УТВЕРЖДЕНА
приказом Национального института
общественного здоровья
от [__.__.2026] № [___]

ИНСТРУКЦИЯ

по краткосрочному прогнозированию заболеваемости ОРВИ и гриппом в Кыргызской Республике

Оглавление

Глава 1. Общие положения
Глава 2. Термины
Глава 3. Ответственные лица и периодичность
Глава 4. Подготовка рабочего места
Глава 5. Входные данные
Глава 6. Подготовка данных
Глава 7. Модели и параметры расчёта
Глава 8. Порядок запуска скрипта
Глава 9. Результаты расчёта
Глава 10. Отчёт и интерпретация результатов
Глава 11. Ограничения модели
Глава 12. Оценка качества прогноза
Глава 13. Техническая проверка расчёта и частые ошибки
Глава 14. Утверждение и пересмотр

Приложения

Приложение 1. Описание шагов скрипта
Приложение 2. Текст скрипта forecast_ari_kg_v4_2.R
Приложение 3. Структура аналитического отчёта и источники данных для его разделов
Приложение 4. Текст шаблона отчёта report_ari_ru.Rmd

Глава 1. Общие положения

1.1. Настоящая Инструкция устанавливает порядок краткосрочного прогнозирования заболеваемости острыми респираторными вирусными инфекциями (ОРВИ) и гриппом в Кыргызской Республике с использованием скрипта прогнозирования forecast_ari_kg_v4_2.R (алгоритм версии 4.2) в среде R.

1.2. Цель прогнозирования — оценка текущей эпидемиологической ситуации и ожидаемой динамики регистрируемой заболеваемости ОРВИ и гриппом на ближайшие 8 недель на республиканском уровне и по территориям.

1.3. Прогнозируется недельное число зарегистрированных случаев ОРВИ и гриппа (абсолютные числа) и интенсивный показатель на 100 тыс. населения. Прогноз строится на 8 недель после последней недели, за которую имеются данные, — по Кыргызской Республике в целом и по 9 территориям: г. Бишкек, г. Ош, Чуйская, Таласская, Иссык-Кульская, Нарынская, Ошская, Баткенская и Жалал-Абадская области.

1.4. Результаты прогнозирования используются для эпидемиологического надзора и могут использоваться Министерством здравоохранения Кыргызской Республики и организациями здравоохранения как дополнительный ориентир при планировании готовности к возможному увеличению обращаемости.

1.5. Инструкция включает четыре приложения: Приложение 1 — описание шагов скрипта; Приложение 2 — текст скрипта; Приложение 3 — структура аналитического отчёта и источники данных для его разделов; Приложение 4 — текст шаблона отчёта. Приложения являются неотъемлемой частью Инструкции.

Глава 2. Термины

2.1. В Инструкции используются следующие термины:

- **Эпидемиологическая неделя** — календарная неделя с понедельника по воскресенье, пронумерованная по стандарту ISO 8601. Скрипт определяет дату окончания недели (воскресенье) по её номеру и году. В году может быть 52 или 53 недели.
- **Абсолютное число случаев** — число случаев ОРВИ и гриппа, зарегистрированных за неделю. Абсолютные числа являются первичными данными, по которым строится прогноз.
- **Интенсивный показатель** — число зарегистрированных случаев на 100 тыс. населения за неделю.
- **Знаменатель (численность населения)** — численность населения, по которой рассчитан официальный интенсивный показатель. Скрипт не задаёт её вручную, а восстанавливает делением абсолютного числа на интенсивный показатель (с умножением на 100 000).
- **Режим знаменателя** — непрерывный отрезок недель, на котором официальный знаменатель не менялся. Официальная численность населения обновляется внутри года (как правило, с 25-й недели), поэтому за год может быть несколько режимов.
- **Официальный и гармонизированный интенсивный показатель** — официальный рассчитывается по знаменателю, действовавшему в данную неделю (со «ступеньками» при его смене); гармонизированный — по единому (опорному) знаменателю территории для всего ряда. Гармонизированный показатель используется для анализа динамики, так как не содержит скачков, вызванных сменой знаменателя.

- **Горизонт прогноза** — число недель после последней недели данных, на которое строится прогноз; в настоящей Инструкции — 8 недель.
- **Медиана (медианный прогноз)** — основной точечный прогноз на каждую неделю: центральное значение прогнозного распределения. Фактическое значение с равной вероятностью окажется выше или ниже медианы.
- **Среднее (ожидаемое) значение** — математическое ожидание прогнозного распределения. Средние значения можно суммировать, поэтому они используются для оценки общего ожидаемого числа случаев за несколько недель и для согласования региональных прогнозов с национальным.
- **Прогнозный интервал (80% и 95%)** — диапазон, в который фактическое значение должно попасть с вероятностью 80% или 95%. Чем дальше неделя от последней недели данных, тем шире интервал.
- **Доверительный интервал (ДИ)** — интервал, характеризующий точность оценки параметра; в Инструкции используется для недельного темпа прироста (95% ДИ).
- **Недельный темп прироста** — среднее относительное изменение числа зарегистрированных случаев за неделю, рассчитанное по последним 6 неделям.
- **Время удвоения (сокращения вдвое)** — число недель, за которое число случаев увеличится (уменьшится) в два раза при сохранении текущего темпа. Если 95% ДИ темпа прироста включает 0, направление не определено (плато).
- **Скользящая кросс-валидация** — проверка точности моделей на прошлых данных: модель многократно обучается на ряде, обрезанном на прошлую неделю (отсечение), прогноз на 8 недель вперёд сравнивается с фактическими данными.
- **Модели** — SNaive — сезонная наивная модель (значение той же недели прошлого года), используется как эталон сравнения; STL+ETS — сезонно-трендовая декомпозиция STL с экспоненциальным сглаживанием ETS; ARIMA с гармониками Фурье — модель ARIMA с годовой сезонностью, заданной гармониками Фурье, и праздничными неделями; TBATS — модель с тригонометрической сезонностью; NNAR — нейросетевая авторегрессия с гармониками Фурье; ансамбль — медиана прогнозов всех рассчитанных моделей.
- **MAE** — средняя абсолютная ошибка прогноза, в случаях.
- **MAPE** — средняя абсолютная процентная ошибка.
- **WAPE** — взвешенная абсолютная процентная ошибка: сумма абсолютных ошибок, делённая на сумму фактических значений, в процентах.
- **MASE** — средняя абсолютная масштабированная ошибка: ошибка модели, делённая на ошибку сезонной наивной модели на обучающих данных. Значение меньше 1 означает, что модель точнее сезонного наивного прогноза.
- **Смещение (Bias)** — систематическое завышение (положительное значение) или занижение (отрицательное значение) прогноза.
- **Покрытие интервала** — доля фактических значений, попавших в прогнозный интервал при кросс-валидации. Исходное (raw) покрытие — до калибровки, калиброванное (calibrated) — после.
- **Коэффициент калибровки k** — множитель ширины прогнозных интервалов, подобранный по ошибкам кросс-валидации так, чтобы покрытие 95% интервала соответствовало заявленному. k больше 1 расширяет интервал, меньше 1 — сужает.
- **Согласование уровней** — приведение суммы региональных прогнозов к национальному прогнозу: для каждой недели вычисляется коэффициент

«национальное среднее / сумма региональных средних», который применяется к региональным медианам, средним и границам интервалов.

- **Анализ чувствительности** — дополнительный расчёт, в котором аномально высокие недели заменяются базовым уровнем, чтобы оценить, насколько от них зависит прогноз. Основным прогнозом не является.
- **Проспективная проверка** — сравнение ранее выпущенного прогноза с фактическими данными, поступившими после его выпуска и не использовавшимися при его построении.

Глава 3. Ответственные лица и периодичность

3.1. Прогнозирование проводит ЦАУРОЗиПЗ.

3.2. Ввод еженедельных данных в файлы данных выполняет эпидемиолог ДПЗиГСЭН, научный сотрудник НИОЗ

3.3. Расчёт и подготовку отчёта выполняет научный сотрудник ЦАУРОЗиПЗ.

3.4. Отчёт утверждает и направляет в Министерство здравоохранения руководитель ЦАУРОЗиПЗ.

3.5. Прогноз пересчитывается еженедельно — после поступления данных за каждую завершившуюся эпидемиологическую неделю. Срок направления отчёта: по запросу ДПЗиГСЭН..

Глава 4. Подготовка рабочего места

4.1. Установка программ (обе программы бесплатные). Сначала устанавливается R, затем RStudio:

Программа	Где скачать	Назначение
R	https://cran.r-project.org	Язык, на котором написан скрипт
RStudio Desktop	https://posit.co/downloads/	Редактор для работы со скриптом

4.2. Для работы запускается только RStudio. Программу R отдельно запускать не нужно: RStudio находит её автоматически.

4.3. Установка пакетов R (выполняется один раз; нужно подключение к интернету). Скрипт использует пакеты pacman, readxl, dplyr, tidyr, purrr, stringr, readr, tibble, ggplot2 и forecast; для формирования отчёта (глава 10) нужны пакеты rmarkdown, knitr и flextable. Все пакеты устанавливаются одной командой: в RStudio скопировать команду в панель Console (слева внизу), нажать Enter и дождаться, когда в консоли снова появится знак «>»:

```
install.packages(c("pacman", "readxl", "dplyr", "tidyr", "purrr", "stringr",
"readr", "tibble", "ggplot2", "forecast", "rmarkdown", "knitr", "flextable"))
```

Если в консоли появится вопрос «Do you want to install from sources the packages which need compilation?», ввести n и нажать Enter. Установка занимает несколько минут. Проверить, что пакеты установлены, можно командой library(forecast); library(flextable) — при успешной установке сообщений об ошибке не появится. При

запуске скрипт сам доустанавливает недостающие пакеты через pacman, но предварительная установка исключает задержку и сбои при первом расчёте.

4.4. Создание папки проекта. На компьютере создаётся папка проекта forecast_flu_kg, в ней — подпапка data. Файлы размещаются так:

Папка / файл	**Что в неё помещается**
forecast_flu_kg/	Папка проекта
forecast_flu_kg/forecast_ari_kg_v4_2.R	Скрипт расчёта (текст — Приложение 2)
forecast_flu_kg/report_ari_ru.Rmd	Шаблон отчёта (текст — Приложение 4)
forecast_flu_kg/report_style.docx	Образец оформления отчёта в Word (шрифт, заголовки, таблицы); не редактируется
forecast_flu_kg/data/ОРВИ_и_грипп_по_КР.xlsx	Файл абсолютных чисел (глава 5)
forecast_flu_kg/data/Интен_показатель_ОРВИ_и_гриппа.xlsx	Файл интенсивных показателей (глава 5)
forecast_flu_kg/output_v4_2/	Создаётся скриптом автоматически; сюда сохраняются результаты (глава 9)
forecast_flu_kg/archive/check_log.csv	Журнал проверок прогноза; пополняется шаблоном отчёта автоматически
forecast_flu_kg/archive/	Архив результатов предыдущих расчётов (пункт 8.2)

4.5. Файлы скрипта, шаблона отчёта и образца оформления берутся из базы ЦАУРОЗиПЗ. Если электронной копии нет, текст переносится из Приложения 2: в RStudio выбрать File → New File → R Script, вставить текст приложения полностью и сохранить через File → Save with Encoding… в кодировке UTF-8 под именем forecast_ari_kg_v4_2.R в папку проекта.

4.6. Создание проекта RStudio (выполняется один раз). Скрипт ищет файлы данных и сохраняет результаты относительно рабочей папки (пути data/… и output_v4_2/), поэтому рабочей папкой должна быть папка проекта. Для этого в RStudio выбрать File → New Project → Existing Directory → указать папку forecast_flu_kg → Create Project. В папке появится файл forecast_flu_kg.Rproj.

4.7. Открытие проекта. Работа всегда начинается с двойного щелчка по файлу forecast_flu_kg.Rproj — запускается RStudio, и рабочей папкой становится папка проекта. В панели Files (справа внизу) должны быть видны папка data и файл скрипта.

Глава 5. Входные данные

5.1. Для расчёта используются два файла Excel, которые лежат в папке data/ и называются точно так:

Файл	Содержание
ОРВИ_и_грипп_по_КР.xlsx	Недельные абсолютные числа зарегистрированных случаев ОРВИ и гриппа — первичные данные для прогноза
Интен_показатель_ОРВИ_и_гриппа.xlsx	Официальные недельные интенсивные показатели на 100 тыс. населения — источник знаменателей и контрольной сверки

5.2. Оба файла имеют одинаковую структуру. Скрипт читает первый лист каждого файла, поэтому данные должны находиться на первом листе.

- На листе одна строка заголовка: для каждой территории — ячейка с её названием, а справа от неё — столбцы по годам с заголовками вида «2023г», «2024г», «2025г», «2026г».
- Под названием территории идёт столбец с номерами недель (1, 2, 3 … или «1 нед.»), под заголовками годов — значения за соответствующую неделю.
- Блоки территорий разделены пустым столбцом. В файлах 9 территорий: Бишкек, Ош, Чуй, Талас, ИК, Нарын, Ош обл, Баткен, Ж-Абад.
- Названия территорий в двух файлах должны побуквенно совпадать.
- Числа вводятся без пробелов-разделителей; дробная часть интенсивного показателя может отделяться точкой или запятой.

5.3. Еженедельное обновление. После завершения эпидемиологической недели в обоих файлах заполняется строка этой недели для всех 9 территорий. В начале нового года в каждый блок территории добавляется столбец нового года (например, «2027г») сразу после столбца предыдущего года.

5.4. Требования к полноте данных:

- недели должны идти без пропусков: скрипт останавливается, если в республиканском ряду пропущена неделя (в году с 53 ISO-неделями должна быть заполнена и 53-я неделя);
- в республиканский ряд включаются только недели, за которые отчитались все территории; если неделя заполнена не по всем территориям, скрипт исключает её и выводит сообщение;
- для каждой недели, по которой есть абсолютное число, в файле интенсивных показателей должно быть значение за ту же неделю (кроме лет, которых в нём нет, — пункт 6.2).

5.5. Данные вносятся из рабочей формы заполненной эпидемиологом ДПЗиГСЭН. Перед расчётом значения последней недели сверяются с исходной отчётностью.

Глава 6. Подготовка данных

6.1. Подготовку данных скрипт выполняет автоматически (Шаги 1–5, Приложение 1):

– переводит оба файла из «широкого» формата в таблицу «территория — год — неделя — значение»;
– восстанавливает знаменатели населения по каждой территории и неделе, выделяет режимы знаменателя и определяет опорный знаменатель территории — последний официальный;
– выполняет контрольную сверку файлов: пересчитывает интенсивный показатель из абсолютных чисел и восстановленных знаменателей и сравнивает его с файлом интенсивных показателей. Допустимое расхождение — 0,5%; при большем расхождении расчёт останавливается (это защищает от перепутанных столбцов и изменённой структуры листа);
– собирает ряды по территориям и республиканский ряд (сумма территорий по неделям, за которые отчитались все территории), рассчитывает официальный и гармонизированный интенсивные показатели;
– проводит диагностику данных: годовая динамика, недели с превышением медианы предыдущих 52 недель более чем в 3 раза (аномальные недели), провалы отчётности в праздничные недели, межтерриториальные различия заболеваемости за последний полный год, недельный темп прироста за последние 6 недель.

6.2. Для года, которого нет в файле интенсивных показателей, знаменатель принимается равным последнему официальному. Это допущение указывается в ограничениях отчёта.

6.3. Аномальные недели из основного ряда не удаляются: их влияние на прогноз оценивается анализом чувствительности (пункт 7.9).

Глава 7. Модели и параметры расчёта

7.1. Моделируются недельные абсолютные числа зарегистрированных случаев. Для стабилизации разброса используется логарифмическое преобразование (если все значения ряда больше нуля); годовая сезонность задаётся периодом 52 недели. Интенсивные показатели рассчитываются из прогноза абсолютных чисел.

7.2. Модели-кандидаты:

Уровень	Модели
Республиканский	SNaive, STL+ETS, ARIMA с гармониками Фурье, TBATS, NNAR с гармониками Фурье, ансамбль
Территориальный	SNaive, STL+ETS, ARIMA с гармониками Фурье, ансамбль

7.3. Каждая модель проверяется скользящей кросс-валидацией с горизонтом оценки 8 недель (равен горизонту прогноза): 52 отсечения на республиканском уровне и 26 — на территориальном. Минимальная длина обучающего ряда — 108 недель (два года и 4 недели). По результатам рассчитываются MAE, RMSE, MAPE, WAPE, MASE, смещение, покрытие 80% и 95% интервалов — в целом и отдельно для каждого горизонта от 1 до 8 недель.

7.4. Правило выбора рабочей модели (отдельно для республики и каждой территории): из моделей, кроме SNaive, отбираются те, у которых MAE не выше, чем у SNaive, и исходное покрытие 95% интервала не ниже 85%; из них выбирается модель с наименьшим MAE. Если ни одна модель условиям не отвечает, рабочей становится

SNaive. Выбор SNaive означает, что содержательные модели не дали прироста точности и прогноз является продолжением прошлогодней сезонной кривой; это указывается в резюме расчёта.

7.5. Точечный прогноз — медиана прогнозного распределения. Среднее значение получается отдельным расчётом той же модели с поправкой на смещение (bias adjustment); для ансамбля — как среднее арифметическое средних значений моделей.

7.6. Калибровка интервалов. По ошибкам кросс-валидации для выбранной модели подбирается коэффициент k, на который умножается ширина 80% и 95% интервалов. Нижняя граница k — 0,6, технический предел — 4,0; при достижении предела выводится предупреждение. Калиброванное покрытие при выборе модели не используется и служит диагностическим показателем.

7.7. Надёжность территориального прогноза оценивается по показателям кросс-валидации выбранной модели:

Оценка	**Условие**
Низкая	WAPE более 40%, или MASE более 1,25, или калиброванное покрытие 95% интервала менее 85%
Ограниченная	WAPE более 25%, или MASE более 1,00, или калиброванное покрытие менее 90% (если не выполнено условие «низкая»)
Приемлемая	Ни одно из указанных условий не выполнено

7.8. Согласование уровней. Прогнозы по территориям рассчитываются независимо от республиканского. Затем для каждой недели вычисляется коэффициент «республиканское среднее / сумма средних по территориям», и на него умножаются медианы, средние и границы интервалов территорий. Расхождение суммы территорий с республиканским прогнозом до согласования сохраняется; если оно превышает 25%, выводится предупреждение. Границы региональных интервалов не суммируются: республиканский интервал берётся из республиканской модели.

7.9. Анализ чувствительности. Если в ряду есть аномальные недели (пункт 6.1), республиканский прогноз дополнительно рассчитывается по ряду, в котором эти недели заменены медианой предыдущих 52 недель. Сравнение с основным прогнозом показывает, насколько прогноз зависит от этих недель.

7.10. Регрессоры: 3 пары гармоник Фурье (годовая сезонность) и праздничные недели 1, 18 и 19, в которые систематически снижается регистрация. Праздничные недели учитываются только моделями ARIMA и NNAR.

7.11. Значения параметров заданы в разделе «1. Параметры» скрипта (Приложение 2) и при расчёте не меняются:

Параметр в скрипте	**Значение**	**Назначение**
H_FORECAST	8	Горизонт прогноза, недель
H_EVAL	8	Горизонт оценки при кросс-валидации, недель

Параметр в скрипте	Значение	Назначение
CV_ORIG_NAT / CV_ORIG_REG	52 / 26	Число отсечений кросс-валидации: республика / территории
FOURIER_K	3	Пар гармоник Фурье
HOLIDAY_WEEKS	1, 18, 19	Праздничные недели
MIN_IMPROVEMENT	0,00	Сложная модель должна быть не хуже SNaive по MAE
MIN_COVER_RAW	85	Минимальное исходное покрытие 95% интервала при выборе модели, %
K_MIN / K_SAFETY_CAP	0,6 / 4,0	Нижняя граница и технический предел коэффициента калибровки
REL_WAPE_ACCEPT / LOW	25 / 40	Пороги WAPE для оценки надёжности, %
REL_MASE_ACCEPT / LOW	1,00 / 1,25	Пороги MASE для оценки надёжности
REL_COV_ACCEPT / LOW	90 / 85	Пороги калиброванного покрытия для оценки надёжности, %
POP_TOL	0,002	Порог смены режима знаменателя (0,2%)
CHECK_TOL	0,005	Допустимое расхождение при сверке файлов (0,5%)
CV_NPATHS	100	Число траекторий NNAR при кросс-валидации (в итоговом прогнозе — 1000)
SEED	2026	Начальное значение генератора случайных чисел: повторный расчёт на тех же данных даёт тот же результат

Глава 8. Порядок запуска скрипта

8.1. Скрипт (Приложение 2) запускается целиком. Он последовательно выполняет 10 шагов (ШАГ 1 — ШАГ 10); что делает каждый шаг и какие сообщения выводит в консоль, описано в Приложении 1.

8.2. Сохранение результатов предыдущего расчёта. Скрипт при каждом запуске перезаписывает файлы в папке output_v4_2/, а результаты предыдущего расчёта нужны для проверки прошлого прогноза (глава 12). Поэтому перед каждым новым запуском выполняется следующее:

– папка output_v4_2 копируется в папку archive/ (в Проводнике: правой кнопкой по папке → Копировать; в папке archive правой кнопкой по пустому месту → Вставить);
– копия переименовывается по номеру и году последней недели данных этого расчёта, латинскими буквами: w39_2026 — расчёт по данным до 39-й недели 2026 г.;
– проверяется, что в копии есть все файлы результатов (глава 9), в том числе 10_прогноз_национальный.csv и 13_прогноз_региональный.csv;
– после проверки папка output_v4_2 удаляется целиком. Скрипт при запуске создаёт её заново, и в ней остаются только результаты нового расчёта. Это необходимо, так как файл 06_аномальные_недели.csv создаётся не при каждом расчёте: без удаления в папке может остаться файл от прошлого расчёта, и отчёт будет сформирован с устаревшими сведениями.

8.3. Перед запуском проверяется, что:

– в оба файла данных внесены данные за последнюю завершившуюся неделю по всем 9 территориям и сверены с исходной отчётностью (глава 5);
– имена файлов данных и их расположение в папке data/ соответствуют пункту 5.1;
– результаты предыдущего расчёта скопированы в папку archive/, копия проверена, папка output_v4_2 удалена (пункт 8.2);
– проект открыт через файл forecast_flu_kg.Rproj (пункт 4.7);
– компьютер подключён к сети питания, спящий режим отключён: расчёт длится около 7 минут (время тестового запуска; на разных компьютерах может отличаться).

8.4. Запуск. В открытом проекте открыть файл скрипта forecast_ari_kg_v4_2.R (панель Files → щелчок по файлу) и нажать кнопку Source в правом верхнем углу окна скрипта (или сочетание клавиш Ctrl+Shift+S). При запуске через Source выполнение сразу останавливается, если скрипт выдал сообщение об ошибке.

8.5. Во время расчёта в консоли (панель Console, слева внизу) последовательно появляются заголовки шагов в рамке из знаков «=»:

Шаг	**Содержание**	**Контрольное сообщение**
1	Загрузка файлов	Число наблюдений, территорий, годы; список территорий
2	Восстановление знаменателей населения	Число режимов знаменателя; опорные знаменатели; сумма по стране
3	Контрольная сверка файлов	«СВЕРКА ПРОЙДЕНА. Максимальное расхождение: …%»
4	Сборка рядов	«Национальный ряд: … недель, … — …»; «Последнее значение: … случаев»
5	Диагностика данных	Годовая динамика; аномальные недели; праздничные недели; межтерриториальные различия; темп прироста

Шаг	Содержание	Контрольное сообщение
6	Национальный прогноз	Метрики кросс-валидации; выбранная модель; «Национальный прогноз построен.»
7	Региональные прогнозы	По каждой территории: кросс-валидация и выбранная модель; карта региональных моделей. Самый длительный шаг
8	Согласование уровней	Расхождение суммы регионов с национальной моделью; ситуация на последнюю неделю; «Контроль сходимости региональной таблицы и национального ряда: ОК»
9	Графики	«Графики сохранены.»
10	Резюме	Текст резюме прогноза; «ГОТОВО»; «Все результаты в каталоге: …»

8.6. Контроль данных в начале расчёта. После Шага 4 в консоли появляются сообщения «Национальный ряд: … недель, ДД.ММ.ГГГГ — ДД.ММ.ГГГГ» и «Последнее значение: … случаев (… на 100 тыс.)». Дата окончания ряда должна соответствовать воскресенью последней внесённой недели, а число случаев — республиканскому итогу этой недели по исходной отчётности. Если они не совпадают, расчёт останавливается кнопкой Stop (красный квадрат над консолью), данные проверяются и скрипт запускается заново.

8.7. Во время расчёта компьютер не выключается и RStudio не закрывается. Расчёт завершён, когда в консоли появились сообщения «ГОТОВО» и «Все результаты в каталоге: …».

8.8. Предупреждения. Скрипт выводит предупреждения (Warning) о ситуациях, которые не останавливают расчёт, но требуют внимания (глава 13). При запуске через Source они выводятся в конце; если в консоли есть строка «There were … warnings», их полный текст выводится командой warnings(), введённой в консоли. Текст предупреждений сохраняется и учитывается при подготовке отчёта.

8.9. Если расчёт остановился с сообщением об ошибке, причина определяется по главе 13; после исправления скрипт запускается заново целиком.

Глава 9. Результаты расчёта

9.1. Результаты сохраняются в папку output_v4_2/:

Файл	Содержание
00_параметры_расчёта.csv	Паспорт расчёта: рабочая модель КР, период кросс-валидации, коэффициенты калибровки, покрытие, горизонт, опорное население
01_знаменатели_населения.csv	Режимы знаменателя по территориям и годам
02_сверка_абс_и_интенс.csv	Результат контрольной сверки двух файлов
03_ряд_национальный.csv	Республиканский недельный ряд
04_ряд_региональный.csv	Недельные ряды по территориям
05_годовая_динамика.csv	Годовые суммы, медиана и максимум недельных значений
06_аномальные_недели.csv	Недели с превышением медианы предыдущих 52 недель более чем в 3 раза (если есть)
07_вариабельность_регистрации.csv	Заболеваемость по территориям за последний полный год
08_темп_прироста.csv	Недельный темп прироста, 95% ДИ, время удвоения (сокращения вдвое)
08b_диагностика_праздничных_недель.csv	Снижение регистрации в праздничные недели
09_метрики_национальные.csv	Метрики кросс-валидации моделей республиканского уровня
09b_метрики_по_горизонту.csv	Те же метрики отдельно по горизонтам 1–8 недель
10_прогноз_национальный.csv	Прогноз по КР на 8 недель: медиана, среднее, 80% и 95% интервалы, показатели на 100 тыс.
11_карта_моделей_региональных.csv	Модель, метрики и оценка надёжности по каждой территории
12_согласование_уровней.csv	Расхождение суммы регионов с республиканским прогнозом и коэффициенты согласования по неделям
12a_региональные_прогнозы_до_согласования.csv	Прогнозы территорий до согласования

Файл	Содержание
12b_… и 13_прогноз_региональный.csv	Прогнозы территорий после согласования (одинаковое содержание): медиана, среднее, интервалы, показатели на 100 тыс., доля в КР
14_ситуация_текущая.csv	Ситуация на последнюю неделю по территориям: случаи, доля, показатели на 100 тыс., динамика к предыдущей неделе
sensitivity_2025.csv	Результат анализа чувствительности
g1_прогноз_национальный.png	График: наблюдаемый ряд за 3 года и прогноз КР с интервалами
g2_регионы_интенсивные.png	Гармонизированные интенсивные показатели по территориям за 2 года
g3_эффект_знаменателя.png	Разность официального и гармонизированного показателей
резюме_прогноза.txt	Текстовое резюме: исходные данные, текущая ситуация, прогноз, региональные модели, ограничения

9.2. Файлы 06_структурные_сдвиги.csv и 07_полнота_регистрации.csv повторяют файлы 06_аномальные_недели.csv и 07_вариабельность_регистрации.csv и сохраняются скриптом для совместимости; в отчёте не используются.

9.3. Таблицы сохраняются в формате CSV с кодировкой, которая корректно открывается в Excel.

Глава 10. Отчёт и интерпретация результатов

10.1. По результатам каждого расчёта составляется аналитический отчёт «Заболеваемость ОРВИ и гриппом в Кыргызской Республике. Аналитический отчёт по краткосрочному прогнозированию на 8 недель». Структура отчёта и файлы, из которых берутся данные для каждого раздела, приведены в Приложении 3.

10.2. Отчёт формируется по шаблону report_ari_ru.Rmd (текст — Приложение 4). Шаблон — это файл R Markdown: в нём текст отчёта объединён с командами R. При формировании отчёта команды читают файлы результатов из папки output_v4_2/ (глава 9), подставляют числа в текст, строят таблицы и вставляют рисунок. Результат — файл Word, оформленный по образцу report_style.docx. Вручную дописываются только выводы эксперта.

10.3. Порядок формирования отчёта (после завершения расчёта по главе 8):

- открыть в проекте файл report_ari_ru.Rmd (панель Files → щелчок по файлу);
- в шапке файла, в строке prev_dir: "", вписать между кавычками путь к копии предыдущего расчёта в папке archive/ (пункт 8.2), например prev_dir:

"archive/w39_2026", и сохранить файл (Ctrl+S). По этой папке шаблон формирует раздел 3 отчёта — проверку предыдущего прогноза по КР и территориям (глава 12). Если строка оставлена пустой, раздел 3 не формируется;

- закрыть в Word отчёт, сформированный ранее, если он открыт, — иначе RStudio не сможет перезаписать файл;
- нажать кнопку Knit над окном файла. Через 1–2 минуты в папке проекта появится файл report_ari_ru.docx, и он откроется в Word;
- сохранить отчёт под именем с номером недели и годом, например Прогноз_грипп_ОРВИ_нед39_2026.docx: при следующем формировании файл report_ari_ru.docx перезаписывается.

10.4. Аналитические выводы. Шаблон формирует их автоматически по результатам расчёта: накопленную проверку прогноза (раздел 3.2) — по журналу проверок archive/check_log.csv, который шаблон сам пополняет при каждом формировании отчёта; оценку начала и возможного пика сезонного подъёма (7.1), риски (7.2), предложения (7.3) и заключение (10) — по правилам, заложенным в шаблоне (Приложение 4). Эксперт просматривает автоматические выводы и при необходимости дополняет или корректирует их в сохранённом файле отчёта, например сведениями лабораторного надзора, данными о госпитализации или решениями МЗ КР.

10.5. Проверка отчёта перед отправкой. Аналитик проверяет, что номер и год недели на титульном листе соответствуют последней неделе данных, таблицы и рисунок на месте, а выводы эксперта соответствуют числам в таблицах.

10.6. Если отчёт не формируется: сообщение о том, что файл в папке output_v4_2/ не найден, означает, что расчёт не выполнен или проект открыт не через forecast_flu_kg.Rproj; сообщение «permission denied» — что файл отчёта открыт в Word; сообщение об отсутствии пакета — что не выполнена установка пакетов (пункт 4.3).

10.7. Интерпретация прогноза по срокам. Точность прогноза снижается с удалением от последней недели данных (файл 09b_метрики_по_горизонту.csv):

- на 1–2 недели расчётные значения используются для текущей оценки ожидаемого числа зарегистрированных случаев;
- на 3–4 недели учитываются не только медиана, но и границы прогнозного интервала и фактические данные, поступающие каждую неделю;
- на 5–8 недель прогноз показывает прежде всего направление изменения заболеваемости — рост, снижение или сохранение уровня; конкретные недельные значения могут существенно измениться после поступления новых данных.

10.8. При интерпретации учитывается следующее:

- медиана используется для оценки каждой отдельной недели, среднее — для оценки общего ожидаемого числа случаев за несколько недель;
- максимальное значение внутри 8-недельного горизонта не является оценкой сезонного пика, так как период после горизонта расчётом не охвачен;
- прогноз числа зарегистрированных случаев не является прямым прогнозом госпитализаций, потребности в койках или других ресурсов;
- различия интенсивных показателей между территориями не трактуются автоматически как различия риска заражения: они могут отражать обращаемость, доступность медицинской помощи и практику регистрации;

- для территорий с ограниченной или низкой надёжностью прогноза особенно важно еженедельное сопоставление прогноза с фактическими данными;
- если рабочей моделью выбрана SNaive, это указывается в отчёте: прогноз в этом случае является продолжением прошлогодней сезонной кривой;
- прогноз интерпретируется совместно с данными эпидемиологического и лабораторного надзора.

10.9. Отчёт направляется в Министерство здравоохранения КР (УОЗ) по запросу. Копия отчёта хранится в папке archive/ вместе с результатами расчёта, по которым он составлен (пункт 8.2).

Глава 11. Ограничения модели

11.1. Ограничения, которые указываются в каждом отчёте:

- ОРВИ и грипп моделируются как единый массив зарегистрированных случаев; вирусологическая структура и штаммовый состав не разделяются.
- Прогнозируется число обратившихся и зарегистрированных в организациях здравоохранения, а не истинная заболеваемость населения.
- Зарегистрированная заболеваемость существенно различается между территориями; вклад эпидемиологических различий, обращаемости, доступности помощи и практики регистрации по этим данным не определяется.
- Официальный знаменатель населения обновляется внутри года, что создаёт «ступеньки» в официальном интенсивном показателе; для анализа динамики используется единый знаменатель.
- Для года, отсутствующего в файле интенсивных показателей, знаменатель принят равным последнему официальному — это допущение.
- Глубина ряда — около трёх сезонов, что является нижней границей для сезонных моделей с годовым циклом.
- Аномально высокие недели основной моделью не удаляются и могут влиять на сезонную компоненту; их влияние показывает анализ чувствительности.
- В модель не включены вирусологические данные, метеорологические факторы и будущие управленческие воздействия.
- Интервалы быстро расширяются с увеличением срока; поздние недельные значения нельзя рассматривать как точные числа случаев.
- Модели используют сезонный период 52 недели; при наличии 53-й ISO-недели скрипт выводит предупреждение.

Глава 12. Оценка качества прогноза

12.1. Качество прогноза оценивается двумя способами: ретроспективно — скользящей кросс-валидацией, которую скрипт выполняет при каждом расчёте (глава 7), и проспективно — сравнением ранее выпущенного прогноза с поступившими фактическими данными.

12.2. Проспективная проверка проводится при каждом расчёте. Прогноз предыдущего расчёта (файлы 10_прогноз_национальный.csv и 13_прогноз_региональный.csv из папки archive/) сравнивается с фактическими данными за недели, которые поступили после его выпуска. Для КР и для каждой территории фиксируются:

- по каждой неделе: прогноз (медиана), факт, разность «прогноз − факт», абсолютная ошибка в процентах;

- за проверяемые недели: сумма прогноза и факта, суммарное завышение или занижение, MAPE, WAPE;
- число фактических значений, попавших в 95% прогнозный интервал;
- верно ли определено направление изменения заболеваемости.

12.3. Результаты проспективных проверок накапливаются: в каждом отчёте приводятся результаты текущей проверки и итог всех предыдущих проверок. Результаты отдельных проверок не считаются устойчивой характеристикой точности модели до накопления достаточного числа проверенных недель.

12.4. Контроль согласованности уровней. При каждом расчёте фиксируется расхождение суммы региональных прогнозов с республиканским до согласования (файл 12_согласование_уровней.csv) по каждой неделе и территории, вносящие наибольший вклад в расхождение.

12.5. Если повышенное расхождение уровней, систематическое завышение или занижение прогноза либо выход фактических значений за 95% интервал повторяются в нескольких последовательных расчётах, проводится дополнительная оценка региональных моделей, стабильности выбора модели и процедуры согласования. Решение о пересмотре алгоритма принимает руководитель ЦАУРОЗиПЗ. По результатам одного расчёта алгоритм не изменяется.

12.6. Проспективную проверку выполняет рководитель ЦАУРОЗиПЗ. Результаты проверок хранятся в ЦАУРОЗиПЗ.

Глава 13. Техническая проверка расчёта и частые ошибки

13.1. После завершения расчёта проверяется, что в папке output_v4_2/ есть файлы, перечисленные в главе 9, а в резюме (резюме_прогноза.txt) указана дата последней недели данных, соответствующая внесённым данным.

13.2. Сообщения, останавливающие расчёт:

Сообщение	**Что делать**
«Файл не найден: data/…»	Проверить, что проект открыт через forecast_flu_kg.Rproj, файлы лежат в папке data/ и их имена совпадают с пунктом 5.1
«Не найдена строка заголовка с годами в …»	Проверить, что данные на первом листе и заголовки годов имеют вид «2023г»
«Файлы не пересекаются по территориям/неделям»	Проверить, что названия территорий и номера недель в двух файлах совпадают
«СВЕРКА НЕ ПРОЙДЕНА: расхождение …%»	Сверить значения последних внесённых недель в двух файлах: перепутаны столбцы, сдвинуты строки или изменена структура листа. Подробности — в файле 02_сверка_абс_и_интенс.csv
«Нет знаменателя для: …»	Для указанной территории и года нет значений в файле интенсивных показателей — заполнить

Сообщение	Что делать
«Разрыв недельной сетки после …»	В республиканском ряду пропущена неделя (в том числе 53-я неделя в году, где она есть) или неделя заполнена не по всем территориям — заполнить
«Недостаточно данных для КВ ряда …»	Ряд короче, чем требует кросс-валидация, — проверить, что в файлы внесены все годы
«Финальная … модель не посчиталась» / «Ни один регион не был посчитан»	Записать точный текст сообщения и обратиться к руководству ДПЗиГСЭН.
«Сумма региональной таблицы … не совпадает с национальным рядом»	Проверить заполнение последней недели по всем территориям

13.3. Предупреждения, не останавливающие расчёт:

Предупреждение	Что делать
«Списки территорий в файлах не совпадают»	Привести названия территорий в двух файлах к одинаковому написанию и запустить заново
«Внутри режима знаменатель нестабилен»	Проверить файл интенсивных показателей
«ВНИМАНИЕ: … недель отчитались не все территории — исключены»	Дозаполнить данные; если это последняя неделя, расчёт фактически выполнен по предыдущей неделе
«Данные заканчиваются … — это … дней назад»	Данные устарели более чем на 21 день — внести данные за последние недели
«Ряд … заканчивается …, а национальный — …» / «Регион пропущен»	По территории не заполнена последняя неделя или есть пропуск — дозаполнить
«Калибровка … достигла технического safety cap»	Интервалы могут оставаться недостаточно широкими — указать в отчёте
«Уровни расходятся до …%»	Расхождение республиканского прогноза и суммы регионов больше 25% — отразить в отчёте и в контроле согласованности (пункт 12.4)
«В ряду присутствует 53-я ISO-неделя»	Учитывается как техническое предупреждение (глава 11)
«История короче трёх сезонов»	Сезонная компонента неустойчива — указать в отчёте

13.4. Если ни одна ситуация не подошла, аналитик записывает точный текст сообщения и номер шага и обращается к руководителю ЦАУРОЗиПЗ.

Глава 14. Утверждение и пересмотр

14.1. Инструкция утверждается приказом НИОЗ.

14.2. Изменение скрипта, параметров или структуры отчёта оформляется. Новой версии скрипта присваивается новый номер; имя папки результатов задаётся в скрипте (OUT_DIR).

14.3. Инструкция пересматривается в указанный срок.

Приложение 1. Описание шагов скрипта

Скрипт forecast_ari_kg_v4_2.R состоит из подготовительной части (раздел «0. Пакеты» — подключение пакетов; раздел «1. Параметры» — значения параметров, глава 7) и 10 шагов, заголовки которых выводятся в консоль. Ниже для каждого шага указано, что он делает, какие сообщения выводит и какие файлы сохраняет.

Шаг	Что делает	Сообщения и файлы
Подготовка: пакеты и параметры	Подключаются пакеты readxl, dplyr, tidyr, purrr, stringr, readr, tibble, ggplot2, forecast; недостающие устанавливаются автоматически. Задаются параметры расчёта и создаётся папка output_v4_2.	—
Шаг 1. Загрузка файлов	Читает первый лист каждого файла данных, находит строку заголовка с годами и переводит данные в таблицу «территория — год — неделя — значение». Проверяет, что списки территорий в двух файлах совпадают.	В консоли: имя файла и листа; «Абсолютные числа : … наблюдений, 9 территорий, годы …»; «Интенсивные пок-ли: …»; «Территории: …».
Шаг 2. Восстановление знаменателей населения	Восстанавливает численность населения по каждой территории и неделе (абсолютное число / интенсивный показатель × 100 000), выделяет режимы знаменателя и определяет опорный знаменатель каждой территории. Для года, которого нет в файле интенсивных показателей, принимает последний официальный знаменатель.	В консоли: «Обнаружено режимов знаменателя: …»; таблица режимов; опорные знаменатели; «Сумма по стране: …». Файл 01.
Шаг 3. Контрольная сверка файлов	Пересчитывает интенсивный показатель из абсолютных чисел и восстановленных знаменателей и сравнивает с файлом интенсивных показателей. При расхождении более 0,5% расчёт останавливается.	В консоли: таблица максимальных расхождений по территориям; «СВЕРКА ПРОЙДЕНА. Максимальное расхождение: …%». Файл 02.

Шаг	Что делает	Сообщения и файлы
Шаг 4. Сборка рядов	Собирает недельные ряды по территориям и республиканский ряд (только недели, за которые отчитались все территории), рассчитывает официальный и гармонизированный интенсивные показатели. Проверяет, что в ряду нет пропущенных недель (иначе расчёт останавливается), есть ли 53-я неделя, достаточна ли длина ряда и не устарели ли данные.	В консоли: «Национальный ряд: … недель, … — …»; «Последнее значение: … случаев (… на 100 тыс.)» — контроль по пункту 8.6. Файлы 03, 04.
Шаг 5. Диагностика данных	Рассчитывает годовую динамику; выявляет аномальные недели (более чем в 3 раза выше медианы предыдущих 52 недель); оценивает снижение регистрации в праздничные недели; сравнивает заболеваемость территорий за последний полный год; рассчитывает недельный темп прироста за последние 6 недель с 95% ДИ и время удвоения (сокращения вдвое).	В консоли: таблица годовой динамики; аномальные недели (если есть); праздничные недели; межтерриториальная вариабельность; «Темп прироста за последние 6 нед.: …% (95% ДИ …)». Файлы 05–08b.
Шаг 6. Национальный прогноз	Выполняет кросс-валидацию 5 моделей и ансамбля на республиканском ряду, рассчитывает метрики и коэффициенты калибровки, выбирает рабочую модель (пункт 7.4), строит прогноз на 8 недель (медиана, среднее, 80% и 95% интервалы, показатели на 100 тыс.) и паспорт расчёта.	В консоли: «КВ 'КР, всего': … отсечений»; таблица метрик; «'КР, всего': выбрана … (MAE … против … у SNaive; raw coverage 95% = …%)»; «Калибровка: raw k=…»; «Национальный прогноз построен.»; медианный сценарий и ожидаемое число за 8 недель. Файлы 00, 09, 09b, 10.
Шаг 7. Региональные прогнозы	Для каждой территории выполняет кросс-валидацию 3 моделей и ансамбля, выбирает рабочую модель, строит прогноз и оценивает надёжность. Территория пропускается с предупреждением, если её ряд заканчивается раньше республиканского или имеет пропуски. Это самый длительный шаг.	В консоли по каждой территории: «--- Территория ---», число отсечений, выбранная модель; в конце — «Карта региональных моделей». Файлы 11, 12a.

Шаг	Что делает	Сообщения и файлы
Шаг 8. Согласование уровней	Сравнивает сумму средних прогнозов территорий с республиканским прогнозом, согласует прогнозы территорий с республиканским, рассчитывает долю каждой территории; выполняет анализ чувствительности (если есть аномальные недели); формирует таблицу ситуации на последнюю неделю и проверяет, что сумма территорий равна республиканскому значению.	В консоли: таблица расхождений по неделям; «Максимальное расхождение уровней: …%» или предупреждение; «Sensitivity analysis 2025: …»; «Ситуация на …»; «Контроль сходимости региональной таблицы и национального ряда: ОК». Файлы 12, 12b, 13, 14, sensitivity_2025.
Шаг 9. Графики	Строит и сохраняет три графика.	В консоли: «Графики сохранены.» Файлы g1, g2, g3.
Шаг 10. Резюме	Формирует текстовое резюме расчёта: исходные данные, текущая ситуация, прогноз, региональные модели, ограничения.	В консоли: текст резюме; «ГОТОВО»; «Все результаты в каталоге: …». Файл резюме_прогноза.txt.

> Номера разделов в тексте скрипта (например, «# ---- 9. УРОВЕНЬ А: национальный прогноз») не совпадают с номерами шагов, которые выводятся в консоль («ШАГ 6. НАЦИОНАЛЬНЫЙ ПРОГНОЗ»). В Инструкции используются номера шагов, выводимые в консоль.

Приложение 2. Текст скрипта forecast_ari_kg_v4_2.R

Текст скрипта приводится без изменений.

```
# ============================================================================
#  ПРОГНОЗ ЗАБОЛЕВАЕМОСТИ ОРВИ И ГРИППОМ ПО КЫРГЫЗСКОЙ РЕСПУБЛИКЕ
#  Версия v4.2: абсолютные числа + интенсивные показатели; усиленная валидация
#  Точечные правки после независимого аудита v3 vs v4.1:
#  1) NNAR: одна обученная сеть -> два forecast() (median/mean);
#  2) выбор модели снова использует RAW coverage как независимый CV-фильтр;
#     calibrated coverage остаётся диагностикой и используется для финальных PI.
#
#  Республиканский центр иммунопрофилактики
#
#  ПРИНЦИП РАБОТЫ
#  --------------
#  Файл абсолютных чисел  -> первичные данные для моделирования.
#  Файл интенсивных пок-й -> источник знаменателей, восстановленных из ОФИЦИАЛЬНЫХ
   интенсивных показателей
#                            и эталон для контрольной сверки.
#
#  Знаменатели не задаются вручную, а восстанавливаются делением
#  (абсолюты / интенсивный показатель * 100000) и проверяются на
#  устойчивость. Обнаруженные режимы знаменателя фиксируются в таблице.
#
#  Моделируются СЧЁТНЫЕ ДАННЫЕ. Интенсивные показатели рассчитываются
#  из прогноза для представления результата в двух формах.
#
#  ВХОД : data/ОРВИ_и_грипп_по_КР.xlsx          (лист с абс. числами)
#         data/Интен_показатель_ОРВИ_и_гриппа.xlsx (лист с показателями)
#  ВЫХОД: output_v4_2/*.csv, output_v4_2/*.png, output_v4_2/резюме_прогноза.txt
# ============================================================================



# ---- 0. Пакеты -------------------------------------------------------------
if (!requireNamespace("pacman", quietly = TRUE)) install.packages("pacman")
pacman::p_load(readxl, dplyr, tidyr, purrr, stringr, readr, tibble, ggplot2, forecast)



# ---- 1. Параметры ----------------------------------------------------------
XLSX_ABS <- "data/ОРВИ_и_грипп_по_КР.xlsx"
XLSX_INT <- "data/Интен_показатель_ОРВИ_и_гриппа.xlsx"
OUT_DIR  <- "output_v4_2"

H_FORECAST <- 8      # горизонт прогноза, недель
H_EVAL     <- 8      # горизонт оценки при кросс-валидации (= горизонту прогноза)
CV_ORIG_NAT <- 52    # число отсечений скользящей КВ, национальный уровень
CV_ORIG_REG <- 26    # то же для регионов (компромисс по времени счёта)
FOURIER_K  <- 3      # пар гармоник Фурье
BIASADJ    <- FALSE  # основной точечный прогноз = медиана на преобразованной шкале
CV_NPATHS  <- 100    # траекторий NNAR при КВ (в финале 1000)
SEED       <- 2026
MIN_IMPROVEMENT <- 0.00  # сложная модель должна быть не хуже SNaive по MAE;
                         # 0.02 означает требование улучшения минимум на 2%
MIN_COVER_RAW <- 85      # независимый CV-фильтр качества исходных 95% интервалов, %
MIN_COVER_CAL <- 90      # только диагностический ориентир после калибровки; в pick_model не
   используется
K_MIN      <- 0.6        # нижняя граница эмпирической калибровки
K_SAFETY_CAP <- 4.0      # только технический предохранитель от экстремальных значений;
                         # достижение лимита всегда маркируется предупреждением
REL_WAPE_ACCEPT <- 25    # операционные пороги территориальной надёжности, %
REL_WAPE_LOW    <- 40
```

```
REL_MASE_ACCEPT <- 1.00
REL_MASE_LOW    <- 1.25
REL_COV_ACCEPT  <- 90
REL_COV_LOW     <- 85
HOLIDAY_WEEKS <- c(1L, 18L, 19L)  # недели с систематическим провалом отчётности
POP_TOL    <- 0.002  # порог смены режима знаменателя (0.2%)
CHECK_TOL  <- 0.005  # допустимое расхождение при обратной сверке файлов (0.5%)

MODELS_NAT <- c("snaive", "stl_ets", "arima_fourier", "tbats", "nnar_fourier")
MODELS_REG <- c("snaive", "stl_ets", "arima_fourier")

dir.create(OUT_DIR, showWarnings = FALSE, recursive = TRUE)
set.seed(SEED)

log_msg <- function(...) message(format(Sys.time(), "[%H:%M:%S] "), ...)
hr <- function(t) message("\n", strrep("=", 78), "\n  ", t, "\n", strrep("=", 78))



# ---- 2. Универсальный парсер широкого формата ------------------------------
# Формат обоих файлов: блоки по территориям, внутри блока колонка номеров
# недель, затем по одной колонке на год ("2023г", "2024г", ...).

to_num <- function(x) {
  x <- as.character(x)
  x <- str_replace_all(x, "[^0-9,.-]", "")
  x <- str_replace_all(x, ",", ".")
  suppressWarnings(as.numeric(x))
}

parse_wide <- function(path, value_col) {
  if (!file.exists(path)) stop("Файл не найден: ", path, call. = FALSE)

  sheets <- readxl::excel_sheets(path)
  sheet  <- sheets[1]
  log_msg("  ", basename(path), " -> лист '", sheet, "'")

  raw <- readxl::read_excel(path, sheet = sheet, col_names = FALSE, col_types = "text")
  raw <- as.data.frame(raw, stringsAsFactors = FALSE)

  # строка заголовка = первая строка, где минимум два года вида "2023г"
  year_hits <- apply(raw, 1, function(r) {
    v <- as.character(r); v[is.na(v)] <- ""
    sum(str_detect(str_trim(v), "^\\d{4}"))
  })
  hdr_row <- which(year_hits >= 2)[1]
  if (is.na(hdr_row)) stop("Не найдена строка заголовка с годами в ", path, call. = FALSE)

  hdr <- as.character(unlist(raw[hdr_row, ], use.names = FALSE))
  hdr[is.na(hdr)] <- ""
  hdr <- str_trim(hdr)
  is_year <- str_detect(hdr, "^\\d{4}")
  is_terr <- nzchar(hdr) & !is_year

  body <- raw[(hdr_row + 1):nrow(raw), , drop = FALSE]

  out <- map_dfr(which(is_terr), function(cc) {
    reg <- hdr[cc]
    ycols <- integer(0)
    j <- cc + 1L
    while (j <= length(hdr) && is_year[j]) { ycols <- c(ycols, j); j <- j + 1L }
    if (!length(ycols)) return(NULL)
```

```
    weeks <- to_num(body[[cc]])
    map_dfr(ycols, function(kk) {
      tibble(
        territory = reg,
        year      = as.integer(str_sub(hdr[kk], 1, 4)),
        week      = as.integer(weeks),
        value     = to_num(body[[kk]])
      )
    })
  })

  out <- out |>
    filter(!is.na(week), week >= 1, week <= 53, !is.na(value)) |>
    arrange(territory, year, week)

  names(out)[names(out) == "value"] <- value_col
  out
}

hr("ШАГ 1. ЗАГРУЗКА ФАЙЛОВ")
abs_long <- parse_wide(XLSX_ABS, "cases")
int_long <- parse_wide(XLSX_INT, "rate_file")

log_msg("Абсолютные числа : ", nrow(abs_long), " наблюдений, ",
        n_distinct(abs_long$territory), " территорий, годы ",
        paste(range(abs_long$year), collapse = "-"))
log_msg("Интенсивные пок-ли: ", nrow(int_long), " наблюдений, ",
        n_distinct(int_long$territory), " территорий, годы ",
        paste(range(int_long$year), collapse = "-"))

terr_abs <- sort(unique(abs_long$territory))
terr_int <- sort(unique(int_long$territory))
if (!setequal(terr_abs, terr_int)) {
  warning("Списки территорий в файлах не совпадают.\n  Только в абс.: ",
          paste(setdiff(terr_abs, terr_int), collapse = ", "),
          "\n  Только в инт.: ", paste(setdiff(terr_int, terr_abs), collapse = ", "))
}
TERRITORIES <- terr_abs
log_msg("Территории: ", paste(TERRITORIES, collapse = ", "))



# ---- 3. Восстановление официальных знаменателей населения -------------------
# Знаменатель обновляется внутри года (как правило с 25-й недели), поэтому
# определяем РЕЖИМЫ: непрерывные отрезки недель с одним знаменателем.

hr("ШАГ 2. ВОССТАНОВЛЕНИЕ ЗНАМЕНАТЕЛЕЙ НАСЕЛЕНИЯ")

pop_raw <- abs_long |>
  inner_join(int_long, by = c("territory", "year", "week")) |>
  filter(rate_file > 0) |>
  mutate(pop_implied = cases / rate_file * 1e5) |>
  arrange(territory, year, week)

if (nrow(pop_raw) == 0) stop("Файлы не пересекаются по территориям/неделям.", call. = FALSE)

pop_regimes <- pop_raw |>
  group_by(territory, year) |>
  arrange(week, .by_group = TRUE) |>
  mutate(
    jump   = c(0, as.integer(abs(diff(pop_implied)) / head(pop_implied, -1) > POP_TOL)),
    regime = cumsum(jump)
```

```
  ) |>
  group_by(territory, year, regime) |>
  summarise(
    week_from = min(week),
    week_to   = max(week),
    population = round(median(pop_implied)),
    n_weeks   = n(),
    cv_pct    = round(sd(pop_implied) / mean(pop_implied) * 100, 4),
    .groups = "drop"
  ) |>
  arrange(territory, year, week_from)

log_msg("Обнаружено режимов знаменателя: ", nrow(pop_regimes))
log_msg("Максимальный разброс внутри режима: ",
        formatC(max(pop_regimes$cv_pct, na.rm = TRUE), format = "f", digits = 4), "%")
if (max(pop_regimes$cv_pct, na.rm = TRUE) > 0.5) {
  warning("Внутри режима знаменатель нестабилен (> 0.5%). ",
          "Возможно, интенсивные показатели рассчитаны не по постоянному населению.")
}

print(as.data.frame(pop_regimes), row.names = FALSE)
write_excel_csv(pop_regimes, file.path(OUT_DIR, "01_знаменатели_населения.csv"))

# Развернуть режимы в таблицу «территория-год-неделя -> население»
pop_by_week <- pop_regimes |>
  rowwise() |>
  mutate(week = list(seq.int(week_from, week_to))) |>
  ungroup() |>
  select(territory, year, week, population) |>
  unnest(week)

# Знаменатель для лет, отсутствующих в интенсивном файле (например 2026):
# берём последний официальный режим территории и держим постоянным.
years_all <- sort(unique(abs_long$year))
years_int <- sort(unique(int_long$year))
years_ext <- setdiff(years_all, years_int)

pop_last <- pop_regimes |>
  group_by(territory) |>
  slice_max(order_by = year * 100 + week_from, n = 1, with_ties = FALSE) |>
  ungroup() |>
  select(territory, population_last = population, ref_year = year, ref_week = week_from)

if (length(years_ext)) {
  log_msg("Годы без интенсивного показателя: ", paste(years_ext, collapse = ", "),
          " -> знаменатель берётся последний официальный (допущение)")
  pop_ext <- crossing(pop_last |> select(territory, population = population_last),
                      year = years_ext, week = 1:53)
  pop_by_week <- bind_rows(pop_by_week, pop_ext |> select(territory, year, week,
   population))
}

# Опорный (гармонизированный) знаменатель: один на всю историю территории.
pop_ref <- pop_last |> select(territory, pop_ref = population_last)

log_msg("Опорные знаменатели (для гармонизированного ряда):")
print(as.data.frame(pop_ref), row.names = FALSE)
log_msg("Сумма по стране: ", format(sum(pop_ref$pop_ref), big.mark = " "))


# ---- 4. КОНТРОЛЬНАЯ СВЕРКА ДВУХ ФАЙЛОВ (стоп-контроль) ----------------------
# Пересчитываем интенсивный показатель из абсолютов и восстановленных
```

```
# знаменателей и сравниваем с исходным файлом. Это защита от подмены колонок.

hr("ШАГ 3. КОНТРОЛЬНАЯ СВЕРКА ФАЙЛОВ")

check <- abs_long |>
  inner_join(int_long, by = c("territory", "year", "week")) |>
  inner_join(pop_by_week, by = c("territory", "year", "week")) |>
  mutate(rate_calc = cases / population * 1e5,
         rel_err   = ifelse(rate_file > 0, abs(rate_calc - rate_file) / rate_file,
   NA_real_))

check_summary <- check |>
  group_by(territory) |>
  summarise(n = n(),
            max_rel_err_pct = round(max(rel_err, na.rm = TRUE) * 100, 4),
            .groups = "drop") |>
  arrange(desc(max_rel_err_pct))

print(as.data.frame(check_summary), row.names = FALSE)
write_excel_csv(check, file.path(OUT_DIR, "02_сверка_абс_и_интенс.csv"))

max_err <- max(check$rel_err, na.rm = TRUE)
if (max_err > CHECK_TOL) {
  stop("СВЕРКА НЕ ПРОЙДЕНА: расхождение ", round(max_err * 100, 2),
       "% превышает допуск ", CHECK_TOL * 100, "%.\n",
       "  Вероятные причины: перепутаны колонки файлов, изменена структура листа,\n",
       "  либо интенсивные показатели рассчитаны по иному знаменателю.\n",
       "  Подробности: ", file.path(OUT_DIR, "02_сверка_абс_и_интенс.csv"), call. = FALSE)
}
log_msg("СВЕРКА ПРОЙДЕНА. Максимальное расхождение: ",
        formatC(max_err * 100, format = "f", digits = 4), "%")
log_msg("Интенсивный файл воспроизводится из абсолютного -> ",
        "первичными данными считаем абсолютные числа.")


# ---- 5. Сборка рабочих рядов ------------------------------------------------
hr("ШАГ 4. СБОРКА РЯДОВ")

iso_week_end <- function(year, week) {
  jan4 <- as.Date(sprintf("%d-01-04", year))
  monday_w1 <- jan4 - (as.integer(format(jan4, "%u")) - 1L)
  monday_w1 + 7L * (week - 1L) + 6L
}

reg_series <- abs_long |>
  left_join(pop_by_week, by = c("territory", "year", "week")) |>
  left_join(pop_ref,     by = "territory") |>
  mutate(
    week_end  = iso_week_end(year, week),
    rate_off  = cases / population * 1e5,   # официальный (со ступеньками знаменателя)
    rate_harm = cases / pop_ref    * 1e5    # гармонизированный (единый знаменатель)
  ) |>
  arrange(territory, week_end)

if (any(is.na(reg_series$population))) {
  bad <- reg_series |> filter(is.na(population)) |> distinct(territory, year)
  stop("Нет знаменателя для: ", paste(bad$territory, bad$year, collapse = "; "), call. =
   FALSE)
}

# Национальный ряд: только недели, где отчитались ВСЕ территории
nat_series <- reg_series |>
```

```
  group_by(year, week, week_end) |>
  summarise(n_terr = n(),
            cases  = sum(cases),
            pop    = sum(population),
            pop_ref_sum = sum(pop_ref),
            .groups = "drop") |>
  arrange(week_end)

incomplete <- nat_series |> filter(n_terr != length(TERRITORIES))
if (nrow(incomplete)) {
  log_msg("ВНИМАНИЕ: ", nrow(incomplete), " недель отчитались не все территории —
   исключены:")
  print(as.data.frame(incomplete |> select(year, week, n_terr)), row.names = FALSE)
}

nat_series <- nat_series |>
  filter(n_terr == length(TERRITORIES)) |>
  mutate(rate_off  = cases / pop * 1e5,
         rate_harm = cases / pop_ref_sum * 1e5)

# Жёсткая проверка непрерывности сетки: молчаливые дыры недопустимы
gaps <- which(as.integer(diff(nat_series$week_end)) != 7L)
if (length(gaps)) {
  stop("Разрыв недельной сетки после ",
       paste(format(nat_series$week_end[gaps], "%d.%m.%Y"), collapse = ", "),
       ". Модели с frequency=52 требуют сплошного ряда.", call. = FALSE)
}

n_nat     <- nrow(nat_series)
last_date <- max(nat_series$week_end)
START_YW  <- c(nat_series$year[1], nat_series$week[1])

# ISO-неделя 53: модели forecast::ts используют frequency=52.
# Для текущего короткого ряда это допустимое приближение, но факт наличия week=53
# обязательно фиксируется как техническое предупреждение для дальнейшей эксплуатации.
if (any(nat_series$week == 53L)) {
  warning("В ряду присутствует 53-я ISO-неделя. Модели используют frequency=52; ",
          "необходимо проверить сезонную фазу при долгосрочной эксплуатации алгоритма.")
}

log_msg("Национальный ряд: ", n_nat, " недель, ",
        format(min(nat_series$week_end), "%d.%m.%Y"), " — ",
        format(last_date, "%d.%m.%Y"))
log_msg("Последнее значение: ", format(nat_series$cases[n_nat], big.mark = " "), " случаев
   (",
        formatC(nat_series$rate_harm[n_nat], format = "f", digits = 1), " на 100 тыс.)")

if (n_nat < 3 * 52) {
  warning("История короче трёх сезонов (", n_nat, " нед.) — ",
          "сезонная компонента будет неустойчивой.")
}

# Предупреждение об устаревании данных
days_stale <- as.integer(Sys.Date() - last_date)
if (days_stale > 21) {
  warning("Данные заканчиваются ", format(last_date, "%d.%m.%Y"), " — это ",
          days_stale, " дней назад. Прогноз будет ретроспективным. ",
          "Обновите файл до продолжения.")
}

write_excel_csv(nat_series, file.path(OUT_DIR, "03_ряд_национальный.csv"))
write_excel_csv(reg_series |> select(territory, year, week, week_end, cases,
```

```
                                  population, pop_ref, rate_off, rate_harm),
                file.path(OUT_DIR, "04_ряд_региональный.csv"))


# ---- 6. Диагностика ------------------------------------------------------
hr("ШАГ 5. ДИАГНОСТИКА ДАННЫХ")

# 6.1 Годовая динамика
yearly <- nat_series |>
  group_by(year) |>
  summarise(недель = n(), сумма = sum(cases),
            медиана_нед = round(median(cases)), макс_нед = max(cases),
            нед_максимума = week[which.max(cases)], .groups = "drop")
log_msg("Годовая динамика (национальный уровень):")
print(as.data.frame(yearly), row.names = FALSE)
write_excel_csv(yearly, file.path(OUT_DIR, "05_годовая_динамика.csv"))

# 6.2 Детекция аномально высокого недельного уровня относительно медианы предыдущих 52
   недель
brk <- nat_series |>
  mutate(base = sapply(seq_len(n()), function(i) {
    if (i <= 52) NA_real_ else median(cases[(i - 52):(i - 1)])
  }),
  ratio = cases / base) |>
  filter(!is.na(ratio), ratio > 3) |>
  select(year, week, week_end, cases, base, ratio)

if (nrow(brk)) {
  log_msg("Недели с превышением годовой медианы более чем в 3 раза:")
  print(as.data.frame(brk |> mutate(ratio = round(ratio, 1), base = round(base))), row.names
   = FALSE)
  log_msg("-> Это диагностический флаг, а не статистический тест structural break.")
  write_excel_csv(brk, file.path(OUT_DIR, "06_аномальные_недели.csv"))
  # legacy-файл для совместимости старого Rmd
  write_excel_csv(brk, file.path(OUT_DIR, "06_структурные_сдвиги.csv"))
} else {
  log_msg("Аномально высоких недель (>3x к медиане предыдущих 52 недель) не обнаружено.")
}

# 6.3 Праздничные провалы отчётности
hol_check <- nat_series |>
  group_by(week) |>
  summarise(медиана = median(cases), .groups = "drop") |>
  mutate(к_соседям = медиана / ((lag(медиана) + lead(медиана)) / 2)) |>
  filter(week %in% HOLIDAY_WEEKS)
log_msg("Заданные праздничные недели (отношение к соседним):")
print(as.data.frame(hol_check |> mutate(медиана = round(медиана),
                                        к_соседям = round(к_соседям, 2))), row.names =
   FALSE)
hol_year <- nat_series |> arrange(year, week) |> group_by(year) |>
  mutate(neighbor_mean = (lag(cases) + lead(cases)) / 2, ratio_neighbors = cases /
   neighbor_mean) |>
  ungroup() |> filter(week %in% HOLIDAY_WEEKS)
hol_diag <- hol_year |> group_by(week) |> summarise(
  сезонов = n(),
  сезонов_со_снижением = sum(ratio_neighbors < 1, na.rm = TRUE),
  медиана_отношения_к_соседям = median(ratio_neighbors, na.rm = TRUE), .groups = "drop")
write_excel_csv(hol_diag, file.path(OUT_DIR, "08b_диагностика_праздничных_недель.csv"))

# 6.4 Межтерриториальная вариабельность зарегистрированной заболеваемости
ref_year_full <- if (length(unique(reg_series$year)) > 1) {
  max(reg_series$year[reg_series$year < max(reg_series$year)])
} else max(reg_series$year)
```

```
completeness <- reg_series |>
  filter(year == ref_year_full) |>
  group_by(territory) |>
  summarise(год = first(year), случаев = sum(cases),
            население = first(pop_ref),
            на_100тыс_в_год = round(sum(cases) / first(pop_ref) * 1e5), .groups = "drop") |>
  arrange(desc(на_100тыс_в_год)) |>
  mutate(кратность_к_минимуму = round(на_100тыс_в_год / min(на_100тыс_в_год), 1))
log_msg("Межтерриториальная вариабельность зарегистрированной заболеваемости (последний
   полный год):")
print(as.data.frame(completeness), row.names = FALSE)
log_msg("-> Различие в ", max(completeness$кратность_к_минимуму),
        " раз может отражать сочетание реальных эпидемиологических различий, обращаемости,
   ",
        "доступности помощи и практики регистрации; вклад факторов по этим данным не
   определяется.")
write_excel_csv(completeness, file.path(OUT_DIR, "07_вариабельность_регистрации.csv"))
# legacy-файл для совместимости старого Rmd
write_excel_csv(completeness, file.path(OUT_DIR, "07_полнота_регистрации.csv"))



# ---- 7. Текущий темп прироста ------------------------------------------
GROWTH_WEEKS <- 6
recent <- nat_series |> slice_tail(n = GROWTH_WEEKS) |> mutate(t = row_number())
gm     <- glm(cases ~ t, family = quasipoisson(link = "log"), data = recent)
r_wk   <- unname(coef(gm)["t"])
r_ci   <- suppressMessages(confint.default(gm)["t", ])

ci_zero <- (r_ci[1] <= 0 && r_ci[2] >= 0)
t_label <- if (ci_zero) "Время изменения, недель" else
  if (r_wk > 0) "Время удвоения, недель" else "Время сокращения вдвое, недель"
t_value <- if (ci_zero) "не определено (плато)" else
  as.character(round(log(if (r_wk > 0) 2 else 0.5) / r_wk, 1))

growth <- tibble(
  показатель = c("Недельный темп прироста, %", "Нижняя граница 95% ДИ, %",
                 "Верхняя граница 95% ДИ, %", t_label, "Период оценки"),
  значение = c(
    formatC((exp(r_wk) - 1) * 100, format = "f", digits = 1),
    formatC((exp(r_ci[1]) - 1) * 100, format = "f", digits = 1),
    formatC((exp(r_ci[2]) - 1) * 100, format = "f", digits = 1),
    t_value,
    paste0(format(min(recent$week_end), "%d.%m.%Y"), " — ", format(last_date, "%d.%m.%Y"))
  )
)
log_msg("Темп прироста за последние ", GROWTH_WEEKS, " нед.: ",
        growth$значение[1], "% (95% ДИ ", growth$значение[2], "…", growth$значение[3], "%)")
write_excel_csv(growth, file.path(OUT_DIR, "08_темп_прироста.csv"))



# ---- 8. Движок прогнозирования -----------------------------------------
# Регрессоры: гармоники Фурье (годовая сезонность) + праздничные недели.

mk_ts <- function(v, start_yw) ts(v, start = start_yw, frequency = 52)

hol_matrix <- function(weeks, keep = NULL) {
  m <- sapply(HOLIDAY_WEEKS, function(w) as.numeric(weeks == w))
  m <- matrix(as.numeric(m), nrow = length(weeks),
              dimnames = list(NULL, paste0("hol_w", HOLIDAY_WEEKS)))
  # столбцы задаются по ОБУЧАЮЩЕЙ выборке; в прогнозе те же столбцы,
  # даже если там все нули (горизонт не попал на праздничную неделю)
  if (is.null(keep)) keep <- colnames(m)[colSums(m) > 0]
```

```
  m[, keep, drop = FALSE]
}

# Номера недель для будущих дат прогноза
future_weeks <- function(dates) as.integer(format(dates, "%V"))

fit_forecast <- function(y_ts, h, model, wk_train, wk_future,
                         K = FOURIER_K, npaths = 1000, quiet = TRUE,
                         biasadj = BIASADJ) {
  lam <- if (min(y_ts, na.rm = TRUE) > 0) 0 else NULL   # лог-шкала только при y > 0
  res <- tryCatch({
    switch(
      model,
      "snaive" = snaive(y_ts, h = h, lambda = lam, biasadj = biasadj),

      "stl_ets" = stlf(y_ts, h = h, lambda = lam, biasadj = biasadj,
                       s.window = "periodic", method = "ets"),

      "arima_fourier" = {
        h_tr  <- hol_matrix(wk_train)
        h_fu  <- hol_matrix(wk_future, keep = colnames(h_tr))
        xr_tr <- cbind(fourier(y_ts, K = K), h_tr)
        xr_fu <- cbind(fourier(y_ts, K = K, h = h), h_fu)
        fit <- auto.arima(y_ts, xreg = xr_tr, seasonal = FALSE,
                          lambda = lam, biasadj = biasadj)
        forecast(fit, xreg = xr_fu, h = h, biasadj = biasadj)
      },

      "tbats" = forecast(tbats(y_ts, use.box.cox = TRUE), h = h, biasadj = biasadj),

      "nnar_fourier" = {
        h_tr  <- hol_matrix(wk_train)
        h_fu  <- hol_matrix(wk_future, keep = colnames(h_tr))
        xr_tr <- cbind(fourier(y_ts, K = K), h_tr)
        xr_fu <- cbind(fourier(y_ts, K = K, h = h), h_fu)
        fit <- nnetar(y_ts, xreg = xr_tr, lambda = lam)
        forecast(fit, xreg = xr_fu, h = h, PI = TRUE, npaths = npaths, biasadj = biasadj)
      },

      stop("Неизвестная модель: ", model)
    )
  }, error = function(e) {
    if (!quiet) log_msg("    сбой модели '", model, "': ", conditionMessage(e))
    attr(e, "failed_model") <- model
    NULL
  })
  res
}

# Парный финальный прогноз для одной и той же обученной модели.
# Ключевой случай — NNAR: nnetar() обучается РОВНО ОДИН РАЗ, после чего
# forecast() вызывается дважды на одном fit: biasadj=FALSE для медианы/PI
# и biasadj=TRUE для математического ожидания. Это исключает рассогласование
# из-за случайной инициализации нейросети.
fit_forecast_pair <- function(y_ts, h, model, wk_train, wk_future,
                              K = FOURIER_K, npaths = 1000, quiet = TRUE) {
  if (model != "nnar_fourier") {
    f_med <- fit_forecast(y_ts, h, model, wk_train, wk_future,
                          K = K, npaths = npaths, quiet = quiet, biasadj = FALSE)
    f_mean <- fit_forecast(y_ts, h, model, wk_train, wk_future,
                           K = K, npaths = npaths, quiet = quiet, biasadj = TRUE)
    if (is.null(f_med) || is.null(f_mean)) return(NULL)
```

```
    return(list(median = f_med, mean = f_mean))
  }

  lam <- if (min(y_ts, na.rm = TRUE) > 0) 0 else NULL
  tryCatch({
    h_tr  <- hol_matrix(wk_train)
    h_fu  <- hol_matrix(wk_future, keep = colnames(h_tr))
    xr_tr <- cbind(fourier(y_ts, K = K), h_tr)
    xr_fu <- cbind(fourier(y_ts, K = K, h = h), h_fu)

    # ОДНА обученная сеть для обеих точечных характеристик.
    # Seed задаётся локально, поэтому результат не зависит от порядка других моделей.
    pair_seed <- SEED + 100000L + length(y_ts) + h
    set.seed(pair_seed)
    fit <- nnetar(y_ts, xreg = xr_tr, lambda = lam)

    # Одинаковый seed перед двумя forecast() синхронизирует симуляционные траектории PI;
    # это не заменяет принцип "один fit", а делает пару median/mean ещё воспроизводимее.
    set.seed(pair_seed + 1L)
    f_med <- forecast(fit, xreg = xr_fu, h = h, PI = TRUE,
                      npaths = npaths, biasadj = FALSE)
    set.seed(pair_seed + 1L)
    f_mean <- forecast(fit, xreg = xr_fu, h = h, PI = TRUE,
                       npaths = npaths, biasadj = TRUE)
    list(median = f_med, mean = f_mean)
  }, error = function(e) {
    if (!quiet) log_msg("    сбой парного прогноза '", model, "': ", conditionMessage(e))
    NULL
  })
}

# Скользящая кросс-валидация одного ряда
run_cv <- function(y_vec, wk_vec, start_yw, models, n_origins, h_eval, label) {
  n <- length(y_vec)
  min_train <- 2 * 52 + 4
  from <- max(min_train, n - h_eval - n_origins + 1)
  to   <- n - h_eval
  if (from > to) stop("Недостаточно данных для КВ ряда '", label, "'.", call. = FALSE)
  origins <- seq.int(from, to)
  log_msg("  КВ '", label, "': ", length(origins), " отсечений (запрошено ", n_origins,
          "), горизонт ", h_eval, " нед.")

  fails <- character(0)
  res <- map_dfr(origins, function(o) {
    y_tr <- mk_ts(y_vec[seq_len(o)], start_yw)
    y_te <- y_vec[(o + 1):(o + h_eval)]
    wk_tr <- wk_vec[seq_len(o)]
    wk_fu <- wk_vec[(o + 1):(o + h_eval)]

    map_dfr(models, function(m) {
      set.seed(SEED + o)
      fc <- fit_forecast(y_tr, h_eval, m, wk_tr, wk_fu, npaths = CV_NPATHS)
      if (is.null(fc)) { fails <<- c(fails, m); return(NULL) }
      lv <- fc$level
      i80 <- which(lv == 80); i95 <- which(lv == 95)
      mase_scale <- if (length(y_tr) > 52) {
        mean(abs(y_tr[(53):length(y_tr)] - y_tr[1:(length(y_tr) - 52)]), na.rm = TRUE)
      } else {
        mean(abs(diff(y_tr)), na.rm = TRUE)
      }
      if (!is.finite(mase_scale) || mase_scale <= 0) mase_scale <- NA_real_
      tibble(origin = o, model = m, h = seq_len(h_eval),
             actual = y_te, pred = as.numeric(fc$mean), mase_scale = mase_scale,
```

```
            lo80 = if (length(i80)) as.numeric(fc$lower[, i80]) else NA_real_,
            hi80 = if (length(i80)) as.numeric(fc$upper[, i80]) else NA_real_,
            lo95 = if (length(i95)) as.numeric(fc$lower[, i95]) else NA_real_,
            hi95 = if (length(i95)) as.numeric(fc$upper[, i95]) else NA_real_)
    })
  })

  if (length(fails)) {
    ft <- table(fails)
    log_msg("  сбоев моделей: ",
            paste(names(ft), "=", as.integer(ft), collapse = ", "),
            " (из ", length(origins), " отсечений)")
  }

  # ансамбль = медиана по успешно посчитанным моделям
  ens <- res |>
    group_by(origin, h, actual) |>
    summarise(model = "ensemble", pred = median(pred), mase_scale = first(mase_scale),
              lo80 = median(lo80, na.rm = TRUE), hi80 = median(hi80, na.rm = TRUE),
              lo95 = median(lo95, na.rm = TRUE), hi95 = median(hi95, na.rm = TRUE),
              .groups = "drop")
  bind_rows(res, ens)
}

cv_metrics <- function(cv) {
  out <- cv |>
    group_by(model) |>
    summarise(
      n        = n(),
      MAE      = mean(abs(actual - pred), na.rm = TRUE),
      RMSE     = sqrt(mean((actual - pred)^2, na.rm = TRUE)),
      MAPE     = mean(abs(actual - pred) / pmax(actual, 1e-9), na.rm = TRUE) * 100,
      WAPE     = sum(abs(actual - pred), na.rm = TRUE) / pmax(sum(actual, na.rm = TRUE), 1e-
   9) * 100,
      # MASE масштабирован средней абсолютной ошибкой сезонного naive с лагом 52
      # на КАЖДОЙ обучающей выборке. Поэтому MASE SNaive на внешнем CV не обязан
      # быть ровно 1: значение 1 относится к историческому in-sample масштабу.
      MASE     = mean(abs(actual - pred) / mase_scale, na.rm = TRUE),
      BIAS     = mean(pred - actual, na.rm = TRUE),
      BIAS_pct = sum(pred - actual, na.rm = TRUE) / pmax(sum(actual, na.rm = TRUE), 1e-9) *
   100,
      COV80_raw = mean(actual >= lo80 & actual <= hi80, na.rm = TRUE) * 100,
      COV95_raw = mean(actual >= lo95 & actual <= hi95, na.rm = TRUE) * 100,
      .groups  = "drop"
    )
  base <- out$MAE[out$model == "snaive"]
  if (length(base) == 1 && is.finite(base)) {
    out <- out |> mutate(MAE_skill_vs_SNaive_pct = (1 - MAE / base) * 100)
  } else {
    out <- out |> mutate(MAE_skill_vs_SNaive_pct = NA_real_)
  }
  out |> arrange(MAE)
}

cv_metrics_horizon <- function(cv) {
  raw <- cv |>
    group_by(model, h) |>
    summarise(
      n        = n(),
      MAE      = mean(abs(actual - pred), na.rm = TRUE),
      RMSE     = sqrt(mean((actual - pred)^2, na.rm = TRUE)),
      MAPE     = mean(abs(actual - pred) / pmax(actual, 1e-9), na.rm = TRUE) * 100,
```

```
    WAPE     = sum(abs(actual - pred), na.rm = TRUE) / pmax(sum(actual, na.rm = TRUE), 1e-
 9) * 100,
    MASE     = mean(abs(actual - pred) / mase_scale, na.rm = TRUE),
    BIAS_pct = sum(pred - actual, na.rm = TRUE) / pmax(sum(actual, na.rm = TRUE), 1e-9) *
 100,
    COV80_raw = mean(actual >= lo80 & actual <= hi80, na.rm = TRUE) * 100,
    COV95_raw = mean(actual >= lo95 & actual <= hi95, na.rm = TRUE) * 100,
    .groups = "drop"
  )

# Добавляем skill относительно SNaive ОТДЕЛЬНО для каждого горизонта.
base_h <- raw |> filter(model == "snaive") |> select(h, MAE_snaive = MAE)
raw <- raw |> left_join(base_h, by = "h") |>
  mutate(MAE_skill_vs_SNaive_pct = (1 - MAE / MAE_snaive) * 100) |>
  select(-MAE_snaive)

# Calibrated coverage по каждому горизонту при том же ГЛОБАЛЬНОМ k,
# который используется для финального прогноза соответствующей модели.
cal_h <- map_dfr(unique(cv$model), function(m) {
  kk <- calib_k_one(cv, m)$k[1]
  cv |> filter(model == m) |> apply_calib_cv(kk) |>
    group_by(model, h) |>
    summarise(
      COV80_cal = mean(actual >= lo80_cal & actual <= hi80_cal, na.rm = TRUE) * 100,
      COV95_cal = mean(actual >= lo95_cal & actual <= hi95_cal, na.rm = TRUE) * 100,
      .groups = "drop"
    )
})

raw |> left_join(cal_h, by = c("model", "h")) |> arrange(model, h)
}

# Эмпирическая калибровка ширины интервалов по остаткам КВ.
# Возвращаются raw k и фактически применённый k; safety cap — только защита
# от численной нестабильности, а не критерий "нормы".
calib_k_one <- function(cv, model_name) {
Z95 <- qnorm(0.975)
d <- cv |> filter(model == model_name, lo95 > 0, hi95 > 0, actual > 0, pred > 0)
if (!nrow(d)) return(tibble(model = model_name, k_raw = 1, k = 1, cap_hit = FALSE))
k_raw <- d |>
  mutate(sd_log = (log(hi95) - log(lo95)) / (2 * Z95),
         z = (log(actual) - log(pred)) / sd_log) |>
  filter(is.finite(z), sd_log > 0) |>
  summarise(k = quantile(abs(z), 0.95, na.rm = TRUE) / Z95) |>
  pull(k)
k_raw <- unname(k_raw)
if (!length(k_raw) || !is.finite(k_raw)) k_raw <- 1
k <- min(max(k_raw, K_MIN), K_SAFETY_CAP)
tibble(model = model_name, k_raw = k_raw, k = k, cap_hit = isTRUE(k_raw >= K_SAFETY_CAP))
}

apply_calib_cv <- function(d, k) {
Z95 <- qnorm(0.975); Z80 <- qnorm(0.90)
d |>
  mutate(
    ctr95 = (log(pmax(hi95, 1e-9)) + log(pmax(lo95, 1e-9))) / 2,
    ctr80 = (log(pmax(hi80, 1e-9)) + log(pmax(lo80, 1e-9))) / 2,
    sd80 = (log(pmax(hi80, 1e-9)) - log(pmax(lo80, 1e-9))) / (2 * Z80),
    sd95 = (log(pmax(hi95, 1e-9)) - log(pmax(lo95, 1e-9))) / (2 * Z95),
    lo80_cal = exp(ctr80 - k * Z80 * sd80),
    hi80_cal = exp(ctr80 + k * Z80 * sd80),
    lo95_cal = exp(ctr95 - k * Z95 * sd95),
    hi95_cal = exp(ctr95 + k * Z95 * sd95)
```

```
    ) |>
    select(-ctr95, -ctr80, -sd80, -sd95)
}

calibration_table <- function(cv) {
  mods <- unique(cv$model)
  map_dfr(mods, function(m) {
    kk <- calib_k_one(cv, m)
    d <- cv |> filter(model == m) |> apply_calib_cv(kk$k[1])
    tibble(
      model = m,
      k_raw = kk$k_raw[1],
      k = kk$k[1],
      cap_hit = kk$cap_hit[1],
      COV80_cal = mean(d$actual >= d$lo80_cal & d$actual <= d$hi80_cal, na.rm = TRUE) * 100,
      COV95_cal = mean(d$actual >= d$lo95_cal & d$actual <= d$hi95_cal, na.rm = TRUE) * 100
    )
  })
}

# Правило отбора v4.2: сложная модель должна не проигрывать SNaive по MAE
# И иметь приемлемое СЫРОЕ (до калибровки) покрытие 95% PI.
# Raw coverage используется как независимый CV-фильтр; calibrated coverage
# не участвует в выборе, потому что k подбирается на тех же CV-остатках.
pick_model <- function(mt, label) {
  base <- mt$MAE[mt$model == "snaive"]
  if (!length(base) || is.na(base)) {
    log_msg("  '", label, "': SNaive не посчитался — беру модель с минимальным MAE")
    return(mt$model[1])
  }
  threshold <- base * (1 - MIN_IMPROVEMENT)
  cand <- mt |>
    filter(model != "snaive", MAE <= threshold, COV95_raw >= MIN_COVER_RAW) |>
    arrange(MAE)
  if (nrow(cand)) {
    log_msg("  '", label, "': выбрана ", cand$model[1],
            " (MAE ", round(cand$MAE[1], 1), " против ", round(base, 1),
            " у SNaive; raw coverage 95% = ", round(cand$COV95_raw[1], 1),
            "%; calibrated = ", round(cand$COV95_cal[1], 1), "%)")
    return(cand$model[1])
  }
  cand2 <- mt |>
    filter(model != "snaive", MAE <= threshold) |>
    arrange(MAE)
  if (nrow(cand2)) {
    log_msg("  '", label, "': кандидат улучшает SNaive по MAE, но raw coverage < ",
            MIN_COVER_RAW, "%. Беру SNaive как более консервативную модель сравнения.")
  } else {
    log_msg("  '", label, "': содержательные модели не улучшили SNaive -> SNaive")
  }
  "snaive"
}


# В v4 математическое ожидание НЕ восстанавливается универсальной логнормальной
# формулой из ширины PI. Для финального прогноза оно получается отдельным запуском
# той же модели с biasadj=TRUE (где это поддерживается forecast), что корректнее
# для Box-Cox моделей и не связывает mean с эмпирической калибровкой PI.
attach_mean <- function(fc_tbl, mean_vec) {
  if (length(mean_vec) != nrow(fc_tbl)) stop("Длина mean_vec не совпадает с горизонтом
   прогноза.")
  fc_tbl |> mutate(среднее = as.numeric(mean_vec))
}
```

```
apply_calib <- function(fc_tbl, k) {
  if (abs(k - 1) < 1e-6) return(fc_tbl)
  Z95 <- qnorm(0.975); Z80 <- qnorm(0.90)
  fc_tbl |>
    mutate(
      ctr80 = (log(pmax(hi80, 1e-9)) + log(pmax(lo80, 1e-9))) / 2,
      ctr95 = (log(pmax(hi95, 1e-9)) + log(pmax(lo95, 1e-9))) / 2,
      sd80 = (log(pmax(hi80, 1e-9)) - log(pmax(lo80, 1e-9))) / (2 * Z80),
      sd95 = (log(pmax(hi95, 1e-9)) - log(pmax(lo95, 1e-9))) / (2 * Z95),
      lo80 = exp(ctr80 - k * Z80 * sd80), hi80 = exp(ctr80 + k * Z80 * sd80),
      lo95 = exp(ctr95 - k * Z95 * sd95), hi95 = exp(ctr95 + k * Z95 * sd95)
    ) |>
    select(-ctr80, -ctr95, -sd80, -sd95)
}


# ---- 9. УРОВЕНЬ A: национальный прогноз ----------------------------------
hr("ШАГ 6. НАЦИОНАЛЬНЫЙ ПРОГНОЗ (абсолютные числа)")

fc_dates <- last_date + 7L * seq_len(H_FORECAST)
fc_wk    <- future_weeks(fc_dates)
fc_yr    <- as.integer(format(fc_dates, "%G"))

y_nat  <- nat_series$cases
wk_nat <- nat_series$week

cv_nat <- run_cv(y_nat, wk_nat, START_YW, MODELS_NAT, CV_ORIG_NAT, H_EVAL, "КР, всего")
mt_nat <- cv_metrics(cv_nat) |> left_join(calibration_table(cv_nat), by = "model") |>
   arrange(MAE)
mt_nat_h <- cv_metrics_horizon(cv_nat)
log_msg("Метрики кросс-валидации (raw и calibrated coverage):")
print(as.data.frame(mt_nat |> mutate(across(where(is.numeric), ~round(.x, 2)))), row.names =
   FALSE)
write_excel_csv(mt_nat, file.path(OUT_DIR, "09_метрики_национальные.csv"))
write_excel_csv(mt_nat_h, file.path(OUT_DIR, "09b_метрики_по_горизонту.csv"))

best_nat <- pick_model(mt_nat, "КР, всего")
k_nat_row <- mt_nat |> filter(model == best_nat) |> slice(1)
k_nat <- k_nat_row$k
if (isTRUE(k_nat_row$cap_hit)) {
  warning("Калибровка национальной модели достигла технического safety cap: raw k=",
          round(k_nat_row$k_raw, 2), ", применено k=", round(k_nat, 2),
          ". Интервал может оставаться недостаточно калиброванным.")
}
log_msg("Калибровка: raw k=", round(k_nat_row$k_raw, 2), "; применено k=", round(k_nat, 2),
        "; coverage95 raw=", round(k_nat_row$COV95_raw, 1),
        "%; calibrated=", round(k_nat_row$COV95_cal, 1), "%")

# Финальный национальный прогноз. Для NNAR медиана и mean всегда получаются
# из ОДНОЙ обученной сети через fit_forecast_pair().
gm_col <- function(f, part, lvl) {
  i <- which(f$level == lvl)
  if (!length(i)) return(rep(NA_real_, H_FORECAST))
  as.numeric(f[[part]][, i])
}

if (best_nat == "ensemble") {
  pairs <- map(MODELS_NAT, ~fit_forecast_pair(mk_ts(y_nat, START_YW), H_FORECAST, .x,
                                              wk_nat, fc_wk, npaths = 1000, quiet = FALSE))
   |>
```

```
    compact()
  if (!length(pairs)) stop("Ни одна компонентная модель ансамбля не посчиталась.", call. =
   FALSE)
  fits <- lapply(pairs, `[[`, "median")
  fits_mean <- lapply(pairs, `[[`, "mean")

  nat_fc <- tibble(
    прогноз = apply(do.call(cbind, lapply(fits, function(f) as.numeric(f$mean))), 1,
   median),
    lo80 = apply(do.call(cbind, lapply(fits, gm_col, "lower", 80)), 1, median, na.rm =
   TRUE),
    hi80 = apply(do.call(cbind, lapply(fits, gm_col, "upper", 80)), 1, median, na.rm =
   TRUE),
    lo95 = apply(do.call(cbind, lapply(fits, gm_col, "lower", 95)), 1, median, na.rm =
   TRUE),
    hi95 = apply(do.call(cbind, lapply(fits, gm_col, "upper", 95)), 1, median, na.rm = TRUE)
  )
  # Equal-weight mixture: E[X] = среднее component means.
  mean_nat <- rowMeans(do.call(cbind, lapply(fits_mean, function(f) as.numeric(f$mean))),
   na.rm = TRUE)
} else if (best_nat == "nnar_fourier") {
  pair_nat <- fit_forecast_pair(mk_ts(y_nat, START_YW), H_FORECAST, best_nat,
                                wk_nat, fc_wk, npaths = 1000, quiet = FALSE)
  if (is.null(pair_nat)) stop("Финальная NNAR-модель не посчиталась.", call. = FALSE)
  f <- pair_nat$median
  f_mean <- pair_nat$mean
  i80 <- which(f$level == 80); i95 <- which(f$level == 95)
  nat_fc <- tibble(
    прогноз = as.numeric(f$mean),
    lo80 = as.numeric(f$lower[, i80]), hi80 = as.numeric(f$upper[, i80]),
    lo95 = as.numeric(f$lower[, i95]), hi95 = as.numeric(f$upper[, i95])
  )
  mean_nat <- as.numeric(f_mean$mean)
} else {
  f <- fit_forecast(mk_ts(y_nat, START_YW), H_FORECAST, best_nat,
                    wk_nat, fc_wk, npaths = 1000, quiet = FALSE, biasadj = FALSE)
  if (is.null(f)) stop("Финальная национальная модель не посчиталась.", call. = FALSE)
  i80 <- which(f$level == 80); i95 <- which(f$level == 95)
  nat_fc <- tibble(
    прогноз = as.numeric(f$mean),
    lo80 = as.numeric(f$lower[, i80]), hi80 = as.numeric(f$upper[, i80]),
    lo95 = as.numeric(f$lower[, i95]), hi95 = as.numeric(f$upper[, i95])
  )
  f_mean <- fit_forecast(mk_ts(y_nat, START_YW), H_FORECAST, best_nat,
                         wk_nat, fc_wk, npaths = 1000, quiet = FALSE, biasadj = TRUE)
  if (is.null(f_mean)) stop("Не удалось получить bias-adjusted mean национальной модели.",
   call. = FALSE)
  mean_nat <- as.numeric(f_mean$mean)
}

pop_nat_fc <- sum(pop_ref$pop_ref)

nat_fc <- nat_fc |>
  attach_mean(mean_nat) |>
  apply_calib(k_nat) |>
  mutate(across(everything(), ~pmax(.x, 0))) |>
  mutate(
    week_end = fc_dates, год = fc_yr, неделя = fc_wk,
    across(c(прогноз, среднее, lo80, hi80, lo95, hi95), round),
    прогноз_на_100тыс = round(прогноз / pop_nat_fc * 1e5, 1),
    lo95_на_100тыс    = round(lo95    / pop_nat_fc * 1e5, 1),
    hi95_на_100тыс    = round(hi95    / pop_nat_fc * 1e5, 1)
  ) |>
```

```
  select(год, неделя, week_end, прогноз, среднее, lo80, hi80, lo95, hi95,
         прогноз_на_100тыс, lo95_на_100тыс, hi95_на_100тыс)

log_msg("Национальный прогноз построен.")
log_msg("  Медианный сценарий за ", H_FORECAST, " нед.: ",
        format(sum(nat_fc$прогноз), big.mark = " "), " случаев")
log_msg("  Ожидаемое (среднее) число за ", H_FORECAST, " нед.: ",
        format(sum(nat_fc$среднее), big.mark = " "), " случаев")
write_excel_csv(nat_fc, file.path(OUT_DIR, "10_прогноз_национальный.csv"))

# Паспорт расчёта — чтобы отчёт не пересчитывал параметры самостоятельно
cv_from_i <- max(2 * 52 + 4, n_nat - H_EVAL - CV_ORIG_NAT + 1)
cv_to_i   <- n_nat - H_EVAL
params <- tibble(
  параметр = c("Рабочая модель (KP)", "Число отсечений кросс-валидации (KP)",
               "Период кросс-валидации (KP)", "Горизонт оценки при отборе, нед.",
               "Горизонт прогноза, нед.", "Коэффициент калибровки raw k (KP)",
               "Коэффициент калибровки применённый k (KP)",
               "Покрытие 95% PI до калибровки, %", "Покрытие 95% PI после калибровки, %",
               "Критерий покрытия при выборе модели", "Точечный прогноз", "Гармоник Фурье
   (пар)",
               "Праздничные недели-регрессоры", "Опорное население KP"),
  значение = c(
    best_nat,
    as.character(cv_to_i - cv_from_i + 1),
    paste0(format(nat_series$week_end[cv_from_i + 1], "%d.%m.%Y"), " — ",
           format(nat_series$week_end[min(cv_to_i + H_EVAL, n_nat)], "%d.%m.%Y")),
    as.character(H_EVAL), as.character(H_FORECAST),
    formatC(k_nat_row$k_raw, format = "f", digits = 2),
    formatC(k_nat, format = "f", digits = 2),
    formatC(k_nat_row$COV95_raw, format = "f", digits = 1),
    formatC(k_nat_row$COV95_cal, format = "f", digits = 1),
    paste0("raw coverage 95% >= ", MIN_COVER_RAW, "% (calibrated coverage — диагностическая
   метрика)"),
    "медиана распределения; mean = bias-adjusted forecast той же модели; для NNAR median и
   mean получаются из одного fit; для ensemble mean = арифметическое среднее component
   means",
    as.character(FOURIER_K),
    paste(HOLIDAY_WEEKS, collapse = ", "),
    format(pop_nat_fc, big.mark = " ")
  )
)
write_excel_csv(params, file.path(OUT_DIR, "00_параметры_расчёта.csv"))


# ---- 10. УРОВЕНЬ B: региональные прогнозы ----------------------------------
hr("ШАГ 7. РЕГИОНАЛЬНЫЕ ПРОГНОЗЫ")
log_msg("Внимание: этот блок считается дольше остальных (9 рядов x ",
        CV_ORIG_REG, " отсечений x ", length(MODELS_REG), " моделей).")

reg_fc_all  <- list()
reg_report  <- list()

for (reg in TERRITORIES) {
  log_msg("--- ", reg, " ---")
  d <- reg_series |> filter(territory == reg) |> arrange(week_end)

  if (max(d$week_end) != last_date) {
    warning("Ряд '", reg, "' заканчивается ", format(max(d$week_end), "%d.%m.%Y"),
            ", а национальный — ", format(last_date, "%d.%m.%Y"), ". Регион пропущен.")
    next
  }
  gap_r <- which(as.integer(diff(d$week_end)) != 7L)
```

```
if (length(gap_r)) {
  warning("Разрыв сетки в ряду '", reg, "'. Регион пропущен.")
  next
}

cv_r <- run_cv(d$cases, d$week, c(d$year[1], d$week[1]),
               MODELS_REG, CV_ORIG_REG, H_EVAL, reg)
mt_r <- cv_metrics(cv_r) |> left_join(calibration_table(cv_r), by = "model") |>
 arrange(MAE)
best_r <- pick_model(mt_r, reg)
k_r_row <- mt_r |> filter(model == best_r) |> slice(1)
k_r <- k_r_row$k
if (isTRUE(k_r_row$cap_hit)) warning("Регион '", reg, "': достигнут safety cap калибровки;
 raw k=", round(k_r_row$k_raw, 2))

if (best_r == "ensemble") {
  fits <- compact(map(MODELS_REG, ~fit_forecast(mk_ts(d$cases, c(d$year[1], d$week[1])),
                                                H_FORECAST, .x, d$week, fc_wk, npaths =
 1000)))
  gm_col <- function(f, part, lvl) {
    i <- which(f$level == lvl)
    if (!length(i)) return(rep(NA_real_, H_FORECAST))
    as.numeric(f[[part]][, i])
  }
  tb <- tibble(прогноз = apply(do.call(cbind, lapply(fits, function(f)
 as.numeric(f$mean))), 1, median),
               lo80 = apply(do.call(cbind, lapply(fits, gm_col, "lower", 80)), 1, median,
 na.rm = TRUE),
               hi80 = apply(do.call(cbind, lapply(fits, gm_col, "upper", 80)), 1, median,
 na.rm = TRUE),
               lo95 = apply(do.call(cbind, lapply(fits, gm_col, "lower", 95)), 1, median,
 na.rm = TRUE),
               hi95 = apply(do.call(cbind, lapply(fits, gm_col, "upper", 95)), 1, median,
 na.rm = TRUE))
} else {
  f <- fit_forecast(mk_ts(d$cases, c(d$year[1], d$week[1])), H_FORECAST, best_r,
                    d$week, fc_wk, npaths = 1000, quiet = FALSE)
  if (is.null(f)) { warning("Финальная модель для '", reg, "' не посчиталась."); next }
  i80 <- which(f$level == 80); i95 <- which(f$level == 95)
  tb <- tibble(прогноз = as.numeric(f$mean),
               lo80 = as.numeric(f$lower[, i80]), hi80 = as.numeric(f$upper[, i80]),
               lo95 = as.numeric(f$lower[, i95]), hi95 = as.numeric(f$upper[, i95]))
}

# Отдельный bias-adjusted запуск для expected mean.
if (best_r == "ensemble") {
  fits_mean_r <- compact(map(MODELS_REG, ~fit_forecast(mk_ts(d$cases, c(d$year[1],
 d$week[1])),
                                                       H_FORECAST, .x, d$week, fc_wk,
 npaths = 1000,
                                                       biasadj = TRUE)))
  # Expected mean equal-weight mixture = арифметическое среднее component means.
  mean_r <- rowMeans(do.call(cbind, lapply(fits_mean_r, function(f) as.numeric(f$mean))),
 na.rm = TRUE)
} else {
  f_mean_r <- fit_forecast(mk_ts(d$cases, c(d$year[1], d$week[1])), H_FORECAST, best_r,
                           d$week, fc_wk, npaths = 1000, quiet = FALSE, biasadj = TRUE)
  if (is.null(f_mean_r)) { warning("Mean-прогноз для '", reg, "' не посчитался; использую
 медиану."); mean_r <- tb$прогноз }
  else mean_r <- as.numeric(f_mean_r$mean)
}

pop_r <- unique(d$pop_ref)[1]
```

```
  tb <- tb |> attach_mean(mean_r) |> apply_calib(k_r) |>
    mutate(across(everything(), ~pmax(.x, 0)))

  reg_fc_all[[reg]] <- tb |>
    mutate(territory = reg, week_end = fc_dates, год = fc_yr, неделя = fc_wk,
           население = pop_r)

  reg_report[[reg]] <- tibble(
    Регион = reg, Модель = best_r,
    MAE = round(k_r_row$MAE, 1),
    MAPE_pct = round(k_r_row$MAPE, 1),
    WAPE_pct = round(k_r_row$WAPE, 1),
    MASE = round(k_r_row$MASE, 2),
    Bias_pct = round(k_r_row$BIAS_pct, 1),
    Покрытие95_raw_pct = round(k_r_row$COV95_raw, 1),
    Покрытие95_cal_pct = round(k_r_row$COV95_cal, 1),
    k_raw = round(k_r_row$k_raw, 2),
    k = round(k_r, 2),
    Надёжность = {
      wp <- k_r_row$WAPE; ms <- k_r_row$MASE; cv <- k_r_row$COV95_cal
      if (!is.finite(wp) || !is.finite(ms) || !is.finite(cv)) "не определена"
      else if (wp > REL_WAPE_LOW || ms > REL_MASE_LOW || cv < REL_COV_LOW) "низкая"
      else if (wp > REL_WAPE_ACCEPT || ms > REL_MASE_ACCEPT || cv < REL_COV_ACCEPT)
   "ограниченная"
      else "приемлемая"
    },
    Средний_уровень = round(mean(tail(d$cases, 52))),
    Население = pop_r
  )
}

if (!length(reg_fc_all)) stop("Ни один регион не был посчитан.", call. = FALSE)

reg_fc   <- bind_rows(reg_fc_all)
write_excel_csv(reg_fc, file.path(OUT_DIR, "12a_региональные_прогнозы_до_согласования.csv"))
reg_tbl  <- bind_rows(reg_report) |> arrange(desc(Средний_уровень))
log_msg("Карта региональных моделей:")
print(as.data.frame(reg_tbl), row.names = FALSE)
write_excel_csv(reg_tbl, file.path(OUT_DIR, "11_карта_моделей_региональных.csv"))



# ---- 11. Согласование уровней ------------------------------------------
# Региональные прогнозы пропорционально сводятся к национальному expected mean.
# Тот же коэффициент применяется к медиане и границам региональных PI, чтобы
# точечный прогноз и интервалы оставались в одном согласованном масштабе.

hr("ШАГ 8. СОГЛАСОВАНИЕ УРОВНЕЙ")

# Согласование выполняется по СРЕДНИМ значениям: только они аддитивны.
# Полученный недельный коэффициент применяется и к медиане, и к среднему.
recon <- reg_fc |>
  group_by(week_end) |>
  summarise(сумма_регионов = sum(среднее), .groups = "drop") |>
  left_join(nat_fc |> select(week_end, национальный = среднее), by = "week_end") |>
  mutate(коэффициент = национальный / сумма_регионов,
         расхождение_pct = round((сумма_регионов / национальный - 1) * 100, 1))

log_msg("Расхождение снизу-вверх против национальной модели:")
print(as.data.frame(recon |> mutate(across(c(сумма_регионов, национальный), round),
                                    коэффициент = round(коэффициент, 3))), row.names =
   FALSE)
```

```
max_disc <- max(abs(recon$расхождение_pct))
if (max_disc > 25) {
  warning("Уровни расходятся до ", max_disc, "%. ",
          "Национальная и региональные модели описывают разную динамику — ",
          "результат требует содержательной проверки эпидемиолога.")
} else {
  log_msg("Максимальное расхождение уровней: ", max_disc, "% — приемлемо.")
}
write_excel_csv(recon, file.path(OUT_DIR, "12_согласование_уровней.csv"))

reg_fc_final <- reg_fc |>
  left_join(recon |> select(week_end, коэффициент), by = "week_end") |>
  mutate(
    прогноз_согл = прогноз * коэффициент,
    среднее_согл = среднее * коэффициент,
    lo80_согл = lo80 * коэффициент, hi80_согл = hi80 * коэффициент,
    lo95_согл = lo95 * коэффициент, hi95_согл = hi95 * коэффициент,
    across(c(прогноз, прогноз_согл, среднее_согл, lo80_согл, hi80_согл, lo95_согл,
   hi95_согл), round),
    прогноз_на_100тыс = round(прогноз_согл / население * 1e5, 1),
    lo95_на_100тыс    = round(lo95_согл / население * 1e5, 1),
    hi95_на_100тыс    = round(hi95_согл / население * 1e5, 1)
  ) |>
  select(territory, год, неделя, week_end, население,
         прогноз_абс = прогноз_согл, среднее_абс = среднее_согл,
         прогноз_абс_несогл = прогноз,
         lo80 = lo80_согл, hi80 = hi80_согл, lo95 = lo95_согл, hi95 = hi95_согл,
         прогноз_на_100тыс, lo95_на_100тыс, hi95_на_100тыс) |>
  arrange(territory, week_end)

# Доля территории в республике — от СУММЫ за весь период прогноза,
# а не среднее недельных долей (среднее долей не согласуется с суммами).
reg_share <- reg_fc_final |>
  group_by(territory) |>
  summarise(сумма_периода = sum(среднее_абс), .groups = "drop") |>
  mutate(доля_в_КР_pct = round(сумма_периода / sum(сумма_периода) * 100, 1))
reg_fc_final <- reg_fc_final |> left_join(reg_share |> select(territory, доля_в_КР_pct),
                                          by = "territory")

write_excel_csv(reg_fc_final, file.path(OUT_DIR,
   "12b_региональные_прогнозы_после_согласования.csv"))
write_excel_csv(reg_fc_final, file.path(OUT_DIR, "13_прогноз_региональный.csv"))

# Приёмочный контроль аддитивности expected mean после reconciliation.
recon_check <- reg_fc_final |> group_by(week_end) |> summarise(reg_mean = sum(среднее_абс),
   .groups = "drop") |>
  left_join(nat_fc |> select(week_end, nat_mean = среднее), by = "week_end") |>
  mutate(diff = reg_mean - nat_mean)
if (max(abs(recon_check$diff), na.rm = TRUE) > length(TERRITORIES) + 1) {
  warning("После округления сумма согласованных региональных средних отличается от
   национальной более чем ожидается.")
}


# ---- 11.1 Sensitivity analysis: аномальный эпизод конца 2025 года -----------
# Это НЕ альтернативный основной прогноз. Цель — оценить, насколько финальная
# траектория зависит от недель, отмеченных диагностическим правилом >3x медианы
# предыдущих 52 недель. В sensitivity-сценарии такие недели заменяются базовой
# rolling-медианой; исходный ряд остаётся неизменным в основном расчёте.
if (nrow(brk)) {
  y_sens <- y_nat
  idx_sens <- match(brk$week_end, nat_series$week_end)
  ok_idx <- which(!is.na(idx_sens))
```

```
if (length(ok_idx)) {
  y_sens[idx_sens[ok_idx]] <- brk$base[ok_idx]
  if (best_nat == "ensemble") {
    sens_pairs <- map(MODELS_NAT, ~fit_forecast_pair(mk_ts(y_sens, START_YW), H_FORECAST,
 .x,
                                                     wk_nat, fc_wk, npaths = 1000, quiet
 = FALSE)) |>
      compact()
    fs <- lapply(sens_pairs, `[[`, "median")
    fsm <- lapply(sens_pairs, `[[`, "mean")
    sens_med <- apply(do.call(cbind, lapply(fs, function(f) as.numeric(f$mean))), 1,
 median)
    sens_mean <- rowMeans(do.call(cbind, lapply(fsm, function(f) as.numeric(f$mean))),
 na.rm = TRUE)
    sens_lo95 <- apply(do.call(cbind, lapply(fs, function(f) { i <- which(f$level == 95);
 as.numeric(f$lower[, i]) })), 1, median, na.rm = TRUE)
    sens_hi95 <- apply(do.call(cbind, lapply(fs, function(f) { i <- which(f$level == 95);
 as.numeric(f$upper[, i]) })), 1, median, na.rm = TRUE)
    sens_ok <- length(fs) > 0
  } else if (best_nat == "nnar_fourier") {
    sens_pair <- fit_forecast_pair(mk_ts(y_sens, START_YW), H_FORECAST, best_nat,
                                   wk_nat, fc_wk, npaths = 1000, quiet = FALSE)
    sens_ok <- !is.null(sens_pair)
    if (sens_ok) {
      f_sens <- sens_pair$median
      f_sens_mean <- sens_pair$mean
      i95s <- which(f_sens$level == 95)
      sens_med <- as.numeric(f_sens$mean)
      sens_mean <- as.numeric(f_sens_mean$mean)
      sens_lo95 <- as.numeric(f_sens$lower[, i95s])
      sens_hi95 <- as.numeric(f_sens$upper[, i95s])
    }
  } else {
    f_sens <- fit_forecast(mk_ts(y_sens, START_YW), H_FORECAST, best_nat,
                           wk_nat, fc_wk, npaths = 1000, quiet = FALSE, biasadj = FALSE)
    f_sens_mean <- fit_forecast(mk_ts(y_sens, START_YW), H_FORECAST, best_nat,
                                wk_nat, fc_wk, npaths = 1000, quiet = FALSE, biasadj =
 TRUE)
    sens_ok <- !is.null(f_sens)
    if (sens_ok) {
      i95s <- which(f_sens$level == 95)
      sens_med <- as.numeric(f_sens$mean)
      sens_mean <- if (!is.null(f_sens_mean)) as.numeric(f_sens_mean$mean) else
 rep(NA_real_, H_FORECAST)
      sens_lo95 <- as.numeric(f_sens$lower[, i95s])
      sens_hi95 <- as.numeric(f_sens$upper[, i95s])
    }
  }
  if (sens_ok) {
    sens_tbl <- tibble(
      week_end = fc_dates,
      год = fc_yr, неделя = fc_wk,
      основной_медиана = nat_fc$прогноз,
      sensitivity_медиана = round(sens_med),
      основной_среднее = nat_fc$среднее,
      sensitivity_среднее = round(sens_mean),
      sensitivity_lo95_raw = round(sens_lo95),
      sensitivity_hi95_raw = round(sens_hi95)
    ) |>
      mutate(
        разница_медиана_pct = round((sensitivity_медиана / pmax(основной_медиана, 1) - 1)
 * 100, 1),
        abs_разница_pct = abs(разница_медиана_pct)
      )
```

```
      write_excel_csv(sens_tbl, file.path(OUT_DIR, "sensitivity_2025.csv"))
      sens_max_diff <- max(sens_tbl$abs_разница_pct, na.rm = TRUE)
      log_msg("Sensitivity analysis 2025: максимальное отличие медианы = ", sens_max_diff,
   "%")
    } else sens_max_diff <- NA_real_
  } else sens_max_diff <- NA_real_
} else {
  sens_max_diff <- 0
  write_excel_csv(tibble(note = "Аномальные недели по правилу >3х не выявлены"),
                  file.path(OUT_DIR, "sensitivity_2025.csv"))
}


# ---- 12. Таблица текущей ситуации по регионам -----------------------------
current <- reg_series |>
  filter(week_end == last_date) |>
  left_join(reg_series |> filter(week_end == last_date - 7L) |>
              select(territory, cases_prev = cases), by = "territory") |>
  mutate(
    доля_в_КР_pct   = round(cases / sum(cases) * 100, 1),
    на_100тыс_офиц  = round(rate_off, 1),
    на_100тыс_гарм  = round(rate_harm, 1),
    динамика_pct    = round((cases / cases_prev - 1) * 100, 1)
  ) |>
  select(Территория = territory, Случаев_абс = cases, Доля_в_КР_pct = доля_в_КР_pct,
         На_100тыс_официальный = на_100тыс_офиц, На_100тыс_гармонизир = на_100тыс_гарм,
         Население = population, Динамика_к_прош_нед_pct = динамика_pct) |>
  arrange(desc(На_100тыс_гармонизир))

log_msg("Ситуация на ", format(last_date, "%d.%m.%Y"), ":")
print(as.data.frame(current), row.names = FALSE)
write_excel_csv(current, file.path(OUT_DIR, "14_ситуация_текущая.csv"))

# КОНТРОЛЬ: сумма по таблице должна совпасть с национальным рядом
if (abs(sum(current$Случаев_абс) - nat_series$cases[n_nat]) > 0.5) {
  stop("Сумма региональной таблицы (", sum(current$Случаев_абс),
       ") не совпадает с национальным рядом (", nat_series$cases[n_nat], ").", call. =
   FALSE)
}
log_msg("Контроль сходимости региональной таблицы и национального ряда: ОК")


# ---- 13. Графики ----------------------------------------------------------
hr("ШАГ 9. ГРАФИКИ")

hist2 <- nat_series |> filter(week_end >= last_date - 7 * 156)

p1 <- ggplot() +
  geom_ribbon(data = nat_fc, aes(week_end, ymin = lo95, ymax = hi95),
              fill = "#2c7fb8", alpha = 0.18) +
  geom_ribbon(data = nat_fc, aes(week_end, ymin = lo80, ymax = hi80),
              fill = "#2c7fb8", alpha = 0.32) +
  geom_line(data = hist2, aes(week_end, cases), colour = "grey20", linewidth = 0.7) +
  geom_line(data = nat_fc, aes(week_end, прогноз), colour = "#08519c", linewidth = 0.9) +
  geom_vline(xintercept = as.numeric(last_date), linetype = "dashed", colour = "grey45") +
  scale_x_date(date_breaks = "3 months", date_labels = "%m.%y") +
  scale_y_continuous(labels = function(x) format(x, big.mark = " ", scientific = FALSE)) +
  labs(title = paste0("ОРВИ и грипп, Кыргызская Республика: прогноз на ", H_FORECAST, "
   нед."),
       subtitle = paste0("Модель: ", best_nat, "; точечная линия — медиана; ",
                         "интервалы откалиброваны по кросс-валидации (k = ",
                         round(k_nat, 2), ")"),
```

```
       x = NULL, y = "Случаев за неделю",
       caption = paste0("Последняя неделя данных: ", format(last_date, "%d.%m.%Y"),
                        ". Штриховая линия — граница наблюдения.")) +
  theme_bw(base_size = 12)
ggsave(file.path(OUT_DIR, "g1_прогноз_национальный.png"), p1, width = 11, height = 5.5, dpi
   = 300)

p2 <- ggplot(reg_series |> filter(week_end >= last_date - 7 * 104),
             aes(week_end, rate_harm)) +
  geom_line(colour = "#08519c", linewidth = 0.5) +
  facet_wrap(~territory, scales = "free_y", ncol = 3) +
  scale_x_date(date_breaks = "6 months", date_labels = "%m.%y") +
  labs(title = "Заболеваемость ОРВИ и гриппом по территориям КР, на 100 тыс. населения",
       subtitle = "Единый (гармонизированный) знаменатель — без административных ступенек",
       x = NULL, y = "На 100 тыс. населения в неделю") +
  theme_bw(base_size = 10)
ggsave(file.path(OUT_DIR, "g2_регионы_интенсивные.png"), p2, width = 12, height = 7, dpi =
   300)

p3 <- reg_series |>
  filter(year >= max(year) - 1) |>
  mutate(разница = rate_off - rate_harm) |>
  ggplot(aes(week_end, разница, colour = territory)) +
  geom_hline(yintercept = 0, colour = "grey60") +
  geom_line(linewidth = 0.6) +
  scale_x_date(date_breaks = "3 months", date_labels = "%m.%y") +
  labs(title = "Влияние смены официального знаменателя на интенсивный показатель",
       subtitle = "Разность «официальный минус гармонизированный», на 100 тыс.",
       x = NULL, y = "Разность", colour = "Территория") +
  theme_bw(base_size = 11)
ggsave(file.path(OUT_DIR, "g3_эффект_знаменателя.png"), p3, width = 11, height = 5.5, dpi =
   300)

log_msg("Графики сохранены.")


# ---- 14. Резюме ------------------------------------------------------------
hr("ШАГ 10. РЕЗЮМЕ")

max_horizon <- nat_fc |> slice_max(прогноз, n = 1, with_ties = FALSE)
cur  <- nat_series$cases[n_nat]

txt <- c(
  "РЕЗЮМЕ КРАТКОСРОЧНОГО ПРОГНОЗА ЗАБОЛЕВАЕМОСТИ ОРВИ И ГРИППОМ",
  "Кыргызская Республика",
  paste0("Дата формирования: ", format(Sys.Date(), "%d.%m.%Y")),
  "",
  "1. ИСХОДНЫЕ ДАННЫЕ",
  paste0("   Период наблюдения: ", format(min(nat_series$week_end), "%d.%m.%Y"), " — ",
         format(last_date, "%d.%m.%Y"), " (", n_nat, " недель, ",
         length(TERRITORIES), " территорий)"),
  paste0("   Актуальность данных: последняя неделя завершилась ", days_stale, " дн. назад"),
  paste0("   Знаменатели населения восстановлены из интенсивных показателей; ",
         "сверка пройдена с расхождением ", formatC(max_err * 100, format = "f", digits =
   4), "%"),
  paste0("   Опорное население КР: ", format(pop_nat_fc, big.mark = " "), " человек"),
  "",
  "2. ТЕКУЩАЯ СИТУАЦИЯ",
  paste0("   Последняя неделя (", nat_series$week[n_nat], "-я нед. ",
         nat_series$year[n_nat], " г.): ", format(cur, big.mark = " "), " случаев (",
         formatC(nat_series$rate_harm[n_nat], format = "f", digits = 1), " на 100 тыс.)"),
  paste0("   Недельный темп прироста: ", growth$значение[1], "% (95% ДИ ",
         growth$значение[2], "…", growth$значение[3], "%)"),
```

```
  paste0("   ", growth$показатель[4], ": ", growth$значение[4]),
  "",
  "3. ПРОГНОЗ",
  paste0("   Рабочая модель (национальный уровень): ", best_nat,
         if (best_nat == "snaive")
           "  [ВНИМАНИЕ: SNaive — эталон сравнения. Его выбор означает, что содержательные
   модели не дали прироста точности; прогноз есть экстраполяция прошлогодней сезонной
   кривой]"
         else ""),
  paste0("   MAE кросс-валидации: ", round(mt_nat$MAE[mt_nat$model == best_nat], 1),
         " случаев; MAPE: ", round(mt_nat$MAPE[mt_nat$model == best_nat], 1), "%",
         "; WAPE: ", round(k_nat_row$WAPE, 1), "%; MASE: ", round(k_nat_row$MASE, 2)),
  paste0("   Покрытие 95% PI: raw ", round(k_nat_row$COV95_raw, 1),
         "%; после калибровки ", round(k_nat_row$COV95_cal, 1), "%"),
  paste0("   Коэффициент калибровки: raw k=", round(k_nat_row$k_raw, 2),
         "; применено k=", round(k_nat, 2),
         if (isTRUE(k_nat_row$cap_hit)) " [ДОСТИГНУТ SAFETY CAP]" else ""),
  paste0("   Горизонт: ", H_FORECAST, " нед., ", format(min(fc_dates), "%d.%m.%Y"), " — ",
         format(max(fc_dates), "%d.%m.%Y")),
  paste0("   Медианный сценарий за период: ", format(sum(nat_fc$прогноз), big.mark = " "),
         " случаев (примерно в половине случаев факт окажется ниже)"),
  paste0("   Ожидаемое (среднее) число за период: ", format(sum(nat_fc$среднее), big.mark =
   " "),
         " случаев — эта оценка пригодна для суммирования и расчёта нагрузки"),
  paste0("   Максимальное медианное значение в пределах 8-недельного горизонта: ",
   max_horizon$неделя, "-я нед. ", max_horizon$год, " г. — ",
         format(max_horizon$прогноз, big.mark = " "), " случаев",
         " (95% интервал ", format(max_horizon$lo95, big.mark = " "), " — ",
         format(max_horizon$hi95, big.mark = " "), ")"),
  paste0("   В интенсивных показателях: ", max_horizon$прогноз_на_100тыс, " на 100 тыс. в
   неделю максимума внутри горизонта"),
  "",
  "4. РЕГИОНАЛЬНЫЕ МОДЕЛИ"
)
for (i in seq_len(nrow(reg_tbl))) {
  txt <- c(txt, paste0("   ", format(reg_tbl$Регион[i], width = 9), " ",
                       format(reg_tbl$Модель[i], width = 14),
                       " MAE ", format(reg_tbl$MAE[i], width = 7),
                       " WAPE ", format(reg_tbl$WAPE_pct[i], width = 5), "%",
                       " MASE ", format(reg_tbl$MASE[i], width = 5),
                       " покрытие(cal) ", format(reg_tbl$Покрытие95_cal_pct[i], width = 5),
   "%",
                       " k=", reg_tbl$k[i], " [", reg_tbl$Надёжность[i], "]"))
}
txt <- c(txt, "",
  paste0("   Максимальное расхождение суммы регионов с национальной моделью: ",
         max_disc, "%"),
  "",
  "5. ОГРАНИЧЕНИЯ",
  "   1) Синдромный учёт: ОРВИ и грипп объединены в один массив; вирусологическая",
  "      структура и штаммовый состав моделью не разделяются.",
  "   2) Фактор обращаемости: прогнозируется число обратившихся в организации",
  "      здравоохранения, а не истинная заболеваемость населения.",
  "   3) Зарегистрированная заболеваемость существенно различается между территориями.",
  "      Различия могут отражать эпидемиологию, обращаемость, доступность помощи и
   практику",
  "      регистрации; вклад факторов по этому ряду не определяется. См.
   07_вариабельность_регистрации.csv.",
  "   4) Официальный знаменатель населения обновляется внутри года (как правило",
  "      с 25-й недели), что создаёт ступеньки в опубликованном интенсивном ряду.",
  "      Для моделирования и анализа динамики использован единый знаменатель.",
  "      См. 01_знаменатели_населения.csv и g3_эффект_знаменателя.png.",
  "   5) Для года, отсутствующего в файле интенсивных показателей, знаменатель",
  "      принят равным последнему официальному — это допущение, а не факт.",
```

```
  "   6) Национальный интервал получен из национальной модели. Границы",
  "      региональных интервалов НЕ суммируются: суммирование предполагает полную",
  "      корреляцию территорий и даёт неоправданно широкий коридор.",
  "   7) Глубина ряда — около трёх сезонов, что является нижней границей для",
  "      сезонных моделей с годовым циклом.",
  "   8) Аномально высокие недели (см. 06_аномальные_недели.csv) основной моделью не
   удаляются",
  "      и могут влиять на сезонную компоненту; sensitivity_2025.csv показывает
   чувствительность.",
  paste0("   9) Праздничные недели (", paste(HOLIDAY_WEEKS, collapse = ", "),
         ") используются как регрессоры только в ARIMA/NNAR с xreg; STL/SNaive/TBATS их
   напрямую не используют."),
  paste0("   10) Максимум внутри 8-недельного горизонта не интерпретируется автоматически
   как сезонный пик."),
  "",
  paste0("Каталог результатов: ", normalizePath(OUT_DIR))
)

writeLines(txt, file.path(OUT_DIR, "резюме_прогноза.txt"), useBytes = TRUE)
cat(txt, sep = "\n")

hr("ГОТОВО")
log_msg("Все результаты в каталоге: ", normalizePath(OUT_DIR))
```

Приложение 3. Структура аналитического отчёта и источники данных для его разделов

Структура составлена по аналитическому отчёту по данным за 39-ю эпидемиологическую неделю 2026 г. (дата формирования расчёта 29.09.2026). Номера файлов — по главе 9.

Раздел отчёта	Источник данных
Титульный лист: номер и год последней недели данных, дата формирования расчёта	03_ряд_национальный.csv (последняя строка); резюме_прогноза.txt
Список сокращений	—
Краткое резюме	Основные выводы разделов 1–6
1. Текущая эпидемиологическая ситуация	—
1.1. Динамика заболеваемости по КР	03_ряд_национальный.csv; 08_темп_прироста.csv
1.2. Заболеваемость по территориям	14_ситуация_текущая.csv
1.3. Интерпретация территориальных различий	07_вариабельность_регистрации.csv; g2_регионы_интенсивные.png
2. Многолетняя динамика и контекст прогноза	05_годовая_динамика.csv; 06_аномальные_недели.csv; sensitivity_2025.csv
3. Проверка предыдущего прогноза по фактическим данным (по КР; 3.1 — по территориям; 3.2 — накопленная проспективная проверка)	Архив предыдущего расчёта: 10_прогноз_национальный.csv, 13_прогноз_региональный.csv; факт — 03_ряд_национальный.csv, 04_ряд_региональный.csv текущего расчёта (глава 12)
4. Прогноз на ближайшие 8 недель	—
4.1. Прогноз по КР (таблица: неделя, дата окончания, медиана, среднее, 95% прогнозный интервал, на 100 тыс.; рисунок)	10_прогноз_национальный.csv; g1_прогноз_национальный.png
4.2. Максимальное значение внутри прогнозируемого периода	10_прогноз_национальный.csv
4.3. Неопределённость прогноза	10_прогноз_национальный.csv (границы интервалов)
5. Прогноз заболеваемости по территориям	13_прогноз_региональный.csv
5.1. Согласованность национального и региональных прогнозов	12_согласование_уровней.csv

Раздел отчёта	Источник данных
6. Насколько можно доверять прогнозу	—
6.1. Национальная модель	00_параметры_расчёта.csv; 09_метрики_национальные.csv
6.2. Ближайшие недели надёжнее дальнего периода	09b_метрики_по_горизонту.csv
6.3. Надёжность территориальных прогнозов	11_карта_моделей_региональных.csv
6.4. Устойчивость к обработке аномального периода	sensitivity_2025.csv
7. Практическая интерпретация для эпидемиологического надзора (7.1 — оценка начала и возможного пика сезонного подъёма; 7.2 — риски и факторы; 7.3 — предложения по усилению профилактических мероприятий и готовности системы здравоохранения)	Результаты разделов 1–6; глава 10 Инструкции
8. Методология прогнозирования	Глава 7 Инструкции; 00_параметры_расчёта.csv
9. Ограничения	Глава 11 Инструкции; раздел «5. Ограничения» файла резюме_прогноза.txt
10. Заключение и дальнейшие действия	Основные выводы отчёта

Приложение 4. Текст шаблона отчёта report_ari_ru.Rmd

Текст шаблона отчёта. Шаблон используется вместе с файлом образца оформления report_style.docx (шрифт Times New Roman, чёрные заголовки, рамки таблиц), который хранится в папке проекта рядом с шаблоном.

```
---
output:
  word_document:
    reference_docx: report_style.docx
params:
  out_dir: "output_v4_2"
  prev_dir: ""
---

```{r setup, include=FALSE}
# ============================================================================
#  ШАБЛОН ОТЧЁТА: краткосрочный прогноз ОРВИ и гриппа (к скрипту forecast_ari_kg_v4_2.R)
#  Отчёт собирается автоматически из файлов результатов скрипта.
#  Перед формированием отчёта (кнопка Knit) заполняется только параметр prev_dir
#  в шапке файла: папка архива ПРЕДЫДУЩЕГО расчёта, например "archive/w39_2026".
#  Если prev_dir пустой, раздел 3 (проверка предыдущего прогноза) не формируется.
# ============================================================================
knitr::opts_chunk$set(echo = FALSE, message = FALSE, warning = FALSE)
for (p in c("readr", "dplyr", "tidyr", "knitr", "flextable")) {
  if (!requireNamespace(p, quietly = TRUE)) install.packages(p)
}
suppressPackageStartupMessages({
  library(readr); library(dplyr); library(tidyr); library(knitr); library(flextable)
})
# Оформление таблиц в Word: Times New Roman 11, тёмно-синяя шапка, чередование строк,
# компактные строки, таблица не разрывается между страницами
NAVY <- "#1F3864"; ZEBRA <- "#EEF2F8"; LINE <- officer::fp_border(color = "#BFBFBF", width =
   0.5)
tab <- function(df, align_left = 1) {
  df <- as.data.frame(df, check.names = FALSE)
  ft <- flextable(df)
  ft <- font(ft, fontname = "Times New Roman", part = "all")
  ft <- fontsize(ft, size = 11, part = "all")
  ft <- bold(ft, part = "header")
  ft <- color(ft, color = "white", part = "header")
  ft <- bg(ft, bg = NAVY, part = "header")
  if (nrow(df) > 1) ft <- bg(ft, i = seq(2, nrow(df), by = 2), bg = ZEBRA, part = "body")
  ft <- align(ft, align = "center", part = "all")   # числа — по центру столбца
  if (length(align_left) && all(align_left > 0)) ft <- align(ft, j = align_left, align =
   "left", part = "all")
  ft <- border_remove(ft)
  ft <- hline(ft, border = LINE, part = "body")
  ft <- hline_top(ft, border = officer::fp_border(color = NAVY, width = 1), part = "header")
  ft <- hline_bottom(ft, border = officer::fp_border(color = NAVY, width = 1), part =
   "body")
  ft <- padding(ft, padding.top = 2, padding.bottom = 2, padding.left = 4, padding.right =
   4, part = "all")
  ft <- line_spacing(ft, space = 1, part = "all")
  ft <- valign(ft, valign = "center", part = "all")
  if (names(df)[1] == "Территория") {
    # Таблицы по территориям: название территории в одну строку, остальные столбцы поровну
    ft <- set_table_properties(ft, layout = "fixed")
    ft <- width(ft, j = 1, width = 1.9)
    ft <- width(ft, j = seq(2, ncol(df)), width = 4.6 / (ncol(df) - 1))
  } else {
    ft <- set_table_properties(ft, layout = "autofit", width = 1)
```
```

```
  }
  paginate(ft, init = nrow(df) <= 8, hdr_ftr = TRUE)
}
# Блок «Основные показатели»: подпись сверху, крупное число снизу
kpi <- function(labels, values) {
  df <- as.data.frame(t(values), stringsAsFactors = FALSE); names(df) <- paste0("v",
   seq_along(values))
  ft <- flextable(df)
  ft <- set_header_labels(ft, values = setNames(as.list(labels), names(df)))
  ft <- font(ft, fontname = "Times New Roman", part = "all")
  ft <- fontsize(ft, size = 10, part = "header"); ft <- fontsize(ft, size = 16, part =
   "body")
  ft <- color(ft, color = "#595959", part = "header"); ft <- color(ft, color = NAVY, part =
   "body")
  ft <- bold(ft, part = "body")
  ft <- bg(ft, bg = ZEBRA, part = "all")
  ft <- align(ft, align = "center", part = "all")
  ft <- border_remove(ft)
  ft <- vline(ft, j = seq_len(length(values) - 1), border = officer::fp_border(color =
   "white", width = 3), part = "all")
  ft <- padding(ft, padding.top = 4, padding.bottom = 4, part = "all")
  ft <- line_spacing(ft, space = 1, part = "all")
  ft <- width(ft, width = 6.5 / length(values))   # ширина текста страницы 16,5 см = 6,5
   дюйма
  set_table_properties(ft, layout = "fixed")
}

OUT  <- params$out_dir
PREV <- params$prev_dir
rd   <- function(f) readr::read_csv(file.path(OUT, f), show_col_types = FALSE)

# Форматирование чисел: пробел между тысячами, запятая в дробной части
fn  <- function(x, d = 0) formatC(as.numeric(x), format = "f", digits = d, big.mark = " ",
   decimal.mark = ",")
fp  <- function(x, d = 1, sign = FALSE) {
  s <- fn(x, d); if (sign) s <- ifelse(as.numeric(x) > 0, paste0("+", s), s); paste0(s, "%")
}
# Согласование слова с числом: 1 случай, 2 случая, 5 случаев; дробные — «случая»
pl <- function(x, f1, f2, f5) {
  x <- as.numeric(x)
  ifelse(x != round(x), f2, {
    n <- abs(round(x)) %% 100; n1 <- n %% 10
    ifelse(n >= 11 & n <= 14, f5, ifelse(n1 == 1, f1, ifelse(n1 >= 2 & n1 <= 4, f2, f5)))
  })
}
cs <- function(x, d = 0) paste(fn(x, d), pl(round(as.numeric(x), d), "случай", "случая",
   "случаев"))
fdate <- function(x) format(as.Date(x), "%d.%m.%Y")
num  <- function(s) as.numeric(gsub(",", ".", s))

# Полные названия территорий
TERR <- c("Бишкек" = "г. Бишкек", "Ош" = "г. Ош", "Чуй" = "Чуйская область",
          "Талас" = "Таласская область", "ИК" = "Иссык-Кульская область",
          "Нарын" = "Нарынская область", "Ош обл" = "Ошская область",
          "Баткен" = "Баткенская область", "Ж-Абад" = "Жалал-Абадская область")
tname <- function(x) ifelse(x %in% names(TERR), TERR[x], x)
# Названия в предложном падеже: «в г. Бишкек», «в Чуйской области»
LOC <- c("Бишкек" = "г. Бишкек", "Ош" = "г. Ош", "Чуй" = "Чуйской области",
         "Талас" = "Таласской области", "ИК" = "Иссык-Кульской области",
         "Нарын" = "Нарынской области", "Ош обл" = "Ошской области",
         "Баткен" = "Баткенской области", "Ж-Абад" = "Жалал-Абадской области")
tloc <- function(x) ifelse(x %in% names(LOC), LOC[x], x)
```

```
# Названия моделей
MOD <- c(snaive = "SNaive", stl_ets = "STL+ETS", arima_fourier = "ARIMA с гармониками
   Фурье",
         tbats = "TBATS", nnar_fourier = "NNAR с гармониками Фурье", ensemble = "ансамбль")
mname <- function(x) ifelse(x %in% names(MOD), MOD[x], x)

# ---- Чтение результатов ----
par   <- rd("00_параметры_расчёта.csv")
pv    <- function(name) par$значение[par$параметр == name]
nat   <- rd("03_ряд_национальный.csv")
reg   <- rd("04_ряд_региональный.csv")
yearly <- rd("05_годовая_динамика.csv")
compl <- rd("07_вариабельность_регистрации.csv")
growth <- rd("08_темп_прироста.csv")
mt    <- rd("09_метрики_национальные.csv")
mth   <- rd("09b_метрики_по_горизонту.csv")
fc    <- rd("10_прогноз_национальный.csv")
rmap  <- rd("11_карта_моделей_региональных.csv")
recon <- rd("12_согласование_уровней.csv")
rfc   <- rd("13_прогноз_региональный.csv")
cur   <- rd("14_ситуация_текущая.csv")
anom  <- if (file.exists(file.path(OUT, "06_аномальные_недели.csv")))
   rd("06_аномальные_недели.csv") else NULL
sens  <- rd("sensitivity_2025.csv")

# ---- Основные величины ----
n     <- nrow(nat)
last  <- nat[n, ]
prev1 <- nat[n - 1, ]
wk    <- last$week; yr <- last$year
ch1   <- (last$cases / prev1$cases - 1) * 100

# Число последовательных недель роста перед последней неделей
streak <- 0
for (i in n:2) { if (nat$cases[i] > nat$cases[i - 1]) streak <- streak + 1 else break }
seq_vals <- nat$cases[(n - streak):n]
ch_streak <- (last$cases / nat$cases[n - streak] - 1) * 100

g_rate <- num(growth$значение[1]); g_lo <- num(growth$значение[2]); g_hi <-
   num(growth$значение[3])
g_period <- growth$значение[5]
g_dir <- if (g_lo > 0) "рост" else if (g_hi < 0) "снижение" else "нет"
g_text <- switch(g_dir,
  "рост"     = "что указывает на сформировавшийся краткосрочный рост регистрируемой
   заболеваемости",
  "снижение" = "что указывает на краткосрочное снижение регистрируемой заболеваемости",
  "нет"      = "устойчивое направление изменения заболеваемости пока не определяется
   (доверительный интервал включает 0)")

best  <- pv("Рабочая модель (КР)")
mbest <- mt[mt$model == best, ][1, ]

fc_first <- fc[1, ]; fc_last <- fc[nrow(fc), ]
fc_max   <- fc[which.max(fc$прогноз), ]
max_is_last <- fc_max$неделя == fc_last$неделя
wk_from <- fc_first$неделя; wk_to <- fc_last$неделя

recon$week_n <- fc$неделя[match(as.Date(recon$week_end), as.Date(fc$week_end))]
disc_max <- recon[which.max(abs(recon$расхождение_pct)), ]
disc_near <- range(recon$расхождение_pct[1:min(3, nrow(recon))])

cur$Терр <- tname(cur$Территория)
```

```
top_abs  <- cur[which.max(cur$Случаев_абс), ]
top_rate <- cur[which.max(cur$На_100тыс_официальный), ]

sens_diff <- if ("abs_разница_pct" %in% names(sens)) max(sens$abs_разница_pct, na.rm = TRUE)
   else NA

# ---- Проверка предыдущего прогноза ----
has_prev <- nzchar(PREV) && file.exists(file.path(PREV, "10_прогноз_национальный.csv"))
if (has_prev) {
  pf  <- readr::read_csv(file.path(PREV, "10_прогноз_национальный.csv"), show_col_types =
   FALSE)
  chk <- pf |> mutate(week_end = as.Date(week_end)) |>
    inner_join(nat |> mutate(week_end = as.Date(week_end)) |> select(week_end, факт =
   cases), by = "week_end") |>
    mutate(разн = прогноз - факт, ош = abs(разн) / факт * 100, в95 = факт >= lo95 & факт <=
   hi95)
  has_prev <- nrow(chk) > 0
}
if (has_prev) {
  ch_mape <- mean(chk$ош); ch_wape <- sum(abs(chk$разн)) / sum(chk$факт) * 100
  ch_bias <- (sum(chk$прогноз) / sum(chk$факт) - 1) * 100
  ch_in   <- sum(chk$в95)
  ch_w    <- paste0(min(chk$неделя), "–", max(chk$неделя))
  prf <- readr::read_csv(file.path(PREV, "13_прогноз_региональный.csv"), show_col_types =
   FALSE)
  rchk <- prf |> mutate(week_end = as.Date(week_end)) |>
    inner_join(reg |> mutate(week_end = as.Date(week_end)) |> select(territory, week_end,
   факт = cases),
               by = c("territory", "week_end")) |>
    group_by(territory) |>
    summarise(прогноз = sum(прогноз_абс), факт_с = sum(факт),
              MAPE = mean(abs(прогноз_абс - факт) / факт * 100),
              в95 = sum(факт >= lo95 & факт <= hi95), n = n(), .groups = "drop")
}

# ---- Журнал проверок: archive/check_log.csv ----
# При каждом формировании отчёта результат текущей проверки записывается в журнал
# (повторное формирование с той же папкой prev_dir обновляет строку, а не дублирует её).
hist_log <- NULL
if (has_prev) {
  log_file <- file.path(dirname(PREV), "check_log.csv")
  rec <- data.frame(prev = basename(PREV), weeks = ch_w, year = max(chk$год), n = nrow(chk),
                    mape = round(ch_mape, 1), wape = round(ch_wape, 1), bias =
   round(ch_bias, 1),
                    in95 = ch_in, made = format(Sys.Date()), stringsAsFactors = FALSE)
  old <- if (file.exists(log_file)) as.data.frame(readr::read_csv(log_file, show_col_types =
   FALSE,
                                                                  col_types = readr::cols(.default =
   "c"))) else NULL
  if (!is.null(old)) old <- old[old$prev != rec$prev, , drop = FALSE]
  hist_log <- rbind(if (!is.null(old)) old[, names(rec)] else NULL, data.frame(lapply(rec,
   as.character)))
  hist_log$w1 <- as.integer(sub("–.*", "", hist_log$weeks))
  hist_log <- hist_log[order(as.integer(hist_log$year), hist_log$w1), names(rec)]
  readr::write_excel_csv(hist_log, log_file)
  for (v in c("n", "in95")) hist_log[[v]] <- as.integer(hist_log[[v]])
  for (v in c("mape", "wape", "bias")) hist_log[[v]] <-
   suppressWarnings(as.numeric(hist_log[[v]]))
}

# Вспомогательное: диапазон недель «48–51-я» и тысячи «8,2 тыс.»
wrange <- function(w) { w <- sort(unique(w)); if (length(w) > 1 && all(diff(w) == 1))
   paste0(min(w), "–", max(w), "-й") else paste0(paste(w, collapse = ", "), "-й") }
```

```
tys <- function(x) paste0(fn(x / 1000, 1), " тыс.")
# Список территорий в предложном падеже: «г. Бишкек, г. Ош, Чуйской и Ошской областях»
loc_list <- function(x) {
  l <- tloc(x); is_obl <- grepl(" области$", l); cities <- l[!is_obl]; obl <- sub("
   области$", "", l[is_obl])
  obl_part <- if (length(obl) == 1) paste0(obl, " области") else if (length(obl) > 1)
    paste0(paste(head(obl, -1), collapse = ", "), " и ", tail(obl, 1), " областях") else
   NULL
  if (length(obl) == 0 && length(cities) > 1) return(paste0(paste(head(cities, -1), collapse
   = ", "), " и ", tail(cities, 1)))
  paste(c(cities, obl_part), collapse = ", ")
}
sens_w <- if ("abs_разница_pct" %in% names(sens)) sens$неделя[sens$abs_разница_pct > 10]
   else integer(0)
sens_wmax <- if ("abs_разница_pct" %in% names(sens))
   sens$неделя[which.max(sens$abs_разница_pct)] else NA
anom_last <- if (!is.null(anom) && nrow(anom)) anom[anom$year == max(anom$year), ] else NULL
scen <- length(sens_w) > 0 && !is.null(anom_last)
stab_w <- if (scen) fc$неделя[!(fc$неделя %in% sens_w)] else fc$неделя
stab_sum <- sum(fc$среднее[fc$неделя %in% stab_w])
stab_max <- if (scen && "abs_разница_pct" %in% names(sens))
   max(sens$abs_разница_pct[sens$неделя %in% stab_w], na.rm = TRUE) else NA
wlist <- function(w) { w <- sort(unique(w)); if (length(w) > 1 && all(diff(w) == 1))
   paste0(min(w), "–", max(w)) else paste(w, collapse = ", ") }
ordw <- function(k) if (k <= 10) c("впервые", "во второй раз", "в третий раз", "в четвёртый
   раз", "в пятый раз", "в шестой раз",
                                    "в седьмой раз", "в восьмой раз", "в девятый раз", "в
   десятый раз")[k] else paste0("в ", k, "-й раз")
```

::: {custom-style="Титул организация"}
НАЦИОНАЛЬНЫЙ ИНСТИТУТ ОБЩЕСТВЕННОГО ЗДОРОВЬЯ\
МИНИСТЕРСТВА ЗДРАВООХРАНЕНИЯ КЫРГЫЗСКОЙ РЕСПУБЛИКИ
:::

::: {custom-style="Титул название"}
Заболеваемость ОРВИ и гриппом\
в Кыргызской Республике
:::

::: {custom-style="Титул подзаголовок"}
Аналитический отчёт по краткосрочному прогнозированию на 8 недель
:::

::: {custom-style="Титул сведения"}
Расчёт по данным по `r wk`-ю эпидемиологическую неделю `r yr` г. включительно
:::

::: {custom-style="Титул сведения"}
Прогнозный период: `r wk_from`–`r wk_to`-я недели (`r fdate(as.Date(fc_first$week_end) - 6)`
   — `r fdate(fc_last$week_end)`)
:::

::: {custom-style="Титул сведения"}
Дата формирования отчёта: `r format(Sys.Date(), "%d.%m.%Y")`
:::

::: {custom-style="Титул место"}
Бишкек, `r format(Sys.Date(), "%Y")`
:::

```{=openxml}


```

```
<w:p><w:r><w:br w:type="page"/></w:r></w:p>
```

Список сокращений

```{r}
data.frame(`Сокращение` = c("ДИ", "КР", "ОРВИ", "ARIMA", "MAE", "MAPE", "MASE", "RMSE",
   "SNaive", "STL+ETS", "TBATS", "WAPE"),
           `Расшифровка` = c("доверительный интервал", "Кыргызская Республика",
                             "острые респираторные вирусные инфекции",
                             "авторегрессионная интегрированная модель скользящего среднего
   (AutoRegressive Integrated Moving Average)",
                             "средняя абсолютная ошибка (Mean Absolute Error)",
                             "средняя абсолютная процентная ошибка (Mean Absolute Percentage
   Error)",
                             "средняя абсолютная масштабированная ошибка (Mean Absolute
   Scaled Error)",
                             "среднеквадратическая ошибка (Root Mean Squared Error)",
                             "сезонная наивная модель (Seasonal Naive)",
                             "модель, сочетающая сезонно-трендовую декомпозицию STL и
   экспоненциальное сглаживание ETS",
                             "модель с тригонометрической сезонностью, преобразованием Бокса
   — Кокса, ARMA-ошибками, трендом и сезонными компонентами",
                             "взвешенная абсолютная процентная ошибка (Weighted Absolute
   Percentage Error)"),
           check.names = FALSE) |> tab(1:2) |> set_table_properties(layout = "fixed") |>
   width(j = 1, width = 1.3) |> width(j = 2, width = 5.2)
```

Основные показатели

```{r}
kpi(c(paste0("Случаев за ", wk, "-ю неделю"), "На 100 тыс. населения", paste0("Изменение к
   ", prev1$week, "-й неделе"), "Недельный темп прироста"),
    c(fn(last$cases), fn(last$rate_harm, 1), fp(ch1, sign = TRUE), fp(g_rate, sign = TRUE)))
```

```{r}
kpi(c(paste0("Прогноз на ", fc$неделя[1], "-ю неделю"), paste0("Прогноз на ", fc$неделя[2],
   "-ю неделю"),
      if (scen) paste0("Ожидаемое число случаев за ", wlist(stab_w), "-ю недели") else
   "Ожидаемое число случаев за 8 недель", if (has_prev) "Ошибка прошлого прогноза (MAPE)"
   else "Рабочая модель"),
    c(fn(fc$прогноз[1]), fn(fc$прогноз[2]), fn(if (scen) stab_sum else sum(fc$среднее)), if
   (has_prev) fp(ch_mape) else mname(best)))
```

::: {custom-style="Примечание"}
Прогнозные значения — медианы; ожидаемое число случаев — сумма средних прогнозных значений`r
   if (scen) paste0(" без сценарных ", wlist(sens_w), "-й недель (раздел 7.1)") else ""`.
:::

Краткое резюме

На `r wk`-й эпидемиологической неделе `r yr` года в Кыргызской Республике зарегистрировано
   `r cs(last$cases)` ОРВИ и гриппа, или `r fn(last$rate_harm, 1)` случая на 100 тыс.
   населения. По сравнению с `r prev1$week`-й неделей число зарегистрированных случаев `r if
   (ch1 >= 0) "увеличилось" else "уменьшилось"` на `r fp(abs(ch1))`. За период `r g_period`
   расчётный недельный темп прироста составляет `r fp(g_rate)` (95% ДИ `r fn(g_lo,1)`–`r
   fn(g_hi,1)`%), `r g_text`.

Наибольшее абсолютное число случаев отмечено в `r tloc(top_abs$Территория)` — `r fn(top_abs$Случаев_абс)`, наиболее высокий интенсивный показатель — в `r tloc(top_rate$Территория)`: `r fn(top_rate$На_100тыс_официальный, 1)` на 100 тыс. населения.

`r if (has_prev) paste0("Проспективная проверка прогноза предыдущего расчёта для ", ch_w, "-й недель: MAPE национального прогноза ", fp(ch_mape), ", суммарное ", if (ch_bias >= 0) "завышение" else "занижение", " — ", fp(abs(ch_bias)), ". Внутри ранее рассчитанных 95% прогнозных интервалов находились ", ch_in, " из ", nrow(chk), " фактических значений по КР.") else ""`

По прогнозу на `r wk_from`–`r wk_to`-ю недели медианное число зарегистрированных случаев составляет `r fn(fc$прогноз[1])` на `r fc$неделя[1]`-й неделе, `r fn(fc$прогноз[2])` на `r fc$неделя[2]`-й и `r fn(fc$прогноз[3])` на `r fc$неделя[3]`-й. Наиболее высокое медианное значение среди прогнозируемых восьми недель приходится на `r fc_max$неделя`-ю неделю — `r cs(fc_max$прогноз)`, или `r fn(fc_max$прогноз_на_100тыс, 1)` на 100 тыс. населения. `r if (scen) paste0("Значения на ", wrange(sens_w), " неделях в значительной мере повторяют необычный подъём ", wrange(anom_last$week), " недель ", anom_last$year[1], " года и рассматриваются как возможный сценарий, а не как ожидаемый уровень (раздел 7.1). Наиболее обоснованным является прогноз на ближайшие две недели.") else paste0("Это значение не следует считать сезонным пиком, поскольку период после ", wk_to, "-й недели текущим расчётом не охвачен.")`

Максимальное расхождение между суммой средних региональных прогнозов и национальным прогнозом до согласования составляет `r fp(disc_max$расхождение_pct)` (`r disc_max$week_n`-я неделя).

1. Текущая эпидемиологическая ситуация

1.1. Динамика заболеваемости по Кыргызской Республике

Последняя включённая в расчёт неделя заканчивается `r fdate(last$week_end)`. За `r wk`-ю неделю зарегистрировано `r cs(last$cases)` ОРВИ и гриппа. Интенсивный показатель составил `r fn(last$rate_harm, 1)` на 100 тыс. населения. По сравнению с `r prev1$week`-й неделей число случаев `r if (ch1 >= 0) "увеличилось" else "уменьшилось"` на `r fp(abs(ch1))`.`r if (streak >= 2) paste0(" Число случаев увеличивается ", streak, " ", pl(streak, "неделю", "недели", "недель"), " подряд: с ", nat$week[n - streak], "-й по ", wk, "-ю неделю — ", paste(fn(seq_vals), collapse = " → "), " случаев (", fp(ch_streak, sign = TRUE), ").") else ""`

За последние шесть недель расчётный недельный темп прироста составляет `r fp(g_rate)` при 95% ДИ `r fn(g_lo,1)`–`r fn(g_hi,1)`%, `r g_text`. `r d4 <- suppressWarnings(num(growth$значение[4])); if (is.na(d4)) "Время удвоения не определено: устойчивого направления изменения нет." else paste0(sub(", недель", "", growth$показатель[4]), " — около ", fn(d4, 1), " ", pl(1.5, "недели", "недели", "недель"), ".")`

1.2. Заболеваемость по территориям

```{r}
cur |> transmute(Территория = Терр, `Случаев, абс.` = fn(Случаев_абс),
                 `На 100 тыс.` = fn(На_100тыс_официальный, 1),
                 `К предыдущей неделе` = fp(Динамика_к_прош_нед_pct, sign = TRUE))  |>
  tab()
```

```{r}
up   <- cur$Терр[cur$Динамика_к_прош_нед_pct > 0]
down <- cur$Терр[cur$Динамика_к_прош_нед_pct < 0]
```

Наибольшее абсолютное число случаев зарегистрировано в `r tloc(top_abs$Территория)` — `r fn(top_abs$Случаев_абс)`. Наиболее высокий интенсивный показатель отмечен в `r tloc(top_rate$Территория)` — `r fn(top_rate$На_100тыс_официальный, 1)` на 100 тыс. населения. `r if (length(down) == 0) "Рост к предыдущей неделе зарегистрирован во всех территориях." else if (length(up) == 0) "Снижение к предыдущей неделе зарегистрировано во всех территориях." else paste0("Рост к предыдущей неделе зарегистрирован в ", length(up), " территориях, снижение — в ", paste(tloc(cur$Территория[cur$Динамика_к_прош_нед_pct < 0]), collapse = ", "), ".")`

1.3. Интерпретация территориальных различий

Интенсивные показатели существенно различаются между территориями: за `r compl$год[1]` год зарегистрированная заболеваемость в территории с наибольшим показателем в `r fn(max(compl$кратность_к_минимуму), 1)` раза выше, чем в территории с наименьшим. Такой разброс нельзя автоматически трактовать как различие истинного риска заражения. Он может одновременно отражать эпидемиологическую ситуацию, обращаемость населения, доступность медицинской помощи и особенности регистрации. Поэтому динамику необходимо оценивать отдельно по каждой территории, сопоставляя текущие показатели с предыдущими неделями.

2. Многолетняя динамика и контекст прогноза

```{r}
full_y <- yearly[yearly$недель >= 52, ]; part_y <- yearly[yearly$недель < 52, ]
```

Временной ряд охватывает период с `r fdate(nat$week_end[1])` по `r fdate(last$week_end)` — `r n` недель наблюдения по `r length(unique(reg$territory))` территориям. `r paste0("В ", full_y$year[1], " году зарегистрировано ", cs(full_y$сумма[1]), if (nrow(full_y) > 1) paste0(", ", paste0("в ", full_y$year[-1], " году — ", fn(full_y$сумма[-1]), collapse = ", ")) else "")`.`r if (nrow(part_y)) paste0(" За первые ", part_y$недель, " ", pl(part_y$недель, "неделю", "недели", "недель"), " ", part_y$year, " года зарегистрировано ", cs(part_y$сумма), ".") else ""`

`r if (!is.null(anom) && nrow(anom)) paste0("В ряду выявлены недели с необычно высокими значениями (более чем в 3 раза выше медианы предыдущих 52 недель): ", paste(sapply(split(anom$week, anom$year), function(w) { w <- sort(w); if (length(w) > 1 && all(diff(w) == 1)) paste0(min(w), "–", max(w), "-я недели ") else paste0(paste0(w, "-я", collapse = ", "), if (length(w) > 1) " недели " else " неделя ") }), names(split(anom$week, anom$year)), "г.", collapse = "; "), " Эти недели сохранены в основном временном ряду, поскольку по имеющимся данным нельзя однозначно определить, в какой мере рост отражал реальное усиление эпидемического процесса, изменение обращаемости или особенности регистрации.") else "Недель с необычно высокими значениями (более чем в 3 раза выше медианы предыдущих 52 недель) не выявлено."` `r if (!is.na(sens_diff)) paste0("Альтернативная обработка этих недель изменяет медианный прогноз на величину до ", fp(sens_diff), if (!is.na(sens_wmax)) paste0(" (наибольшее различие — на ", sens_wmax, "-й неделе)") else "", ".") else ""`

3. Проверка предыдущего прогноза по фактическим данным

```{r, results='asis'}
if (!has_prev) {
  cat("Папка предыдущего расчёта не указана (параметр prev_dir) или не содержит недель, по которым уже поступили фактические данные. Проверка предыдущего прогноза в этом отчёте не проводилась.\n")
} else {
  cat("Прогноз предыдущего расчёта сопоставлен с фактическими данными за ", ch_w,
      "-ю недели. Эти фактические значения не использовались при построении предыдущего прогноза.\n\n", sep = "")
  flextable_to_rmd(tab(chk |> transmute(Неделя = неделя, `Предыдущий прогноз` = fn(прогноз), Факт = fn(факт),
                                        `Прогноз – факт` = ifelse(разн > 0, paste0("+", fn(разн)), fn(разн)),
```

```
                              `Абс. ошибка, %` = fp(ош),
                              `В 95% интервале` = ifelse(в95, "да", "нет")), 0))
cat("\nЗа ", ch_w, "-ю недели прогнозировалось ", cs(sum(chk$прогноз)), ", фактически
 зарегистрировано ",
    fn(sum(chk$факт)), ". Суммарное ", if (ch_bias >= 0) "завышение" else "занижение", "
 составило ",
    cs(abs(sum(chk$разн))), ", или ", fp(abs(ch_bias)), " относительно факта. MAPE
 составила ", fp(ch_mape),
    ", WAPE — ", fp(ch_wape), ". Внутри ранее рассчитанных 95% прогнозных интервалов
 находились ", ch_in,
    " из ", nrow(chk), " фактических значений.\n\n", sep = "")
cat("## 3.1. Проверка по территориям\n\n")
flextable_to_rmd(tab(rchk |> arrange(desc(факт_с)) |>
             transmute(Территория = tname(territory), `Прогноз` = fn(прогноз), Факт =
 fn(факт_с),
                       Разница = ifelse(прогноз - факт_с > 0, paste0("+", fn(прогноз -
 факт_с)), fn(прогноз - факт_с)),
                       `Разница, %` = fp((прогноз - факт_с) / факт_с * 100, sign = TRUE),
                       MAPE = fp(MAPE), `В 95% интервале` = paste0(в95, "/", n))) |>
                  width(j = 1, width = 1.9) |> width(j = 2:6, width = 0.72) |> width(j =
 7, width = 1.0))
cat("\nВнутри ранее рассчитанных 95% прогнозных интервалов находились ", sum(rchk$в95), "
 из ", sum(rchk$n),
    " фактических территориальных значений.\n\n", sep = "")
cat("## 3.2. Накопленная проспективная проверка\n\n")
h <- hist_log
item <- paste0(fp(h$mape), " для ", h$weeks, "-й недель",
               ifelse(is.na(h$bias), "", ifelse(h$bias >= 0, paste0(" (с завышением на ",
 fp(abs(h$bias)), ")"),
                                                paste0(" (с занижением на ",
 fp(abs(h$bias)), ")"))))
cat("Проспективная проверка, то есть сравнение ранее выпущенного прогноза с поступившими
 фактическими данными, ",
    if (nrow(h) == 1) "проведена впервые. " else paste0("проведена по ", nrow(h), "
 последовательным ", pl(nrow(h), "циклу", "циклам", "циклам"), " расчёта. "),
    "Средняя ошибка национального прогноза (MAPE) составила ", paste(item, collapse = ",
 "), ".\n\n", sep = "")
tot_n <- sum(h$n); tot_in <- sum(h$in95)
p2 <- if (tot_in == tot_n) paste0("За все ", tot_n, " проверенных ", pl(tot_n, "неделю",
 "недели", "недель"),
                                  " фактическое число случаев по КР находилось в пределах
 95% прогнозного интервала, то есть диапазона, в который фактическое значение должно
 попасть с вероятностью 95%. Интервалы на дальних горизонтах широкие, поэтому попадание в
 них не означает высокой точности центрального прогноза.")
     else paste0("Из ", tot_n, " проверенных недель фактическое число случаев по КР
 находилось в пределах 95% прогнозного интервала (диапазона, в который фактическое
 значение должно попасть с вероятностью 95%) в ", tot_in, " случаях.")
if (nrow(h) >= 2) {
  m <- h$mape; k <- nrow(h)
  trend <- if (k >= 3 && all(diff(m) > 0)) "ошибка прогноза увеличивалась от проверки к
 проверке (для вывода об устойчивом ухудшении модели число проведённых проверок пока
 недостаточно)" else
           if (m[k] > m[k - 1] + 2) "ошибка прогноза увеличилась по сравнению с предыдущей
 проверкой" else
           if (m[k] < m[k - 1] - 2) "ошибка прогноза уменьшилась по сравнению с предыдущей
 проверкой" else
           "ошибка прогноза сопоставима с предыдущей проверкой"
  b <- h$bias
  flip <- !is.na(b[k]) && !is.na(b[k - 1]) && sign(b[k]) != sign(b[k - 1])
  p2 <- paste0(p2, " Вместе с тем ", trend,
               if (flip && b[k] < 0) ", а её направление сменилось с завышения на
 занижение: на последних неделях рост заболеваемости опережал расчётный" else
               if (flip && b[k] > 0) ", а её направление сменилось с занижения на
 завышение: на последних неделях заболеваемость росла медленнее расчётной" else "", ".")
```

```
  }
  rr <- rchk |> mutate(rel = (прогноз - факт_с) / факт_с * 100)
  under <- rr |> filter(rel <= -20) |> arrange(rel) |> head(2)
  over  <- rr |> filter(rel >= 20) |> arrange(desc(rel)) |> head(2)
  outi  <- rr |> filter(в95 < n)
  cat(p2, "\n\n", sep = "")
  big <- rr |> filter(MAPE >= 50)
  terr <- c(
    if (nrow(under) || nrow(over)) paste0("Смещение центрального прогноза (занижение или
   завышение суммы за проверяемый период; для вывода о систематическом смещении нужны
   последующие проверки): ",
      paste(c(if (nrow(under)) paste0("занижение — в ", paste0(tloc(under$territory), " (на
   ", fp(abs(under$rel)), ")", collapse = " и ")),
              if (nrow(over)) paste0("завышение — в ", paste0(tloc(over$territory), " (на ",
   fp(over$rel), ")", collapse = " и "))), collapse = "; "), "."),
    if (nrow(big)) paste0("Наибольшая средняя ошибка недельных значений (MAPE 50% и более):
   ",
                          paste0(tname(big$territory), " (", fp(big$MAPE), ")", collapse =
   ", "), "."),
    if (nrow(outi)) paste0("Покрытие прогнозных интервалов: фактические значения вышли за
   пределы 95% интервала ",
                           paste0("в ", tloc(outi$territory), " — ", outi$n - outi$в95, " из
   ", outi$n, collapse = "; "), ".")
             else "Покрытие прогнозных интервалов: все фактические территориальные значения
   находились в пределах 95% интервалов."
  )
  cat("Результаты проверки по территориям:\n\n", paste0("- ", terr, collapse = "\n"), "\n",
   sep = "")
}
```

4. Прогноз на ближайшие 8 недель

4.1. Прогноз динамики заболеваемости по Кыргызской Республике

```{r}
fc |> transmute(Неделя = неделя, `Дата окончания` = fdate(week_end), Медиана = fn(прогноз),
                Среднее = fn(среднее), `95% прогнозный интервал` = paste0(fn(lo95), "–",
   fn(hi95)),
                `На 100 тыс.` = fn(прогноз_на_100тыс, 1))  |>
  tab()
```

После фактических `r cs(last$cases)` на `r wk`-й неделе медианный прогноз на `r
   fc$неделя[1]`-ю неделю составляет `r cs(fc$прогноз[1])`. Сумма медианных недельных
   значений за восемь недель составляет `r cs(sum(fc$прогноз))`, сумма средних прогнозных
   значений — `r fn(sum(fc$среднее))``r if (scen) paste0(" (в том числе за ", wlist(stab_w),
   "-ю недели — ", fn(stab_sum), "; остальное приходится на сценарные ", wlist(sens_w), "-ю
   недели)") else ""`. Медиана показывает центральное расчётное значение числа случаев для
   каждой недели. Среднее прогнозное значение учитывает возможные значения и их вероятности
   и используется для оценки общего ожидаемого числа зарегистрированных случаев за весь
   период.

```{r, fig.cap = "Рисунок 1. Наблюдаемая динамика и национальный прогноз на 8 недель",
   out.width = "6.5in"}
# Копия рисунка с латинским именем: так рисунок надёжно вставляется в Word
g1 <- file.path(tempdir(), "g1_forecast.png")
invisible(file.copy(file.path(OUT, "g1_прогноз_национальный.png"), g1, overwrite = TRUE))
knitr::include_graphics(g1)
```

4.2. Максимальное значение внутри прогнозируемого периода

Наиболее высокое медианное значение среди прогнозируемых восьми недель приходится на `r fc_max$неделя`-ю эпидемиологическую неделю — `r cs(fc_max$прогноз)`, или `r fn(fc_max$прогноз_на_100тыс, 1)` на 100 тыс. населения. `r if (scen) "Это значение в значительной мере повторяет необычный подъём прошлого года и рассматривается как сценарий (раздел 7.1). " else ""``r if (max_is_last) paste0(if (scen) "Кроме того, текущий" else "Это значение не является оценкой сезонного пика: текущий", " расчёт заканчивается ", fdate(fc_last$week_end), " и не охватывает последующий период сезона", if (scen) ", поэтому срок сезонного пика по нему не определяется." else ".") else "Максимальное значение внутри восьминедельного периода не является оценкой сезонного пика."`

4.3. Неопределённость прогноза

По мере удаления от последней фактической недели прогнозный диапазон расширяется. Для `r fc_first$неделя`-й недели медиана составляет `r cs(fc_first$прогноз)`, 95% прогнозный интервал — `r fn(fc_first$lo95)`–`r fn(fc_first$hi95)`. Для `r fc_last$неделя`-й недели медиана составляет `r cs(fc_last$прогноз)`, а 95% интервал — `r fn(fc_last$lo95)`–`r fn(fc_last$hi95)`. Поэтому конкретные значения на поздних неделях нельзя использовать как точные ожидаемые числа; они прежде всего показывают возможное направление дальнейшего изменения заболеваемости.`r if (fc_last$hi95 / max(fc_last$lo95, 1) > 10) paste0(" Для ", fc_last$неделя, "-й недели верхняя граница интервала более чем в 10 раз превышает нижнюю, поэтому такой диапазон для количественного планирования ресурсов непригоден.") else ""`

5. Прогноз заболеваемости по территориям

```{r}
rt <- rfc |> group_by(territory) |>
  summarise(first = прогноз_абс[1], sum_med = sum(прогноз_абс),
            mx = max(прогноз_абс), mx_w = неделя[which.max(прогноз_абс)], .groups = "drop")
   |>
  arrange(desc(sum_med))
rt |> transmute(Территория = tname(territory), !!paste0(wk_from, "-я неделя") := fn(first),
                !!paste0("Сумма медиан ", wk_from, "–", wk_to) := fn(sum_med),
                `Макс. медиана` = fn(mx), Неделя = mx_w)  |>
  tab()
```

Наибольшее число зарегистрированных случаев в течение прогнозируемого периода ожидается в следующих территориях: `r paste(tname(rt$territory[1:3]), collapse = ", ")`.`r near4s <- rfc |> filter(неделя %in% fc$неделя[1:4]) |> group_by(territory) |> summarise(s = sum(прогноз_абс), .groups = "drop") |> arrange(desc(s)); if (scen) paste0(" Эти суммы включают сценарные значения ", wrange(sens_w), " недель (раздел 7.1); на ближайшие четыре недели наибольшее число случаев ожидается в ", loc_list(near4s$territory[1:3]), ".") else ""` `r w_mode <- as.integer(names(which.max(table(rt$mx_w)))); paste0("Для большинства территорий наиболее высокие медианные значения внутри периода приходятся на ", w_mode, "-ю неделю.")` `r oth <- rt[rt$mx_w != w_mode,]; if (nrow(oth)) paste0("Исключения: ", paste0(tname(oth$territory), " (", oth$mx_w, "-я неделя)", collapse = ", "), ".") else ""` Сумма медиан в таблице приводится для сравнения территорий между собой и не является ожидаемым совокупным числом случаев: для такой оценки используется сумма средних прогнозных значений (раздел 4.1).

5.1. Согласованность национального и региональных прогнозов

```{r}
recon |> transmute(Неделя = week_n, `Сумма регионов до согласования` = fn(сумма_регионов),
                   `Национальный прогноз` = fn(национальный), Расхождение =
   fp(расхождение_pct))  |>
  tab()
```

Национальный прогноз и прогнозы по территориям рассчитываются независимо, поскольку для разных временных рядов могут выбираться разные модели. После этого средние региональные прогнозные значения согласуются с национальным прогнозом.

`r if (disc_max$week_n %in% fc$неделя[1:3]) paste0("В текущем расчёте расхождение на ближайшие три недели составляет от ", fp(disc_near[1]), " до ", fp(disc_near[2]), "; наибольшее значение за весь прогнозный период (", fp(disc_max$расхождение_pct), ") приходится на ", disc_max$week_n, "-ю неделю.") else paste0("В текущем расчёте расхождение на ближайшие три недели составляет от ", fp(disc_near[1]), " до ", fp(disc_near[2]), "; максимальное расхождение — ", fp(disc_max$расхождение_pct), " на ", disc_max$week_n, "-й неделе.")` `r if (abs(disc_max$расхождение_pct) > 25) "Расхождение превышает 25% и требует содержательной проверки эпидемиолога." else ""` Величина расхождения фиксируется при каждом расчёте; если повышенное расхождение повторяется в нескольких последовательных расчётах, проводится дополнительная оценка региональных моделей и процедуры согласования.

6. Насколько можно доверять прогнозу

6.1. Национальная модель

Рабочей моделью на национальном уровне является `r mname(best)`. По результатам скользящей ретроспективной проверки MAE составляет `r fn(mbest$MAE, 1)` случая, MAPE — `r fp(mbest$MAPE)`, WAPE — `r fp(mbest$WAPE)`, MASE — `r fn(mbest$MASE, 2)`. `r if (best == "snaive") "Выбор SNaive означает, что содержательные модели не дали прироста точности; прогноз является продолжением прошлогодней сезонной кривой." else paste0("По MAE модель лучше сезонной наивной модели SNaive на ", fp(mbest$MAE_skill_vs_SNaive_pct), ".")` Исходное покрытие 95% прогнозного интервала составляет `r fp(mbest$COV95_raw)`, после калибровки — `r fp(mbest$COV95_cal)`.

6.2. Ближайшие недели надёжнее дальнего периода

```{r}
hb <- mth |> filter(model == best) |> arrange(h)
pos <- hb$h[hb$MAE_skill_vs_SNaive_pct > 0]
```

`r if (best == "snaive") "Рабочая модель — SNaive, сравнение с ней самой не проводится." else paste0("При прогнозе на одну неделю ошибка ", mname(best), " по MAE ", if (hb$MAE_skill_vs_SNaive_pct[1] > 0) paste0("на ", fp(hb$MAE_skill_vs_SNaive_pct[1]), " ниже") else paste0("на ", fp(abs(hb$MAE_skill_vs_SNaive_pct[1])), " выше"), ", чем у SNaive, на две недели — ", if (hb$MAE_skill_vs_SNaive_pct[2] > 0) paste0("на ", fp(hb$MAE_skill_vs_SNaive_pct[2]), " ниже") else paste0("на ", fp(abs(hb$MAE_skill_vs_SNaive_pct[2])), " выше"), ". ", if (length(pos)) paste0("Преимущество перед SNaive по MAE сохраняется при прогнозе на ", if (length(pos) > 1) paste0(paste(head(pos, -1), collapse = ", "), " и ", tail(pos, 1)) else pos, " ", pl(max(pos), "неделю", "недели", "недель"), " вперёд.") else "Преимущества перед SNaive по MAE ни на одном горизонте нет."` Наиболее уверенно следует использовать прогноз на ближайшие 1–2 недели. На 3–4 недели необходимо учитывать прогнозный интервал и поступающие фактические данные, а на 5–8 недель — прежде всего направление изменения заболеваемости.

6.3. Надёжность территориальных прогнозов

```{r}
rmap |> transmute(Территория = tname(Регион), Модель = mname(Модель), WAPE = fp(WAPE_pct),
                  MASE = fn(MASE, 2), Оценка = Надёжность)  |>
  tab(1:2) |> set_table_properties(layout = "fixed") |>
  width(j = 1, width = 1.8) |> width(j = 2, width = 2.1) |> width(j = 3:4, width = 0.7) |>
   width(j = 5, width = 1.2)
```

`r ok <- tname(rmap$Регион[rmap$Надёжность == "приемлемая"]); lim <- tname(rmap$Регион[rmap$Надёжность == "ограниченная"]); low <- tname(rmap$Регион[rmap$Надёжность == "низкая"]); paste0(if (length(ok)) paste0("Приемлемая надёжность прогноза — ", paste(ok, collapse = ", "), ". ") else "", if

(length(lim)) paste0("Ограниченная — ", paste(lim, collapse = ", "), ". ") else "", if (length(low)) paste0("Низкая — ", paste(low, collapse = ", "), ". ") else "", if (length(lim) + length(low)) paste0("Для территорий с ", if (length(lim) && length(low)) "ограниченной и низкой" else if (length(low)) "низкой" else "ограниченной", " надёжностью особенно важно еженедельное сопоставление прогноза с фактическими данными.") else "")`

6.4. Устойчивость к обработке необычно высоких недель

```{r}
if (scen) sens |> transmute(Неделя = неделя, `Основной прогноз (медиана)` = fn(основной_медиана),
                            `При альтернативной обработке` = fn(sensitivity_медиана),
                            `Различие, %` = fp(разница_медиана_pct, sign = TRUE)) |> tab()
```

`r if (!is.na(sens_diff) && scen && length(stab_w)) paste0("Альтернативная обработка необычно высоких значений практически не влияет на прогноз на ", wlist(stab_w), "-ю недели (различия не более ", fp(stab_max), "), тогда как на ", wlist(sens_w), "-й неделях результаты существенно расходятся — до ", fp(sens_diff), " на ", sens_wmax, "-й неделе. Поэтому прогноз на эти недели рассматривается как сценарий (раздел 7.1).") else if (!is.na(sens_diff)) paste0("Альтернативная обработка необычно высоких значений изменяет медианный прогноз на величину до ", fp(sens_diff), if (!is.na(sens_wmax)) paste0(" (наибольшее различие — на ", sens_wmax, "-й неделе)") else "", ". ", if (sens_diff < 5) "Текущий прогноз устойчив к выбранному способу обработки этого периода." else if (scen && length(stab_w)) paste0("Прогноз на ", wlist(stab_w), "-ю недели относительно устойчив к альтернативной обработке этих наблюдений (различия не более ", fp(stab_max), "), тогда как на ", wlist(sens_w), "-й неделях результаты существенно расходятся (до ", fp(sens_diff), ").") else "Прогноз заметно зависит от способа обработки этого периода; это учитывается при интерпретации.") else "Необычно высоких недель не выявлено; анализ чувствительности не требовался."`

7. Практическая интерпретация для эпидемиологического надзора

Прогноз необходимо пересчитывать после поступления данных за каждую завершившуюся эпидемиологическую неделю. Одновременно с новым расчётом следует проверять предыдущий прогноз на уже поступивших фактических данных отдельно по Кыргызской Республике и каждой территории.

На ближайшие 1–2 недели расчётные значения можно использовать для текущей оценки ожидаемого числа зарегистрированных случаев. На 3–4 недели точность ниже: при оценке ситуации необходимо смотреть не только на центральное расчётное значение, но и на границы прогнозного интервала и фактические данные, поступающие каждую неделю. На 5–8 недель прогноз показывает прежде всего, ожидается рост, снижение или сохранение уровня заболеваемости и, соответственно, возможное изменение нагрузки на организации здравоохранения.

Для организаций здравоохранения прогноз может использоваться как дополнительный ориентир при планировании готовности к возможному увеличению обращаемости. Прогноз числа зарегистрированных случаев не является прямым прогнозом госпитализаций, потребности в койках или других ресурсов.

7.1. Оценка начала и возможного пика сезонного подъёма

```{r}
dbl <- suppressWarnings(num(growth$значение[4]))
n_up <- sum(cur$Динамика_к_прош_нед_pct > 0, na.rm = TRUE)
sens_w <- if ("abs_разница_pct" %in% names(sens)) sens$неделя[sens$abs_разница_pct > 10] else integer(0)
anom_last <- if (!is.null(anom) && nrow(anom)) anom[anom$year == max(anom$year), ] else NULL
near2 <- fc[1:2, ]
```

`r if (g_dir == "рост") paste0("Имеющиеся данные свидетельствуют ", if (streak >= 3) "о начале сезонного подъёма заболеваемости" else "о росте заболеваемости", ". Рост регистрируется ", streak, " ", pl(streak, "неделю", "недели", "недель"), " подряд, недельный темп прироста составляет ", fp(g_rate), " (95% ДИ ", fn(g_lo, 1), "–", fn(g_hi, 1), "%)", if (!is.na(dbl)) paste0(", что соответствует удвоению числа случаев примерно за ", fn(dbl, 0), " ", pl(round(dbl), "неделю", "недели", "недель")) else "", ". Рост отмечен в ", n_up, " из ", nrow(cur), " территорий.") else if (g_dir == "снижение") paste0("Имеющиеся данные свидетельствуют о снижении регистрируемой заболеваемости: недельный темп изменения составляет ", fp(g_rate), " (95% ДИ ", fn(g_lo, 1), "–", fn(g_hi, 1), "%)", if (!is.na(dbl)) paste0(", что соответствует сокращению числа случаев вдвое примерно за ", fn(dbl, 0), " ", pl(round(dbl), "неделю", "недели", "недель")) else "", ". Снижение к предыдущей неделе отмечено в ", nrow(cur) - n_up, " из ", nrow(cur), " территорий.") else "Устойчивое направление изменения заболеваемости в настоящее время не определяется: доверительный интервал недельного темпа прироста включает 0."`

`r if (length(sens_w) && !is.null(anom_last)) paste0("Расчётные значения на ", wrange(sens_w), " неделях (наиболее высокое медианное значение — ", cs(fc_max$прогноз), " на ", fc_max$неделя, "-й неделе) в значительной мере обусловлены необычно высокими значениями ", wrange(anom_last$week), " недель ", anom_last$year[1], " года, которые модель воспроизводит как сезонную закономерность. При альтернативной обработке этих недель медианный прогноз изменяется до ", fp(sens_diff), ". В связи с этим значения на ", wrange(sens_w), " неделях следует рассматривать как сценарий повторения прошлогоднего подъёма, а не как ожидаемый уровень.") else paste0("В пределах текущего прогноза наиболее высокое медианное значение приходится на ", fc_max$неделя, "-ю неделю — ", cs(fc_max$прогноз), ". ", if (max_is_last) "Эта неделя является последней неделей прогноза, поэтому значение нельзя считать оценкой сезонного пика." else "Максимальное значение внутри восьминедельного периода не является оценкой сезонного пика.")` Наиболее обоснованным является прогноз на ближайшие две недели: около `r tys(near2$прогноз[1])` случаев на `r near2$неделя[1]`-й неделе и `r tys(near2$прогноз[2])` на `r near2$неделя[2]`-й. Срок сезонного пика по краткосрочному прогнозу не определяется; оценка уточняется при каждом еженедельном пересчёте.

7.2. Риски и факторы, способные повлиять на эпидемиологическую ситуацию

На точность прогноза могут повлиять изменения циркулирующих респираторных вирусов, интенсивности лабораторного выявления, профилактических и противоэпидемических мероприятий, а также другие факторы, которые непосредственно не включены в используемую статистическую модель. Поэтому прогноз должен интерпретироваться совместно с текущими данными эпидемиологического и лабораторного надзора.

На развитие эпидемиологической ситуации в течение сезона могут влиять как собственно сезонная динамика респираторных инфекций, так и изменения обращаемости населения, доступности медицинской помощи и регистрации случаев. Поэтому изменение числа зарегистрированных случаев не во всех ситуациях эквивалентно изменению истинного числа инфекций в населении.

```{r, results='asis'}
top3 <- cur[order(-cur$На_100тыс_официальный), ][1:3, ]
jump <- cur[!is.na(cur$Динамика_к_прош_нед_pct) & cur$Динамика_к_прош_нед_pct > 100, ]
risks <- c(
  if (g_dir == "рост") paste0(if (g_rate >= 10) "Высокий темп роста заболеваемости" else "Продолжающийся рост заболеваемости",
                              " (", fp(g_rate), " в неделю)", if (streak >= 4) " при уже достигнутом высоком уровне регистрации" else "", "."),
  if (has_prev && ch_bias <= -10) paste0("Занижение роста в последней проверке прогноза (на ", fp(abs(ch_bias)),
                                         "): фактическая нагрузка на организации здравоохранения может превысить центральные прогнозные значения."),
  paste0("Наиболее высокие интенсивные показатели зарегистрированы в ",
    tloc(top3$Территория[1]), " (",
         fn(top3$На_100тыс_официальный[1], 1), " на 100 тыс. населения), ",
    tloc(top3$Территория[2]), " (",
```

```
        fn(top3$На_100тыс_официальный[2], 1), ") и ", tloc(top3$Территория[3]), " (",
    fn(top3$На_100тыс_официальный[3], 1), ")."),
  if (nrow(jump)) paste0("Увеличение числа зарегистрированных случаев за неделю в ",
                         paste0(tloc(jump$Территория), " (в ", fn(1 +
    jump$Динамика_к_прош_нед_pct / 100, 1), " раза)", collapse = ", "),
                         " требует уточнения: оно может отражать как реальный рост, так и
    запоздалую регистрацию."),
  if (abs(disc_max$расхождение_pct) > 25) paste0("Расхождение национального прогноза и суммы
    региональных прогнозов достигает ",
                                                 fp(abs(disc_max$расхождение_pct)), " (",
    disc_max$week_n, "-я неделя), что снижает надёжность территориальных оценок на дальних
    неделях.")
)
cat(paste0("- ", risks, collapse = "\n"), "\n")
```

7.3. Предложения по усилению профилактических мероприятий и готовности системы здравоохранения

В рамках компетенции НИОЗ предлагается продолжить еженедельное обновление прогноза после поступления данных за завершившуюся эпидемиологическую неделю, одновременно проводить проспективную проверку предыдущего прогноза и контроль согласованности национального и региональных расчётов.

```{r, results='asis'}
lim_r <- rmap$Регион[rmap$Надёжность != "приемлемая"]
near4 <- rfc |> filter(неделя %in% fc$неделя[1:4]) |> group_by(territory) |> summarise(s =
    sum(прогноз_абс), .groups = "drop") |> arrange(desc(s))
props <- c(
  paste0(if (g_dir == "рост") "Обеспечить готовность организаций здравоохранения к росту
    обращаемости в ближайшие 2–4 недели, в первую очередь в " else
         "Поддерживать готовность организаций здравоохранения к обращаемости в ближайшие 2–4
    недели, в первую очередь в ",
         loc_list(near4$territory[1:4]), ", где в этот период ожидается наибольшее число
    случаев."),
  if (length(lim_r)) paste0("Усилить еженедельный мониторинг в ", loc_list(lim_r), ", где
    надёжность прогноза ограничена."),
  if (nrow(jump)) paste0("Уточнить полноту и своевременность регистрации случаев в ",
    loc_list(jump$Территория), "."),
  "Сопоставлять результаты прогнозирования с данными лабораторного надзора и сведениями о
    госпитализации."
)
cat(paste0("- ", props, collapse = "\n"), "\n")
```

8. Методология прогнозирования

Алгоритм моделирует недельные абсолютные числа зарегистрированных случаев ОРВИ и гриппа. Для стабилизации дисперсии используется логарифмическое преобразование, годовая сезонность задаётся периодом 52 недели. На национальном уровне сравниваются сезонная наивная модель SNaive, STL+ETS, ARIMA с гармониками Фурье, TBATS, нейросетевая авторегрессия с гармониками Фурье и ансамбль моделей; на территориальном — SNaive, STL+ETS, ARIMA с гармониками Фурье и ансамбль.

Качество моделей проверяется скользящей кросс-валидацией (число отсечений на национальном уровне — `r pv("Число отсечений кросс-валидации (КР)")`, период `r pv("Период кросс-валидации (КР)")`) с отдельной оценкой горизонтов от 1 до `r pv("Горизонт оценки при отборе, нед.")` недель. Используются показатели MAE, RMSE, WAPE, MASE, смещение и фактическое покрытие прогнозных интервалов. При выборе модели учитывается исходное покрытие 95% прогнозного интервала не ниже 85%; калиброванное покрытие используется как диагностический показатель. Коэффициент калибровки интервалов национальной модели — `r gsub("\\.", ",", pv("Коэффициент калибровки применённый k (КР)"))`.

Основной недельный прогноз представлен медианой прогнозного распределения. Для оценки общего ожидаемого числа случаев за несколько недель используется среднее прогнозное значение. Прогнозы по территориям сначала рассчитываются независимо, после чего их средние значения и прогнозные интервалы согласуются с общим прогнозом по Кыргызской Республике.

9. Ограничения

- Моделируется объединённый массив зарегистрированных случаев ОРВИ и гриппа без разделения по этиологическим агентам и штаммам.
- Прогноз отражает зарегистрированную обращаемость, а не истинное число инфекций в населении.
- Межтерриториальные различия могут быть связаны не только с эпидемиологической ситуацией, но и с обращаемостью, доступностью медицинской помощи и практикой регистрации.
- Временной ряд содержит около `r floor(n / 52)` полных сезонных `r pl(floor(n / 52), "цикла", "циклов", "циклов")`, что ограничивает устойчивость оценки многолетней сезонности.
- Для года, отсутствующего в файле интенсивных показателей, знаменатель населения принят равным последнему официальному.
- В модель не включены вирусологические данные, метеорологические факторы и будущие управленческие воздействия.
- По мере увеличения срока прогнозирования интервалы быстро расширяются; поздние недельные значения нельзя рассматривать как точные числа случаев.
- Максимальное значение внутри восьминедельного периода не является оценкой сезонного пика.
- Расхождение между национальным прогнозом и суммой независимых региональных прогнозов в текущем расчёте достигает `r fp(abs(disc_max$расхождение_pct))`. Итоговые региональные значения согласуются с национальным прогнозом.

10. Заключение и дальнейшие действия

По состоянию на `r wk`-ю эпидемиологическую неделю `r yr` года в Кыргызской Республике зарегистрировано `r cs(last$cases)` ОРВИ и гриппа, или `r fn(last$rate_harm, 1)` на 100 тыс. населения; по сравнению с предыдущей неделей число случаев `r if (ch1 >= 0) "увеличилось" else "уменьшилось"` на `r fp(abs(ch1))`. Недельный темп прироста за последние шесть недель — `r fp(g_rate)` (95% ДИ `r fn(g_lo,1)`–`r fn(g_hi,1)`%).

Прогноз на `r wk_from`–`r wk_to`-ю недели: медиана `r cs(fc_first$прогноз)` на `r fc_first$неделя`-й неделе и `r fn(fc_last$прогноз)` на `r fc_last$неделя`-й`r if (scen && fc_last$неделя %in% sens_w) " (сценарное значение, см. раздел 7.1)" else ""`; ожидаемое число случаев за весь период — `r fn(sum(fc$среднее))``r if (scen) paste0(". Эта сумма включает сценарные значения ", wlist(sens_w), "-й недель и поэтому также существенно зависит от развития заболеваемости в конце периода; ожидаемое число случаев за ", wlist(stab_w), "-ю недели — ", fn(stab_sum)) else ""`.

После поступления каждой новой эпидемиологической недели выполняются три взаимосвязанные процедуры: обновление восьминедельного прогноза; проспективная проверка предыдущего прогноза по КР и всем территориям; контроль согласованности национального и региональных прогнозов. Особое внимание уделяется территориям с ограниченной надёжностью и случаям систематического завышения или занижения прогноза.

`r paste0("Таким образом, ", if (g_dir == "рост") (if (streak >= 3) "сезонный подъём заболеваемости ОРВИ и гриппом начался и продолжается" else "регистрируется рост заболеваемости ОРВИ и гриппом") else if (g_dir == "снижение") "регистрируется снижение заболеваемости ОРВИ и гриппом" else "устойчивое направление изменения заболеваемости ОРВИ и гриппом не определяется", ". В ближайшие две недели ожидается около ", tys(near2$прогноз[1]), " случаев на ", near2$неделя[1], "-й неделе и около ", tys(near2$прогноз[2]), " на ", near2$неделя[2], "-й (медианный прогноз; фактически на ", wk, "-й неделе зарегистрировано ", tys(last$cases), "). ", if (length(sens_w) && !is.null(anom_last)) paste0("Дальнейшая динамика на ", wrange(sens_w), " неделях в текущем расчёте во многом определяется необычным подъёмом прошлого года и будет уточняться при еженедельном пересчёте. ") else "Оценка на более поздние недели будет уточняться при еженедельном пересчёте. ", if (has_prev && ch_bias <= -10) "С учётом того,

```
что в последней проверке фактическая заболеваемость превысила прогноз, при планировании ресурсов на ближайшие недели целесообразно предусматривать резерв сверх центральных прогнозных значений. Верхняя граница 95% прогнозного интервала характеризует неблагоприятный вариант развития ситуации и не является плановой потребностью в ресурсах." else if (has_prev && ch_bias >= 10) "С учётом того, что в последней проверке прогноз превысил фактическую заболеваемость, центральные значения следует рассматривать как верхнюю оценку." else "")`
```