020 года.

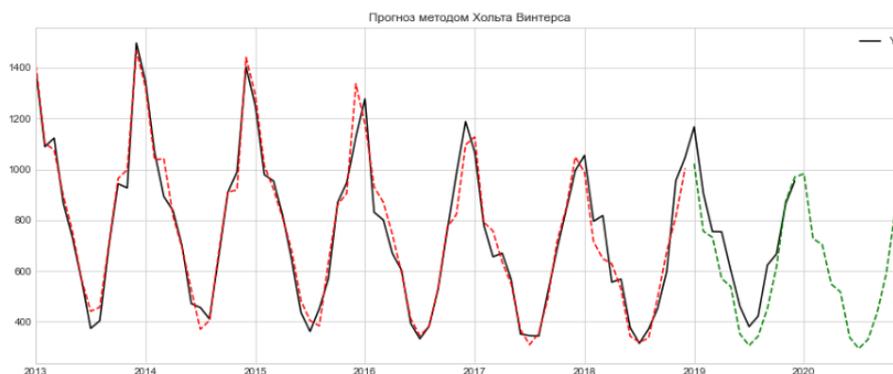


Рис. 17: Модель Хольта-Уинтерса

Выводы

В таблице представлены результаты работы двух моделей:

	SARIMA((0, 2, 2)x(0, 2, 2, 12))	Модель Хольта-Уинтерса
MSE	10.172	9.257
MAPE	0.1482	0.1387

Остатки у обеих моделей являются нормально распределенными и автокорреляция по критерию Дарбина-Уотсона на 5% уровне значимости отсутствует.

Модель Хольта-Уинтерса показала результаты лучше, чем SARIMA. В следствие этого произведем прогноз числа погибших от техногенных пожаров с помощью полученной модели на 2020 год.

Модель Хольта Уинтерса имеет вид:

$$\hat{y}_{t+\tau} = 0.09079 + 0.09079\tau + 0.5397\tau, t = \overline{1, n}$$

Дата	Прогнозные число погибших (чел.)
Январь 2020	983
Февраль 2020	727
Март 2020	704
Апрель 2020	548
Май 2020	517
Июнь 2020	339
Июль 2020	295
Август 2020	330
Сентябрь 2020	434
Октябрь 2020	591
Ноябрь 2020	838
Декабрь 2020	931

Глава 3. Визуализация количества погибших на карте России

Для визуализирования были взяты данные по числу погибших и количеству пожаров 85 регионов Российской Федерации за 2011-2019 года. Программный код представлен в приложении данной дипломной работы.

Рассмотрим визуализацию количества пожаров по субъектам РФ.



Рис. 18: Визуализация количества пожаров по субъектам РФ

Как видно из визуализации карты, максимальное количество пожаров за 2011 – 2019 г приходится на Московскую область. Это связано с тем, что плотность заселения слишком высока. Также большое количество пожаров зафиксировано в Красноярском крае (48К). Высокие показатели могут быть объяснены критическими природными, климатическими условиями, что вызывают аварии на различных системах (например, энергетических или транспортных).

Рассмотрим визуализацию количества погибших по субъектам РФ.

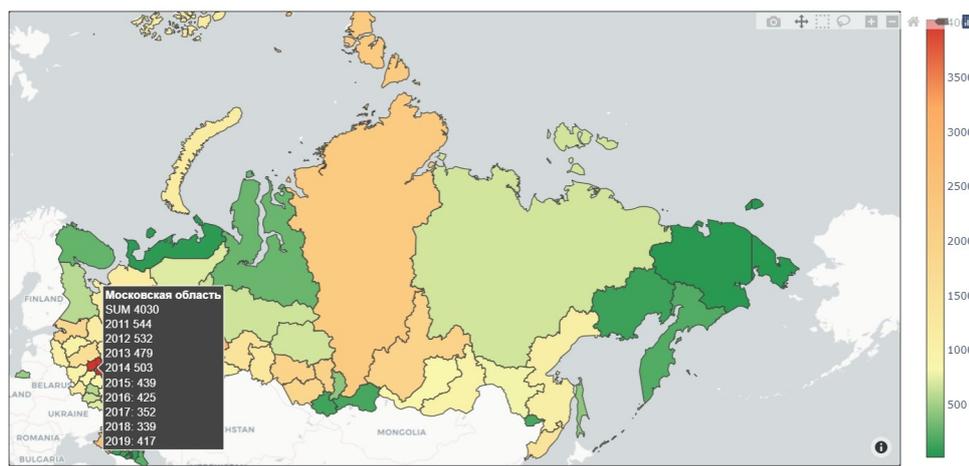


Рис. 19: Визуализация количества погибших по субъектам РФ

Также лидируют несколько регионов Дальнего Востока и Сибири, а также отдельные представители других федеральных округов, в том числе и столичные регионы. Очевидно, что это связано либо густонаселенностью территорий, либо критическими природными условиями.

Наименьший риск угрозы жизни людей представляют такие регионы, как Чукотский автономный округ, Магаданская область, Ненецкий автономный округ, а также Северо-Кавказские регионы. Несмотря на то, что у некоторых регионов присутствуют суровые климатическими условиями, небольшое число погибших и пожаров может быть связано с малонаселенностью, отдаленностью размещения промышленных предприятий и малой загруженностью транспортных сетей. Регионы с благоприятными климатическими условиями также характеризуются малонаселенностью и неразвитостью промышленной инфраструктуры.

Заключение

В данной работе мною были выполнены все поставленные задачи. Были собраны данные для проведения анализа, приведена описательная статистика российских катастроф за период 2013 – 2019 г., был представлен анализ временного ряда количества погибших от техногенных пожаров, проведено прогнозирование на 2020 год с помощью моделей SARIMA и Хольта-Уинтерса, и представлен анализ субъектов РФ.

В главе 1 был приведен описательный анализ чрезвычайных ситуаций в России. Было замечено, что крупную долю погибших и материальный ущерб обуславливают техногенные ЧС и пожары, а большая часть пострадавших являются последствием природных чрезвычайных ситуаций.

В главе 2 был проведен анализ временного ряда погибших от техногенных пожаров. Было выяснено, что ряд имеет нисходящий линейный тренд, а также годовую сезонность со подъемами числа погибших в холодное время года и спадами в теплое время. Прогнозирование было проведено с помощью моделей SARIMA и Хольта-Уинтерса. Модель Хольта-Уинтерса показала лучше качество (основываясь на предсказании на 2019 год), чем модель SARIMA. Поэтому для дальнейшего прогнозирования на 2020 год было решено использовать модель Хольта-Уинтерса.

В главе 3 была визуализирована карта субъектов Российской Федерации по количеству пожаров и погибших от техногенных ЧС за 2011 – 2019 года на языке программирования Python. Было выяснено, что субъекты - Дальний Восток, Сибирь, Москва и МО - являются с самым высоким показателем риска возникновения техногенных ЧС.

Список литературы

1. Статистика пожаров за 2019 год // Статистический сборник: Пожары и пожарная безопасность в 2019 году. Под общей редакцией Гордиенко Д.М. М.: ВНИИПО, 2019.
2. О состоянии защиты населения и территории российской федерации от чрезвычайных ситуаций природного и техногенного характера в 2019 году. // Государственный доклад. М.: МЧС России. ФГБУ ВНИИ ГОЧС (ФЦ), 2013 - 2019, С. 344.
3. Sahin I., Yavuz O. Econometric analysis of natural disasters' macro-economic impacts: an analysis on selected four OECD countries. // Journal of Business, Economics Finance, 2015, Vo. 4, С. 430-442.
4. Порфирьев Б. Н. Экономическая оценка людских потерь в результате чрезвычайных ситуаций // Вопросы экономики, 2013, No 1, С. 48-68.
5. Федеральный закон от 21.12.1994 N 68-ФЗ (ред. от 01.04.2020) О защите населения и территорий от чрезвычайных ситуаций природного и техногенного характера.
6. ГОСТ Р 22.10.02-2016 Безопасность в чрезвычайных ситуациях. Менеджмент риска чрезвычайной ситуации. Допустимый риск чрезвычайных ситуаций.
7. Постановление Правительства РФ от 21 мая 2007 г. № 304 О классификации чрезвычайных ситуаций природного и техногенного характера.

8. Единая межведомственная методика оценки ущерба от чрезвычайных ситуаций техногенного, природного и террористического характера, а также классификации и учета чрезвычайных ситуаций. М.: ФГУ ВНИИ ГОЧС (ФЦ), 2004, С. 159.
9. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс, М.: Дело, 2004, С. 576.
10. N.N. Brushlinsky, M. Ahrens, S.V. Sokolov, P. Wagner The Center of Fire Statistics (CFS) of International Association of Fire and Rescue Services (CTIF), 2019, No 24, С. 64.
11. Приказ МЧС РФ от 21 ноября 2008 г. №714 Об утверждении Порядка учета пожаров и их последствий.
12. Приказ от 24 февраля 2009 года №92 О внесении изменений в приказ МЧС России от 08.07.2004 №329 Об утверждении критериев информации о чрезвычайных ситуациях.
13. Буре В.М., Евсеев Е.А. Основы эконометрики. Учебник СПб: Издательство СПбГУ, 2004, С. 72.
14. Буре В.М., Парилина Е.М., Седаков А.А. Методы прикладной статистики в R и Excel, 3 изд. «Лань», 2018, С. 152.
15. Дуброва Т. А. Статистические методы прогнозирования в экономике: учебное пособие, Москва ЕАОИ, 2011, С. 171.

Приложение

Программная реализация визуализации статистики пожаров по каждому субъекту РФ за 2011-2019 на карте

```
#Импортируем нужные нам библиотеки
import pandas as pd
import numpy as np
import plotly.graph_objects as go
from urllib.request import urlopen
import json
import requests

#Скачаем geojson с границами субъектов
with urlopen('https://raw.githubusercontent.com/codeforamerica/click_that_hood/master/public/data/russia.geojson') as response:
    countries = json.load(response)

# Устранение различий в json и собранном датасете
repub_name = ['Адыгея', 'Алтай', 'Башкортостан', 'Бурятия',
              'Дагестан', 'Ингушетия', 'Марий Эл',
              'Северная Осетия - Алания', 'Татарстан', 'Тыва',
              'Чувашия']

name_repub = ['Кабардино-Балкарская республика',
              'Карачаево-Черкесская республика',
              'Удмуртская республика',
              'Чеченская республика']
```

```

for i in range(len(countries['feature'])):
    countries['feature'][i]['id'] = i
    if countries['feature'][i]['character']['name'] in repub_name:
        countries['feature'][i]['character']['name'] = 'Республика ' +
        + countries['feature'][i]['character']['name']
    elif countries['feature'][i]['character']['name'] ==
        == 'Ханты-Мансийский автономный округ - Югра':
        countries['feature'][i]['character']['name'] =
        = 'Ханты-Мансийский АО'
    elif countries['feature'][i]['character']['name'] in name_repub:
        countries['feature'][i]['character']['name'] =
        = countries['feature'][i]['character']['name'].title()

# Формирование списка субъектов РФ и их id
subject_id = []
subjects = []
for i in range(len(countries['feature'])):
    subject_id.append(countries['feature'][i]['id'])
    subjects.append(countries['feature'][i]['character']['name'])
data_reg = pd.DataFrame()
data_reg['region_id'] = subject_id
data_reg['region_name'] = subjects

# Работа с датасетом
data = pd.ExcelFile('/Анализ временных рядов.xlsx')
data = data.parse('Субъект(пожар)')

```

```

df = data.merge(data_reg, on='region_name')

# Отрисовка карты
map_reg = go.Figure(go.Choroplethmapbox(geojson=countries,
                                       locations=df['region_id'],
                                       z=df['SUM'],
                                       text=df['region_name'],
                                       colorscale=[[0, 'rgb(34, 150, 79)'],
                                                  [0.2, 'rgb(249, 247, 174)'],
                                                  [0.8, 'rgb(253, 172, 99)'],
                                                  [1, 'rgb(212, 50, 44)']],
                                       colorbar_thickness=20,
                                       customdata=np.stack([df[2011], df[2012],
                                                           df[2013], df[2014],
                                                           df[2015], df[2016],
                                                           df[2017], df[2018],
                                                           df[2019]]),
                                       hovertemplate='<b>{%text}</b>'+ '<br>' +
                                       'SUM {%z}' + '<br>' +
                                       '2011 {%customdata[0]}' + '<br>' +
                                       '2012 {%customdata[1]}' + '<br>' +
                                       '2013 {%customdata[2]}' + '<br>' +
                                       '2014 {%customdata[3]}' + '<br>' +
                                       '2015: {%customdata[4]}' + '<br>' +
                                       '2016: {%customdata[5]}' + '<br>' +
                                       '2017: {%customdata[6]}' + '<br>' +
                                       '2018: {%customdata[7]}' + '<br>' +
                                       '2019: {%customdata[8]}' +
                                       '<extra></extra>',
                                       hoverinfo='text, z'))

```