

УДК 519.8:528.8:504.064:630*907.1
DOI 10.17513/snt.40930



CC BY 4.0

**РАЗРАБОТКА И ПРОГРАММНАЯ РЕАЛИЗАЦИЯ
АЛГОРИТМА АВТОМАТИЗИРОВАННОГО АНАЛИЗА
МНОГОЛЕТНЕЙ ДИНАМИКИ ЛЕСНОГО ПОКРОВА
НА ОСНОВЕ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ
(НА ПРИМЕРЕ БАССЕЙНА РЕКИ СУНЖА)**

^{1,2}**Никифорова А. А. ORCID ID 0000-0002-2741-4153,**

^{1,2}**Табунщик В. А. ORCID ID 0000-0003-3555-6087,**

^{1,2}**Горбунова Т. Ю. ORCID ID 0000-0003-2155-6502**

¹*Федеральное государственное бюджетное учреждение науки Федеральный исследовательский центр «Институт биологии южных морей имени А. О. Ковалевского» Российской академии наук, Севастополь, Российская Федерация, e-mail: nikiforova_a@ibss-ras.ru;*

²*Федеральное государственное бюджетное образовательное учреждение высшего образования «Грозненский государственный нефтяной технический университет имени академика М. Д. Миллионщикова», Грозный, Российская Федерация*

Цель исследования – разработка и апробация комплекса алгоритмов и программных средств для автоматизированной количественной оценки многолетней динамики лесного покрова на основе глобального продукта Hansen Global Forest Change (v1.13) в границах произвольных водосборных бассейнов. В качестве модельной территории выбран бассейн р. Сунжа (Северный Кавказ, площадь 12 130,7 км²). Обработка состоит из двух этапов: выгрузка растровых слоев через Google Earth Engine и последующая статистическая обработка в среде R с использованием пакетов «terra» и «sf». Вычислительный эксперимент показал, что полный цикл обработки (от импорта до генерации итоговых таблиц и карт плотности) занимает 18 с на 16-ядерной машине с 128 Гб оперативной памяти, что подтверждает пригодность решения для оперативного мониторинга. Годовые потери леса в бассейне р. Сунжа варьировали от 0,5 км² (2015) до 42,8 км² (2016), с дополнительными пиками в 2004 г. (26,4 км²) и 2025 г. (29,7 км²). Пространственно потери локализованы преимущественно в горных и предгорных ландшафтах центральной части бассейна и в прирусловых зонах. Полученные результаты могут служить основой для моделирования углеродного и водного баланса, а также для разработки адаптивных стратегий лесопользования в регионе.

Ключевые слова: геоинформационные системы, дистанционное зондирование Земли, Hansen Global Forest Change, бассейн р. Сунжа, лесной покров, временные ряды Landsat, пространственный анализ, R-программирование

**DEVELOPMENT AND SOFTWARE IMPLEMENTATION OF AN ALGORITHM
FOR AUTOMATED ANALYSIS OF THE LONG-TERM DYNAMICS
OF FOREST COVER BASED ON REMOTE SENSING DATA
(USING THE SUNZHA RIVER BASIN AS AN EXAMPLE)**

^{1,2}**Nikiforova A. A. ORCID ID 0000-0002-2741-4153,**

^{1,2}**Tabunschik V. A. ORCID ID 0000-0003-3555-6087,**

^{1,2}**Gorbunova T. Yu. ORCID ID 0000-0003-2155-6502**

¹*Federal State Budgetary Institution of Science Federal Research Center “Institute of Biology of the Southern Seas named after A. O. Kovalevskiy of the Russian Academy of Sciences”, Sevastopol, Russian Federation, e-mail: nikiforova_a@ibss-ras.ru;*

²*Federal State Budgetary Educational Institution of Higher Education “Grozny State Petroleum Technical University named after Academician M. D. Millionschikov”, Grozny, Russian Federation*

The aim of the study is to develop and test a set of algorithms and software tools for the automated quantitative assessment of long-term forest cover dynamics based on the Hansen Global Forest Change (v1.13) global product within the boundaries of arbitrary catchment basins. The Sunzha River basin (North Caucasus, area 12,130.7 km²) was selected as the model territory. The processing consists of two stages: downloading raster layers via Google Earth Engine and subsequent statistical processing in the R environment using the “terra” and “sf” packages. The computational experiment showed that a complete processing cycle (from import to generation of final tables and density maps) takes 18 seconds on a 16-core machine with 128 gigabyte of Random Access Memory, which confirms the solution’s suitability for operational monitoring. Annual forest losses in the Sunzha River basin varied from 0.5 km² (2015) to 42.8 km² (2016), with additional peaks in 2004 (26.4 km²) and 2025 (29.7 km²). The losses are spatially localized mainly in the mountainous and foothill landscapes of the central part of the basin and in the riverbank zones. The obtained results can serve as a basis for modeling carbon and water balances, as well as for developing adaptive forest management strategies in the region.

Keywords: Geographic Information Systems, Earth remote sensing, Hansen Global Forest Change, Sunzha River basin, forest cover, Landsat time series, spatial analysis, R programming

Введение

Лесные экосистемы много значат в регуляции климата, сохранении биоразнообразия, хранении углерода и обеспечении водных ресурсов [1, 2]. Вклад лесного покрова в смягчение климатических изменений обусловлен двумя ключевыми аспектами углеродного цикла: интенсивной секвестрацией (поглощением) диоксида углерода на уровне 15,6 млрд т ежегодно и одновременным предотвращением высвобождения углерода за счет сдерживания процессов обезлесения, при котором эмиссия от пожаров, рубок и деградации составляет около 8,1 млрд т CO₂ в год [3].

Современные проблемы обезлесения включают необходимость повышения точности обнаружения небольших участков обезлесения и улучшения временного разрешения мониторинга [4]. Развитие систем мониторинга в режиме, близком к реальному времени, открывает новые возможности для оперативного реагирования на случаи незаконной вырубке лесов [5]. Традиционные методы мониторинга лесов основывались на наземных наблюдениях и статистических данных, которые часто были неполными и несопоставимыми между странами [4]. Прорыв в этой области произошел с развитием дистанционного зондирования Земли и созданием глобальных наборов данных о лесном покрове [6]. В современной практике анализа динамики типов наземного покрова и лесопокрытых площадей широкое применение находят методы дистанционного зондирования Земли (ДЗЗ), в том числе геоинформационные системы (ГИС) и данные космической съемки [7]. В этом контексте особый интерес представляет глобальный набор данных Hansen Global Forest Change, содержащий сведения о сокращении площади лесов, начиная с 2001 г. по настоящее время. В то же время данные наборы содержатся в сыром, необработанном виде и существует проблема получения этих данных для проведения исследований в границах водосборных бассейнов. Особенно это актуально в случаях, когда анализируются водосборные бассейны средних и больших рек, что требует огромных затрат для обработки этих данных. В то же время сейчас использование платформы облачных вычислений Google Earth Engine и языка программирования R! позволяет выгружать и анализировать огромные массивы геопространственной информации.

Информационные технологии, математические методы и высокопроизводительное программное обеспечение лежат в основе современных географических исследований, особенно в области мониторинга

динамики лесного покрова. Особую значимость в решении подобных задач приобретает междисциплинарный подход, органично объединяющий достижения экологии, геоинформатики, математического моделирования, дистанционного зондирования и наук о данных. Комплексная динамика лесных экосистем не может быть адекватно описана исключительно средствами одной научной дисциплины: для выявления трендов, аномалий и причинно-следственных связей необходимо совместное использование спутниковых данных, численных методов обработки временных рядов, статистического анализа и географических знаний о ландшафтной структуре территорий [8].

В отечественных исследованиях часто используются математические методы для решения геоэкологических задач. В своем исследовании Барталев и соавт. [9] доказали, что спутниковое картографирование растительного покрова не имеет альтернативных методов получения пространственно-однородных данных, однако его полноценная автоматизация принципиально недостижима без применения современных вычислительных технологий и математического аппарата обработки многомерных сигналов. В создании алгоритмов классификации, устойчивых к значительной пространственной вариабельности спектрально-отражательных характеристик необходимо применение комплексного использования методов машинного обучения, статистического анализа временных рядов и высокопроизводительных вычислений, что выводит задачу за рамки сугубо географических или экологических подходов и ставит ее в центр междисциплинарных исследований на стыке дистанционного зондирования, прикладной математики и информатики. В работе Лапко с соавт. [10] рассматриваются модификации непараметрического алгоритма распознавания, основанного на критерии максимального правдоподобия с введением дополнительных решающих функций. Синтез алгоритмов выполняется на основе анализа отношений оценок плотностей вероятности распределения случайных величин в классах и их функционалов с использованием пороговых значений, выбираемых в зависимости от условий классификации. Разработанные решения применяются для оценки состояния лесных насаждений по данным дистанционного зондирования, что подтверждает их практическую значимость в задачах обработки многомерной геопространственной информации. Также учеными [11] разрабатывалась информационная система для оценивания состояния природных объектов

(тематического дешифрирования лесных массивов) по данным дистанционного зондирования, ядром которой выступают непараметрические алгоритмы распознавания образов с градациями преимущества, оптимизируемые по коэффициентам размытости ядерных функций. Программный комплекс реализован в среде Delphi и обеспечивает полный цикл обработки. Таким образом разработка программных продуктов и средств автоматических вычислений является чрезвычайно важной задачей для геоэкологического мониторинга.

Цель исследования – разработка и исследование комплекса моделей, алгоритмов и программных средств для автоматизированной обработки многолетних данных дистанционного зондирования Земли (на примере глобального набора Hansen Global Forest Change) для выявления пространственно-временных закономерностей динамики лесного покрова в бассейне р. Сунжа за 2001–2025 гг.

Материал и методы исследования

Район исследования. Бассейн р. Сунжа расположен в восточной части Северного Кавказа (рис. 1). Исследуемая территория, являясь главной водной артерией Чеченской Республики и Республики Ингушетия [12], характеризуется значительным ландшафт-

ным разнообразием и контрастностью природных условий, что позволяет проследить пространственную дифференциацию процессов сокращения лесов в широком диапазоне высот – от равнинных участков (27 м) до высокогорных массивов (4464 м). Выбор бассейна р. Сунжа в качестве модельного региона для анализа динамики лесного покрова обусловлен комплексом физико-географических, экологических и антропогенных факторов, определяющих его высокую репрезентативность для изучения процессов деградации лесных экосистем на востоке Северного Кавказа. Бассейн р. Сунжа играет ключевую роль в водоснабжении более 1,5 млн чел. В его пределах расположены крупные населенные пункты: Грозный, Назрань, Гудермес, Магас, Аргун и Урус-Мартан.

Река берет начало из ледников на горном массиве Ушкорт, на Главном Кавказском хребте (Ингушетия), в дальнейшем впадает в Терек. «Водосборный бассейн реки Сунжа, рассчитанный по данным ЦМР SRTM 30m, характеризуется площадью 12130,7 км² и периметром 583,8 км. Бассейн обладает значительным перепадом высот с абсолютными отметками от 27 до 4464 м при средней высоте 1037,3 м, что обуславливает резко расчлененный рельеф с максимальным уклоном поверхности 79,1° и средним уклоном 13°» [12].

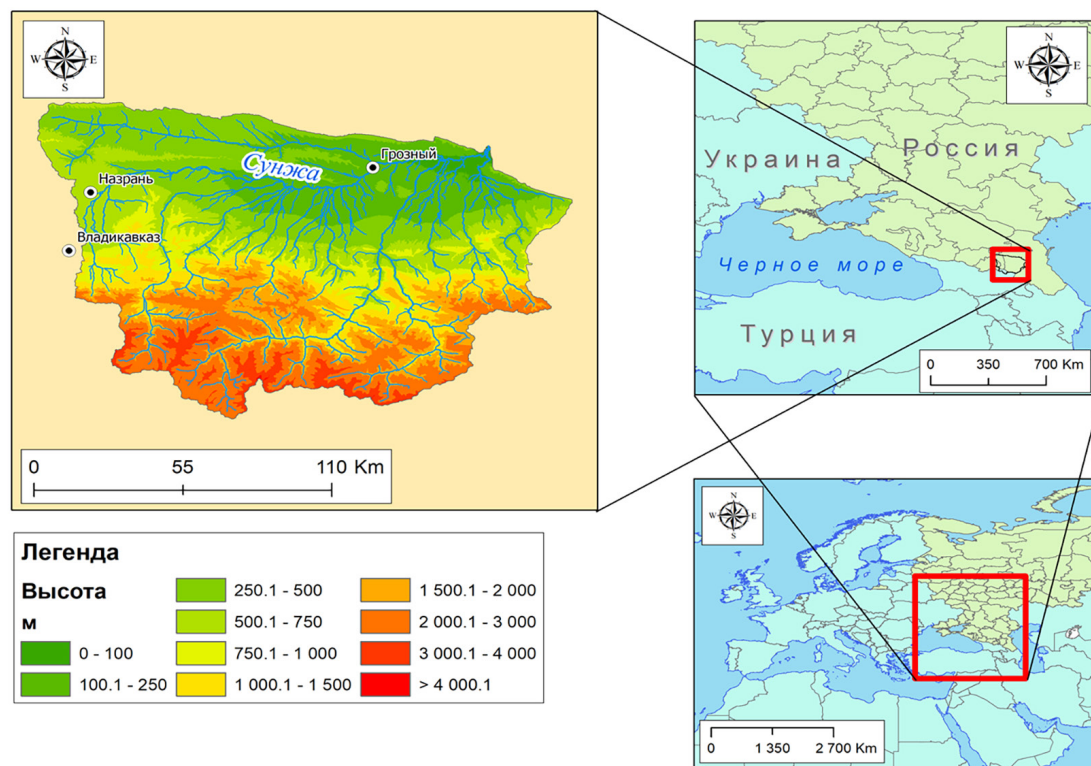


Рис. 1. Бассейн р. Сунжа
Примечание: составлен авторами на основе источника [12]

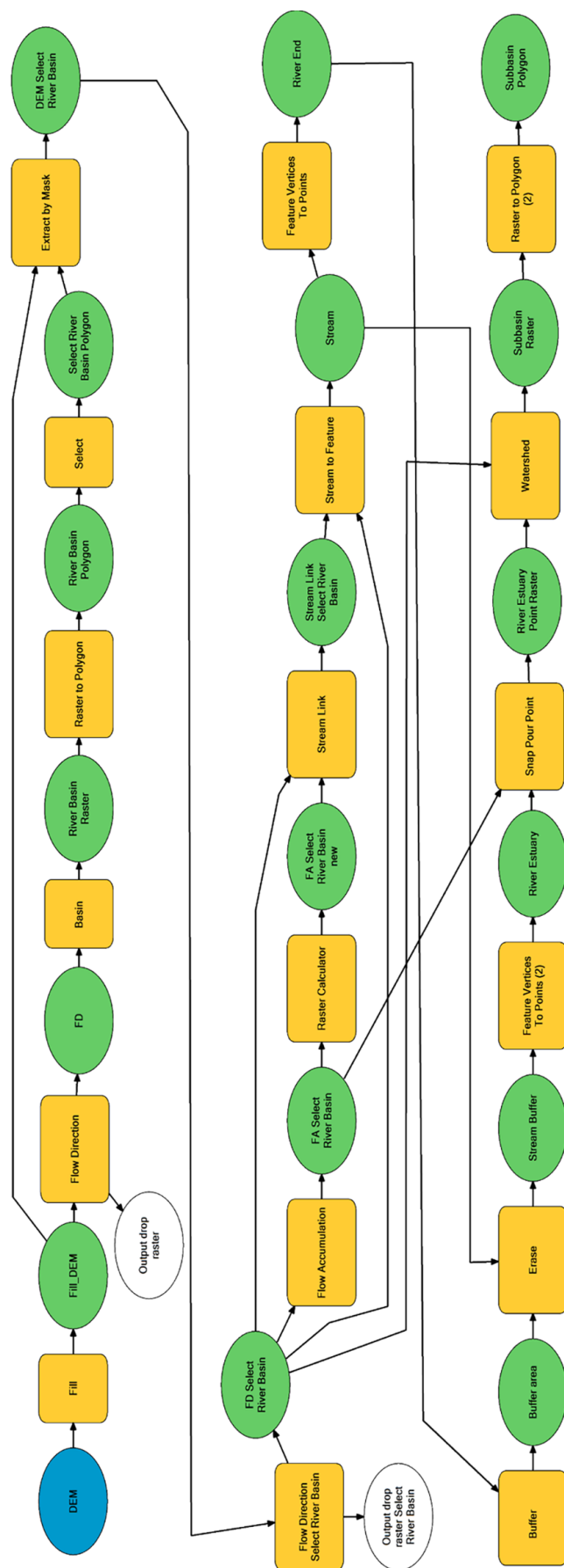


Рис. 2. Модель выделения бассейна р. Сунжа с использованием ArcGIS ModelBuilder
Примечание: составлен авторами по результатам данного исследования

Выделение границы водосборного бассейна р. Сунжа произведено с использованием программного комплекса ArcGIS 10.8 с использованием среды ModelBuilder, для автоматического выполнения выделения водосборного бассейна и расчета его характеристик с использованием цифровой модели рельефа SRTM (рис. 2).

В основе настоящего исследования лежит глобальный набор данных Hansen Global Forest Change (версия 1.13, 2001–2025) [13], сформированный по результатам обработки временных серий спутниковых снимков Landsat с пространственным разрешением 30 м. Лесной покров в данной продукции определяется как древесная растительность высотой более 5 м при проективном покрытии не менее 10 % (показатель *treecover2000*). Потери леса (показатель *lossyear*) регистрируются в случае полной смены лесного покрова на нелесные категории (вследствие рубок, пожаров или иных нарушений). Первичные данные представлены в виде бинарных масок (0/1) и ежегодных растровых слоев (1–25), агрегированных в медианные композиты для минимизации влияния облачности; валидация результатов выполнена на основе сопоставления с данными Landsat. Исходные материалы были загружены из открытого доступа [13] с использованием платформы облачных вычислений Google Earth Engine, с использованием стандартного скрипта Google Earth Engine (приложение А).

Обработка и анализ полученных данных производился с использованием ArcGIS 10.8 и языка программирования R!. Для автоматизированной обработки данных Hansen Global Forest Change и расчета площадей потерь лесного покрова разработан скрипт на языке R. Выбор R в качестве среды реализации обусловлен наличием специализированных пакетов («terra», «sf») для работы с растровыми (данные об обезлесении) и векторными геоданными (границы водосборных бассейнов).

Разработанный скрипт включает следующее:

1. Функция загрузки и предобработки данных:

1.1. импорт растровых слоев Hansen Global Forest Change из локальных источников (GeoTIFF);

1.2. чтение векторной маски бассейна из Shapefile с использованием пакета *sf*;

1.3. приведение всех слоев к единой проекции и пространственному разрешению.

2. Функция обрезки по маске:

2.1. обрезка каждого растрового слоя по границам векторной маски с использованием пакета *terra*;

2.2. маскирование пикселей, расположенных за пределами области интереса.

3. Функция расчета ежегодных потерь:

3.1. формирование бинарной маски потерь;

3.2. вычисление площади для каждого пикселя маски с учетом его фактической площади в квадратных метрах;

3.3. суммирование площадей всех пикселей маски в пределах области интереса;

3.4. агрегация результатов в ежегодную таблицу потерь.

4. Функция экспорта результатов (*export_results*):

4.1. генерация CSV-файла с ежегодными площадями потерь;

4.2. создание растровой карты плотности потерь в формате GeoTIFF;

4.3. формирование сводной статистики.

Разработанный скрипт (приложение Б) позволяет загружать полученные данные об обезлесении и границах водосборных бассейнов и получать на выходе статистические данные о пространственно-временной изменчивости показателя. Это делает скрипт универсальным для анализа обезлесения на территории любых водосборных бассейнов.

Расчеты выполнялись на стационарном компьютере с характеристиками: процессор AMD Ryzen 9 9950X3D OEM; 128 Гб оперативной памяти. Дополнительное сравнение проводилось на Intel Core i5-9300H; 32 Гб оперативной памяти.

Пространственный анализ потерь лесного покрова выполнялся в среде R с использованием пакетов *terra* и *sf*. Исходный растр года потери леса (Hansen *lossyear*) был обрезан и замаскирован по границе бассейна р. Сунжа. На его основе создавался бинарный растр наличия потерь (*lossyear* > 0), который затем агрегировался методом суммирования с коэффициентом агрегации, обеспечивающим оптимальное число пространственных единиц.

Скрипт выполняет предварительную проверку входных данных: наличие растра и векторного слоя, наличие системы координат, исправимость геометрии бассейна и факт пространственного пересечения слоев. При географической системе координат растр перепроецируется в зону UTM, определяемую по центроиду бассейна; если исходный растр уже задан в метрической проекции, перепроецирование не выполняется. Расчетная область формируется операциями «*сгор*» и «*mask*»: ячейки вне полигона получают значение NA и исключаются из анализа. При неполном покрытии бассейна растром выдается предупреждение с указанием доли площади, оставшейся вне сцены. Размер ячейки агрегации подбирается

так, чтобы число пространственных единиц было не ниже заданного порога, необходимого для статистических тестов. При пустой маске, отсутствии кодов потерь, нулевой площади бассейна либо невозможности набрать минимальное число наблюдений выполнение прерывается и выдается ошибка в консоль, что требует в дальнейшем исправления ошибок и повторного запуска скрипта.

Пространственная структура потерь оценивалась на регулярной сетке после агрегации исходных пикселей (30 м) в ячейки километрового масштаба. На исходном разрешении 30 м число наблюдений составляет сотни тысяч, а автокорреляция отражает соседство отдельных пикселей, а не очагов утраты полога. Ячейка 1 км соответствует характерному размеру детектируемых нарушений и дает выборку достаточного объема ($n = 12\,720$) при полной восьмизвязной окрестности у большинства единиц. Коэффициент агрегации есть отношение целевого размера ячейки к разрешению раstra после перепроецирования в UTM 38N: $(1000/30) = 33$; фактический размер ячейки составил 990 м. Устойчивость вывода к масштабу проверена на сетках 500 и 2000 м.

Полученный растр плотности потерь использовался для расчета глобального индекса Морана и локального индикатора (LISA) (приложение Б). Глобальный индекс Морана вычислялся функцией `terra::autocor()` с матрицей пространственных весов `queen's case` (8 соседей). Статистическая значимость индекса оценивалась методом Монте-Карло: значения раstra случайным образом перемешивались 999 раз, для каждой перестановки повторно рассчитывался индекс Морана, после чего вычислялся p -value как доля симуляций, в которых значение индекса было больше или равно наблюдаемому. Локальные индикаторы пространственной ассоциации (LISA) рассчитывались той же функцией `autocor()` в режиме `global = FALSE`. Для каждой ячейки определялся тип кластера (HH, LL, HL, LH) на основе сравнения значения ячейки и пространственного лага со средним значением раstra. Статистическая значимость локальных значений оценивалась с помощью Monte-Carlo теста (99 перестановок). В основной спецификации использован критерий `queen`. В контрольной спецификации использован критерий `rook`. Ячейки вне полигона границы бассейна реки в расчет не входят. Гипотеза пространственной случайности отклонялась при $p < 0,05$. Для глобального индекса Морана дополнительно сопоставлялись спецификации `queen/rook` и разрешения 500, 1000 и 2000 м; локальные кластеры картировались только при $p < 0,05$.

Тестирование производительности выполнялось в границах бассейна р. Сунжа. Протокол включал прогревочный прогон, три независимых повтора для каждой конфигурации и согласованное ограничение потоков. Для этапов фиксировалось стенное время; по повторам рассчитывались среднее, стандартное отклонение, диапазон, коэффициент вариации и ускорение относительно однопоточного режима (приложение В).

Корректность расчета годовых площадей потерь проверяли на синтетическом растре, воспроизводящем геометрию Hansen Global Forest Change: регулярная сетка в системе координат EPSG:4326. На таком входе исполняются те же операции, что и на рабочих данных: учет площади ячейки и отсечение по полигону. Растр 2000×2000 ячеек задавали в окне с заданными координатами. Значение ячейки кодировало год потери: коды 1, ..., 25 соответствуют 2001–2025 гг. Для каждого года фиксировали число пикселей и размещали их тремя фрагментами в разных широтных поясах, чтобы площадь ячейки менялась вдоль сетки. Полигон-маска был прямоугольным и выровнен по границам ячеек. За его пределами задавали контрольные пиксели потерь; они не включались в расчет.

Сопоставление с другими программными продуктами выполнялось с двумя распространенными средствами расчета зональной статистики раstra в ГИС: QGIS / GRASS GIS (Zonal Statistics, `r.mask`, `r.stats`) и ArcGIS 10.8 (Extract by Mask, Zonal Statistics as Table). Для всех вариантов использовались идентичные входной растр `lossyear` и векторная маска бассейна.

Результаты исследования и их обсуждение

Сравнительные испытания производительности проведены на двух вычислительных узлах при неизменном алгоритме расчета площадей потерь лесного покрова и едином протоколе измерений (холодный прогон, прогревочный прогон, три повтора, контроль совпадения годовых площадей). Первый узел характеризовался 16 физическими (32 логическими) ядрами и 128 Гб оперативной памяти, второй характеризовался 4 физическими (8 логическими) ядрами и 32 Гб оперативной памяти. Тестировались конфигурации 1–32 и 1–8 потоков соответственно. Полное время обработки при холодном прогреве бассейна площадью $12\,130,7 \text{ км}^2$ по данным Hansen GFC v1.13 составило от 17,8 до 19,9 с на первой системе и от 17,5 до 18,5 с на второй, при этом максимальное ускорение относи-

тельно однопоточного запуска не превысило 6 % (на 4-ядерной машине при 8 потоках), тогда как на 16-ядерной при 32 потоках зафиксирована деградация производительности (ускорение 0,90). Доминирующий этап «Цикл по годам» занимал до 85 % общего времени и практически не ускорялся при росте числа потоков; единственный этап, показавший заметный выигрыш (обрезка и маскирование, в 2,2 раза), имел малый вклад в итоговое время, что свидетельствует о вводно-выводном характере узкого места (чтение растровых данных и работа с памятью) и подтверждает, что разработанный модуль эффективно работает без применения высокопараллельных вычислительных ресурсов. Таким образом, разработанный скрипт подходит для ускоренных вычислений обезлесения на территории бассейнов крупных и средних речных бассейнов. Пиковое потребление памяти оставалось неизменным во всех экспериментах, что исключает ограничения по ОЗУ.

После прогрева среднее полное время выполнения составило 8,39 с на первом узле и 15,80 с на втором (отношение 1,88). Структура затрат оказалась идентичной: на последовательный цикл суммирования по 25 годовым маскам приходилось около 91 % стенного времени (7,67 и 14,39 с). Соотношение времен обрезки с маскированием и расчета площади ячейки между узлами лежало в том же диапазоне (1,75–1,85). Годовые площади потерь совпали во всех прогонах (максимальное расхождение 0 км²); оценка занимаемой памяти для цикла также совпала, что подтверждает тождественность объема данных и схемы расчета. Масштабирование по числу потоков не выявлено. Ускорение полного конвейера относительно однопоточного режима составило 0,98–1,00 на первом узле и 1,00–1,01 на втором; эффективность параллелизации снижалась пропорционально числу потоков. Полученный результат воспроизводится на обеих аппаратных конфигурациях: средства многопоточности *terra* и *OpenMP* не сокращают время реализованного конвейера, и для определения площадей потерь достаточно 1–4 потоков.

Время выполнения расчетов в сравнении с другими программными продуктами с использованием QGIS составило 55 с, а ArcGIS – 46 с. При этом в сравнении с разработанным скриптом QGIS и ArcGIS используются десктопные версии, которые предполагают использование большего количества ручных шагов выполнения процесса (запуск, добавление растровых и векторных данных, создание пирамидальных слоев, сохранение таблиц результатов

и пр.). Прямое сопоставление исходных кодов с выбранными аналогами невыполнимо по объективным причинам. Использование этих программных продуктов написано на других языках программирования, поэтому сравнение с реализацией на R измеряло бы различия языков и сред исполнения, а не свойства предложенных алгоритмических решений. Исходный код QGIS открыт, однако алгоритмы зональной статистики реализованы на C++ в составе библиотеки анализа и частично делегируются внешним провайдерам обработки (GDAL, GRASS), а также неотделимы от каркаса приложения вследствие чего не могут быть выделены в самостоятельную процедуру, сопоставимую со скриптом. Дополнительно следует учитывать, что чтение и запись растровых данных во всех сравниваемых средствах, включая разработанный скрипт, выполняются одной и той же библиотекой GDAL, поэтому наблюдаемые различия во времени характеризуют организацию вычислительного конвейера, а не реализацию расчета площадей. В связи с изложенным единицей сравнения выбран не исходный код, а решение прикладной задачи целиком: идентичные входные данные, совпадение полученных численных результатов, измеряемое время выполнения и число операций, требующих участия оператора. Таким образом, использование разработанного скрипта существенно упрощает задачу обработки больших объемов данных.

Синтетический тест эффективности показал что расчетные годовые площади совпали с аналитическим эталоном на эллипсоиде WGS84. Сумма за 2001–2025 гг. составила 10,992059 км² при том же значении эталона; относительная разность сумм – $5,8 \times 10^{-9}$ %. Абсолютные разности по всем годам были меньше 10⁻⁶ км². Таким образом, на входе с геометрией Hansen GFC реализация воспроизводит площадь ячейки на WGS84, корректно маскирует растр и агрегирует потери по 25 годам. Тест относится к программной реализации и не заменяет расчет по фактическому контуру бассейна.

Пространственное распределение потерь лесного покрова в бассейне р. Сунжа, выявленных в ходе настоящего исследования, представлено на рис. 3 и в таблице.

Анализ количественных данных об утра- те лесного покрова в бассейне р. Сунжа за 2001–2025 гг. свидетельствует о существенной временной неоднородности деградационных процессов, при этом максимальные значения зафиксированы в 2016 г. (42,8 км²) и 2025 г. (29,7 км²), что свидетельствует о наличии внешних факторов повышенной лесопотери.

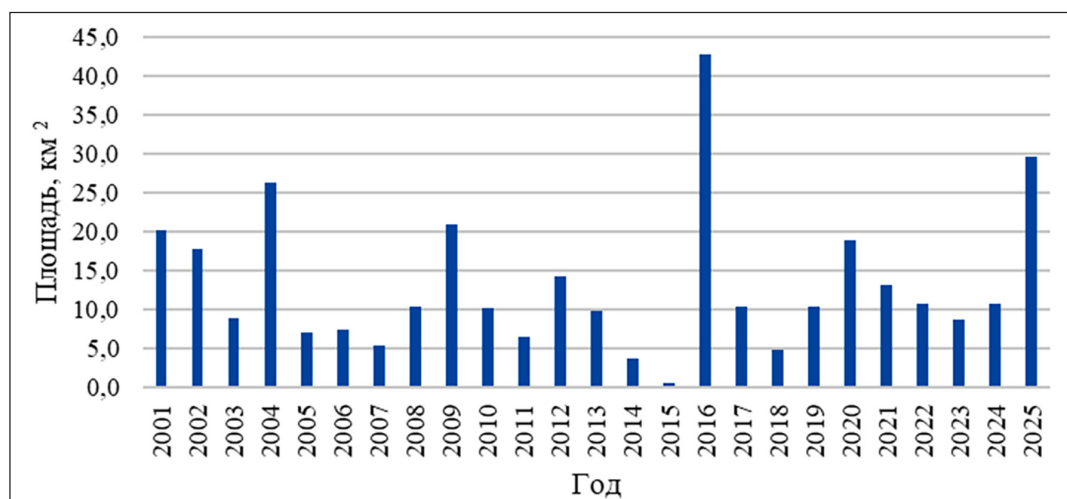


Рис. 3. Сокращение площади лесов в бассейне р. Сунжа, 2001–2025 гг.
Примечание: составлен авторами по результатам данного исследования

Точные значения сокращения площади лесов в бассейне р. Сунжа, 2001–2025 гг.

Год	Сокращение площади лесов в бассейне р. Сунжа, км ²	Год	Сокращение площади лесов в бассейне р. Сунжа, км ²	Год	Сокращение площади лесов в бассейне р. Сунжа, км ²
2001	20,2	2010	10,3	2019	10,4
2002	17,8	2011	6,6	2020	18,9
2003	9,0	2012	14,3	2021	13,2
2004	26,4	2013	9,9	2022	10,7
2005	7,1	2014	3,7	2023	8,8
2006	7,5	2015	0,5	2024	10,7
2007	5,4	2016	42,8	2025	29,7
2008	10,4	2017	10,4		
2009	21,0	2018	5,0		

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Второй по значимости пик сокращения приходится на 2004 г. (26,4 км²). Напротив, минимальные значения утрат лесного покрова, не превышающие 0,5 км² и 3,7 км² соответственно, наблюдаются в 2015 и 2014 гг., что позволяет рассматривать этот двухлетний интервал как период относительной стабилизации лесного фонда, характеризующийся снижением интенсивности антропогенного и техногенного воздействия на лесные экосистемы водосборного бассейна.

По материалам картографического анализа динамики лесного покрова в бассейне р. Сунжа за 2001–2025 гг. (рис. 4) установлено, что процессы сокращения леса носят очаговый характер и локализуются преимущественно в пределах предгорных и горных ландшафтов центральной части бассейна, а также в прирусловых зонах речной сети.

Пространственная приуроченность нарушений обусловлена главным образом

антропогенной нагрузкой: если в верховьях бассейна, где расположены крупные особо охраняемые природные территории, естественные ландшафты сохраняются, то в среднем и нижнем течении доминируют сельскохозяйственные угодья и урбанизированные территории. При этом весомый вклад в сокращение лесов оказывают и геоморфологические процессы [14].

Пространственный анализ выполнен в равноугольной проекции WGS 84 / UTM zone 38N (EPSG:32638). Исходный растр приведен к разрешению 30 м методом ближайшего соседа, после чего потери агрегированы в ячейки целевого размера 1000 м. Коэффициент агрегации задан как отношение целевого размера ячейки к разрешению раstra: $\text{round}(1000/30) = 33$, фактический размер ячейки составил 990 м. Число пространственных единиц равно 12 720. В ячейке фиксировалось число исходных пикселей потери.

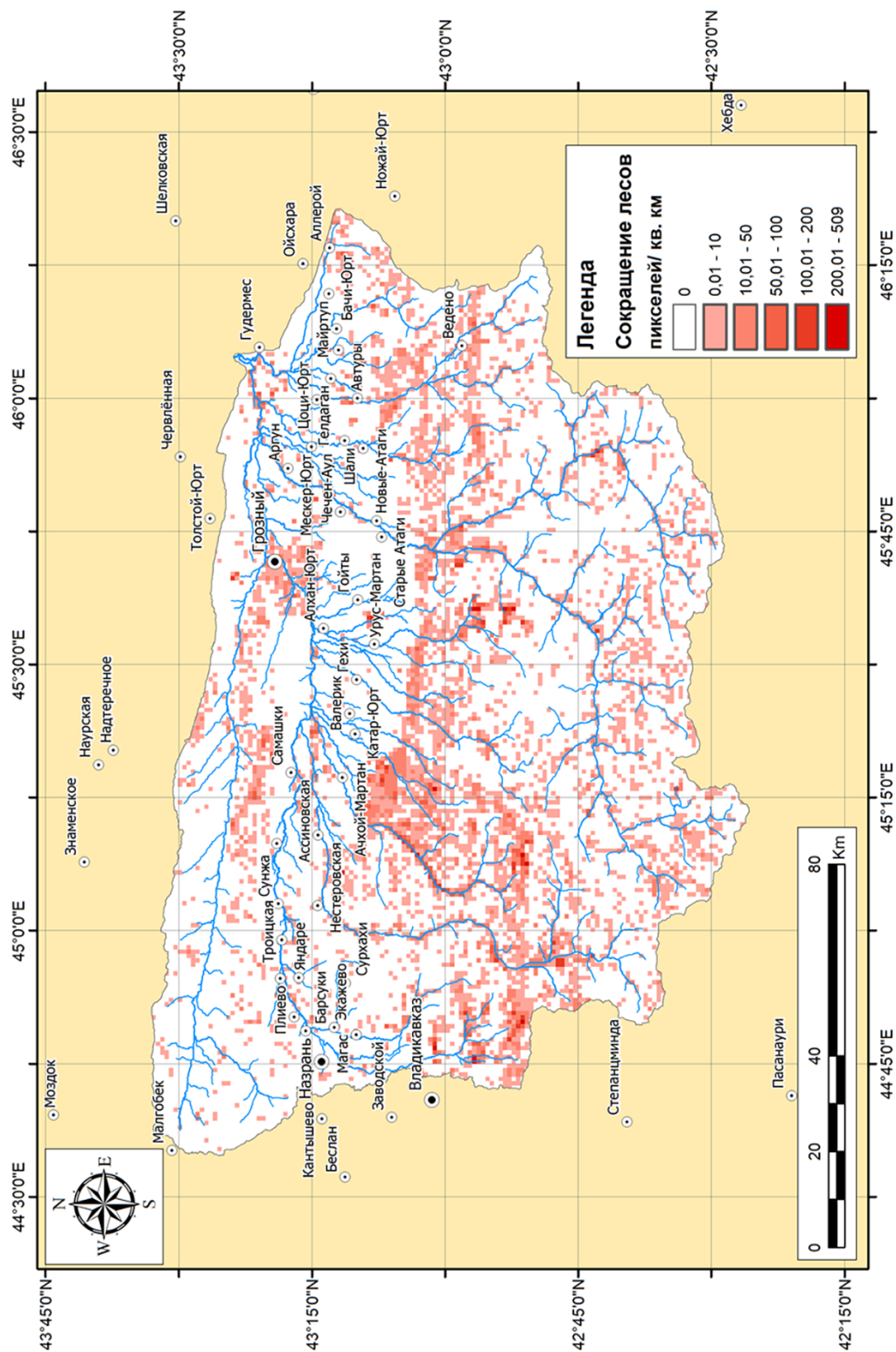


Рис. 4. Картографический анализ сокращения площади лесов в бассейне р. Сунжа с 2001 по 2025 г.
Примечание: составлен авторами по результатам данного исследования

Пространственные веса заданы как бинарное соседство первого порядка на регулярной сетке (окно 3×3 , критерий queen); для глобального индекса Морана веса стандартизированы по строкам.

Глобальный индекс Морана для агрегированной плотности потерь равен $I = 0,39$. Результаты Monte-Carlo теста (999 случайных перестановок) показали высокую статистическую значимость наблюдаемой автокорреляции ($p = 0,001$). Полученное сочетание $I > 0$ и $p < 0,05$ позволяет отклонить гипотезу пространственной случайности и указывает на положительную пространственную автокорреляцию, так как ячейки с повышенной потерей лесного покрова статистически чаще соседствуют друг с другом.

Устойчивость глобального вывода проверена двумя способами. Во-первых, расчет повторен только для внутренних ячеек с полным набором из восьми соседей, то есть краевой эффект не определяет результат. Во-вторых, глобальный индекс пересчитан при разрешениях 500, 1000 и 2000 м для соседства queen и rook (199 перестановок). Во всех спецификациях индекс оставался в интервале 0,2–0,4, знак коэффициента был положительным, а вывод о значимости на уровне 0,05 не изменялся. Несколько более высокие значения I при соседстве rook согласуются с тем, что четырехсвязная схема опирается на меньшее число соседей и сильнее подчеркивает локальные контрасты. Таким образом, заключение о кластеризации потерь не зависит от выбора масштаба агрегации в диапазоне 0,5–2 км и от альтернативы queen/rook. Локальный анализ пространственной автокорреляции (LISA) позволил уточнить характер этой кластеризации. Все значимые кластеры относились исключительно к типу HH (High-High) зоны высоких потерь, окруженные соседними зонами также высоких потерь. Таким образом, процесс обезлесения в бассейне р. Сунжа носит ярко выраженный очаговый характер.

Совместно временной и пространственный анализы показывают, что утрата лесного покрова в бассейне р. Сунжа за 2001–2025 гг. была умеренной по суммарной площади, неравномерной по годам и пространственно кластеризованной. Статистически подтверждается наличие горячих точек обезлесения, а не диффузное рассеяние потерь по всей площади водосбора. Полученные оценки характеризуют факт и местоположение детектируемой потери древесного полога и не позволяют непосредственно разделять причины изменений (рубки, пожары, ветровалы, вспышки вредителей) без привлечения дополнительных данных.

В то же время сокращение площади лесов на Кавказе представляет собой сложный и многогранный процесс, обусловленный историческими, социально-экономическими и экологическими факторами. Исторические исследования показывают, что лесной покров Кавказа значительно сократился за последние тысячелетия. В Армении, например, лесной покров уменьшился с примерно 35 % территории в период 4000–1000 гг. до н. э. до 8,1 % в 1950-х гг., 11,2 % в 1980-х гг. и 7–8 % в настоящее время [15]. Это сокращение связано с многовековой историей использования лесных ресурсов, изменениями в землепользовании и демографическими процессами. При этом распад Советского Союза в 1991 г. стал переломным моментом и точкой социокультурного разрыва и, как следствие, изменения землепользования на всех постсоветских территориях [16]. Исследование, основанное на анализе спутниковых снимков Landsat с 1987 по 2015 г., показывает значительные изменения в землепользовании в Кавказских горах [17]. По данным спутникового мониторинга [18], общая площадь лесов на Северном Кавказе в 2000 г. составляла около 28 % территории региона. В переходный период после распада СССР произошло ослабление контроля за лесопользованием, что привело к увеличению нелегальных рубок и деградации лесных массивов. В Западном Кавказе были зафиксированы случаи вырубki лесного полога, особенно в национальном парке Сочи в связи со строительством объектов для Олимпийских игр [18]. Анализ спутниковых снимков показал, что в Грузии после 1990 г. произошла незначительная потеря лесного покрова – примерно 0,8 % [19]. Основной причиной считается нелегальная заготовка древесины для бытовых нужд, а не крупномасштабные сплошные рубки.

Точность исходных данных Hansen GFC уже валидирована: согласно независимым исследованиям, пользовательская точность продукта составляет 88–98 % в зависимости от региона, а общая точность достигает 99,34 % [20–22]. Отметим, что процедура обрезки по маске и суммирования площадей не модифицирует значения пикселей, а лишь агрегирует их, поэтому ошибка конечного результата не превышает ошибки исходного продукта. Проведение сравнения с другими спутниковыми снимками, наоборот, порождает неопределенности в силу того, что различные спутниковые снимки используют разную ширину каналов спектральной съемки [23]. В то же время пространственное разрешение 30 м не позволяет детектировать потери площадью менее

0,09 га (один пиксель). Для задач выявления точечных рубок это ограничение может быть критическим; перспективным направлением является интеграция с данными Sentinel-2 (10 м).

Заключение

Разработан комплексный алгоритм и программный скрипт на языке программирования R! для автоматизированного расчета ежегодных площадей потерь лесного покрова на основе данных Hansen Global Forest Change v1.13, позволяющий учитывать фактическую площадь пикселей и работать с произвольной векторной маской водосборного бассейна.

Экспериментальная оценка производительности на высокопроизводительной рабочей станции (16 ядер, 128 Гб оперативной памяти) показала, что полный цикл обработки при холодном прогоне для бассейна площадью около 12 тыс. км² и разрешением 30 м выполняется за 18 с, а при прогревочном прогоне – около 8 с. Это подтверждает пригодность модуля для оперативного мониторинга и пакетной обработки нескольких бассейнов. На одном и том же наборе Hansen GFC v1.13 и одной маске бассейна скрипт сократил цепочку «обрезка – суммирование – таблица с данными» относительно интерактивной обработки в QGIS/ArcGIS до одного запуска.

Проведенное исследование динамики лесного покрова бассейна р. Сунжа за 2001–2025 гг. на основе данных дистанционного зондирования Земли (Hansen Global Forest Change) позволило установить, что за 25-летний интервал площадь лесов в пределах водосборного бассейна сократилась на 330,8 км², при этом деструкционные процессы носят выраженный очаговый характер и пространственно приурочены преимущественно к предгорным и долинным ландшафтам средней и нижней частей бассейна, где антропогенная нагрузка, связанная с сельскохозяйственным освоением, линейной инфраструктурой и урбанизированными территориями, достигает максимальных значений.

Практическая ценность разработанного подхода заключается в его универсальности (возможность применения к любым водосборным бассейнам), открытости исходного кода и низких вычислительных затратах, что позволяет использовать его в региональных системах экологического мониторинга и при научных исследованиях обезлесения в различных регионах мира.

Список литературы

1. Curtis P. G., Slay C. M., Harris N. L., Tyukavina A., Hansen M. C. Classifying drivers of global forest loss // *Science*. 2018. Vol. 361. P. 1108–1111. DOI: 10.1126/science.aau3445.

2. Guria R., Mishra M., Baraj B. Examining the drivers of forest cover change and deforestation susceptibility in Northeast India using multicriteria decision-making models // *Environmental Monitoring and Assessment*. 2024. Vol. 196. Art. 13172. DOI: 10.1007/s10661-024-13172-2.

3. Newton P., Kinzer A. T., Miller D. C., Oldekop J. A., Agrawal A. The Number and Spatial Distribution of Forest-Proximate People Globally // *One Earth*. 2020. Vol. 3. Is. 3. P. 363–370. DOI: 10.1016/j.oneear.2020.08.016.

4. Finer M., Novoa S., Weisse M. J., Petersen R., Mascaro J., Souto T., Stearns F., Martinez R. G. Combating deforestation: From satellite to intervention // *Science*. 2018. Vol. 360. P. 1303–1305. DOI: 10.1126/science.aat1203.

5. Popkin G. Satellite alerts track deforestation in real time // *Nature*. 2016. Vol. 530. № 7591. P. 392–393. DOI: 10.1038/530392a.

6. Kim D.-H., Sexton J. O., Noojipady P., Huang C., Anand A., Channan S., Feng M., Townshend J. R. Global, Landsat-based forest-cover change from 1990 to 2000 // *Remote Sensing of Environment*. 2014. Vol. 155. P. 178–193. DOI: 10.1016/j.rse.2014.08.017.

7. Silva C., Olthoff A. Exploración de tendencias de cambio forestal en reservas de la biosfera, Tabasco, desde el espacio // *Kuxulkab*. 2024. Vol. 30. Art. e6391. DOI: 10.19136/kuxulkab.a30n68.6391.

8. Песегов Д. А. Математическое и программное обеспечение географических информационно-аналитических систем регионального уровня (на примере Красноярского края): дис. ... канд. техн. наук. Красноярск, 2003. 180 с.

9. Барталев С. А., Егоров В. А., Ершов Д. В., Исаев А. С., Лупян Е. А., Плотноков Д. Е., Уваров И. А. Спутниковое картографирование растительного покрова России по данным спектрорадиометра MODIS // *Современные проблемы дистанционного зондирования Земли из космоса*. 2011. Т. 8. № 4. С. 285–302.

10. Лапко А. В., Лапко В. А. Непараметрические алгоритмы оценивания состояний природных объектов // *Автоматика*. 2018. Т. 54. № 5. С. 33–39. DOI: 10.15372/AUT20180504.

11. Лапко А. В., Лапко В. А. Информационная система оценивания состояний природных объектов по данным дистанционного зондирования на основе непараметрических алгоритмов принятия решений // *Информатика и системы управления*. 2019. № 1 (59). С. 106–117. DOI: 10.22250/isu.2019.59.106-117.

12. Гелагаев А. Ш., Никифорова А. А., Табунщик В. А. Основные морфометрические параметры бассейна р. Сунжа // *Грозненский естественно-научный бюллетень*. 2026. Т. 11. № 1 (43). С. 9–15. DOI: 10.25744/genb.2026.43.1.002.

13. Hansen M. C., Potapov P. V., Moore R., Hancher M., Turubanova S. A., Tyukavina A., Thau D., Stehman S. V., Goetz S. J., Loveland T. R., Kommareddy A., Egorov A., Chini L., Justice C. O., Townshend J. R. G. High-Resolution Global Maps of 21st-Century Forest Cover Change // *Science*. 2013. Vol. 342. P. 850–853. DOI: 10.1126/science.1244693.

14. Гакаев Р. А., Гаджиев Н. Г., Исраилов М. З., Эскиев М. М. Пространственный анализ структуры и антропогенной трансформации лесных экосистем горной части Чеченской Республики // *Устойчивое развитие горных территорий*. 2026. Т. 18. № 2 (68). С. 1077–1090. DOI: 10.21177/1998-4502-2026-18-2-1077-1090. EDN: ONVHMX.

15. Moreno-Sanchez R., Sayadyan H. Y. Evolution of the forest cover in Armenia // *International Forestry Review*. 2005. Vol. 7. Is. 2. P. 113–127. DOI: 10.1505/for.2005.7.2.113.

16. Горбунов Р. В., Табунщик В. А., Вольхин Д. А., Горбунова Т. Ю. Трансформация ландшафтно-экологических и социокультурных процессов в Горном Крыму в постсоветский период // *Вопросы географии*. 2024. № 158. С. 537–555. DOI: 10.24057/probl.geogr.158.19. EDN: PDTHJQ.

17. Buchner J., Yin H., Frantz D. J., Kuemmerle T., Askerov E., Bakuradze T., Bleyhl B., Elizbarashvili N., Komarova A., Lewińska K., Rizayeva A., Sayadyan H. Y., Tan B., Tepanosyan G., Zazanashvili N., Radeloff V. C. Land-cover change in the Caucasus Mountains since 1987 based on the

topographic correction of multi-temporal Landsat composites // Remote Sensing of Environment. 2020. Vol. 248. Art. 111967. DOI: 10.1016/j.rse.2020.111967.

18. Bragina E. V., Radeloff V. C., Baumann M., Wendland K., Kuemmerle T., Pidgeon A. M. Effectiveness of protected areas in the Western Caucasus before and after the transition to post-socialism // Biological Conservation. 2015. Vol. 184. P. 456–464. DOI: 10.1016/j.biocon.2015.02.013.

19. Olofsson P., Torchinava P., Woodcock C. E., Baccini A., Houghton R. A., Ozdogan M., Zhao F., Yang X. Implications of land use change on the national terrestrial carbon budget of Georgia // Carbon Balance and Management. 2010. Vol. 5. Is. 1. Art. 4. DOI: 10.1186/1750-0680-5-4.

20. Galitsatos N., Donoghue D. N. M., Watt P., Bholanath P., Pickering J., Hansen M. C., Mahmood A. R. J. An Assessment of Global Forest Change Datasets for National Forest Monitor-

ing and Reporting // Remote Sensing. 2020. Vol. 12. Art. 1790. DOI: 10.3390/rs12111790.

21. Linke J., Fortin M.-J., Courtenay S. C., Cormier R. High-resolution global maps of 21st-century annual forest loss: Independent accuracy assessment and application in a temperate forest region of Atlantic Canada // Remote Sensing of Environment. 2017. Vol. 188. P. 164–176. DOI: 10.1016/j.rse.2016.10.040.

22. Shimizu K., Ota T., Mizoue N. Accuracy Assessments of Local and Global Forest Change Data to Estimate Annual Disturbances in Temperate Forests // Remote Sensing. 2020. Vol. 12. Is. 15. Art. 2438. DOI: 10.3390/rs12152438.

23. Gonsamo A., Chen J. M. Spectral Response Function Comparability Among 21 Satellite Sensors for Vegetation Monitoring // IEEE Transactions on Geoscience and Remote Sensing. 2013. Vol. 51. Is. 3. P. 1319–1335. DOI: 10.1109/tgrs.2012.2198828.

Конфликт интересов: Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest: The authors declare that there is no conflict of interest.

Финансирование: Работа выполнена при финансовой поддержке Проекта Российской Федерации в лице Минобрнауки России: грант в форме субсидий на проведение крупных научных проектов по приоритетным направлениям научно-технического развития (Соглашение № 075-15-2024-644).

Financing: This work was supported by the Russian Federation Ministry of Science and Higher Education (Grant Agreement No. 075-15-2024-644) under the program for large scientific projects in priority areas of scientific and technological development.

Приложение А

```
Map.centerObject(Sunzha, 7);
Map.addLayer(roi, {color: 'red'}, 'roi');
Map.addLayer(Sunzha, {color: 'blue'}, 'Sunzha');
var hansen = ee.Image('UMD/hansen/global_forest_change_2025_v1_13');
var bands = ['treecover2000', 'loss', 'gain', 'lossyear'];
function exportRasters(geometry, name) {
  bands.forEach(function(band) {
    Export.image.toDrive({
      image: hansen.select(band).clip(geometry),
      description: name + '_' + band,
      folder: 'GEE_Hansen/' + name,
      fileNamePrefix: name + '_' + band,
      region: geometry,
      scale: 30,
      crs: 'EPSG:4326',
      maxPixels: 1e13
    });
  });
}
// roi
exportRasters(roi, 'roi');

// Sunzha
exportRasters(Sunzha, 'Sunzha');
```

Приложение Б

```
library(terra)
library(sf)

raster_path <- "D:/Sunzha_lossyear.tif"
shape_path <- "D:/01_Sunzha.shp"
output_dir <- "D:/Results"

nsim_global <- 999
nsim_local <- 99
nsim_sens <- 199
target_res_m <- 1000
sens_res_m <- c(500, 1000, 2000)
min_cells_target <- 100
max_cells_warn <- 5000
```

```

w_queen <- matrix(c(1, 1, 1,
                    1, 0, 1,
                    1, 1, 1), nrow = 3)
w_rook <- matrix(c(0, 1, 0,
                   1, 0, 1,
                   0, 1, 0), nrow = 3)

if (!file.exists(raster_path)) stop("Raster not found: ", raster_path)
if (!file.exists(shape_path)) stop("Shapefile not found: ", shape_path)
if (!dir.exists(output_dir)) dir.create(output_dir, recursive = TRUE)

utm_epsg from sf <- function(x) {
  x4326 <- st_transform(st_union(x), 4326)
  xy <- st_coordinates(st_centroid(x4326))
  if (nrow(xy) == 0 || any(!is.finite(xy[, c("X", "Y")]))) {
    stop("Cannot determine UTM zone: basin centroid is undefined.")
  }
  zone <- floor((xy[, "X"] + 180) / 6) + 1
  if (zone < 1 || zone > 60) stop("UTM zone out of range: ", zone)
  hemi <- if (xy[, "Y"] >= 0) 32600 else 32700
  as.integer(hemi + zone)
}

r <- rast(raster_path)
if (nlyr(r) < 1 || ncell(r) < 1) stop("Raster is empty.")
if (nlyr(r) > 1) r <- r[[1]]
if (is.na(crs(r)) || crs(r) == "") stop("Raster has no coordinate reference system.")

shp <- st_read(shape_path, quiet = TRUE)
if (nrow(shp) == 0) stop("Shapefile contains no features.")
if (is.na(st_crs(shp))) stop("Shapefile has no coordinate reference system.")

if (any(!st_is_valid(shp))) {
  shp <- st_make_valid(shp)
  if (any(!st_is_valid(shp))) stop("Invalid basin geometry could not be repaired.")
}
shp <- shp[!st_is_empty(shp), ]
if (nrow(shp) == 0) stop("All basin geometries are empty.")

geom_type <- unique(as.character(st_geometry_type(shp, by_geometry = TRUE)))
if (!any(geom_type %in% c("POLYGON", "MULTIPOLYGON", "GEOMETRYCOLLECTION"))) {
  stop("Basin must be a polygon layer. Found: ", paste(geom_type, collapse = ", "))
}

crs_r <- st_crs(crs(r))
crs_s <- st_crs(shp)
if (is.na(crs_r$epsg) || is.na(crs_s$epsg) || crs_r$epsg != crs_s$epsg) {
  if (!isTRUE(crs_r$wkt == crs_s$wkt)) {
    warning("Raster and shapefile CRS differ; shapefile is reprojected to the raster CRS.")
  }
}

shp_native <- st_transform(shp, crs(r))
if (any(!st_is_valid(shp_native))) shp_native <- st_make_valid(shp_native)

v_basin <- vect(shp_native)
if (nrow(v_basin) == 0) stop("Basin vector is empty after transformation.")

e_r <- ext(r)
e_s <- ext(v_basin)
if (is.null(intersect(e_r, e_s))) {
  stop("Basin does not intersect the raster extent.")
}
if (!isTRUE(related(e_s, e_r, relation = "within"))) {
  warning("Basin is only partially covered by the raster. Results apply to the overlapping part only.")
}

area_all <- sum(expanse(v_basin, unit = "m"), na.rm = TRUE)
if (!is.finite(area_all) || area_all <= 0) stop("Basin area is zero or undefined.")

v_overlap <- crop(v_basin, e_r)
if (nrow(v_overlap) == 0) stop("No basin geometry remains after clipping to the raster extent.")
area_overlap <- sum(expanse(v_overlap, unit = "m"), na.rm = TRUE)
if (is.finite(area_all) && area_all > 0 && area_overlap / area_all < 0.99) {
  warning(sprintf(
    "Partial overlap: %.1f %% of the basin area lies outside the raster.",
    100 * (1 - area_overlap / area_all)
  ))
}

```

```

r_crop <- crop(r, v_basin)
if(ncell(r_crop) < 1) stop("Cropped raster is empty.")

if (is.lonlat(r_crop)) {
  epsg <- utm_epsg_from_sf(shp)
  r_crop <- project(r_crop, paste0("EPSG:", epsg), method = "near", res = 30)
  shp <- st_transform(shp, epsg)
  if (any(!st_is_valid(shp))) shp <- st_make_valid(shp)
} else {
  shp <- shp_native
}

r_mask <- mask(r_crop, vect(shp))
n_mask <- as.numeric(global(r_mask, "notNA")[1, 1])
if (!is.finite(n_mask) || n_mask < 1) {
  stop("Empty mask: no valid raster cells inside the basin.")
}
if (n_mask == 1) {
  stop("Only one valid pixel inside the basin; spatial autocorrelation is undefined.")
}

all_vals <- freq(r_mask)
if (is.null(all_vals) || nrow(all_vals) == 0) {
  stop("Empty mask: frequency table contains no values.")
}
loss_vals <- all_vals[all_vals$value > 0 & all_vals$value <= 25, ]
if (nrow(loss_vals) == 0) {
  stop("No forest-loss codes (1-25) found inside the basin.")
}
if (any(all_vals$value < 0 | all_vals$value > 25, na.rm = TRUE)) {
  warning("Raster contains values outside the Hansen lossyear range 0-25.")
}

cell_area_km2 <- cellSize(r_mask, unit = "km")
if (as.numeric(global(cell_area_km2, "min", na.rm = TRUE)[1, 1]) <= 0) {
  stop("Non-positive cell area after projection.")
}

years <- 1:25
r_years <- segregate(r_mask, classes = years, other = NA, keep = FALSE)
loss_by_year <- as.data.frame(global(r_years * cell_area_km2, "sum", na.rm = TRUE))
loss_by_year$year <- 2000 + years
names(loss_by_year)[1] <- "loss_km2"
loss_by_year$loss_km2[is.na(loss_by_year$loss_km2)] <- 0
loss_by_year <- loss_by_year[, c("year", "loss_km2")]
if (sum(loss_by_year$loss_km2) <= 0) stop("Total forest loss is zero.")

write.csv(
  loss_by_year,
  file.path(output_dir, "annual_loss_sunzha_2001_2025_km2.csv"),
  row.names = FALSE
)

loss_all <- ifel(r_mask > 0, 1, 0)
pixel_res_m <- mean(res(r_mask))
working_crs <- crs(r_mask, describe = TRUE)

test_agg <- function(x, af) {
  if (af <= 1) {
    list(r = x, n = as.numeric(global(x, "notNA")[1, 1]), af = 1)
  } else {
    r_test <- aggregate(x, fact = af, fun = "sum", na.rm = TRUE)
    list(r = r_test, n = as.numeric(global(r_test, "notNA")[1, 1]), af = af)
  }
}

agg_factor_from_res <- function(res_m, pixel_m) {
  max(1, as.integer(round(res_m / pixel_m)))
}

moran_mc <- function(x, w, nsim, seed = 123) {
  I_obs <- as.numeric(autocor(x, w = w, method = "moran",
    global = TRUE, standardize = TRUE))
  n <- as.numeric(global(x, "notNA")[1, 1])

```

```

vals <- values(x)
vals <- vals[!is.na(vals)]
set.seed(seed)
I_sim <- numeric(nsim)
for (i in seq_len(nsim)) {
  r_perm <- x
  v <- values(r_perm)
  v[!is.na(v)] <- sample(vals)
  values(r_perm) <- v
  I_sim[i] <- as.numeric(autocor(r_perm, w = w, method = "moran",
    global = TRUE, standardize = TRUE))
}
p <- (sum(abs(I_sim) >= abs(I_obs), na.rm = TRUE) + 1) / (nsim + 1)
list(I = I_obs, p = p, n = n, Ei = -1 / (n - 1))
}

agg_factor <- agg_factor_from_res(target_res_m, pixel_res_m)
if (!is.finite(agg_factor) || agg_factor < 1) {
  stop("Aggregation factor is invalid. Check that the raster is in a projected CRS.")
}
result <- test_agg(loss_all, agg_factor)

while (result$n < min_cells_target && result$af > 1) {
  new_af <- max(1, floor(result$af / 2))
  if (new_af == result$af) new_af <- 1
  result <- test_agg(loss_all, new_af)
}

if (!is.finite(result$n) || result$n < min_cells_target) {
  stop(sprintf("Too few valid cells ( %s); need >= %d.", as.character(result$n), min_cells_target))
}
if (result$n < 3) {
  stop("Fewer than 3 spatial units; Moran's I is not defined.")
}

agg_factor <- result$af
actual_res_m <- round(agg_factor * mean(res(r_mask)))
loss_density <- result$R
names(loss_density) <- "loss_count"
if (as.numeric(global(loss_density, "notNA")[1, 1]) < 3) {
  stop("Aggregated raster has fewer than 3 valid cells.")
}
if (as.numeric(global(loss_density, "sd", na.rm = TRUE)[1, 1]) == 0) {
  stop("Loss values have zero variance; spatial autocorrelation cannot be computed.")
}

if (result$n > max_cells_warn && nsim_local > 0) {
  warning(sprintf("n = %d > %d; local Monte-Carlo may be slow.", result$n, max_cells_warn))
}

writeRaster(
  loss_density,
  file.path(output_dir, paste0("loss_density_", actual_res_m, ".tif")),
  overwrite = TRUE
)

summary_path <- file.path(output_dir, "summary_statistics_km2.txt")
sink(summary_path, split = TRUE)
on.exit(if (sink.number() > 0) sink(), add = TRUE)

cat("Forest cover loss, Sunzha River basin, 2001-2025\n\n")
cat("Total loss:", round(sum(loss_by_year$loss_km2), 2), "km2\n")
cat("Mean annual loss:", round(mean(loss_by_year$loss_km2), 2), "km2\n")
cat("Median annual loss:", round(median(loss_by_year$loss_km2), 2), "km2\n")
cat("Minimum:", loss_by_year$year[which.min(loss_by_year$loss_km2)],
  "-", round(min(loss_by_year$loss_km2), 2), "km2\n")
cat("Maximum:", loss_by_year$year[which.max(loss_by_year$loss_km2)],
  "-", round(max(loss_by_year$loss_km2), 2), "km2\n")
cat("Coefficient of variation:",
  round(sd(loss_by_year$loss_km2) / mean(loss_by_year$loss_km2) * 100, 1), "%\n")
cat("Working CRS:", working_crs$code, working_crs$name, "\n")
cat("Pixel size after projection:", round(pixel_res_m, 2), "m\n")
cat("Target cell size:", target_res_m, "m\n")
cat("Aggregation factor: round(", target_res_m, "/", round(pixel_res_m, 2),

```

```

") = ", agg_factor, "\n", sep = "")
cat("Actual aggregation resolution:", actual_res_m, "\n")
cat("Number of spatial units:", result$n, "\n")
cat("Weights: binary first-order queen (3x3), row-standardized for global I.\n")

nb_count <- focal(
  ifel(!is.na(loss_density), 1, NA),
  w = w_queen, fun = "sum", na.rm = TRUE
)
nb_vals <- values(nb_count)
nb_vals <- nb_vals[!is.na(values(loss_density))]
cat("Queen neighbors: min =", min(nb_vals), ", mean =",
  round(mean(nb_vals), 2), ", max =", max(nb_vals), "\n")

mc_main <- moran_mc(loss_density, w_queen, nsim_global, seed = 123)
I_obs <- mc_main$I
p_value <- mc_main$p
n_valid <- mc_main$n
Ei <- mc_main$Ei
if (!is.finite(I_obs)) stop("Global Moran's I is not finite.")
if (n_valid < 3) stop("Fewer than 3 valid cells for Moran's I.")

loss_interior <- ifel(nb_count == 8, loss_density, NA)
n_interior <- as.numeric(global(loss_interior, "notNA")[1, 1])
if (is.finite(n_interior) && n_interior >= 3 &&
  isTRUE(as.numeric(global(loss_interior, "sd", na.rm = TRUE)[1, 1]) > 0)) {
  mc_int <- moran_mc(loss_interior, w_queen, nsim_sens, seed = 123)
} else {
  mc_int <- list(I = NA_real_, p = NA_real_, n = n_interior)
}

cat("\nGlobal Moran's I:", round(I_obs, 4), "\n")
cat("E(I):", round(Ei, 4), "\n")
cat("Monte-Carlo p-value (", nsim_global, " permutations): ",
  sprintf("%.4f", p_value), "\n", sep = "")
if (p_value < 0.05 && I_obs > 0) {
  cat("Significant positive spatial autocorrelation.\n")
} else if (p_value < 0.05 && I_obs < 0) {
  cat("Significant negative spatial autocorrelation.\n")
} else {
  cat("Spatial autocorrelation is not significant (p >= 0.05).\n")
}
cat("Interior cells only (8 queen neighbors): n = ", mc_int$n,
  ", I = ", ifelse(is.finite(mc_int$I), sprintf("%.4f", mc_int$I), "NA"),
  ", p = ", ifelse(is.finite(mc_int$p), sprintf("%.4f", mc_int$p), "NA"),
  "\n", sep = "")

local_I <- autocor(loss_density, w = w_queen, method = "moran", global = FALSE)
names(local_I) <- "local_moran_I"
writeRaster(
  local_I,
  file.path(output_dir, paste0("lisa_local_I_", actual_res_m, ".tif")),
  overwrite = TRUE
)

mean_z <- as.numeric(global(loss_density, "mean", na.rm = TRUE)[1, 1])
lag_z <- focal(loss_density, w = w_queen, fun = "mean", na.rm = TRUE)
mean_lag <- as.numeric(global(lag_z, "mean", na.rm = TRUE)[1, 1])
if (!is.finite(mean_z) || !is.finite(mean_lag)) {
  stop("Cannot classify LISA quadrants: mean or spatial lag is undefined.")
}

vals_z <- values(loss_density)
vals_lag <- values(lag_z)
vals_I <- values(local_I)

quad <- rep(0L, length(vals_z))
quad[vals_z > mean_z & vals_lag > mean_lag] <- 1L
quad[vals_z < mean_z & vals_lag < mean_lag] <- 2L
quad[vals_z > mean_z & vals_lag < mean_lag] <- 3L
quad[vals_z < mean_z & vals_lag > mean_lag] <- 4L
quad[is.na(vals_z)] <- NA_integer_

run_local_mc <- nsim_local > 0

```

```

if(run_local_mc) {
  valid_idx <- which(!is.na(vals_z))
  I_obs_vec <- vals[I[valid_idx]]
  I_sim_mat <- matrix(NA_real_, nrow = length(valid_idx), ncol = nsim_local)
  set.seed(123)
  for(s in seq_len(nsim_local)) {
    r_perm <- loss_density
    v <- values(r_perm)
    v[valid_idx] <- sample(vals_z[valid_idx])
    values(r_perm) <- v
    I_sim_mat[, s] <- values(autocor(
      r_perm, w = w_queen, method = "moran", global = FALSE
    ))[valid_idx]
  }
  p_local <- (rowSums(abs(I_sim_mat) >= abs(I_obs_vec), na.rm = TRUE) + 1) /
    (nsim_local + 1)
  sig_mask <- rep(FALSE, length(vals_z))
  sig_mask[valid_idx] <- p_local < 0.05
  quad[!sig_mask] <- 0L
  quad[is.na(vals_z)] <- NA_integer_

  cat("Significant LISA cells (p < 0.05): ",
    sum(sig_mask, na.rm = TRUE), " / ", length(valid_idx), "\n", sep = "")

  r_pval <- setValues(loss_density, NA)
  values(r_pval)[valid_idx] <- p_local
  names(r_pval) <- "lisa_pvalue"
  writeRaster(
    r_pval,
    file.path(output_dir, paste0("lisa_pvalue_", actual_res_m, "m.tif")),
    overwrite = TRUE
  )
} else {
  cat("Local Monte-Carlo significance was not computed.\n")
}

r_cluster <- setValues(loss_density, quad)
rat <- data.frame(
  id = 0:4,
  cluster = c("NS", "HH", "LL", "HL", "LH"),
  description = c(
    "Not significant",
    "High-High",
    "Low-Low",
    "High-Low",
    "Low-High"
  )
)
levels(r_cluster) <- rat
names(r_cluster) <- "lisa_cluster"
writeRaster(
  r_cluster,
  file.path(output_dir, paste0("lisa_clusters_", actual_res_m, "m.tif")),
  overwrite = TRUE
)

tab <- table(
  factor(quad, levels = 0:4, labels = c("NS", "HH", "LL", "HL", "LH")),
  useNA = "no"
)
cat("\nLISA cluster counts:\n")
print(tab)
write.csv(rat, file.path(output_dir, "lisa_cluster_legend.csv"), row.names = FALSE)

sens_rows <- list()
for(res_m in sens_res_m) {
  af <- agg_factor_from_res(res_m, pixel_res_m)
  agg <- test_agg(loss_all, af)
  if(!is.finite(agg$sn) || agg$sn < 3) next
  dens <- agg$sr
  if(isTRUE(as.numeric(global(dens, "sd", na.rm = TRUE)[1, 1]) == 0)) next
  for(nb in c("queen", "rook")) {
    w <- if(nb == "queen") w_queen else w_rook
    mc <- moran_mc(dens, w, nsim_sens, seed = 123)
  }
}

```

```

sens_rows[[length(sens_rows) + 1]] <- data.frame(
  target_res_m = res_m,
  agg_factor = af,
  actual_res_m = round(af * pixel_res_m),
  neighborhood = nb,
  n = mc$n,
  moran_I = mc$I,
  expected_I = mc$Ei,
  p_value = mc$p,
  nsim = nsim_sens,
  stringsAsFactors = FALSE
)
}
sens_df <- do.call(rbind, sens_rows)
write.csv(sens_df, file.path(output_dir, "sensitivity_moran.csv"), row.names = FALSE)

cat("\nSensitivity of global Moran's I:\n")
print(sens_df, row.names = FALSE)

if (sink.number() > 0) sink()

```

Приложение В

```

library(terra)
library(sf)
library(dplyr)

raster_path <- "D:/GIS/Обезлесение/Hanson Forest Change/Sundzha_lossyear.tif"
shape_path <- "D:/Новая папка/00 Шейпы границы бассейнов/01_Sunzha.shp"
output_dir <- "D:/Results_Hansen"
if (!dir.exists(output_dir)) dir.create(output_dir, recursive = TRUE)

max_cores <- 32L
available_cores <- min(parallel::detectCores(logical = TRUE), max_cores)
physical_cores <- tryCatch(parallel::detectCores(logical = FALSE),
  error = function(e) NA_integer_)
mem_total_gb <- 128
n_reps <- 3L
warmup <- TRUE
loss_tol_km2 <- 1e-6
years_code <- 1:25
years_ad <- 2001:2025
agg_factor <- 33L
threads_list <- unique(c(1L, 2L, 4L, 8L, 16L,
  as.integer(min(32L, available_cores))))

env_backup <- Sys.getenv(c("GDAL_NUM_THREADS", "OMP_NUM_THREADS",
  "OMP_WAIT_POLICY"), unset = NA_character_)
terra_opt_backup <- tryCatch(terraOptions(print = FALSE), error = function(e) NULL)
`%||` <- function(a, b) if (is.null(a)) b else a

set_threads <- function(n) {
  n <- as.integer(n)
  terraOptions(parallel = TRUE, threads = n, progress = 0)
  Sys.setenv(GDAL_NUM_THREADS = n, OMP_NUM_THREADS = n,
    OMP_WAIT_POLICY = "PASSIVE")
  invisible(n)
}

reset_threads <- function() {
  for (nm in names(env_backup)) {
    v <- unname(env_backup[nm])
    if (is.na(v) || !nzchar(v)) Sys.unsetenv(nm)
    else do.call(Sys.setenv, stats::setNames(list(v), nm))
  }
  if (!is.null(terra_opt_backup)) {
    try(terraOptions(parallel = isTRUE(terra_opt_backup$parallel),
      threads = terra_opt_backup$threads %||% 16,
      progress = terra_opt_backup$progress %||% 3),
      silent = TRUE)
  }
}

```

```

time_mem <- function(expr) {
  expr <- substitute(expr); env <- parent.frame()
  gc(full = TRUE)
  mem0 <- gc()[, "used"]
  t0 <- Sys.time()
  out <- eval(expr, env)
  dt <- as.numeric(difftime(Sys.time(), t0, units = "secs"))
  mem1 <- sum(gc()[, "used"] - mem0, na.rm = TRUE)
  list(result = out, time_sec = dt, mem_mb = max(0, mem1))
}

run_once <- function(threads) {
  set_threads(threads)
  load <- time_mem(terra::rast(raster_path))
  r <- load$result
  shp <- time_mem({
    s <- sf::st_read(shape_path, quiet = TRUE)
    sf::st_transform(s, crs = terra::crs(r))
  })
  cropm <- time_mem({
    terra::mask(terra::crop(r, shp$result), shp$result)
  })
  r_mask <- cropm$result
  if (!inherits(r_mask, "SpatRaster") || terra::ncell(r_mask) == 0)
    stop("Empty raster after crop/mask.")
  area <- time_mem(terra::cellSize(r_mask, unit = "km"))
  cell_km2 <- area$result
  loss <- time_mem({
    df <- data.frame(year = years_ad, loss_km2 = NA_real_)
    for (i in seq_along(years_code)) {
      m <- r_mask == years_code[i]
      df$loss_km2[i] <- terra::global(m * cell_km2, "sum", na.rm = TRUE)[1, 1]
    }
    df
  })
  dens <- time_mem({
    terra::aggregate(r_mask > 0, fact = agg_factor, fun = "sum", na.rm = TRUE)
  })
  rm(r, r_mask, cell_km2); gc(full = TRUE)
  stages <- c("Загрузка раstra", "Загрузка шейпа", "Обрезка+маскирование",
    "Расчет площади", "Цикл по годам", "Агрегация плотности")
  tm <- list(load, shp, cropm, area, loss, dens)
  perf <- data.frame(
    stage = stages,
    time_sec = vapply(tm, `[`, 0, "time_sec"),
    mem_mb = vapply(tm, `[`, 0, "mem_mb"),
    threads = threads,
    stringsAsFactors = FALSE
  )
  list(perf = perf, loss = loss$result, density = dens$result)
}

if (warmup) invisible(run_once(max(threads_list)))

raw <- list(); losses <- list(); k <- 0L; density_last <- NULL
for (th in threads_list) {
  for (rep in seq_len(n_reps)) {
    res <- run_once(th)
    k <- k + 1L
    raw[[k]] <- transform(res$perf, rep = as.integer(rep),
      timestamp = format(Sys.time(), "%Y-%m-%d %H: %M: %S"))
    losses[[k]] <- res$loss
    density_last <- res$density
  }
}

perf_raw <- do.call(rbind, raw)
tot <- perf_raw %>%
  group_by(threads, rep, timestamp) %>%
  summarise(time_sec = sum(time_sec), mem_mb = sum(mem_mb), .groups = "drop") %>%
  mutate(stage = "Все этапы") %>%
  select(stage, time_sec, mem_mb, threads, rep, timestamp)

```

```

perf_raw <- bind_rows(perf_raw, tot)

summ <- perf_raw %>%
  group_by(stage, threads) %>%
  summarise(
    n_reps = n(),
    time_mean_sec = mean(time_sec), time_sd_sec = sd(time_sec),
    time_min_sec = min(time_sec), time_max_sec = max(time_sec),
    time_cv_pct = 100 * sd(time_sec) / mean(time_sec),
    mem_mean_mb = mean(mem_mb),
    groups = "drop"
  ) %>%
  group_by(stage) %>%
  mutate(
    speedup_mean = time_mean_sec[threads == 1][1] / time_mean_sec,
    efficiency_pct = 100 * speedup_mean / threads
  ) %>%
  ungroup()

ref_loss <- losses[[length(losses)]]
dmax <- max(vapply(losses, function(x)
  max(abs(x$loss_km2 - ref_loss$loss_km2), na.rm = TRUE), 0))

write.csv(perf_raw, file.path(output_dir, "performance_measurements_raw.csv"),
  row.names = FALSE, fileEncoding = "UTF-8")
write.csv(summ, file.path(output_dir, "performance_summary_table.csv"),
  row.names = FALSE, fileEncoding = "UTF-8")
write.csv(ref_loss, file.path(output_dir, "annual_loss_sunzha_2001_2025_km2.csv"),
  row.names = FALSE, fileEncoding = "UTF-8")

sink(file.path(output_dir, "performance_log.txt"))
cat("Hansen GFC, Sunzha | cores", available_cores, "/", physical_cores,
  " | RAM", mem_total_gb, "GB | reps", n_reps, " | warmup", warmup, "\n")
cat("threads:", paste(threads_list, collapse = ", "), "\n")
cat("max |d| loss_km2:", format(dmax, scientific = TRUE),
  if (dmax > loss_tol_km2) "OUT OF TOLERANCE\n" else "OK\n")
cat("\n")
for (st in c(unique(perf_raw$stage[perf_raw$stage != «Все этапы»]), «Все этапы»)) {
  cat("[", st, "]\n", sep = "")
  print(as.data.frame(summ[summ$stage == st,
    c("threads", "n_reps", "time_mean_sec", "time_sd_sec",
      "time_min_sec", "time_max_sec", "time_cv_pct",
      "speedup_mean", "efficiency_pct")]),
    row.names = FALSE, digits = 4)
  cat("\n")
}
sink()

set_threads(max(threads_list))
if (is.null(density_last)) {
  last <- run_once(max(threads_list))
  density_last <- last$density
  if (is.null(ref_loss)) ref_loss <- last$loss
}
terra::writeRaster(density_last, file.path(output_dir, "loss_density_1km.tif"),
  overwrite = TRUE)

sink(file.path(output_dir, "summary_statistics_km2.txt"))
cat("Forest-cover loss, Sunzha basin, 2001-2025 (km^2)\n")
cat("total: ", sum(ref_loss$loss_km2), "\n", sep = "")
cat("mean: ", round(mean(ref_loss$loss_km2), 2), "\n", sep = "")
cat("median: ", round(median(ref_loss$loss_km2), 2), "\n", sep = "")
cat("min: ", ref_loss$year[which.min(ref_loss$loss_km2)], " ",
  round(min(ref_loss$loss_km2), 2), "\n", sep = "")
cat("max: ", ref_loss$year[which.max(ref_loss$loss_km2)], " ",
  round(max(ref_loss$loss_km2), 2), "\n", sep = "")
cat("CV %: ", round(sd(ref_loss$loss_km2) / mean(ref_loss$loss_km2) * 100, 1),
  "\n", sep = "")
sink()

reset_threads()

```

Приложение Г

```

#!/usr/bin/env Rscript
suppressPackageStartupMessages(library(terra))
terraOptions(progress = 0)

OUT_DIR <- "synthetic_test"
dir.create(OUT_DIR, showWarnings = FALSE, recursive = TRUE)
YEARS <- 2001:2025
NY <- length(YEARS)
N <- 2000L
RES <- 0.00025
XMIN <- 45.00
YMAX <- 43.50
MASK_R1 <- 50L; MASK_R2 <- N - 50L
MASK_C1 <- 50L; MASK_C2 <- N - 50L

A_WGS <- 6378137.0
F_WGS <- 1 / 298.257223563
E2 <- 2 * F_WGS - F_WGS^2
EW <- sqrt(E2)
F_lat <- function(phi_deg) {
  s <- sin(phi_deg * pi / 180)
  A_WGS^2 * (1 - E2) * (s / (2 * (1 - E2 * s^2))) +
    log((1 + EW * s) / (1 - EW * s)) / (4 * EW)
}
row_cell_area_m2 <- function(row) {
  phi_top <- YMAX - (row - 1L) * RES
  phi_bot <- phi_top - RES
  (RES * pi / 180) * (F_lat(phi_top) - F_lat(phi_bot))
}
rect_from_cells <- function(r, r1, r2, c1, c2) {
  e <- ext(r); rs <- res(r)
  x1 <- xmin(e) + (c1 - 1) * rs[1]; x2 <- xmin(e) + c2 * rs[1]
  y2 <- ymax(e) - (r1 - 1) * rs[2]; y1 <- ymax(e) - r2 * rs[2]
  vect(cbind(c(x1, x2, x2, x1), c(y1, y1, y2, y2, y1))),
  type = "polygons", crs = crs(r)
}
fmt <- function(x, d = 6) formatC(x, format = "f", digits = d)

make_test <- function() {
  message("Генерация растра ", N, " x ", N, " в EPSG:4326, шаг ", RES, "...")
  r <- rast(nrows = N, ncols = N,
    xmin = XMIN, xmax = XMIN + N * RES,
    ymin = YMAX - N * RES, ymax = YMAX,
    crs = "EPSG:4326", vals = 0L)
  counts <- 60L * seq_len(NY)
  band_row0 <- c(150L, 900L, 1600L)
  band_col0 <- c(200L, 650L, 1150L)
  area_km2 <- numeric(NY)
  frag_log <- list()
  for (y in seq_len(NY)) {
    n <- counts[y]
    n1 <- n %/% 3L; n2 <- n %/% 3L; n3 <- n - n1 - n2
    ns <- c(n1, n2, n3)
    rows_y <- band_row0 + 2L * y
    tot <- 0
    for (k in 1:3) {
      if (ns[k] == 0L) next
      cc <- band_col0[k]:(band_col0[k] + ns[k] - 1L)
      stopifnot(max(cc) <= MASK_C2, min(cc) >= MASK_C1,
        rows_y[k] >= MASK_R1, rows_y[k] <= MASK_R2)
      r[cellFromRowCol(r, rep(rows_y[k], length(cc)), cc)] <- y
      tot <- tot + ns[k] * row_cell_area_m2(rows_y[k])
    }
    area_km2[y] <- tot / 1e6
    frag_log[[y]] <- data.frame(year = YEARS[y], row = rows_y, n = ns)
  }
  leak <- data.frame(year = YEARS, area_km2 = 0)
  add_leak <- function(code, row, cols) {
    r[cellFromRowCol(r, rep(row, length(cols)), cols)] <- code
    leak$area_km2[code] <- leak$area_km2[code] +
      length(cols) * row_cell_area_m2(row) / 1e6
  }
}

```

```

add_leak(16L, 10L, 10:509)
add_leak(1L, 1000L, 1:40)
add_leak(25L, N - 10L, 100:399)
add_leak(7L, 1200L, (N - 39L):N)
p <- rect from cells(r, MASK_R1, MASK_R2, MASK_C1, MASK_C2)
f_tif <- file.path(OUT_DIR, "B_lossyear_geo.tif")
f_shp <- file.path(OUT_DIR, "B_basin_geo.shp")
writeRaster(r, f_tif, datatype = "INT1U", overwrite = TRUE)
writeVector(p, f_shp, overwrite = TRUE)
expected <- data.frame(year = YEARS, n_pixels = counts, area_km2 = area_km2)
write.csv(expected, file.path(OUT_DIR, "B_expected.csv"), row.names = FALSE)
write.csv(do.call(rbind, frag_log), file.path(OUT_DIR, "B_fragments.csv"),
  row.names = FALSE)
write.csv(leak, file.path(OUT_DIR, "B_leak_signature.csv"), row.names = FALSE)
message(" пикселей внутри маски: ", sum(counts),
  ", эталонная сумма: ", fmt(sum(area_km2), 6), " км²")
message(" пикселей за границей маски: «,
  500 + 40 + 300 + 40, « (подлежат исключению)»)
message(" средняя площадь ячейки по аналитике: «,
  fmt(row_cell_area_m2(1000L), 3), " м² на широте «,
  fmt(YMAX - 999 * RES, 4), "°»)
list(tif = f_tif, shp = f_shp, expected = expected, leak = leak)
}

MODE <- "reference"
call_your_pipeline <- function(rast_path, mask_path) {
  stop("Задайте вызов вашей функции в call_your_pipeline()»)
}
CSV_PATH <- file.path(OUT_DIR, "result_B.csv")
CSV_YEAR_COL <- "year"
CSV_AREA_COL <- "area_km2"
CSV_UNITS <- "km2"
read_pipeline_csv <- function(path) {
  if (!file.exists(path)) stop("Файл не найден: ", path)
  df <- read.csv(path, stringsAsFactors = FALSE, check.names = FALSE)
  if (!CSV_YEAR_COL %in % names(df) || !CSV_AREA_COL %in % names(df))
    stop("Нет столбцов '", CSV_YEAR_COL, "' / '", CSV_AREA_COL,
      "' . Найденны: ", paste(names(df), collapse = ", "))
  k <- switch(CSV_UNITS, km2 = 1, ha = 0.01, m2 = 1e-6, stop("CSV_UNITS"))
  yr <- as.numeric(df[[CSV_YEAR_COL]])
  if (all(yr %in % 1:25, na.rm = TRUE)) yr <- 2000 + yr
  data.frame(year = yr, area_km2 = as.numeric(df[[CSV_AREA_COL]]) * k)
}
reference_pipeline <- function(rast_path, mask_path) {
  r <- rast(rast_path); v <- vect(mask_path)
  if (crs(v) != crs(r)) v <- project(v, crs(r))
  r <- mask(crop(r, v), v)
  names(r) <- "lossyear"
  a <- cellSize(r, unit = "m"); names(a) <- "area_m2"
  z <- zonal(a, r, fun = "sum", na.rm = TRUE); names(z) <- c("code", "area_m2")
  out <- data.frame(year = YEARS, area_km2 = 0)
  for (y in seq_len(NY)) {
    hit <- z$area_m2[z$code == y]
    if (length(hit) == 1) out$area_km2[y] <- hit / 1e6
  }
  out
}
run_pipeline <- function(rast_path, mask_path) {
  switch(MODE,
    reference = reference_pipeline(rast_path, mask_path),
    "function" = call_your_pipeline(rast_path, mask_path),
    csv = read_pipeline_csv(CSV_PATH),
    stop("MODE: 'reference', 'function' или 'csv'"))
}

compare <- function(expected, actual, leak, tol_rel_pc = 1e-6) {
  cat("\n", strep("=", 76), "\n", sep = "")
  cat("ТЕСТ В. Площади потерь в EPSG:4326, эталон – эллипсоид WGS84\n»)
  cat(strep("=", 76), "\n", sep = "")
  m <- merge(expected, actual, by = "year", suffixes = c("_exp", "_act"))
  if (nrow(m) != nrow(expected))
    cat("ВНИМАНИЕ: сопоставлено «, nrow(m), « лет из «, nrow(expected), «\n», sep = «»)
  m$diff_km2 <- m$area_km2_act - m$area_km2_exp
  m$rel_diff_pc <- 100 * m$diff_km2 / m$area_km2_exp
}

```

```

print(data.frame(
  Год = m$year,
  Пикселей = m$n_pixels,
  Эталон = fmt(m$area_km2_exp),
  Скрипт = fmt(m$area_km2_act),
  Разность = fmt(m$diff_km2),
  Отн_проц = formatC(m$rel_diff_pc, format = "e", digits = 2),
  check.names = FALSE
), row.names = FALSE)
s_exp <- sum(m$area_km2_exp); s_act <- sum(m$area_km2_act)
max_abs <- max(abs(m$diff_km2)); max_rel <- max(abs(m$rel_diff_pc))
cat("\nСумма эталон: ", fmt(s_exp), " км²\n", sep = "")
cat("Сумма скрипт: ", fmt(s_act), " км²\n", sep = "")
cat("Разность: ", fmt(s_act - s_exp), " км² (",
  formatC(100 * (s_act - s_exp) / s_exp, format = "e", digits = 2), " %)\n", sep = "")
cat("Макс. абс. разность: ", fmt(max_abs), " км²\n", sep = "")
cat("Макс. отн. разность: ", formatC(max_rel, format = "e", digits = 2), " %\n", sep = "")
ok <- max_rel <= tol_rel_pc
cat("\nРЕЗУЛЬТАТ: «, if(ok) «ПРОЙДЕН» else «НЕ ПРОЙДЕН»,
  (допуск ", formatC(tol_rel_pc, format = "e", digits = 1), " %)\n", sep = "")
lk <- merge(m[, c("year", "diff_km2")], leak, by = "year")
lk <- lk[lk$area_km2 > 0, ]
hit <- abs(lk$diff_km2 - lk$area_km2) < 0.05 * lk$area_km2
cat("\nПроверка маскирования: за границей полигона размещены пиксели\n»)
cat("потерь с кодами ", paste(lk$year, collapse = ", ", ".\n", sep = ""))
if (any(hit)) {
  cat("«ПРОТЕЧКА МАСКИ: отклонения по годам «,
    paste(lk$year[hit], collapse = ", ", ".\n", sep = «»)
  } else {
    cat("«Протечки не обнаружено: посторонние пиксели в расчет не вошли.\n»)
  }
if (!ok) {
  cat("«\nДиагностика:\n»)
  if (all(m$diff_km2 > 0) || all(m$diff_km2 < 0)) {
    cat("« Отклонение одностороннее, в среднем «,
      formatC(mean(abs(m$rel_diff_pc)), format = "e", digits = 2),
      " %, разброс «,
      formatC(diff(range(m$rel_diff_pc)), format = "e", digits = 2), " %.\n", sep = ""))
    cat("« Это подпись не арифметической ошибки, а конвенции расчета\n»)
    cat("« площади: суммирование в плоскости равноугольной проекции (UTM)\n»)
    cat("« вместо истинной площади на эллипсоиде. Величина смещения около\n»)
    cat("« 0.07–0.08 %, знак отрицательный. Устраняется расчетом площади\n»)
    cat("« через terra::cellSize() на исходном растре в EPSG:4326 либо\n»)
    cat("« переходом на равновеликую проекцию.\n»)
  } else {
    cat("« Отклонение разнознаковое: вероятно ошибка в учете площади\n»)
    cat("« ячейки либо потеря пикселей при перепроецировании.\n»)
  }
}
write.csv(m, file.path(OUT_DIR, "B_comparison.csv"), row.names = FALSE)
con <- file(file.path(OUT_DIR, "B_table.md"), open = "w", encoding = "UTF-8")
writeLines(c(
  "| Год | Задано пикселей | Аналитическая площадь, км² | Расчет скрипта, км² | Δ, км² |",
  "|---|---|---|---|---|",
  sprintf("| %d | %d | %s | %s | %s |", m$year, m$n_pixels,
    fmt(m$area_km2_exp, 4), fmt(m$area_km2_act, 4), fmt(m$diff_km2, 4)),
  sprintf("| **Сумма** | ** %d** | ** %s** | ** %s** | ** %s** |",
    sum(m$n_pixels), fmt(s_exp, 4), fmt(s_act, 4), fmt(s_act - s_exp, 4))
), con)
close(con)
invisible(list(table = m, ok = ok, max_abs = max_abs, max_rel = max_rel))
}

cat(strrep("=", 76), "\n»)
cat("«ВЕРИФИКАЦИЯ НА СИНТЕТИЧЕСКИХ ДАННЫХ. Тест В. Режим: ", MODE, "\n", sep = "")
if (MODE == "reference")
  cat("«ВНИМАНИЕ: демонстрационный режим. Проверяется стенд, а не ваш код.\n»)
cat(strrep("=", 76), "\n\n")
TB <- make_test()
cat("\nЗапуск конвейера...\n»)
res <- run_pipeline(TB$tif, TB$shp)
chk <- compare(TB$expected, res, TB$leak)
cat("\nФайлы для статьи: ", OUT_DIR, "/B_comparison.csv и B_table.md\n", sep = "")

```