

АО «Национальный холдинг «QazBioPharm»

РГП на ПХВ «Национальный центр общественного здравоохранения»
Министерства Здравоохранения Республики Казахстан

Абдураимов Е.О., Алдынгуров Д.К., Смагул М.А., Ахметова К.М.

**СОЗДАНИЕ СИСТЕМЫ ЭПИДЕМИОЛОГИЧЕСКОГО
ПРОГНОЗИРОВАНИЯ (МОДЕЛИРОВАНИЯ) И РЕАГИРОВАНИЯ НА
ОСОБО ОПАСНЫЕ ИНФЕКЦИИ**

(методические рекомендации)

**Астана
2024**

УДК: 616.013.1:614.45

С40

ББК: 51

Рецензенты:

1. **Сихаева Н.С.** – PhD, Директор департамента науки, Руководитель представительства АО «Национальный холдинг «QazBioPharm» в г.Алматы

2. **Токмурзиева Г.Ж.** – д.м.н., Заместитель генерального директора по науке и лабораторно-диагностической службе РГП на ПХВ «Национальный научный центр особо опасных инфекций имени Масгута Айкимбаева» МЗ РК

Авторы:

Абдураимов Е.О. – д.в.н., Генеральный директор АО «Национальный холдинг «QazBioPharm»

Алдынгулов Д.К. – PhD, Вр. и.о. заместителя генерального директора по биобезопасности АО «Национальный холдинг «QazBioPharm»

Смагул М.А. – магистр здравоохранения, Заместитель Председателя Правления РГП на ПХВ «Национальный центр общественного здравоохранения» МЗ РК

Ахметова К.М. – PhD, Директор департамента биологической безопасности АО «Национальный холдинг «QazBioPharm»

Разработка системы эпидемиологического прогнозирования (моделирования) и реагирования на особо опасные инфекции: Методические рекомендации / Абдураимов Е.О., Алдынгулов Д.К., Смагул М.А., Ахметова К.М.//Астана: АО «Национальный холдинг «QazBioPharm», 2024. – 47 с.

ISBN 978-601-305-625-8

Настоящие методические рекомендации описывают создание системы эпидемиологического прогнозирования и реагирования на особо опасные инфекции, используя современные методы математического моделирования, позволяющие анализировать и предсказывать динамику эпидемий, где подробно описаны этапы разработки системы, включая сбор данных, анализ, прогнозирования (моделирования) и реагирования. Разработка и внедрение эффективной системы прогнозирования и реагирования позволит оперативно выявлять потенциальные угрозы и принимать меры по их предотвращению или минимизации последствий.

Данные методические рекомендации будут полезны специалистам в области эпидемиологии, медицины и общественного здравоохранения, сотрудникам медицинских организаций и уполномоченных органов, отвечающим за обеспечение санитарно-эпидемиологического благополучия населения, а также для реализации мер по предотвращению распространения особо опасных инфекций и обеспечению биологической безопасности.

УДК: 616.013.1:614.45

С40

ББК: 51

Утверждено и разрешено к изданию типографским способом РГП «Республиканский центр развития здравоохранения» (протокол заседания Департамента развития медицинской науки и образования РГП РЦРЗ» № 491 от «30» сентября 2024 года)

© Абдураимов Е.О., Алдынгулов Д.К., Смагул М.А., Ахметова К.М., 2024

Содержание

Перечень сокращений и обозначений	4
Введение	5
1. Создание системы прогнозирования особо опасных инфекций	6
1.1.Методы прогнозирования временных рядов: модели ARIMA и SARIMA	6
1.2.Оценка факторов риска особо опасных инфекций	22
2. Создание системы реагирования особо опасных инфекций	26
Заключение	28
Список использованных источников	30
Приложение А - Результаты линейного регрессионного анализа сибирской язвы	33
Приложение Б - Матрица коэффициентов корреляции Спирмена между переменными и статистика коллинеарности при сибирской язве	34
Приложение В - Результаты отрицательной биномиальной регрессии сибирской язвы с основными переменными.	36
Приложение Г - Результаты отрицательной биномиальной регрессии сибирской язвы с переменными задержек.	37
Приложение Д - Результаты регрессионного анализа ККГЛ.	38
Приложение Е - Сводная статистика погодных параметров в Жамбылской, Кызылординской и Туркестанской областях.	40
Приложение Ж - Матрица коэффициентов корреляции Спирмена между переменными и статистика коллинеарности при ККГЛ.	41
Приложение З - Результаты отрицательной биномиальной регрессии ККГЛ с основными переменными.	43
Приложение И - Результаты отрицательной биномиальной регрессии ККГЛ с переменными задержек.	44
Приложение К - Матрица коэффициентов корреляции Спирмена при бруцеллезе.	45
Приложение Л - Результаты отрицательной биномиальной регрессии бруцеллеза	47

Перечень сокращений и обозначений

В настоящем отчете применяют следующие сокращения и обозначения:

ББ	Биологическая безопасность
ВОЗ	Всемирная организация здравоохранения
ГИС	Государственная информационная система
ИНС	Искусственные нейронные сети
ККГЛ	Конго-Крымская геморрагическая лихорадка
КРС	Крупный рогатый скот
МЗ РК	Министерство здравоохранения Республики Казахстан
НПЦСЭЭиМ	Научно-практический центр санитарно-эпидемиологической экспертизы и мониторинга
ООИ	Особо опасные инфекции
РБП	Роз-Бенгал проба
РЦЭЗ	Республиканский центр электронного здравоохранения
РХ	Реакция агглютинации пластинчатая Хеддельсона
ACF	Функция автокорреляции (Autocorrelation function)
ARIMA	Авторегрессионная интегрированная скользящая средняя (Auto Regressive Integrated Moving Average)
AIC	Информационный критерий Акаике (Akaike's information criterion)
AUC	Площади под кривой (Area Under the Curve)
BIC	Байесовский информационный критерий (Bayesian information criterion)
ETS	Ошибка-Тренд-Сезонность модель (Errors – Trend-Seasonality)
MAE	Средняя абсолютная ошибка (Mean Absolute Error)
MSE	Среднеквадратическая ошибка (Mean Squared Error)
MAPE	Средняя абсолютная процентная ошибка (mean percentage absolute error)
PACF	Функция частичной автокорреляции (Partial Autocorrelation function)
RMSE	Среднеквадратическая ошибка (Root Mean Square Error)
SARIMA	Сезонная авторегрессионная скользящая средняя (Seasonal Auto Regressive Integrated Moving Average)
SMA	Сезонная скользящая средняя (Seasonal Moving Average)
VAR	Метод векторной авторегрессии (Vector AutoRegression)

Введение

В условиях современного мира, когда инфекционные заболевания представляют собой серьёзную угрозу для здоровья населения и национальной безопасности, разработка и внедрение эффективной системы эпидемиологического прогнозирования и реагирования на особо опасные инфекции становится приоритетной задачей для Казахстана.

Основная цель создания системы эпидемиологического прогнозирования (моделирования) и реагирования на особо опасные инфекции заключается в обеспечении своевременного и эффективного реагирования на их вспышки, минимизации рисков распространения заболеваний и улучшении качества медицинской помощи.

В Казахстане природные очаги сибирской язвы, Конго-Крымской геморрагической лихорадки и бруцеллеза занимают значительные площади, где ежегодно регистрируются случаи заражения этими инфекциями. Заболевания отличаются высокой степенью контагиозности и способностью вызывать тяжелые формы. Особо опасные инфекции требуют внимательного подхода к прогнозированию и реагированию на вспышки. В Казахстане за 2022-2023 годы зарегистрировано множество случаев данных заболеваний (ККГЛ – 41 и 40 случаев соответственно, сибирская язва – 20 и 37 случаев, и бруцеллёз – 681 и 693 случая). Эти данные свидетельствуют о том, что ситуация остаётся сложной и актуальной социально-экономической проблемой государства. В связи с этим были проведены расчёты прогнозирования по данным инфекциям.

Создание системы эпидемиологического прогнозирования (моделирования) и реагирования на особо опасные инфекции позволит прогнозировать распространение особо опасных инфекций, оперативно реагировать на вспышки заболеваний, оптимизировать ресурсы здравоохранения, что позволит разработать эффективные стратегии профилактики и контроля распространения инфекций, повысить уровень информированности населения о рисках и мерах предосторожности. Поэтому, создание эффективной системы эпидемиологического прогнозирования и реагирования является важной задачей в системе здравоохранения, которая требует комплексного подхода и сотрудничества между различными организациями и специалистами.

Разработанные методические рекомендации помогут внедрить современные методы прогнозирования (моделирования) распространения особо опасных инфекционных заболеваний, создадут условия для обеспечения безопасности граждан и минимизировать экономический ущерб. Применение системы прогнозирования и реагирования на практике позволит заранее предвидеть вероятные вспышки особо опасных инфекционных заболеваний и своевременно на них реагировать.

1. Создание системы прогнозирования распространения особо опасных инфекций

Создание системы прогнозирования распространения инфекционных заболеваний является важнейшей задачей для обеспечения общественного здоровья и безопасности. Предупредительные меры помогают не только минимизировать последствия эпидемий, но и эффективно использовать ресурсы здравоохранения, тем самым снижая нагрузку на медицинскую инфраструктуру и улучшая качество жизни населения.

Для создания системы прогнозирования и реагирования инфекционных заболеваний, в частности, сибирской язвы, ККГЛ и бруцеллёза, нужно применить комплексный подход и учитывать несколько основных аспектов.

1) *Сбор данных*: для осуществления расчетов по прогнозированию и проведению мониторинга на первом этапе необходимо осуществить сбор достоверных данных. В Республике Казахстан сбор данных по инфекциям проводится филиалом НПЦСЭЭиМ НЦОЗ МЗ РК.

2) *Аналитические методы и моделирование*: определение распространения заболеваний необходимо проводить на основе математических моделей и аналитических методов с использованием данных, полученных от НПЦСЭЭиМ.

3) *Сотрудничество и информационный обмен*: эффективная работа системы прогнозирования подразумевает взаимодействие между разными уровнями системы здравоохранения и эпидемиологической службой. В случае возникновения угрозы распространения заболеваний эти организации должны обмениваться информацией и координировать свои действия. Для налаживания взаимной работы, необходимо интегрировать процесс сбора данных с государственной информационной системой «Биологическая безопасность».

Таким образом, для повышения точности прогнозирования необходимо учитывать и опираться на анализ факторов риска особо опасных инфекций с целью выявления возможных тенденций развития эпидемической ситуации.

1.1. Методы прогнозирования временных рядов: модели ARIMA и SARIMA

В современном мире особо опасные инфекции, характеризующиеся высокой смертностью, устойчивостью к традиционным методам лечения и способностью быстро распространяться, представляют серьёзную угрозу для здоровья и жизни человека. В связи с этим возникает необходимость эффективного контроля и прогнозирования распространения особо опасных инфекций.

Методы прогнозирования временных рядов могут помочь в этой задаче, обеспечивая более точные и своевременные прогнозы. Это позволит принять необходимые меры по предотвращению распространения заболеваний и снижению возможных последствий.

Методы ARIMA и SARIMA являются одними из наиболее популярных инструментов анализа временных рядов. В прогнозировании распространении ООИ эти методы могут быть адаптированы для учёта специфических характеристик этих заболеваний, таких как сезонность, циклические колебания и другие факторы.

Модель ARIMA – это сочетание моделей авторегрессии (AR) и скользящего среднего (MA). В AR прогноз проводится на основе линейной комбинации предыдущих значений переменной, а в MA – на основе линейной комбинации ошибок предыдущих прогнозов. Буква «I» в названии модели означает взятие разности между текущим и предыдущим значением данных. Это позволяет стабилизировать данные и уменьшить автокорреляцию.

Модель SARIMA означает сезонную авторегрессионную интегрированную скользящую среднюю. Является расширением ARIMA за счёт добавления сезонной составляющей. Она учитывает линейную комбинацию прошлых сезонных значений и ошибок прогноза.

Для осуществления статистических расчетов и моделирования используется программное обеспечение RStudio. RStudio – это интегрированная среда разработки для R, языка программирования для статистических вычислений и графики [1].

Чтобы применить модели ARIMA и SARIMA в целях осуществления прогноза необходимо использовать следующие пакеты (библиотеки):

1. Пакет `forecast` – содержит функции для прогнозирования временных рядов и линейных моделей. Позволяет отображать и анализировать одномерные прогнозы временных рядов, включая экспоненциальное сглаживание с помощью моделей в пространстве состояний и автоматическое моделирование ARIMA.

2. Пакет `tseries` – предназначен для анализа временных рядов и вычислительных финансов.

3. Пакет `lmtest` – включает тесты, наборы данных и примеры для диагностической проверки в моделях линейной регрессии. Предоставляет инструменты для вывода в параметрических моделях.

Данные пакеты (библиотеки) можно установить непосредственно в программном обеспечении через вкладку «Packages».

Прогнозирование временных рядов ООИ проводят в программе RStudio, где сначала необходимо загрузить и преобразовать данные.

Для успешного моделирования временных рядов с использованием моделей ARIMA и SARIMA необходимо убедиться, что данные являются стационарными.

Стационарные данные во временных рядах – это набор данных, который не изменяет свои статистические свойства со временем. Он является ключевым условием для эффективной работы моделей ARIMA и SARIMA. Стационарность включает в себя два основных компонента:

- 1) Стационарность в среднем: среднее значение временного ряда остаётся постоянным. Если есть тенденция или систематические закономерности, то до моделирования следует преобразовать данные, чтобы дифференцировать их

(вычесть предыдущее значение из текущего). Это поможет стабилизировать среднее значение.

2) Стационарность в дисперсии: изменчивость временного ряда также остаётся постоянной. Чтобы стабилизировать дисперсию, можно использовать логарифмические преобразования.

Для определения соответствующих параметров модели часто проводят исследование автокорреляционных и частичных автокорреляционных функций разностного ряда. Процесс подбора подходящей модели ARIMA или SARIMA обычно итеративный, требующий многочисленных проб и ошибок. Окончательная модель должна пройти диагностические проверки на соответствие.

Перечисленные методы можно использовать при прогнозировании заболеваний, в частности изученными нами сибирской язвы, ККГЛ, бруцеллеза.

В Республике Казахстан в период 2003 – 2022 годы всего было зарегистрировано 165 случаев сибирской язвы. Отмечается, что 88% всех зарегистрированных случаев зафиксированы в июне, июле, августе и сентябре. Данный факт позволяет использовать модель SARIMA для прогнозирования заболевания в эти месяцы.

Отображение временного ряда. На графике (рисунок 1) показаны функция автокорреляции (ACF) и функция частичной автокорреляции (PACF). Функция ACF измеряет и отображает среднюю корреляцию между точками данных во временном ряду и предыдущими значениями ряда, измеренными для различной длины задержки. Этот график подтверждает наличие сезонности в данных. PACF аналогичен ACF, но каждая частичная корреляция контролирует любую корреляцию между наблюдениями с более короткой задержкой. На верхнем графике показаны случаи заболеваемости в динамике по годам.

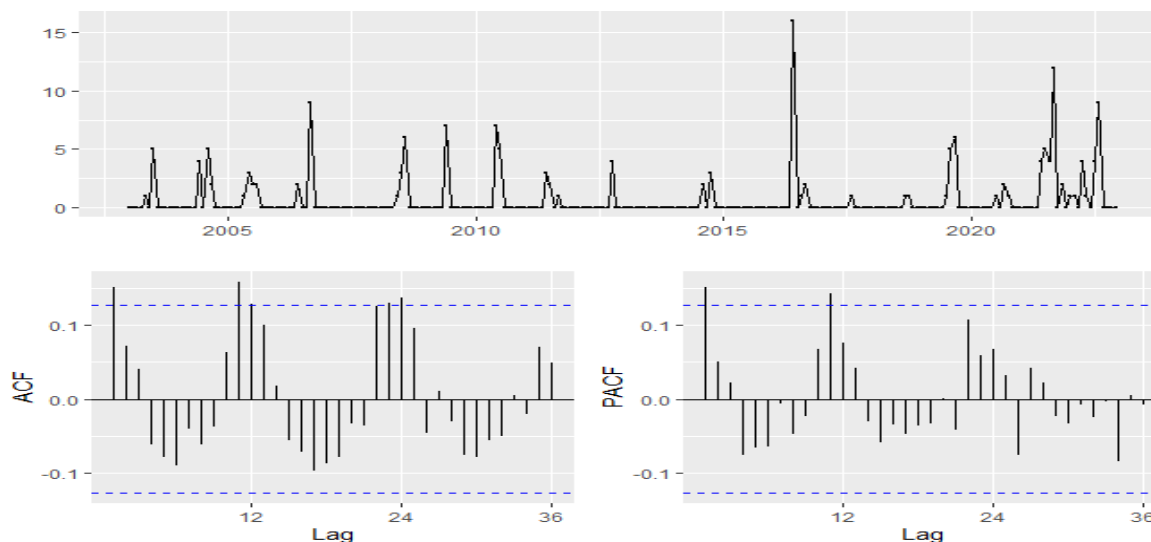


Рисунок 1 – Результаты временного ряда с использованием функции автокорреляции (ACF) и частичной автокорреляции (PACF)

Определение коэффициентов для модели и подбор лучшей модели ARIMA к одномерному временному ряду проводится с помощью функции `auto.arima`.

Функция `auto.arima` позволяет определить лучшую модель ARIMA на основе значений AIC, AICс или BIC, а также подбирает возможную модель в пределах предоставленных ограничений порядка.

С помощью данной функции были подобраны следующие коэффициенты (рисунок 2): ARIMA (0,0,0) (0,0,2).

```
Series: y1
ARIMA(0,0,0)(0,0,2)[12] with non-zero mean

Coefficients:
          sma1      sma2      mean
          0.1183    0.1973    0.7129
s.e.       0.0651    0.0779    0.1602

sigma^2 = 3.737:  log likelihood = -497.77
AIC=1003.53    AICc=1003.7    BIC=1017.46
```

Рисунок 2 – Результаты лучшей модели ARIMA к одномерному временному ряду

Для анализа временного ряда `y1` была использована модель ARIMA (0, 0, 0) (0, 0, 2) [2], полученная с помощью функции `auto.arima`.

Спецификация модели:

- ARIMA (0, 0, 0) – это несезонная часть модели. Здесь нет компонента авторегрессии (AR) (0), разностей (0) и скользящего среднего (MA) (0).

- (0, 0, 2) – сезонная часть модели со скользящим средним (SMA) порядков 1 и 2 и сезонным периодом 12.

- «mean» указывает на ненулевое среднее значение в модели.

Для модели были рассчитаны следующие коэффициенты:

- `sma1` (сезонное скользящее среднее при задержке 1) = 0,1183;

- `sma2` (сезонное скользящее среднее при задержке 2) = 0,1973;

- среднее (ненулевое среднее) = 0,7129.

Здесь `s.e.` – стандартные ошибки оценок коэффициентов.

$\text{Sigma}^2 = 3,737$ – это расчётная дисперсия белого шума или ошибки в модели. Значение логарифмического правдоподобия составляет – 497,77. Большие отрицательные значения указывают на лучшее соответствие.

В качестве информационных критериев были использованы:

- AIC (информационный критерий Акаике) = 1003,53;

- AICс (исправленный информационный критерий Акаике) = 1003,7;

- BIC (байесовский информационный критерий) = 1017,46.

Эти критерии используются для оценки компромисса между сложностью

модели и степенью соответствия. Более низкие значения указывают на более подходящую модель.

Полученные результаты могут быть полезны при прогнозировании и анализе временных рядов. Они дают представление о том, как модель учитывает сезонные колебания и как влияет на её структуру.

Проверка на шум и на остатки.

Все точки находятся внутри круга, таким образом, выбранные коэффициенты можно использовать для дальнейшей работы (рисунок 3). Это позволяет убедиться в том, что данные были обработаны корректно.

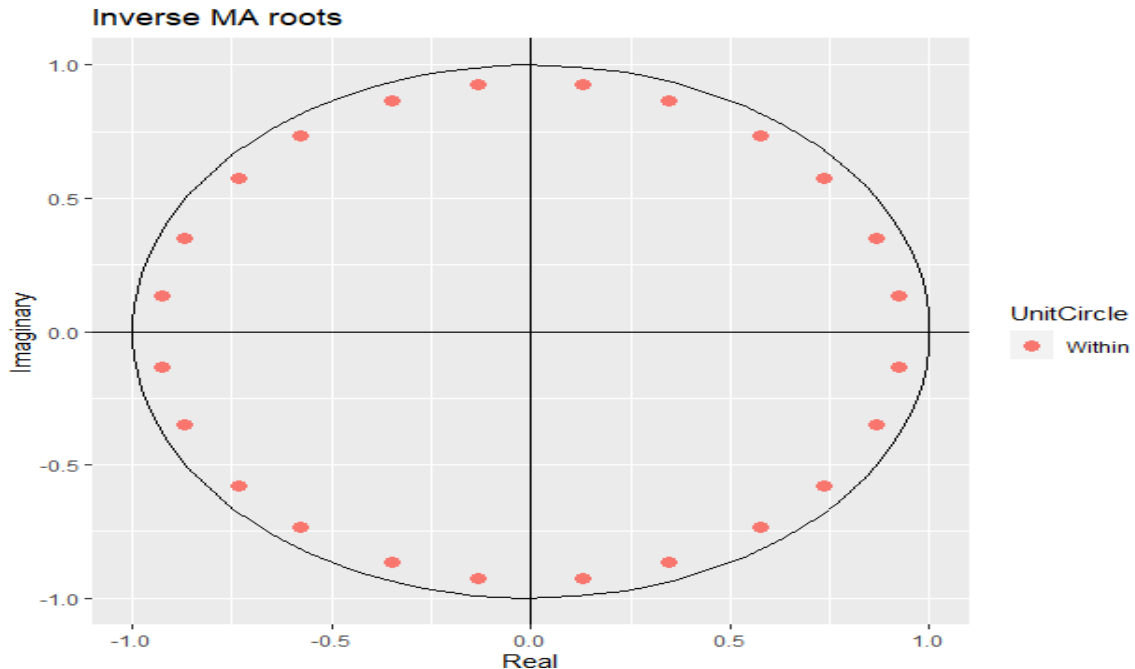


Рисунок 3 – Результаты проверки на шум и на остатки ошибок

Далее используется функция *coefrest*. Эта функция используется для выполнения z- и (квази-) t-тестов Вальда оцениваемых коэффициентов. Она также позволяет вычислить соответствующие доверительные интервалы Вальда (рисунок 4). В результате проверки мы получаем p-значения для каждого коэффициента. Если значение p меньше 0,05, это говорит о статистической значимости данных.

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)
sma1	0.118339	0.065149	1.8164	0.06930 .
sma2	0.197275	0.077865	2.5336	0.01129 *
intercept	0.712920	0.160234	4.4492	8.617e-06 ***

 signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Рисунок 4 – Результаты теста Вальда

При помощи *checkresiduals* проводится проверка, чтобы убедиться, что остатки модели временных рядов выглядят как белый шум. В случае, если тест не является ложным, распечатываются результаты теста Лjungа или теста Бреуша-Годфри (рисунок 5).

```
Ljung-Box test

data:  Residuals from ARIMA(0,0,0)(0,0,2)[12] with non-zero mean
Q* = 24.509, df = 22, p-value = 0.3211

model df: 2.    total lags used: 24
```

Рисунок 5 – Результаты теста Лjungа-Бокса

В ходе анализа были получены следующие результаты для переменных *sma1*, *sma2* и точки пересечения.

Для переменной *sma1* расчётный коэффициент равен 0,118339, значение *z* равно 1,8164. Соответствующее значение *p* составляет 0,06930 – это немного выше общего уровня значимости 0,05, но меньше 0,1. Это означает, что коэффициент *sma1* не имеет статистической значимости, хотя потенциально может влиять на переменную ответа.

Значение переменной *sma2*: расчётный коэффициент – 0,197275, значение *z* – 2,5336. Значение *p* при этом – 0,01129 (**). Значит, переменная *sma2* имеет статистическую значимость на уровне 0,05. Она оказывает заметное влияние на переменную отклика модели.

Точка пересечения имеет расчётный коэффициент 0,712920, значение *z* – 4,4492. Также у неё очень низкое значение *p* – 8,617e-06 (***), что указывает на статистическую значимость. Эта точка вносит значительный вклад в модель.

Результаты говорят о том, что переменные *sma2* и точка пересечения имеют статистическую значимость и оказывают заметное влияние на переменную ответа в модели временных рядов. Переменная *sma1* незначима, но её воздействие менее определённо и требует дальнейшей оценки.

С помощью *checkresiduals(fit2)* определяем, есть ли серийная корреляция в остатках модели ARIMA. Результат: значение *p* = 0,3211. Это относительно высокое значение. Оно говорит об отсутствии убедительных доказательств серийной корреляции в остатках, что является хорошим знаком. Он указывает на то, что модель успешно уловила основные закономерности в данных, а остатки ведут себя как белый шум.

Таким образом, результаты теста Лjungа-Бокса показывают, что модель ARIMA разумно подходит для данных, поскольку нет никаких существенных доказательств серийной корреляции в остатках.

На следующем этапе было проведено моделирование ККГЛ. ККГЛ является

эндемичной инфекцией, с 2003 по 2022 годы в Казахстане зарегистрировано 285 случаев. Заболеваемость ККГЛ в Республике Казахстан наиболее высока в период с мая по август месяцы. На эти четыре месяца приходится 88% всех случаев заболевания ККГЛ в стране. Учитывая данные особенности, можно использовать модель SARIMA для моделирования (прогнозирования) случаев ККГЛ.

Процесс моделирования и прогнозирования с использованием модели SARIMA для ККГЛ аналогичен процессу моделирования, описанному ранее. Разница заключается только в исходных данных и числовых значениях, в то время как пошаговый алгоритм (принцип работы) остаётся неизменным и проводится также в программе R Studio с загрузкой и преобразованием данных.

Для определения стационарности временного ряда использовали статистический тест, основанный на вычислении p -value. Полученное значение p -value оказалось меньше критического уровня значимости 0,05, что позволяет отвергнуть нулевую гипотезу о нестационарности данных и предположить их стационарность.

Отображение временного ряда. На графике (рисунок 6) представлены функции автокорреляции (ACF) и частичной автокорреляции (PACF). Функция автокорреляции ACF показывает среднюю корреляцию между точками данных во временном ряду и предыдущими значениями ряда, измеренными для различной длины задержки. График ACF также подтверждает наличие сезонности в данных. Функция частичной автокорреляции, PACF, аналогична ACF, но измеряет корреляцию после учёта влияния всех других лагов. То есть каждая частичная корреляция контролирует любую корреляцию между наблюдениями с более короткой задержкой.

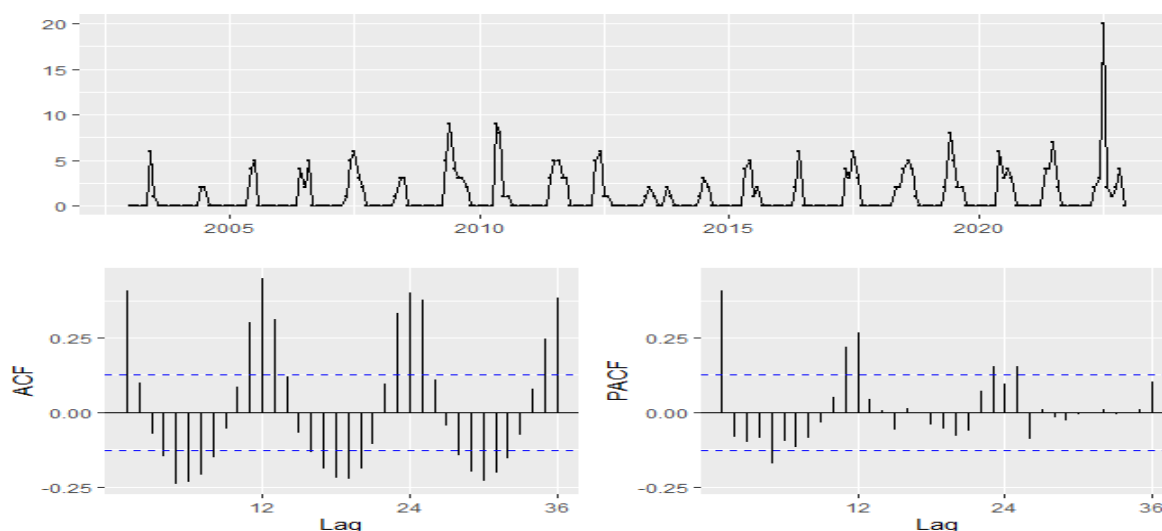


Рисунок 6 – Результаты временного ряда с использованием функции автокорреляции (ACF) и частичной автокорреляции (PACF)

Определение коэффициентов для модели и подбор лучшей модели ARIMA к одномерному временному ряду используется при помощи функции *auto.arima*. Функция *auto.arima* в автоматическом режиме выбирает лучшую модель ARIMA на основе одного из информационных критериев (AIC, AICс или BIC) в заданных пользователем пределах порядка модели.

С помощью данной функции подобраны следующие коэффициенты (рисунок 7): ARIMA (0,0,1) (2,0,0)

```
> auto.arima(y2)
Series: y2
ARIMA(0,0,1)(2,0,0)[12] with non-zero mean

Coefficients:
      ma1      sar1      sar2      mean
      0.1707   0.4406   0.2675   1.2575
s.e.      0.0688   0.0718   0.0782   0.4057

sigma^2 = 3.328:  log likelihood = -486.41
AIC=982.82      AICc=983.08      BIC=1000.23
```

Рисунок 7 – Результаты лучшей модели ARIMA к одномерному временному ряду

Для анализа временного ряда «у2» с помощью функции *auto.arima*, была подобрана модель ARIMA (0, 0, 1)(2, 0, 0)[2]. Это означает, что она включает несезонную часть ARIMA (0,0,1) без компонента авторегрессии (AR), без дифференциала (I) и компонента скользящего среднего (MA) порядка 1, и сезонную часть (2,0,0) с сезонным периодом 12.

Рассчитаны следующие коэффициенты:

- *ma1* (срок скользящего среднего при задержке 1): 0,1707;
- *sar1* (сезонный период авторегрессии при задержке 1): 0,4406;
- *sar2* (сезонная авторегрессия при задержке 2): 0,2675;
- среднее (ненулевое среднее): 1,2575.

Вместе с тем, рассчитаны стандартное отклонение и стандартные ошибки оценок коэффициентов. Информационные критерии модели указывают на то, насколько хорошо она соответствует данным:

- AIC (информационный критерий Акаике): 982,82;
- AICс (исправленный информационный критерий Акаике): 983,08;
- BIC (байесовский информационный критерий): 1000,23.

Обычно предпочтительны более низкие значения этих критериев, так как они указывают на более подходящую модель.

Таким образом, подобранная модель предназначена для фиксации временных закономерностей и сезонности в данных временного ряда «у2». Она может использоваться для прогнозирования и анализа временных рядов, так как учитывает сезонные и несезонные компоненты временного ряда.

Проверка на шум и на остатки

Так как все точки находятся внутри круга, то выбранные коэффициенты могут быть использованы далее (рисунок 8).

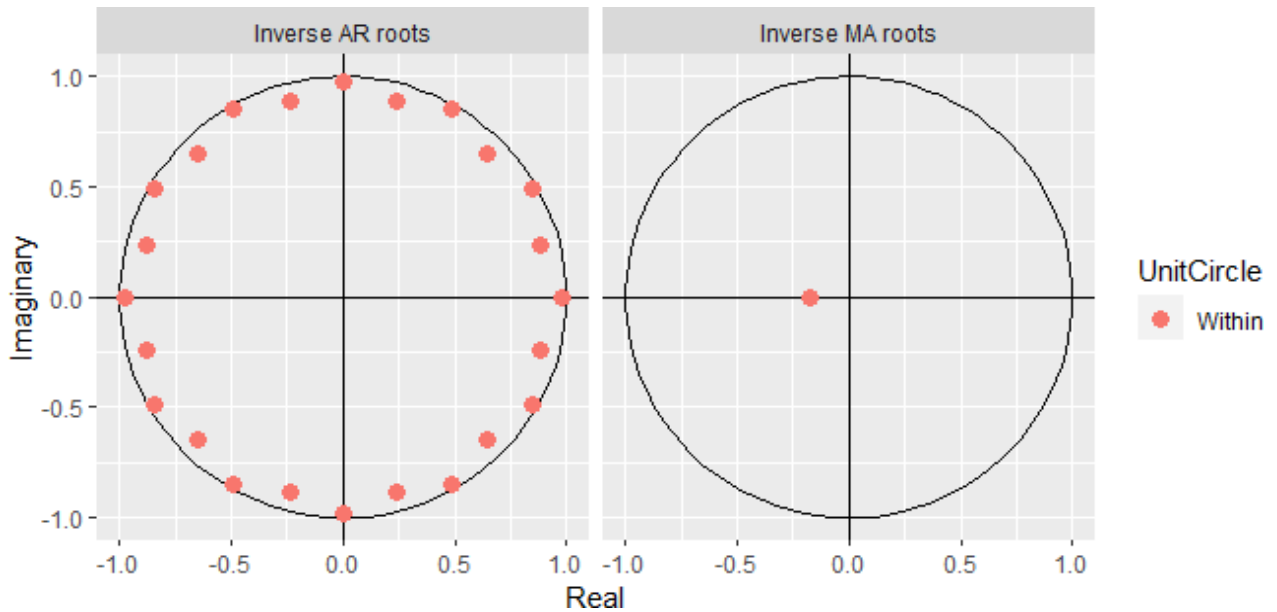


Рисунок 8 – Результаты проверки на шум и на остатки ошибок

С помощью функции `coefstest` можно провести z- и (квази-) t-тесты Вальда для оцениваемых коэффициентов. Функция `coefsci` помогает рассчитать соответствующие доверительные интервалы Вальда. Проведя расчёты, мы выяснили, что полученные значения p для сезонных средних скользящих меньше 0,05. Это позволяет сделать вывод о статистической значимости этих данных.

Проверка, что остатки модели временных рядов выглядят как белый шум.

При `plot=TRUE` формируется график, отражающий зависимость остатков от временного лага и их частотное распределение. Если тест не является ложным, отображаются результаты теста Лjung-Бокса (рисунок 9) или теста Бреуша-Годфри. Интерпретация результатов:

- Для «ma1» расчётный коэффициент равен 0,170722, значение z равно 2,4815. Значение p составляет 0,0130849, что меньше 0,05 («*»). Это указывает на статистическую значимость коэффициента «ma1», то есть его заметное влияние на переменную ответа.

- Для «sar1» расчётный коэффициент – 0,440642, значение z – 6,1382, а значение p – 8,344e-10 («***»). Эти значения указывают на высокую статистическую значимость «sar1», который оказывает сильное воздействие на переменную ответа.

- Коэффициент «sar2» также имеет значительное влияние: расчётный коэффициент – 0,267515, значение z – 3,4194, значение p – 0,0006277 («***»).


```

> checkresiduals(fit2)

Ljung-Box test

data:  Residuals from ARIMA(0,0,1)(2,0,0)[12] with non-zero mean
Q* = 14.601, df = 21, p-value = 0.8424

Model df: 3.    Total lags used: 24

```

Рисунок 9 – Результаты теста Лjung-Бокса

Термин пересечения имеет расчетный коэффициент 1,257496, значение z 3,0993 и значение p 0,0019401 («*»), что указывает на то, что он статистически значим и вносит вклад в модель.

Таким образом, результаты показывают, что все коэффициенты («ma1», «sar1», «sar2» и точка пересечения) статистически значимы, то есть они оказывают значимое влияние на переменную ответа в модели временных рядов. Эта информация имеет решающее значение для понимания взаимосвязей между переменными-предикторами и переменной отклика в модели.

P -значение = 0,8424: значение p , связанное с тестовой статистикой, используется для определения того, являются ли остатки серийно коррелированными. Высокое значение p (близкое к 1) предполагает отсутствие признаков серийной корреляции в остатках. В данном случае значение p составляет 0,8424, что довольно высоко.

В результате теста Лjung-Бокса, высокое значение p (0,8424) предполагает отсутствие убедительных доказательств серийной корреляции в остатках модели ARIMA. Это хороший знак, поскольку он указывает на то, что модель успешно уловила основные закономерности в данных, а остатки ведут себя как белый шум, что является желательным свойством для хорошо подобранной модели.

Таким образом, результаты теста Лjung-Бокса показывают, что модель ARIMA разумно подходит для данных, поскольку нет никаких существенных доказательств серийной корреляции в остатках.

Вместе с тем, была смоделирована модель прогнозирования бруцеллеза. Заболевания бруцеллёзом в Казахстане за 2003-2022 годы составили 31842 зарегистрированных случая. Для прогнозирования использовались данные за период с 2011 по 2022 годы (13 192 зарегистрированных случая), поскольку они доступны с разбивкой по месяцам. Так, с мая по август регистрируется больше всего случаев заболевания бруцеллёзом – 7075, что составляет 54% от общего числа случаев. Также много случаев фиксируется в апреле и в период с сентября по октябрь. Всего с апреля по октябрь зарегистрировано 10380 случаев, что соответствует 79% всех случаев заболевания бруцеллезом в Республике Казахстан.

Процесс моделирования и прогнозирования с использованием модели ARIMA для бруцеллеза аналогичен процессу моделирования, описанному ранее. Разница

заключается только в исходных данных и числовых значениях, в то время как пошаговый алгоритм (принцип работы) остаётся неизменным и проводится также в программе R Studio с загрузкой и преобразованием данных.

Отображение временного ряда на рисунке 10 представлен график ACF, который измеряет и показывает среднюю корреляцию между точками данных во временном ряду и предыдущими значениями ряда, измеренными для различной длины задержки. Этот график также позволяет увидеть наличие сезонности в данных. Также на верхнем графике показаны случаи заболеваемости в динамике по годам. PACF похож на ACF, но здесь каждая частичная корреляция контролирует любую корреляцию между наблюдениями с более короткой задержкой.

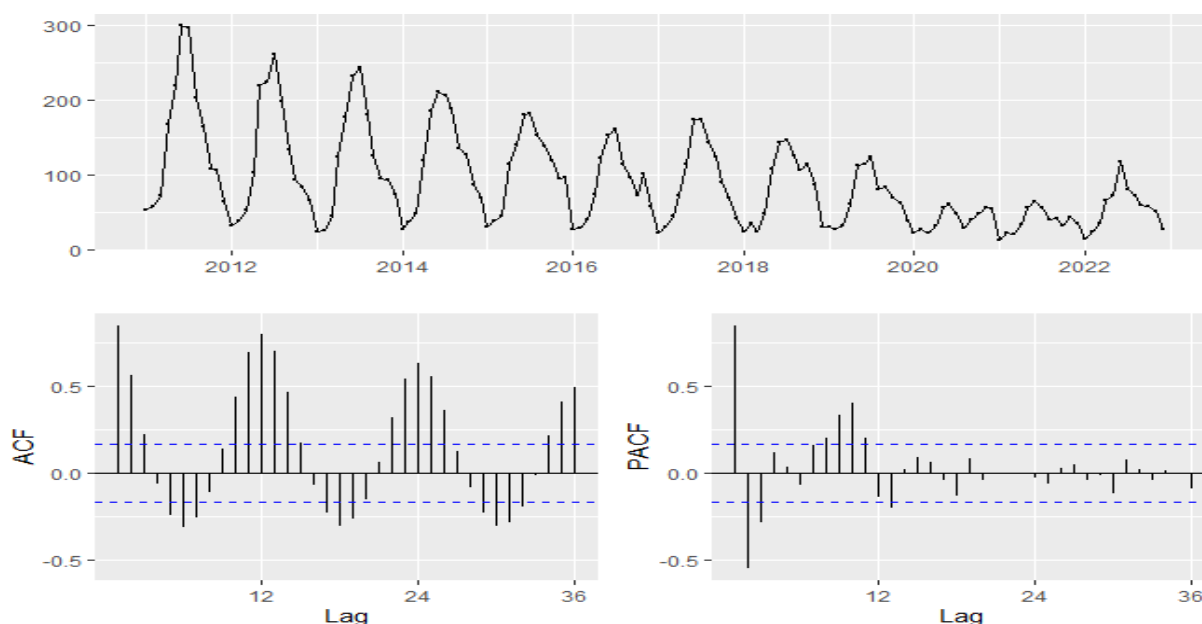


Рисунок 10 – Результаты временного ряда с использованием функции автокорреляции (ACF) и частичной автокорреляции (PACF)

На графике (рисунок 10) наблюдается тенденция к уменьшению количества случаев, поэтому необходимо использовать специальную команду для стабилизации данных и выравнивания их вокруг значения 0. В R для вычисления различий между элементами числового вектора или временного ряда используется функция `diff()`. Она определяет разницу между последовательными элементами, что помогает преобразовать данные и сделать их стационарными или проанализировать изменения во временном ряду (рисунок 11).


```
> bruc.diff=diff(y3)
> bruc.diff
```

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2011		6	14	95	52	80	-4	-93	-39	-55	-2	-42
2012	-32	6	15	49	116	5	37	-63	-64	-41	-9	-17
2013	-43	2	19	79	53	56	10	-62	-55	-30	-2	-20
2014	-46	9	11	71	66	27	-5	-19	-52	-9	-40	-18
2015	-38	8	6	70	25	40	3	-30	-14	-20	-24	2
2016	-70	3	10	34	49	31	7	-47	-17	-24	29	-44
2017	-36	9	14	27	43	60	-1	-30	-19	-35	-21	-27
2018	-17	11	-12	25	59	35	4	-21	-20	9	-27	-57
2019	0	-3	5	28	52	2	9	-43	3	-14	-7	-24
2020	-16	5	-5	10	23	5	-13	-18	11	8	7	-1
2021	-42	10	-2	13	22	8	-8	-16	2	-9	10	-7
2022	-22	11	10	32	6	45	-38	-7	-13	-2	-6	-24

Рисунок 11 – Результаты теста для определения разницы между последовательными элементами, чтобы сделать их стационарными

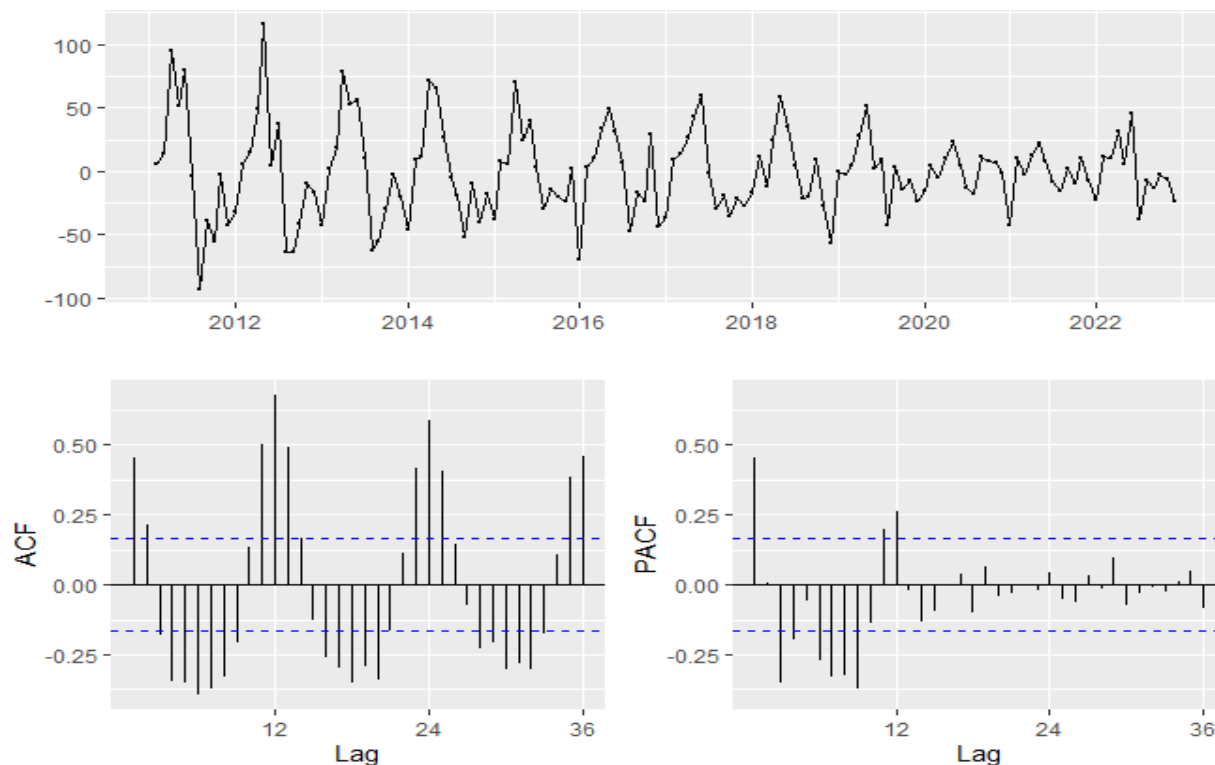


Рисунок 12 – Результаты временного ряда с использованием функции автокорреляции (ACF) и частичной автокорреляции (PACF) с данными после определения разницы между последовательными элементами

На рисунке 12 можно наблюдать, что теперь данные стационарные и для проверки осуществляется тест Дики-Фуллера (рисунок 13).

```
> adf.test(bruc.diff)

Augmented Dickey-Fuller Test

data: bruc.diff
Dickey-Fuller = -8.2687, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

Рисунок 13 – Результаты расширенного теста Дикки-Фуллера

Поскольку p -значение меньше 0,05, можно сделать вывод, что нулевая гипотеза о нестационарности данных отвергается и данные признаются стационарными.

Подбор лучшей модели ARIMA к одномерному временному ряду возможно применить при помощи функций *auto.arima*. Функция возвращает лучшую модель ARIMA на основе критериев AIC, AICс или BIC. Функция осуществляет поиск модели в заданных пределах порядка. Для быстрой оценки моделей для большого количества временных рядов можно использовать значения по умолчанию. Подобраны следующие коэффициенты (рисунок 14): ARIMA (2,0,2) (0,1,1)

```
> auto.arima(y3)
Series: y3
ARIMA(2,0,2)(0,1,1)[12] with drift

Coefficients:
          ar1          ar2          ma1          ma2          sma1          drift
          1.5997      -0.8241      -1.1197      0.523      -0.5551      -0.7674
s.e.        0.1013        0.1008        0.1515        0.117        0.0997        0.1118

sigma^2 = 297.6: log likelihood = -562.55
AIC=1139.1   AICc=1140   BIC=1159.28
```

Рисунок 14 – Результаты лучшей модели ARIMA к одномерному временному ряду

Но так как было осуществлено преобразование данных с помощью функции *diff* на 1, то для модели будет добавлена единица к коэффициенту d .

Предоставленные данные являются результатом подбора модели ARIMA к ряду данных временного ряда «у3»:

- Модель ARIMA (2, 1, 2) (0, 2, 1) [2] со смещением имеет следующую структуру:

- ARIMA (2,1,2): представляет собой несезонную часть модели. Это модель ARIMA с:

- Авторегрессионный компонент второго порядка (ar1 и ar2);
- Дифференциация порядка 1;
- Компонент скользящей средней второго порядка (ma1 и ma2). (0,2,1)
- Сезонная часть модели (0, 2, 1) [2]:
- сезонный период равен 12;

- компонент сезонного скользящего среднего порядка 1 (sma1).
- «drift»: это указывает на наличие дрейфа в модели.
- Значения коэффициентов для модели следующие:
 - ar1: 1,5997 – член авторегрессии при задержке 1;
 - ar2: -0,8241 – член авторегрессии при задержке 2;
 - ma1: -1,1197 – срок скользящего среднего при задержке 1;
 - ma2: 0,5234 – срок скользящего среднего при задержке 2;
 - sma1: -0,5551 – сезонное скользящее среднее с задержкой 12.
- Значение дрейфа: -0,7674. *s.e* – стандартные ошибки этих оценок коэффициентов.
- σ^2 – расчётная дисперсия белого шума или ошибки в модели, равная 297,6.
- Логарифмическое правдоподобие: значение логарифмического правдоподобия составляет -562,55. Это показатель того, насколько хорошо модель соответствует данным. Как правило, более подходящим считается большее отрицательное значение этого показателя.
- Информационные критерии:
 - AIC (информационный критерий Акаике): 1139,1;
 - AICc (исправленный информационный критерий Акаике): 1140;
 - BIC (байесовский информационный критерий): 1159,28.

Информационные критерии используются для оценки качества модели и помогают найти баланс между сложностью модели и степенью её соответствия данным. Обычно предпочтительны более низкие значения этих критериев, так как это указывает на то, что модель лучше соответствует данным. Модель ARIMA предназначена для фиксации временных закономерностей, сезонности и дрейфа в данных временного ряда «уз». Оценочные коэффициенты позволяют определить вклад каждого компонента в модель.

Проверка на шум и на остатки. Так как все точки находятся внутри круга, то выбранные коэффициенты могут быть использованы далее (рисунок 15).

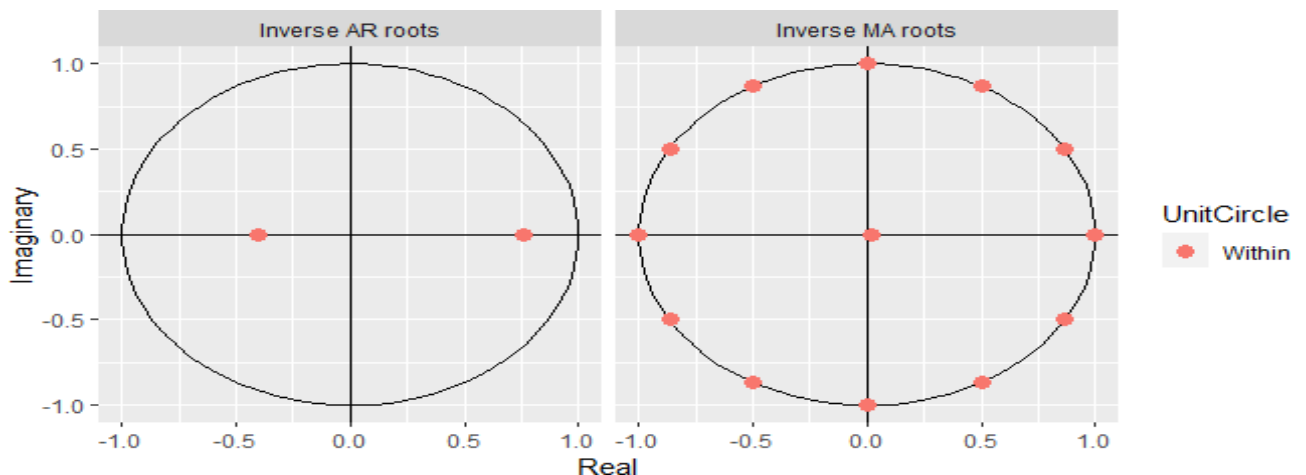


Рисунок 15 – Результаты проверки на шум и на остатки ошибок

Функция *coeftest* – это общая функция для выполнения z- и (квази-)t-тестов Вальда оцениваемых коэффициентов. *coefci* вычисляет соответствующие доверительные интервалы Вальда (рисунок 16).

```
> coeftest(fit3)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1    0.357143   0.214804   1.6626   0.09638 .
ar2    0.309104   0.129691   2.3834   0.01715 *
ma1   -1.019356   0.226270  -4.5050  6.636e-06 ***
ma2    0.021207   0.215898   0.0982   0.92175
sma1  -0.994998   0.099571  -9.9928 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Рисунок 16 – Результаты теста Вальда

Согласно данной команде полученные сезонные средние скользящие имеет значение p меньше, чем 0.05, это свидетельствует о том, что данные значения статистически значимые.

После используется команда *checkresiduals*. Проверка, что остатки модели временных рядов выглядят как белый шум. Описание: если *plot=TRUE*, создается временной график остатков, соответствующий ACF, и гистограмма. Если тест не ЛОЖНЫЙ, распечатываются результаты теста Лjung-Бокса (рисунок 17) или теста Бреуша-Годфри.

```
> checkresiduals(fit3)

Ljung-Box test

data:  Residuals from ARIMA(2,1,2)(0,2,1)[12]
Q* = 24.766, df = 19, p-value = 0.1684

model df: 5.    Total lags used: 24
```

Рисунок 17 – Результаты теста Лjung-Бокса

Интерпретация результатов

В результате проведенного анализа были получены следующие данные:

- Для «ar1» расчётный коэффициент равен 0,357143. Значение z – 1,6626, значение p – 0,09638. Эти значения говорят о том, что «ar1» не является статистически значимым на уровне значимости 0,05.

- Для «ar2» расчётный коэффициент – 0,309104. Значение z – 2,3834, значение p – 0,01715 (помечается символом '*'). Это указывает на то, что «ar2» статистически значим на уровне значимости 0,05.

- Для «ma1» значение расчётного коэффициента составляет -1,019356. Его значение z равно -4,5050. Значение p очень низкое и составляет 6,636e-06 («***»). Эти значения указывают на высокую статистическую значимость «ma1».

- Для «ma2» значение расчётного коэффициента равно 0,021207, значение z – 0,0982, значение p – 0,92175. Они свидетельствуют о том, что «ma2» не является статистически значимым.

- Для «sma1» расчётный коэффициент составил -0,994998. Его значение z равнялось -9,9928, а значение p было меньше 2,2e-16 («***»). Это говорит о высокой статистической значимости «sma1».

Таким образом, в результате исследования было выявлено, что переменные «ar2», «ma1», и «sma1» являются статистически значимыми, в то время как «ar1» и «ma2» таковыми не являются. Значения коэффициентов отражают их важность в модели ARIMA.

Проанализировав полученное значение p , можно сделать вывод, что уровень корреляции остатков в данной модели достаточно низкий, он равен 0,1684. Это означает, что в остатках модели отсутствует серийная корреляция. Такое заключение является положительным результатом, поскольку модель успешно уловила основные закономерности в данных, а остатки ведут себя как белый шум, что является желательным свойством для хорошо подобранной модели.

Таким образом, результаты теста Лjung-Бокса показывают, что модель ARIMA разумно подходит для данных, поскольку нет никаких существенных доказательств серийной корреляции в остатках.

В данной работе представлен пошаговый алгоритм создания системы прогнозирования на особо опасных инфекции, а именно сибирская язва, ККГЛ и бруцеллез. В алгоритме указаны все необходимые шаги, включая статистическое программное обеспечение, программные библиотеки (пакеты) и команды (коды), данные с указанием источника, результаты осуществленного прогнозирования.

В качестве моделей были применены ARIMA и SARIMA. Эти модели позволили провести анализ временных рядов данных по каждому заболеванию, выявить закономерности в изменении заболеваемости и оценить будущие тенденции распространения сибирской язвы, бруцеллёза и ККГЛ.

Модели показали свою эффективность и могут быть использованы для дальнейшего развития системы прогнозирования инфекционных заболеваний, мониторинга текущей ситуации и оперативного реагирования на новые случаи заболевания. Это позволит предотвратить распространение инфекции и обеспечить более эффективное управление ресурсами здравоохранения. Кроме того, для прогнозирования инфекционных заболеваний, в том числе и ООИ, необходимо оценить факторы риска.

1.2. Оценка факторов риска особо опасных инфекций

Особо опасные инфекции (сибирская язва, ККГЛ, бруцеллез) характеризуются определенными факторами риска, которые могут играть значимую роль в распространении инфекций и воздействии на здоровье населения.

Оценка факторов риска позволяет заблаговременно выявлять потенциальные угрозы и принимать меры для предотвращения распространения заболевания. Поэтому при исследовании факторов риска особо опасных инфекций можно использовать корреляцию для анализа связи между инфекциями и различными факторами.

Для изучения факторов риска корреляция помогает оценить связь между двумя или более факторами. Изучение факторов риска также может включать в себя регрессионный анализ. Он позволяет количественно оценить влияние различных факторов на вероятность возникновения инфекции. Это помогает понять, какие факторы наиболее значимы в формировании риска развития инфекции, и разработать соответствующие меры профилактики и контроля.

В качестве ключевых факторов риска были рассмотрены следующие показатели: число заболевших, сведения о годовой заболеваемости животных сибирской язвой и бруцеллезом, а также сезонной активности клещей, являющихся переносчиками ККГЛ. Кроме того, немаловажным фактором являются климатические параметры, к которым относятся средняя температура, атмосферное давление, уровень осадков, относительная влажность и другие факторы, влияющие на распространение возбудителя инфекции. На основе изучения мирового опыта определены различные факторы риска, влияющие на заболеваемость сибирской язвы, бруцеллеза и ККГЛ.

При сибирской язве основными факторами риска заболеваемости являются: убой инфицированных сибирской язвой животных, употребление в пищу зараженного мяса и принадлежность к мужскому полу [3]. Кроме того, на распространение сибирской язвы влияет изменение климатических условий [4, 5]. Чтобы определить взаимосвязь между количеством поголовья больного скота и случаями заболевания сибирской язвой, была проведена корреляция Пирсона. Результаты корреляции Пирсона позволяют оценить силу и направление линейной связи между двумя непрерывными переменными, т.е. между количеством поголовья больного скота и случаями заболевания сибирской язвы.

Следующим этапом работы является проведение линейного регрессионного анализа, так как имеются две непрерывные переменные: количество случаев заболевания сибирской язвой среди людей – это зависимая переменная, и количество больного скота – независимая переменная. Результаты проведенного линейного регрессионного анализа позволяют сделать вывод о существовании статистически значимой положительной связи между количеством заболевших животных и случаями сибирской язвы среди людей. Результаты представлены в приложении А.

Климатические факторы проанализированы методом корреляции Спирмена, чтобы определить их взаимосвязь со случаями заболевания (приложение Б).

Результаты отрицательной биномиальной регрессии (приложение В) показывают влияние температуры на заболеваемость сибирской язвой.

Анализ сезонных задержек (приложение Г) указывает на увеличение количества случаев сибирской язвы если температура за прошедший месяц повышается и уменьшается если повышается влажность.

Анализ климатических факторов позволяет выявить, что средняя температура и её колебания влияют на рост заболеваний сибирской язвой. Это подтверждается положительной статистически значимой корреляцией, а также результатами биномиальной регрессии. Другие климатические показатели, такие как давление, количество осадков и влажность не являются статистически значимыми предикторами распространения сибирской язвы.

Таким образом, анализ сезонных колебаний температуры и влажности показал их влияние на частоту случаев сибирской язвы за предыдущий месяц. Отмечена сильная положительная линейная взаимосвязь между количеством заболевших людей и животных, которая указывает на важность контроля заболеваемости скота для предотвращения распространения инфекции.

Далее была рассмотрена ККГЛ, где высокому риску заражения подвержены работники сельскохозяйственных профессий, скотобоен и ветеринары. Изучая факторы риска, было выявлено что важным звеном в распространении инфекции являются их источники, по этой причине изучена связь между климатическими факторами и активностью клещей.

С этой целью применяется регрессионная модель. В программе Stata, используя уровень значимости $p < 0,05$ проверяется каждая зависимая переменная (активность клещей, климатические факторы) на наличие статистической связи с количеством случаев ККГЛ.

Проводим регрессионный анализ по определению влияния активности клещей на распространение и возникновение случаев ККГЛ. В результате выявлена связь между видовым составом клещей и количеством подтверждённых случаев ККГЛ. Результаты представлены в приложении Д. При анализе была изучена сезонная активность клещей, собранных из эндемичных регионов Казахстана по ККГЛ (Жамбылская, Кызылординская, Туркестанская области и г. Шымкент).

Затем методом линейной регрессии количественно проведена оценка корреляционной связи между количеством клещей, собранными в природных очагах и количеством случаев ККГЛ. В результате проведенного анализа выявлена прямая связь между количеством клещей вида *H. asiaticum* и числом случаев ККГЛ. Статистически значимая зависимость показывает, что увеличение количества клещей *H. asiaticum* на одну единицу приводит к росту числа случаев ККГЛ примерно на 0,0049. При этом 21,8% изменений в количестве случаев ККГЛ зависит от численности клещей вида *H. asiaticum*. Увеличение количества случаев заболевания ККГЛ связан с ростом численности других видов клещей, но в

меньшей степени, в среднем на 0,003 случая. Эта связь статистически значима ($p < 0,05$), а доля изменений, зависящая от количества клещей других видов, составляет 21,47%.

Кроме того, для эндемичных регионов по ККГЛ проведен анализ факторов риска, где были учтены климатические параметры. Результаты представлены в приложении Е.

Для анализа взаимосвязи климатических факторов со случаями заболеваемости ККГЛ использован метод корреляции Спирмена (приложение Ж).

Далее проводим анализ биномиальной регрессии. Результаты анализа можно найти в приложении З, по которым можно сделать вывод о том, что повышение температуры приводит к увеличению суточной частоты случаев ККГЛ.

После проводим анализ биномиальной регрессии (с переменными задержек), где при увеличении температуры за прошлый период ожидается повышение ожидаемого уровня случаев ККГЛ. Полный результат анализа описан в приложении И.

Таким образом, исследование зависимости заболеваемости ККГЛ от активности клещей и климатических факторов выявило чёткую связь между количеством особей вида *H. asiaticum* и случаями ККГЛ. Также обнаружена зависимость между частотой заболеваемости ККГЛ и климатическими условиями, которые включают температуру, атмосферное давление и количество осадков. Биномиальная регрессия подтвердила статистическую значимость влияния температуры и давления на показатель заболеваемости. Анализ данных показал, что эти взаимосвязи сохраняются в течение нескольких месяцев. Исследование доказывает важную роль климатических условий и активности клещей в распространении ККГЛ.

Также проведен анализ случаев заболевания **бруцеллёзом**, при котором были выявлены факторы риска, повышающие вероятность заражения среди работников сельского хозяйства (фермеры, доярки), которые непосредственно контактируют с животными [6].

На первом этапе проведена работа по определению взаимосвязи между случаями заболевания бруцеллёзом животных и климатическими параметрами с использованием метода корреляции Спирмена (приложение К). Отмечается, что некоторые климатические параметры влияют на случаи выявления заболевания бруцеллёзом животных.

Далее проводим анализ биномиальной регрессии, где повышение среднесуточной температуры на 1°C, среднего давления воздуха на 1 гПа и увеличение количества осадков на 1 мм связано с увеличением суточной частоты случаев бруцеллёза, а повышение среднесуточной влажности на 1% незначительно уменьшает частоту случаев бруцеллёза. Результаты представлены в приложении Л.

Таким образом, в результате проведенного корреляционного анализа и биномиальной регрессии было выявлено влияние климатических факторов на случаи заболеваний бруцеллёзом. Определена взаимосвязь между случаями

заболевания бруцеллезом среди животных и людей. Такие факторы как, высокая температура, наибольшее количество солнечных дней (часов) и низкая скорость ветра способствуют распространению бруцеллеза среди животных.

В проведенных исследованиях [7, 8], где изучали зависимость заболеваемости бруцеллезом животных от климатических факторов, выявлено, что средняя температура воздуха, относительная влажность, низкая скорость ветра, продолжительность солнечных дней и атмосферное давление оказывают прямое на возникновение инфекции.

В исследовании [9], было определено, что, основными факторами, позволяющие снизить риск заболевания бруцеллеза, являются низкая температура и минимальное количество солнечных дней (часов). При этом высокая температура, максимальное количество солнечных дней (часов) и низкая скорость ветра, напротив, оказывали существенное влияние на возникновение данной инфекции и высокую вероятность развития заболевания.

Учитывая тот факт, что заражение человека бруцеллезом происходит при контакте с заболевшим животным, при употреблении инфицированного мяса, молока и молочных продуктов, отмечается высокая вероятность заболевания людей. Поэтому, выявлена прямая зависимость заболеваемости бруцеллезом людей от заболеваемости животных данной инфекцией. Следовательно, увеличение степени заболеваемости бруцеллезом животных способствует росту числа заболевших среди людей.

Таким образом, с помощью методов корреляций Пирсона и Спирмена, а также отрицательной биномиальной регрессии проанализирована связь между климатическими параметрами и уровнем заболеваемости указанными инфекциями. При исследовании учтены временные задержки, что позволило определить отсроченное влияние климатических условий на распространение инфекций. При этом применение этих методов не позволяет предсказать в полной мере сезонность и цикличность распространения инфекционных заболеваний, что может снизить качество оценки прогноза (моделирования). Для более точного прогнозирования могут применяться другие методы. Поэтому для более точного прогнозирования актуальными является применение вышеописанных методик ARIMA и SARIMA.

2. Создание системы реагирования особо опасных инфекций

Система реагирования представляет собой комплекс мероприятий, направленных на предотвращение и контроль возможных инфекционных заболеваний. При регистрации вспышек особо опасных инфекций, таких как сибирская язва, ККГЛ и бруцеллёз, необходимо незамедлительно принимать скоординированные меры реагирования. Ключевыми мерами реагирования являются:

Межведомственное взаимодействие: здравоохранение, ветеринарные службы и другие органы должны скоординировать свои усилия. Важно разработать общие комплексные планы действий, включающие мероприятия со стороны системы здравоохранения (санитарно-эпидемиологическая служба, лечебные учреждения) и ветеринарной службы, МИО и других ответственных органов. Эти планы должны действовать как на национальном уровне, так и на областном и районном уровнях. Также они должны быть разработаны внутри ведомств. Необходимы единые алгоритмы и методики исследований при заболеваниях.

Быстрая диагностика и изоляция: важно обеспечить оперативное выявление случаев заражения и изоляцию больных для предотвращения распространения болезни среди контактных лиц.

Эпидемиологическое расследование: проведение тщательного расследования для выявления источника вспышки. Это включает идентификацию заражённых людей, определение источников заражения (например, контакт с заражённым скотом, употребление заражённых продуктов) и анализ распространения болезни.

Экстренное лечение и медицинская помощь: предоставление высококвалифицированной медицинской помощи заражённым, включая применение соответствующих противовирусных или антибактериальных препаратов. Контроль за распространением – установление мер контроля за заражёнными животными, включающие изоляцию, карантин, тестирование животных и вакцинацию стада.

Идентификация и отслеживание контактов: выявление всех лиц, контактировавших с заражёнными для их мониторинга и, при необходимости, изоляция или лечение.

Общественное информирование: важным является предоставление населению точной информации о путях передачи инфекций, симптомах, методах профилактики и доступных способах лечения. Это помогает предотвратить панику и побуждает людей предпринимать правильные действия для защиты своего здоровья. Не менее важно проведение образовательных кампаний для общественности и медицинского персонала о способах передачи болезней, симптомах и профилактике.

Мониторинг и оценка: после проведения мер контроля необходимо отслеживать показатели заболеваемости, чтобы оценить эффективность принятых мер и понять, требуются ли дополнительные действия. Результаты санитарно-

эпидемиологического мониторинга используются в работе территориальных подразделений и подведомственных организаций государственного органа. По результатам санитарно-эпидемиологического мониторинга:

- составляются сводки, доклады, рекомендации, научные прогнозы, диаграммы, таблицы, характеризующие динамику, направленность и интенсивность развития изменений.

- принимаются управленческие решения в целях устранения нарушений законодательства Республики Казахстан в области обеспечения санитарно-эпидемиологического благополучия населения на территории Республики Казахстан.

- Результаты санитарно-эпидемиологического мониторинга размещаются на официальном интернет-ресурсе государственного органа по итогам полугодия, года и заслушиваются на совещании государственного органа по итогам года, в случаях превышения показателей заболеваемости, ухудшения показателей состояния объектов окружающей среды на совещаниях уполномоченного органа в области охраны окружающей среды.

Постоянное обучение: специалистам различных служб необходимо регулярно проходить обучение, чтобы повысить свою готовность к реагированию на вспышки инфекционных заболеваний и другие чрезвычайные ситуации в сфере общественного здравоохранения. Обучение должно быть доступно для специалистов всех уровней: районного, областного и республиканского.

Таким образом, создание системы реагирования на особо опасные инфекции необходимо для обеспечения готовности и адекватного ответа всех структур на возможные угрозы эпидемиологического характера, что является ключевым условием для обеспечения безопасности и здоровья населения страны.

Заключение

Прогнозирование эпидемиологического процесса помогает оценить эффективность мер по контролю инфекций, улучшить готовность медицинских систем и способствовать улучшению биологической безопасности Республики Казахстан. Оно позволяет оперативно выявлять потенциальные угрозы, оценивать риски и принимать меры для предотвращения распространения инфекций.

В рамках настоящего исследования была разработана система эпидемиологического прогнозирования (моделирования) и реагирования на особо опасные инфекции, предназначенная для обеспечения своевременного и эффективного реагирования на чрезвычайные ситуации, связанные с инфекционными заболеваниями. Прогнозирование (моделирование) были проведены на примере сибирской язвы, ККГЛ и бруцеллеза, так как, указанные инфекционные заболевания являются эндемичными и широко распространены на территории Казахстана.

Анализ различных методов прогнозирования, с использованием корреляции Пирсона и Спирмена, негативной биномиальной регрессии, включая ARIMA и SARIMA, позволили выделить их применимость для осуществления прогноза, позволяющие предвидеть заболеваемость в будущем. В работе были проанализированы исторические данные о заболеваемости, а также изучены различные факторы, которые могут влиять на распространение инфекции.

При помощи корреляций Пирсона и Спирмена, а также отрицательной биномиальной регрессии проведен анализ связи между климатическими факторами и уровнем заболеваемости животных при сибирской язве и бруцеллезе, а также активностью переносчиков клещей при ККГЛ. Это позволило определить отсроченное влияние климатических условий на распространение данных заболеваний. Но в то же время применение перечисленных методов не позволяет в полной мере предсказать сезонность и цикличность распространения инфекционных заболеваний, что могло снизить точность прогнозирования (моделирования). Поэтому, для получения достоверных прогнозов и проведения оценки потенциальных рисков распространения особо опасных инфекций были рассмотрены модели ARIMA и SARIMA.

Исследование показало, что модели ARIMA и SARIMA являются эффективными инструментами для анализа и прогнозирования эпидемических процессов. Полученные результаты подчеркивают важность постоянного мониторинга и анализа эпидемиологической ситуации, а также необходимость разработки и внедрения эффективных мер профилактики и контроля инфекционных заболеваний. Это особенно важно в контексте распространения новых штаммов вирусов и возрастающей угрозы возникновения эпидемий.

Несмотря на все преимущества использования моделей ARIMA и SARIMA для эпидемиологического моделирования, важно помнить, что их применение требует тщательного анализа данных и учёта всех возможных факторов, влияющих на

распространение заболевания. Это позволяет получить наиболее точные и надёжные прогнозы. Кроме того, результаты исследования должны быть интерпретированы с учётом специфики конкретной страны или региона. Также необходимо учитывать возможные изменения в эпидемиологической обстановке и своевременно корректировать модели.

Таким образом, модели ARIMA и SARIMA представляют собой эффективный инструмент для прогнозирования распространения опасных инфекций и выявления потенциальных угроз. Использование данных методов может помочь в разработке эффективных стратегий по снижению рисков и обеспечению безопасности населения. Дальнейшее усовершенствование этих моделей и внедрение их в систему эпидемиологического прогнозирования в Казахстане будет способствовать повышению эффективности реагирования на эпидемические угрозы и улучшению эпидемиологической обстановки в стране.

В заключение необходимо отметить, что система эпидемиологического прогнозирования и реагирования на особо опасные инфекции играет ключевую роль в обеспечении биологической безопасности общества. Она помогает эффективно выявлять, предотвращать и контролировать распространение инфекционных заболеваний, тем самым способствуя сохранению здоровья населения и минимизации экономических потерь. Дальнейшее развитие и совершенствование данной системы будет способствовать укреплению национальной и глобальной безопасности.

В условиях постоянно меняющейся эпидемиологической обстановки важно продолжать развивать и совершенствовать эту систему, внедрять новые методы и технологии для более точного и своевременного прогнозирования, а также повышать уровень подготовки специалистов, ответственных за реагирование на чрезвычайные ситуации. Это позволит обеспечить эффективную защиту населения от возможных эпидемий и пандемий.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. RStudio. (n.d.). RStudio Desktop. Retrieved October 1, 2023, from <https://posit.co/download/rstudio-desktop/>
2. Yule, G.U. (1927). On a method of investigating periodicities disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society A*, 226(636– 646), 267–298.
3. Richard Migisha et al. (2021). Risk factors for human anthrax outbreak in Kiruhura District, Southwestern Uganda: a population-based case-control study. *PAMJ - One Health*, 5, 13. [doi: 10.11604/pamj-oh.2021.5.13.29385]
4. Otieno, F. T., Gachohi, J., Gikuma-Njuru, P., Kariuki, P., Oyas, H., Canfield, S. A., Bett, B., Njenga, M. K., & Blackburn, J. K. (2021). Modeling the Potential Future Distribution of Anthrax Outbreaks under Multiple Climate Change Scenarios for Kenya. *International journal of environmental research and public health*, 18(8), 4176. <https://doi.org/10.3390/ijerph18084176>
5. 79. Walsh, M.G., de Smalen, A.W. & Mor, S.M. (2018). Climatic influence on anthrax suitability in warming northern latitudes. *Sci Rep*, 8, 9269. <https://doi.org/10.1038/s41598-018- 27604-w>
6. Nguna, J., Dione, M., Apamaku, M., Majalija, S., Mugizi, D. R., Odoch, T., Kato, C. D., Tumwine, G., Kabaasa, J. D., Curtis, K., Graham, M., Ejobi, F., & Graham, T. (2019). Seroprevalence of brucellosis and risk factors associated with its seropositivity in cattle, goats and humans in Iganga District, Uganda. *The Pan African medical journal*, 33, 99. <https://doi.org/10.11604/pamj.2019.33.99.16960>
7. Long-Ting Cao, Hong-Hui Liu, J. Li, Xiao-Dong Yin, Y. Duan, J. Wang. Relationship of meteorological factors and human brucellosis in Hebei province, China 2020 Feb 10:703:135491. doi: 10.1016/j.scitotenv.2019.135491.
8. H. Chen, Meng-Xuan Lin, Li-Ping Wang, Yin-Xiang Huang, Y. Feng, Li-Qun Fang, L. Wang, Hong-Bin Song, Li-Gui Wang. Driving role of climatic and socioenvironmental factors on human brucellosis in China: machine-learning-based predictive analyses. 2023 Apr 12;12(1):36. doi: 10.1186/s40249-023-01087-y.
9. S. Huang, H. Wang, Z. Li, Z. Wang, T. Ma, R. Song, M. Lu, X. Han, Y. Zhang, Y Wang, Q Zhen, T Shui. Risk effects of meteorological factors on human brucellosis in Jilin province, China, 2005-2019, 2024 Apr 15;10(8):e29611. doi:10.1016/j.heliyon.2024.e29611.
10. Каримов С.К., Дурумбетов Е.Е., Казаков С.В. Экологические и эпидемиологические аспекты Крымской-Конго геморрагической лихорадки // Монография. – Алматы, 2003. с. 137. URL: <https://cloud.mail.ru/public/kXwe/qNQQZ6xhc>
11. Talgat Nurmakhanov, Yerlan Sansyzbaev, Bakhyt Atshabar and all. Crimean-Congo haemorrhagic fever virus in Kazakhstan (1948-2013) //International Journal of

Infectious Diseases. September 2015. Volume 38, Pages 19–23. [Электронный ресурс] <https://www.sciencedirect.com/science/article/pii/S1201971215001745>

12. Приказ Министра здравоохранения Республики Казахстан от 12 ноября 2021 года № ҚР ДСМ-114 «Об утверждении Санитарных правил "Санитарно-эпидемиологические требования к организации и проведению санитарно-противоэпидемических, санитарно-профилактических мероприятий по предупреждению особо опасных инфекционных заболеваний". [Электронный ресурс]. – URL: <https://adilet.zan.kz/rus/docs/V2100025151>

13. Мониторинг за эпизоотической и эпидемической ситуацией по бруцеллёзу в Республике Казахстан за 2006-2022 годы.

14. Постановление Главного Государственного санитарного врача Министерства здравоохранения Республики Казахстан от 11 февраля 2005 года № 3 «О проведении санитарно-противоэпидемических (профилактических) мероприятий и мониторинга эпизоотолого-эпидемической ситуации, в природных очагах Конго-Крымской геморрагической лихорадки Республики Казахстан». [Электронный ресурс]. – URL: https://online.zakon.kz/Document/?doc_id=30142452&pos=3;-90#pos=3;-90

15. Ерубаев, С.В. Казаков, Айкимбаев А.М., З.С. Турлиев, Г.М. Сайрамбекова, У.Б. Усенов, Т.И. Нурмаханов Исторические сведения о заболеваемости Крамской Конго геморрагической лихорадкой в странах Европы, Африки и Азии а период 1944-2022 годы // Ж. Особо опасные инфекции и биобезопасность. Алматы. №5. - 2023. - С. -38-46

16. Айкимбаев А.М., Казаков С.В., Касымканова Л.С. Конго-Крымская геморрагическая лихорадка. Алматы, 2010. 80 с. URL:<https://cloud.mail.ru/public/xbY2/p2XyJLzNf>

17. Профилактика сибирской язвы в Казахстане.- Алматы, 2020 г./Лухнова Л.Ю., Айкимбаев А.М., Ерубаев Т.К., Избанова У.А.- 256 С.

18. АТЛАС распространения особо опасных инфекций в Республике Казахстан. Издание третье, дополненное и переработанное. Под общей редакцией д.м.н. Ерубаева Т.К. // Алматы.- 2021 г., 217 с.

19. Прислегина Д.А., Дубянский В.М., Платонов А.Е., Малецкая О.В. Влияние климатических факторов на эпидемиологическую ситуацию по природно-очаговым инфекциям // Инфекция и иммунитет. – 2021. – Т. 11, № 5. – С.820-836. [Электронный ресурс]. <https://cyberleninka.ru/article/n/vliyanie-prirodno-klimaticheskikh-faktorov-na-epidemiologicheskuyu-situatsiyu-po-prirodno-ochagovym-infektsiyam/viewer>

20. Сведения о заболеваемости карантинными и особо опасными инфекциями в мире по состоянию 2022 год. Национальный научный центр особо опасных инфекций им. М. Айкимбаева МЗ РК [Электронный ресурс]. – URL: <https://cloud.mail.ru/public/nim3/FH1CNpTxb>

21. Hanslik T., Boelle P. Y. & Flahault A. The control chart: an epidemiological tool for public health monitoring. Public Health 115, 277–281 (2001).

22. Егембердиева Р. А., Ермуханова Н. Т., Утепбергенова Г. А. и др. Внутрибольничное инфицирование Крымской геморрагической лихорадкой // Гигиена, эпидемиология и иммунобиология. 2009. № 1 (39). С. 82—85.
23. Кадастр стационарно неблагополучных по сибирской язве населенных пунктов Республики Казахстан (1935-2018 гг.) / под общей редакцией д.м.н. Л.Ю. Лухновой. – Алматы: КНЦКЗИ, 2019. – 460 с.
24. Woods, C. W., Ospanov, K., Myrzabekov, A., Favorov, M., Plikaytis, B., & Ashford, D. A. (2004). Risk factors for human anthrax among contacts of anthrax-infected livestock in Kazakhstan. *The American journal of tropical medicine and hygiene*, 71(1), 48–52
25. Liu, K., Yang, Z., Liang, W. et al. (2020). Effect of climatic factors on the seasonal fluctuation of human brucellosis in Yulin, northern China. *BMC Public Health*, 20, 506. <https://doi.org/10.1186/s12889-020-08599-4>
26. Ansari, H., Mansournia, M., Izadi, S., Zeinali, M., Mahmoodi, M., & Holakouie-Naieni, K. (2015). Predicting CCHF incidence and its related factors using time-series analysis in the southeast of Iran: Comparison of SARIMA and Markov switching models. *Epidemiology & Infection*, 143(4), 839-850. doi:10.1017/S0950268814001113
27. Akermi, S.E., L'Hadj, M., & Selmane, S. (2022). Epidemiology and time series analysis of human brucellosis in Tebessa province, Algeria, from 2000 to 2020. *Journal of Research in Health Sciences*, 22(1), e00544.
28. Wu, W., An, S.Y., Guan, P., et al. (2019). Time series analysis of human brucellosis in mainland China by using Elman and Jordan recurrent neural networks. *BMC Infectious Diseases*, 19(1), 414. <https://doi.org/10.1186/s12879-019-4028-x>

Приложение А

Результаты линейного регрессионного анализа сибирской язвы “cases” – количеством случаев сибирской язвой и “scot” – количеством заболевшего скота

reg cases scot

Source	SS	df	MS	Number of obs	=	10
Model	637.620558	1	637.620558	F(1, 8)	=	21.95
Residual	232.379442	8	29.0474302	Prob > F	=	0.0016
				R-squared	=	0.7329
				Adj R-squared	=	0.6995
Total	870	9	96.6666667	Root MSE	=	5.3896

cases	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
scot	1.005711	.2146573	4.69	0.002	.5107099	1.500711
_cons	3.36802	2.085604	1.61	0.145	-1.44139	8.177431

Данные о заболевших животных сибирской язвой за период с 2013 г. по 2022 г. были получены от филиала НПЦСЭЭиМ НЦОЗ МЗ РК в виде абсолютных значений. За 2013–2022 годы абсолютное число заболевших животных сибирской язвой составило 62 головы (КРС и МРС) и 90 подтвержденных случаев заболевания среди людей.

На основе полученных данных был проведен анализ Пирсона, где коэффициент корреляции $R = 0,86$ показал сильную положительную линейную взаимосвязь ($p < 0,002$) между количеством поголовья больного скота и случаями заболевания сибирской язвой людей. Это означает, что по мере увеличения поголовья больного скота число случаев сибирской язвы также имеет тенденцию к увеличению.

Коэффициент равный 1,01 показывает, что при увеличении количества заболевших животных на одну единицу, можно ожидать увеличения случаев сибирской язвы среди людей примерно на такую же величину. Значение $p = 0,002$ подтверждает статистическую значимость этой взаимосвязи.

Скорректированный коэффициент детерминации R^2 равен 0,69. Это означает, что 69% изменений в количестве случаев сибирской язвы можно объяснить влиянием количества заболевшего скота.

Приложение Б

Матрица коэффициентов корреляции Спирмена между переменными и статистика коллинеарности при сибирской язве

	cases	temper-e	hum	pressure	precip-n	lag1_t-p	lag2_t-p	lag3_t-p	lag1_hum	lag2_hum	lag3_hum	lag1_p-s	lag2_p-s	lag3_p-s	lag1_p-p	lag2_p-p	lag3_p-p
cases	1.0000																
temperature	0.4042 0.0000	1.0000															
hum	-0.3392 0.0000	-0.8363 0.0000	1.0000														
pressure	-0.0843 0.0000	-0.2316 0.0000	0.2729 0.0000	1.0000													
precipitation	-0.0683 0.0000	-0.0877 0.0000	0.2739 0.0000	-0.2704 0.0000	1.0000												
lag1_temp	0.4049 0.0000	0.8834 0.0000	-0.7455 0.0000	-0.2130 0.0000	-0.0010 0.9433	1.0000											
lag2_temp	0.4055 0.0000	0.8937 0.0000	-0.7439 0.0000	-0.1870 0.0000	-0.0468 0.0012	0.8834 0.0000	1.0000										
lag3_temp	0.4048 0.0000	0.8701 0.0000	-0.7527 0.0000	-0.2036 0.0000	-0.0625 0.0000	0.8937 0.0000	0.8835 0.0000	1.0000									
lag1_hum	-0.3392 0.0000	-0.7473 0.0000	0.7121 0.0000	0.1278 0.0000	0.0924 0.0000	-0.8365 0.0000	-0.7462 0.0000	-0.7442 0.0000	1.0000								
lag2_hum	-0.3403 0.0000	-0.7523 0.0000	0.6964 0.0000	0.1203 0.0000	0.0913 0.0000	-0.7473 0.0000	-0.8366 0.0000	-0.7463 0.0000	0.7126 0.0000	1.0000							
lag3_hum	-0.3407 0.0000	-0.7246 0.0000	0.6783 0.0000	0.1644 0.0000	0.0685 0.0000	-0.7524 0.0000	-0.7478 0.0000	-0.8365 0.0000	0.6964 0.0000	0.7131 0.0000	1.0000						
lag1_press	-0.0835 0.0000	-0.2864 0.0000	0.2482 0.0000	0.1686 0.0000	0.0497 0.0006	-0.2313 0.0000	-0.2126 0.0000	-0.1866 0.0000	0.2722 0.0000	0.1276 0.0000	0.1196 0.0000	1.0000					
lag2_press	-0.0852 0.0000	-0.3006 0.0000	0.2025 0.0000	-0.1987 0.0000	0.1070 0.0000	-0.2869 0.0000	-0.2315 0.0000	-0.2129 0.0000	0.2493 0.0000	0.2727 0.0000	0.1285 0.0000	0.1689 0.0000	1.0000				
lag3_press	-0.0861 0.0000	-0.1156 0.0000	0.0835 0.0000	-0.1029 0.0000	-0.0304 0.0355	-0.3001 0.0000	-0.2866 0.0000	-0.2311 0.0000	0.2024 0.0000	0.2497 0.0000	0.2731 0.0000	-0.1989 0.0000	0.1691 0.0000	1.0000			

Таблица 1 – Расшифровка переменных, использованных в корреляции и регрессии

Наименование переменных	Описание переменных
cases	количество случаев сибирской язвы
scot	количество заболевшего скота
temperature	средняя ежемесячная температура
hum	средняя ежемесячная влажность
pressure	среднее ежемесячное атмосферное давление
precipitation	средние ежемесячные осадки
lag1_temp	задержка температуры на 1 месяц
lag2_temp	задержка температуры на 2 месяца
lag3_temp	задержка температуры на 3 месяца
lag1_hum	задержка влажности на 1 месяц
lag2_hum	задержка влажности на 2 месяца
lag3_hum	задержка влажности на 3 месяца
lag1_press	задержка атмосферного давления на 1 месяц
lag2_press	задержка атмосферного давления на 2 месяца
lag3_press	задержка атмосферного давления на 3 месяца
lag1_precip	задержка осадков на 1 месяц

Климатические параметры были получены с официального сайта Национальной гидрометеорологической службы Республики Казахстан РГП

«Казгидромет».

Климатические параметры имеют умеренную связь со случаями сибирской язвой. В результате проведенной корреляции Спирмена выявлена положительная корреляция между случаями сибирской язвой и средней температурой, задержками по температуре на 1, 2 и 3 месяца, где все результаты являются статистически значимыми ($p < 0,001$). В остальных случаях наблюдалась отрицательная корреляция со случаями сибирской язвы.

Проведённый анализ выявил прямую зависимость между случаями заболевания сибирской язвой и определёнными температурными показателями. Коэффициент корреляции R для этих параметров равен 0,40.

Приложение В

Результаты отрицательной биномиальной регрессии сибирской язвы с основными переменными

```
Negative binomial regression      Number of obs      =      4,794
LR chi2(4)                      =      788.65
Dispersion      = mean          Prob > chi2         =      0.0000
Log likelihood = -3931.1243      Pseudo R2          =      0.0912
```

cases	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
temperature	1.114882	.0082619	14.67	0.000	1.098806	1.131194
hum	.9996972	.0001325	-2.28	0.022	.9994376	.9999569
pressure	1.002978	.0016447	1.81	0.070	.9997592	1.006207
precipitation	1.003668	.0020234	1.82	0.069	.9997105	1.007642
_cons	.0091889	.0143093	-3.01	0.003	.0004343	.1944375
/lnalpha	1.604593	.0500912			1.506416	1.70277
alpha	4.975834	.2492454			4.510536	5.48913

Note: Estimates are transformed only in the first equation.

Note: _cons estimates baseline incidence rate.

LR test of alpha=0: `chibar2(01) = 4490.10` Prob >= `chibar2` = 0.000

Повышение среднесуточной температуры на 1°C приводит к увеличению суточной частоты случаев заболевания сибирской язвой на 11,14%. Относительный риск (IRR) составляет 1,11 с 95% доверительным интервалом от 1,09 до 1,13. Статистическая значимость выявленной взаимосвязи подтверждается значением $p < 0,001$. Это означает, что влияние температуры на заболеваемость сибирской язвой является статистически значимым.

Приложение Г

Результаты отрицательной биномиальной регрессии сибирской язвы с переменными задержек

Negative binomial regression	Number of obs	=	4,793
	LR chi2(4)	=	790.20
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -3930.1311	Pseudo R2	=	0.0913

cases	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
lag1_temp	1.116988	.008313	14.87	0.000	1.100813	1.1334
lag1_hum	.999737	.000132	-1.99	0.046	.9994783	.9999959
lag1_press	1.002785	.0016378	1.70	0.089	.9995801	1.006
lag1_precip	1.003595	.0020124	1.79	0.074	.9996583	1.007547
_cons	.0103721	.0160919	-2.94	0.003	.0004958	.2170046
/lnalpha	1.601295	.0501458			1.503011	1.699579
alpha	4.95945	.2486955			4.495203	5.471642

Note: Estimates are transformed only in the first equation.

Note: _cons estimates baseline incidence rate.

LR test of alpha=0: chibar2(01) = 4476.01 Prob >= chibar2 = 0.000

Если температура за прошедший месяц повышается на 1°C, то ожидается увеличение количества случаев сибирской язвы на 11,7%. Повышение влажности за прошедший месяц на 1% приводит к небольшому уменьшению частоты случаев сибирской язвы – примерно на 0,03%.

Приложение Д

Результаты регрессионного анализа ККГЛ

```
. reg CCHFcases catnumunfavtickscollect ixodid4 abundance alltype1 alltype6 biteblood time1 age4 age5 agricwork unemploy singleca
> ses numofperscont noassociat hwc activities1 activities3 characttreatml center1 tvchannel leaflets cpdoeder
```

Source	SS	df	MS	Number of obs	=	26
Model	610.190243	22	27.7359201	F(22, 3)	=	30.02
Residual	2.77129549	3	.923765164	Prob > F	=	0.0084
				R-squared	=	0.9955
				Adj R-squared	=	0.9623
Total	612.961538	25	24.5184615	Root MSE	=	.96113

CCHFcases	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
catnumunfavtickscollect	.1325337	.070884	1.87	0.158	-.0930509 .3581183
ixodid4	-.0036636	.0008201	-4.47	0.021	-.0062736 -.0010536
abundance	.0598095	.1705212	0.35	0.749	-.4828652 .6024842
alltype1	-.0241619	.0053434	-4.52	0.020	-.041167 -.0071568
alltype6	-.002799	.0013049	2.14	0.121	-.001354 .0069519
biteblood	.0978749	.0570589	1.72	0.185	-.083712 .2794619
time1	.0493302	.3047321	0.16	0.882	-.9204635 1.019124
age4	-.0813486	.4261737	-0.19	0.861	-1.437624 1.274926
age5	-.5664103	.3989152	-1.42	0.251	-1.835936 .7031159
agricwork	1.138746	.6383294	1.78	0.172	-.8927031 3.170195
unemploy	1.17584	.4855967	2.42	0.094	-.369545 2.721226
singlecases	1.537023	.5533375	2.78	0.069	-.2239434 3.29799
numofperscont	-.024844	.0257634	-0.96	0.406	-.1068346 .0571467
noassociat	.0039494	.0279456	0.14	0.897	-.0849859 .0928847
hwc	.0146276	.0236227	0.62	0.580	-.0605502 .0898054
activities1	-.1834618	.6671582	-0.27	0.801	-2.306657 1.939733
activities3	-.5028276	.2974495	-1.69	0.190	-1.449445 .4437894
characttreatml	.0074744	.0799701	0.09	0.931	-.2470261 .2619749
center1	-.0969146	.031941	-3.03	0.056	-.198565 .0047358
tvchannel	.0590383	.0607295	0.97	0.403	-.1342302 .2523068
leaflets	-.0000803	.000027	-2.97	0.059	-.0001662 5.69e-06
cpdoeder	.6393623	.2079611	3.07	0.054	-.0224628 1.301187
_cons	-3.830811	1.161307	-3.30	0.046	-7.526608 -.1350142

Информации форм мониторинга ККГЛ, полученные от НПЦСЭЭиМ НЦОЗ МЗ РК. В исследование включены данные о 157 зарегистрированных случаях заболевания ККГЛ за период с 2016 по 2022 годы.

Результаты регрессионного анализа показывают, что 96% ($R^2 = 0,96$, $p < 0,001$) случаев распространения ККГЛ происходит в результате воздействия клещей.

Таблица 2 – Расшифровка переменных, использованных в регрессии анализа факторов, связанные с количеством случаев ККГЛ

Наименование переменных	Описание переменных
catnumunfavtickscollect	Количество неблагополучных населенных пунктов, из которых собраны клещи
ixodid4	Клещи собраны с очага
abundance	Показатель частоты обилия
alltype1	H. Asiaticum (вид клещей)
alltype6	Другие виды
biteblood	Укус клеща или контакт с кровью клеща
time1	Время госпитализации в течение 1-3 дней
age4	Распределение больных с диагнозом ККГЛ по возрасту 18-29
age5	Распределение больных с диагнозом ККГЛ по возрасту 30+
agricwork	Распределение больных с диагнозом ККГЛ по организационным группам сельскохозяйственные рабочие
Unemploy	Безработные

Продолжение таблицы 2

singlecases	Все очаги зарегистрированы по ККГЛ. В том числе очаги, в которых зарегистрирован 1 случай
numofperscont	Общее количество лиц, имевших контакт с больным ККГЛ
noassociat	По типу связи с пациентом с ККГЛ: нет тесной связи с пациентом
hcw	По типу связи с больным ККГЛ: В том числе медицинским работникам
activities1	Мероприятия, проводимые в центре ККГЛ: Места, где были взяты пробы и проведены лабораторные исследования
activities3	Мероприятия, проводимые в центре ККГЛ: места дезинфекции
characttreatm1	Проведено лечение: Количество населенных пунктов
center1	Лечение проводится в центрах, где зарегистрирован ККГЛ: двор, количество сараев
tvchannel	Профилактические мероприятия интервью телеканалу
leaflets	Профилактические мероприятия листовки были розданы
cpdoeder	Приказы, изданные Департаментом по защите прав потребителей

Приложение Е

Сводная статистика погодных параметров в Жамбылской, Кызылординской и Туркестанской областях

Переменные	Среднее	Стандартное отклонение	min	P25	P50	P75	max
Жамбылская область							
Средняя температура, (°C)	11,46	10,49	-11,25	1,86	11,83	21,22	28,3
Среднее атмосферное давление, (г Па)	941,86	5,14	931,92	937,22	942,09	946,1	951,6
Средняя относительная влажность, (%)	60,08	17,36	29	42,48	62,5	74,58	92
Среднее количество осадков, (мм)	28,35	24,23	0	10,5	23,75	39,7	130,1
Кызылординская область							
Средняя температура, (°C)	11,76	13,01	-17,29	-0,4	11,97	24,36	31,7
Среднее атмосферное давление, (г Па)	1002,4	6,82	989,3	996,89	1002,96	1008,14	1017
Средняя относительная влажность, (%)	53,55	18,61	25,37	35,33	50,46	71,68	87,18
Среднее количество осадков, (мм)	11,53	11,38	0	2,3	8,05	18,1	51
Туркестанская область							
Средняя температура, (°C)	13,87	11,67	-13,6	3,3	13,9	24,9	32,1
Среднее атмосферное давление, (г Па)	992,28	6,95	979,32	985,7	992,74	998,17	1005,6
Средняя относительная влажность, (%)	94,23	58,03	1	41,5	93,5	146,5	190
Среднее количество осадков, (мм)	18,56	17,89	0	2,2	15,3	28,95	74,4

Приложение Ж

Матрица коэффициентов корреляции Спирмена между переменными и статистика коллинеарности при ККГЛ

	cases	temperature	humidity	pressure	precipitation	lag1_temp	lag2_temp	lag1_hum	lag2_hum	lag1_press	lag1_precip	lag3_precip
cases	1.0000											
temperature	0.5289	1.0000										
	0.0000											
humidity	-0.4902	-0.9096	1.0000									
	0.0000	0.0000										
pressure	-0.1502	-0.4126	0.2589	1.0000								
	0.0001	0.0000	0.0000									
precipitation	-0.3270	-0.5022	0.6715	-0.0794	1.0000							
	0.0000	0.0000	0.0000	0.0342								
lag1_temp	0.4536	0.8412	-0.8238	-0.3102	-0.5731	1.0000						
	0.0000	0.0000	0.0000	0.0000	0.0000							
lag2_temp	0.2774	0.4836	-0.5318	-0.1247	-0.5054	0.8406	1.0000					
	0.0000	0.0000	0.0000	0.0009	0.0000	0.0000						
lag1_hum	-0.4208	-0.7304	0.7827	0.1562	0.5679	-0.9089	-0.8239	1.0000				
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000					
lag2_hum	-0.2475	-0.3678	0.4458	-0.0252	0.4528	-0.7287	-0.9094	0.7829	1.0000			
	0.0000	0.0000	0.0000	0.5019	0.0000	0.0000	0.0000	0.0000				
lag1_press	-0.1197	-0.3695	0.2581	0.9429	-0.0124	-0.4121	-0.3075	0.2551	0.1518	1.0000		
	0.0014	0.0000	0.0000	0.0000	0.7408	0.0000	0.0000	0.0000	0.0000			
lag1_precip	-0.1712	-0.2793	0.4134	-0.1870	0.4611	-0.5033	-0.5759	0.6735	0.5725	-0.0832	1.0000	
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0265		
lag3_precip	0.0966	0.3142	-0.1862	-0.4610	0.0210	0.0206	-0.2696	0.0987	0.4071	-0.3297	0.2338	1.0000
	0.0100	0.0000	0.0000	0.0000	0.5753	0.5829	0.0000	0.0085	0.0000	0.0000	0.0000	

Выяснилось, что средняя температура, а также более высокие температуры в последующие два месяца и осадки с задержкой на три месяца положительно коррелируют с случаями ККГЛ, где все результаты были статистически значимы $p < 0,001$. В остальных случаях корреляция между климатическими параметрами и случаями заболевания была отрицательной.

Таблица 3 – Расшифровка переменных, климатических параметров корреляции Спирмена

Наименование переменных	Описание переменных
cases	количество случаев ККГЛ
temperature	средняя ежемесячная температура
hum	средняя ежемесячная влажность
pressure	среднее ежемесячное атмосферное давление
precipitation	средние ежемесячные осадки
lag1_temp	задержка температуры на 1 месяц

Продолжение таблицы 3

lag2_temp	задержка температуры на 2 месяца
lag1_hum	задержка влажности на 1 месяц
lag2_hum	задержка влажности на 2 месяца
lag1_press	задержка атмосферного давления на 1 месяц
lag1_precip	задержка осадков на 1 месяц
lag3_precip	задержка осадков на 3 месяца

Приложение 3

Результаты отрицательной биномиальной регрессии ККГЛ с основными переменными

```
Negative binomial regression      Number of obs      =      718
LR chi2(4)                       =      236.34
Dispersion = mean                Prob > chi2         =      0.0000
Log likelihood = -425.36301      Pseudo R2          =      0.2174
```

cases	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
temperature	1.20468	.0340144	6.60	0.000	1.139824	1.273226
humidity	1.000116	.0024515	0.05	0.962	.9953231	1.004933
pressure	1.011797	.0041557	2.86	0.004	1.003684	1.019975
precipitation	1.013004	.0077095	1.70	0.090	.9980063	1.028228
_cons	7.13e-08	3.07e-07	-3.82	0.000	1.55e-11	.000328
/lnalpha	.1915043	.2105082			-.2210842	.6040927
alpha	1.21107	.2549402			.8016492	1.829592

Note: Estimates are transformed only in the first equation.

Note: _cons estimates baseline incidence rate.

LR test of alpha=0: $\chi^2(01) = 84.23$

Prob >= $\chi^2_{bar2} = 0.000$

Повышение среднесуточной температуры на 1°C приводит к увеличению суточной частоты случаев ККГЛ на 20,47% с индексом относительного риска (IRR) 1,204. Это значение статистически значимо, так как находится в пределах 95% доверительного интервала от 1,14 до 1,27 и $p < 0,001$. Также выявлено, что увеличение среднего давления воздуха на 1 гПА связано с увеличением количества случаев ККГЛ на 1,18%, при этом IRR равен 1,012, что также является статистически значимым предиктором.

Приложение И

Результаты отрицательной биномиальной регрессии ККГЛ с переменными задержек

```
Negative binomial regression          Number of obs      =          716
LR chi2(4)                          =          217.94
Dispersion      = mean              Prob > chi2         =          0.0000
Log likelihood = -437.12318         Pseudo R2          =          0.1995
```

cases	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
lag1_temp	1.249454	.0246992	11.27	0.000	1.20197	1.298814
lag2_hum	1.005349	.0011978	4.48	0.000	1.003004	1.007699
lag1_press	1.022675	.0045241	5.07	0.000	1.013847	1.031581
lag1_precip	1.02352	.0062748	3.79	0.000	1.011295	1.035892
_cons	3.55e-13	1.62e-12	-6.31	0.000	4.80e-17	2.63e-09
/lnalpha	.4436421	.1888074			.0735865	.8136977
alpha	1.558373	.2942322			1.076362	2.256236

Note: Estimates are transformed only in the first equation.

Note: _cons estimates baseline incidence rate.

LR test of alpha=0: `chibar2(01) = 115.00`

`Prob >= chibar2 = 0.000`

- при увеличении температуры за прошлый период (1 месяц назад) на 1°C ожидается увеличение ожидаемого уровня случаев ККГЛ на 24,95%, а индекс относительного риска составит 1,249, при этом $p < 0,001$;
- увеличение влажности за позапрошлый период (2 месяца назад) на 1% связано с небольшим увеличением частоты случаев ККГЛ – 0,53% ($p < 0,001$);
- увеличение давления за прошедший период (1 месяц назад) на 1 гПА приводит к росту частоты случаев ККГЛ на 2,27%, при этом $p < 0,001$;
- повышение уровня осадков за прошлый период (также 1 месяц назад) на 1 мм увеличивает частоту случаев ККГЛ на 2,35%.

Приложение К

Матрица коэффициентов корреляции Спирмена при бруцеллезе

	cases	temper-e	hum	pressure	precip-n	lag1_t~p	lag2_t~p	lag3_t~p	lag1_hum	lag2_hum	lag3_hum	lag1_p~s	lag2_p~s	lag3_p~s	lag1_p~p
cases	1.0000														
temperature	0.6729	1.0000													
	0.0000														
hum	-0.5626	-0.8426	1.0000												
	0.0000	0.0000													
pressure	-0.1569	-0.2406	0.2563	1.0000											
	0.0000	0.0000	0.0000												
precipitation	0.0375	-0.1161	0.3115	-0.2196	1.0000										
	0.0449	0.0000	0.0000	0.0000											
lag1_temp	0.6707	0.8806	-0.7524	-0.2282	-0.0101	1.0000									
	0.0000	0.0000	0.0000	0.0000	0.5891										
lag2_temp	0.6688	0.8936	-0.7477	-0.1973	-0.0559	0.8805	1.0000								
	0.0000	0.0000	0.0000	0.0000	0.0028	0.0000									
lag3_temp	0.6654	0.8686	-0.7622	-0.2213	-0.0577	0.8936	0.8803	1.0000							
	0.0000	0.0000	0.0000	0.0000	0.0020	0.0000	0.0000								
lag1_hum	-0.5608	-0.7525	0.7386	0.1064	0.1193	-0.8427	-0.7523	-0.7479	1.0000						
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000							
lag2_hum	-0.5627	-0.7447	0.6942	0.1469	0.0838	-0.7526	-0.8428	-0.7524	0.7381	1.0000					
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000						
lag3_hum	-0.5627	-0.7259	0.6913	0.1882	0.0681	-0.7445	-0.7526	-0.8426	0.6942	0.7380	1.0000				
	0.0000	0.0000	0.0000	0.0000	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000					
lag1_press	-0.1555	-0.3006	0.2885	0.1719	0.1000	-0.2396	-0.2261	-0.1963	0.2549	0.1047	0.1463	1.0000			
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000				
lag2_press	-0.1578	-0.3003	0.2268	-0.1753	0.0944	-0.3038	-0.2418	-0.2287	0.2909	0.2583	0.1076	0.1711	1.0000		
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000			
lag3_press	-0.1586	-0.1257	0.0735	-0.1003	-0.0511	-0.3000	-0.3040	-0.2413	0.2271	0.2909	0.2587	-0.1766	0.1714	1.0000	
	0.0000	0.0000	0.0001	0.0000	0.0062	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		
lag1_precip	0.0375	-0.0323	0.0891	-0.1685	0.2300	-0.1166	-0.0106	-0.0566	0.3124	0.1201	0.0846	-0.2179	0.1004	0.0945	1.0000
	0.0447	0.0838	0.0000	0.0000	0.0000	0.0000	0.5697	0.0024	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	

В результате корреляционного анализа была выявлена положительная связь между случаями заболевания бруцеллёзом и средними показателями температуры, количеством осадков, а также температурными задержками в 1, 2 и 3 месяца и осадками за 1 месяц. Коэффициент корреляции R: средняя температура = 0,67, задержка по температуре на 1 месяц = 0,67, на 2 месяца = 0,67, на 3 месяца = 0,66, задержка по осадкам на 1 месяц = 0,04.

**Таблица 4 – Расшифровка переменных, использованных в корреляции
Спирмена и регрессии**

Наименование переменных	Описание переменных
cases	количество случаев бруцеллеза
temperature	средняя ежемесячная температура
hum	средняя ежемесячная влажность
pressure	среднее ежемесячное атмосферное давление
precipitation	средние ежемесячные осадки
lag1_temp	задержка температуры на 1 месяц
lag2_temp	задержка температуры на 2 месяца
lag3_temp	задержка температуры на 3 месяца
lag1_hum	задержка влажности на 1 месяц
lag2_hum	задержка влажности на 2 месяца
lag3_hum	задержка влажности на 3 месяца
lag1_press	задержка атмосферного давления на 1 месяц
lag2_precip	задержка атмосферного давления на 2 месяца
lag3_precip	задержка атмосферного давления на 3 месяца
lag1_precip	задержка осадков на 1 месяц

Приложение Л

Результаты отрицательной биномиальной регрессии бруцеллёза

Negative binomial regression			Number of obs = 2,876		
Dispersion = mean			LR chi2(4) = 1945.06		
Log likelihood = -14470.321			Prob > chi2 = 0.0000		
			Pseudo R2 = 0.0630		
cases	IRR	Std. Err.	z	P> z	[95% Conf. Interval]
temperature	1.033577	.0013831	24.68	0.000	1.03087 1.036292
hum	.9999078	.0000263	-3.50	0.000	.9998562 .9999594
pressure	1.001157	.0003375	3.43	0.001	1.000496 1.001819
precipitation	1.00368	.0004553	8.10	0.000	1.002788 1.004573
_cons	20.70597	6.75075	9.29	0.000	10.92896 39.22944
/lnalpha	-1.479214	.0270944			-1.532318 -1.42611
alpha	.2278167	.0061726			.2160343 .2402417

Note: Estimates are transformed only in the first equation.
 Note: _cons estimates baseline incidence rate.
 LR test of alpha=0: $\chi^2(01) = 4.9e+04$ Prob >= $\chi^2 = 0.000$

Исследование показало, что повышение среднесуточной температуры на 1°C связано с увеличением суточной частоты случаев бруцеллёза на 3,35% с IRR (95% ДИ 1,03-1,04, $p < 0,001$). С повышением среднего давления воздуха на 1 гПа частота случаев бруцеллёза увеличивается на 0,1% с IRR (95% ДИ 1,00-1,001, $p < 0,001$). А увеличение количества осадков на 1 мм приводит к увеличению случаев бруцеллёза на 0,3% с IRR (95% ДИ 1,002-1,004, $p < 0,01$). Повышение среднесуточной влажности на 1% незначительно уменьшает частоту случаев бруцеллёза, примерно на 0,03% IRR 0,99 (95% ДИ 0,998-0,999, $p < 0,001$). Влажность также имеет статистически значимое влияние на заболеваемость.

Абдураимов Е.О., Алдынгуров Д.К., Смагул М.А., Ахметова К.М.

**СОЗДАНИЕ СИСТЕМЫ ЭПИДЕМИОЛОГИЧЕСКОГО
ПРОГНОЗИРОВАНИЯ (МОДЕЛИРОВАНИЯ) И РЕАГИРОВАНИЯ НА
ОСОБО ОПАСНЫЕ ИНФЕКЦИИ**

(Методические рекомендации)

Подписано в печать 28.11.2024 г.

Формат 148х210

Тираж 50 экз.

ISBN 978-601-305-625-8

