

УДК 004.9:551.311.8:004.75

ОПЫТ РАЗРАБОТКИ ГЕОИНФОРМАЦИОННОЙ СИСТЕМЫ МОНИТОРИНГА СЕЛЕЙ ДЛЯ МЧС КЫРГЫЗСТАНА: ВОСЕМЬ ПРИКЛАДНЫХ ЗАДАЧ ПРОСТРАНСТВЕННОГО АНАЛИЗА

Божокоев М.М., Орозобекова А.К.

Кыргызский государственный технический университет
им. И. Раззакова

В статье обобщён опыт разработки геоинформационной системы мониторинга селей на территории Кыргызской Республики. Рассмотрены восемь прикладных задач пространственного анализа, возникших в ходе проектирования, и предложенные технические решения. Особое внимание уделено результатам нагрузочного тестирования: показано, что зональная статистика по учётным водотокам без соответствующих оптимизаций приводит к 18% отказов под нагрузкой и требует двух последовательных мер оптимизации для снижения этого показателя до нуля. Установлено также, что перенос расчёта тепловой карты на сторону клиента (браузера) обеспечивает более высокую производительность по сравнению с серверным расчётом при сопоставимых аппаратных ресурсах.

Ключевые слова: геоинформационная система, мониторинг природных рисков, сели, Кыргызстан, зональная статистика, GIST-индекс, пространственный анализ, PostGIS, Redis.

КЫРГЫЗСТАНДЫН ӨЗГӨЧӨ КЫРДААЛДАР МИНИСТРЛИГИ ҮЧҮН СЕЛДИ МОНИТОРИНГДӨӨ ҮЧҮН ГЕОИНФОРМАЦИЯЛЫК СИСТЕМАНЫ ИШТЕП ЧЫГУУ ТАЖРЫЙБАСЫ: МЕЙКИНДИК АНАЛИЗДИН СЕГИЗ КОЛДОНУЛГАН МАСЕЛЕЛЕРИ

Божокоев М.М., Орозобекова А.К.

И. Раззаков атындагы Кыргыз мамлекеттик техникалык университети

Бул макалада Кыргыз Республикасында селди мониторингдөө үчүн геомаалыматтык системаны иштеп чыгуу тажрыйбасы кыскача

баяндалган. Долбоорлоо процессинде пайда болгон сегиз колдонмо мейкиндик анализинин маселелери жана сунушталган техникалык чечимдер каралат. Жүктү сыноонун жыйынтыктарына өзгөчө көңүл бурулат: тиешелүү оптималдаштыруусуз өлчөнгөн агымдар үчүн зоналык статистика жүк астында 18% бузулуу көрсөткүчүнө алып келери жана бул көрсөткүчтү нөлгө чейин төмөндөтүү үчүн эки удаалаш оптималдаштыруу чарасын талап кылаары көрсөтүлгөн. Ошондой эле жылуулук картасын эсептөөнү кардарга (браузерге) өткөрүп берүү салыштырмалуу аппараттык ресурстар менен сервер тарабындагы эсептөөлөргө салыштырмалуу жогорку көрсөткүчтөрдү камсыз кылаары аныкталды.

Баштапкы сөздөр: географиялык маалымат системасы, жаратылыш кырсыктарын мониторингдөө, сел жүрүүлөрү, Кыргызстан, зоналык статистика, GIST индекси, мейкиндик анализи, PostGIS, Redis.

EXPERIENCE IN DEVELOPING A GEOINFORMATION SYSTEM FOR MUDFLOW MONITORING FOR THE MINISTRY OF EMERCOM OF KYRGYZSTAN: EIGHT APPLIED PROBLEMS OF SPATIAL ANALYSIS

Bozhokoev M.M., Orozbekova A.K.

Kyrgyz State Technical University named after I. Razzakov

This article summarizes the experience of developing a geoinformation system for mudflow monitoring in the Kyrgyz Republic. Eight applied spatial analysis problems that arose during the design process and the proposed technical solutions are considered. Particular attention is paid to the results of load testing: it is shown that zonal statistics for metered streams without appropriate optimizations leads to an 18% failure rate under load and requires two consecutive optimization measures to reduce this rate to zero. It was also found that transferring heat map calculation to the client (browser) provides higher performance compared to server-side calculations with comparable hardware resources.

Keywords: geographic information system, natural hazard monitoring, mudflows, Kyrgyzstan, zonal statistics, GIST index, spatial analysis, PostGIS, Redis.

Введение.

Кыргызстан является одной из стран, наиболее подверженных стихийным бедствиям. Опасные природные геологические процессы (известные как «геоопасность»), в основном оползни, сели и

крупномасштабные речные эрозии, являются наиболее частыми и опасными процессами окружающей среды в Кыргызстане, наносящими большой ущерб как с точки зрения потери жизни и средств к существованию, так и перебоев в движении транспорта и осуществлении экономической деятельности.

Сель представляет собой кратковременный паводковый или прорывной водный поток с высоким содержанием грязекаменного материала на горных реках, состоящий из полужидких грязекаменных масс и обладающий колоссальной разрушительной силой. Причины их возникновения — интенсивные и продолжительные ливни, бурное таяние снега и льда в горах, прорывы ледниковых озёр. В результате прорыва перемычки запрудного озера или перенасыщения влагой грунта в верховьях реки образуется поток, в который по мере его продвижения вниз по долине вовлекается смываемый материал. Процесс развивается лавинообразно: сотни тысяч кубических метров пульпы и обломков горных пород представляют реальную опасность, и даже селезащитные сооружения не всегда выполняют удерживающую и отводящую функцию. Нередко они вызывают катастрофические последствия и представляют большую опасность для населения горных районов [14].

Постановка задачи

Когда сотрудник Департамента мониторинга МЧС Кыргызской Республики получает сигнал о сходе селя, у него нет инструмента, который в одном интерфейсе показал бы: какой именно водоток, насколько он опасен исторически, сколько людей живёт ниже по течению, есть ли там объекты СПЛМ и какова текущая метеообстановка. Вся эта информация хранится разрозненно — в таблицах Excel, бумажных паспортах, отдельных архивах Кыргызгидромета. Настоящая работа посвящена созданию платформы, которая объединяет эти данные в единой геоинформационной системе с аналитикой в реальном времени.

Масштаб задачи определяется статистикой: по материалам Атласа селе-паводковой опасности КР 2025 года [31], за 1990–2024 годы в республике произошло 1 990 природных ЧС, связанных с селями. На учёте стоят 3 103 селеопасных водотока, около 370 прорывоопасных озёр и 198 объектов противоселевой защиты. Всё это — огромный массив данных, который сейчас практически не анализируется в пространственном контексте.

На учёте МЧС Кыргызской Республики состоят 3 103 реки с признаками селеопасности и около 370 горных озёр с риском прорыва [1,6].

Оценка лавинной и селевой опасности горных территорий имеет многолетнюю методическую традицию [3]. При постановке задачи объединения этих данных в единой геоинформационной системе предполагалось, что основная сложность заключается в консолидации разнородных источников. Однако данный этап оказался лишь частью проблемы: наиболее существенные трудности проявились при обеспечении работоспособности системы в условиях многопользовательской нагрузки.

Целью работы является разработка и практическая апробация методов пространственного анализа, обеспечивающих работоспособность геоинформационной системы мониторинга селей в условиях многопользовательской нагрузки.

Научная новизна работы заключается в эмпирическом обосновании выбора между серверными и клиентскими вычислениями на основе результатов нагрузочного тестирования, а также в подтверждении недостаточности применения только пространственных индексов для обеспечения требуемой производительности зональной статистики.

Практическая значимость определяется тем, что предложенные решения внедрены в систему, эксплуатируемую в интересах МЧС Кыргызской Республики, и могут быть применены при создании аналогичных систем мониторинга природных рисков, в том числе в условиях разнородности систем координат исходных данных, характерной для постсоветских стран.

Ниже приведено описание каждой из рассмотренных задач в отдельном разделе. Последовательность изложения в целом соответствует порядку их решения в ходе разработки.

Задача 1. Согласование данных в трёх системах координат:

Исходные данные поступили в трёх различных системах координат. Картографические материалы советского периода оцифрованы в системе Пулково 1942, зона 13 (EPSG:28413); данные Кыргызгидромета представлены в проекции UTM, зона 42N (EPSG:32642); современные GPS-измерения - в системе WGS84 (EPSG:4326). Совмещение этих наборов без перепроецирования невозможно, поскольку взаимное смещение достигает 40 км.

Рассматривались два подхода:

Первый - предварительная конвертация всех файлов в единую систему координат - предполагает выполнение постоянной ручной операции при добавлении новых данных и сопряжён с риском ошибок преобразования.

Второй - хранение данных в исходных системах координат с автоматическим перепроецированием на этапе выполнения запроса - требует больших вычислительных ресурсов, однако исключает ручной труд. Был выбран второй подход. В дальнейшем к системе было подключено ещё четыре источника данных, ни один из которых не потребовал изменения схемы базы данных.

Задача 2. Буферный анализ в угловых единицах формирует эллипс, а не окружность

Первоначальная реализация буферного анализа выполнялась в угловых единицах. Анализ результатов на карте показал, что построенные «окружности» имеют вытянутую форму. На широте Бишкека один градус долготы соответствует приблизительно 81 км, а один градус широты - 111 км. При радиусе буфера 5 км погрешность составила 1,85 км, что достаточно для ошибочного отнесения ряда населённых пунктов к зоне поражения.

Расчёт был переработан в метрических единицах по поверхности эллипсоида WGS84. Данный подход примерно втрое более ресурсоёмок по сравнению с расчётом в угловых единицах. По результатам нагрузочного тестирования (150 одновременных пользователей) среднее время отклика составило 213 мс, значение 95-го перцентиля (P95) - 387 мс, что соответствовало предъявляемым требованиям. Применение GIST-индекса для таблицы населённых пунктов позволило снизить исходное время отклика с 640 мс до указанного значения.

Задача 3. Обеспечение строгого разграничения доступа к данным по территориальному признаку

Организационная структура МЧС построена по территориальному принципу: каждый специалист отвечает за закреплённый за ним район, и предоставление ему данных по сопредельным районам нежелательно. Первоначальный вариант предполагал фильтрацию данных на уровне прикладного кода после получения полной выборки из базы данных. Недостаток такого подхода состоит в том, что при прямом обращении к API в обход пользовательского интерфейса данная фильтрация может быть обойдена.

Был применён иной подход: идентификаторы разрешённых районов включаются в JWT-токен через коды СОАТЭ (система обозначений административно-территориальных единиц Кыргызской Республики). Сервер извлекает эти идентификаторы из токена и непосредственно включает их в SQL-запрос. Таким образом, база данных не возвращает строки за пределами разрешённого территориального контура - не

вследствие логики приложения, а в силу отсутствия соответствующих строк в результате запроса. Среднее время отклика составило 87 мс, P95 - 142 мс.

Задача 4. Несоответствие зональной статистики требованиям нагрузочного тестирования

Из восьми рассмотренных методов зональная статистика оказалась единственным, не удовлетворившим требованиям первого нагрузочного теста. При отсутствии оптимизации среднее время отклика составило 892 мс, P95 - 1240 мс, доля отказов - 18%. Подобная доля отказов свидетельствует о неработоспособности метода, а не только о его низком быстродействии.

Применение GIST-индекса для поля границ районов [4] позволило снизить среднее время отклика до 534 мс, а долю отказов - до 3%, что также не соответствовало требованиям. Динамика обоих показателей по мере добавления мер оптимизации представлена на рисунке 1.

Дополнительно было реализовано кэширование в Redis с временем хранения результата запроса по конкретному району 300 секунд. Повторный запрос обслуживается из кэша за 156 мс при нулевой доле отказов. Снижение среднего времени отклика относительно исходного составило 5,7 раза.

Существенным аспектом является необходимость немедленной инвалидации кэша при регистрации нового события, без ожидания истечения 300-секундного интервала. В противном случае специалист получает устаревшую статистику в наиболее критический момент - непосредственно после фиксации нового инцидента.

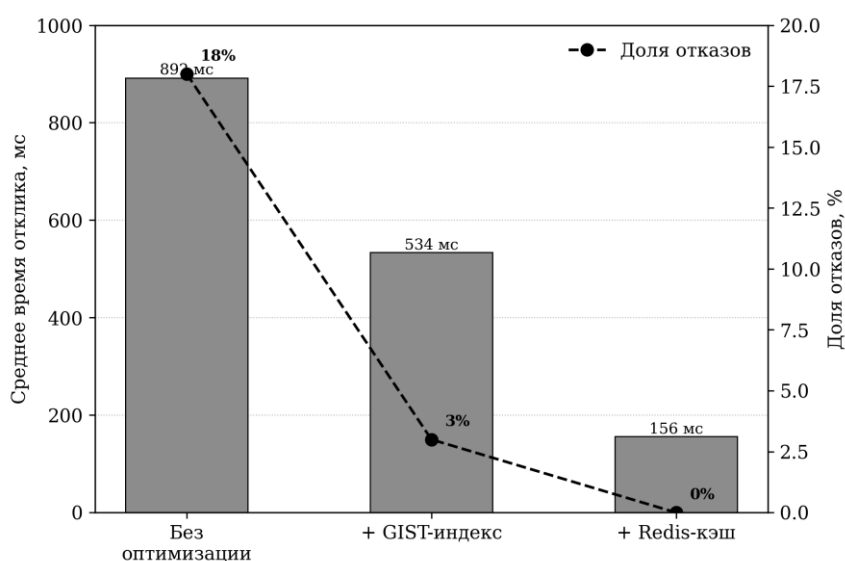


Рис. 1. Влияние мер оптимизации на среднее время отклика и долю отказов при выполнении зональной статистики

Задача 5. Оценка численности населения в зоне риска: три уровня сложности

На первый взгляд задача представляется простой: определить населённые пункты, попадающие в зону поражения, и просуммировать численность их жителей. Однако крупный населённый пункт может входить в зону лишь частично, и суммирование всего его населения приводит к завышению оценки. В связи с этим была реализована методика расчёта доли перекрытия площадей с пропорциональным взвешиванием численности населения. Операция пересечения двух полигонов вычислительно более затратна, чем определение принадлежности точки полигону, однако при выполнении по явному запросу (а не при каждом действии пользователя) полученное время отклика 340 мс является приемлемым.

Второй уровень сложности связан с форматом исходных данных: сведения о границах населённых пунктов предоставлены Нацстаткомом КР в виде полигонов, а не точек. Это повышает точность расчёта, однако требует применения операции ST_Intersection вместо более простой ST_Within.

Третий уровень сложности носит организационный характер. Сотрудники МЧС нуждались в возможности ручной корректировки результатов автоматического расчёта, учитывающей фактическую обстановку (частичная эвакуация населения, размещение временного лагеря в зоне риска и т.п.). Для этого потребовалось хранить не только итоговое значение, но и журнал изменений с фиксацией автора, времени и причины каждой корректировки. Автоматический расчёт и ручная правка функционируют совместно, не замещая друг друга.

Задача 6. Преимущество клиентского расчёта тепловой карты над серверным

Для отображения исторической концентрации событий первоначально применялся серверный подход с формированием растрового изображения. При его использовании выявлен ряд недостатков: растровое изображение не масштабируется при изменении масштаба карты без потери качества; генерация для всей территории КР занимает 3–8 секунд; изменение временного диапазона требует повторного обращения к серверу.

Альтернативный подход предполагает передачу сервером только массива из приблизительно 3000 точек с весовыми коэффициентами, при этом построение тепловой карты выполняется в браузере средствами

WebGL. Время серверного отклика составило 178 мс (P95 = 312 мс). Пересчёт при изменении масштаба или временного диапазона осуществляется на стороне клиента за 40–80 мс без обращения к серверу. Сравнение временных характеристик обоих подходов приведено на рисунке 2.

Примечательно, что данный эндпоинт при передаче около 3000 точек обеспечил более высокую производительность, чем буферный анализ (P95 = 387 мс), несмотря на существенно больший объём передаваемых данных. Это объясняется тем, что сервер выполняет лишь выборку данных, тогда как буферный анализ дополнительно вычисляет геометрические пересечения.

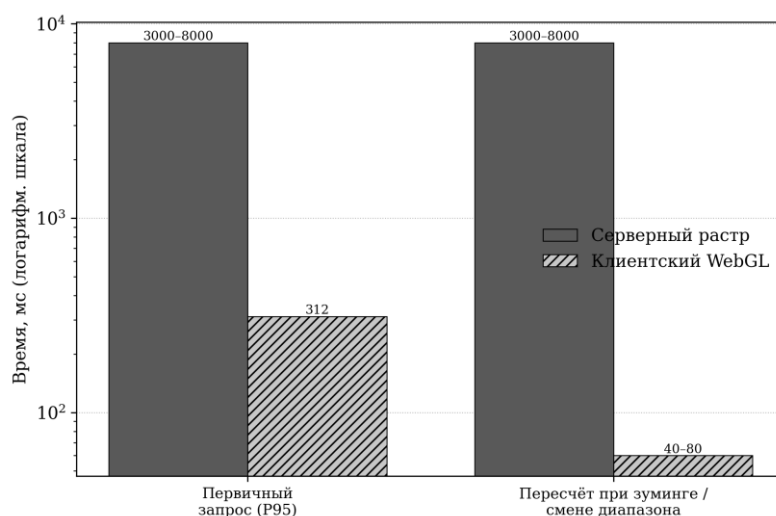


Рис. 2 – Сравнение времени построения тепловой карты при серверном (растровом) и клиентском (WebGL) подходах (логарифмическая шкала)

Задача 7. Функция, не предусмотренная техническим заданием, но наиболее востребованная

В ходе демонстрации прототипа один из сотрудников МЧС запросил возможность оценки крутизны склона в указанной точке карты. Функции отображения уклона и высоты в исходном техническом задании предусмотрено не было.

В систему был добавлен растровый слой SRTM с разрешением 30 м и реализован расчёт производных рельефа по клику. Уклон 25° и более является пороговым значением при оценке вероятности схода грязекаменного потока [2]. Среднее время отклика составило 134 мс благодаря применению растрового GIST-индекса.

На практике данная функция наиболее востребована при планировании эвакуационных маршрутов: специалист указывает предполагаемый маршрут и непосредственно получает информацию о наличии участков с опасным уклоном.

Задача 8. Обеспечение устойчивости к нестабильности внешнего API спутниковых данных

Платформа Google Earth Engine (GEE) [7] предоставляет доступ к данным NDVI, данным о затоплениях по радарным снимкам Sentinel-1 (SAR), а также к данным о влажности почвы. Применение спутниковых данных для оценки опасности схода грязекаменных потоков, инициируемых осадками, рассматривается в ряде работ [5]. Данные SAR имеют особое значение, поскольку радарная съёмка осуществляется сквозь облачный покров, тогда как в период активных осадков оптические снимки горных территорий малопригодны.

Выполнение прямых запросов к API GEE при каждом обращении пользователя оказалось нецелесообразным: время отклика варьировалось от 1,5 до 9 секунд в зависимости от текущей нагрузки на платформу GEE, что при 150 пользователях неприемлемо.

Был реализован отдельный планировщик, обращающийся к GEE по расписанию (NDVI - раз в трое суток, данные о затоплениях - ежедневно) и сохраняющий результаты в Redis. Обращения пользователей обслуживаются из кэша: среднее время отклика составило 42 мс, P95 - 98 мс.

Дополнительным положительным эффектом данной архитектуры стала повышенная отказоустойчивость: за три месяца тестирования зафиксировано два случая недоступности платформы GEE, при которых система продолжала отображать последние сохранённые в кэше данные. Все запросы к GEE географически ограничены ограничивающим прямоугольником территории Кыргызстана [68,0°–81,0° в.д.; 39,0°–43,5° с.ш.], что предотвращает ошибки при некорректном задании параметров. Дальнейшим направлением развития системы является автоматизированное выявление зон проявления склоновых процессов по спутниковым снимкам, методы которого активно разрабатываются [8].

Результаты измерений

В таблице 1 приведена сводка результатов нагрузочного тестирования всех восьми методов при нагрузке 150 виртуальных пользователей (инструмент k6, рост нагрузки с 10 до 150 пользователей за 5 минут). Сравнительное представление среднего времени отклика и значения P95 приведено на рисунке 3.

Таблица 1. – Производительность методов пространственного анализа

Метод	Среднее, мс	P95, мс	Ошибки	Что помогло
Буферный анализ	213	387	0%	Расчёт в метрах по эллипсоиду + GIST
Пересечение по СОАТЭ	87	142	0%	Фильтр в SQL-запросе, не в коде
Зональная статистика	156	289	0%	GIST + Redis; без них 892 мс и 18% ошибок
Население в зоне риска	340	510	0%	Только по явному запросу
KDE-тепловая карта	178	312	0%	Расчёт в браузере через WebGL
Рельеф SRTM	134	267	0%	Растровый GIST-индекс
Данные ДЗЗ через GEE	42	98	0%	Redis-кэш, GEE по расписанию
Гидропосты «Түндүк»	31	67	0%	REST + Redis 60 с

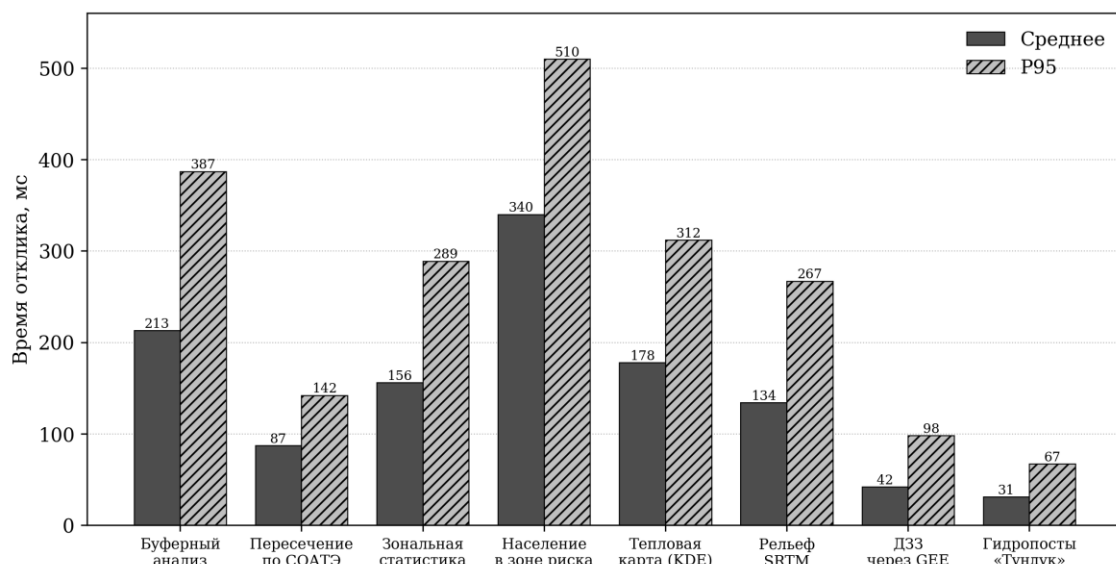


Рис. 3 – Среднее время отклика и значение P95 для восьми методов пространственного анализа при нагрузке 150 виртуальных пользователей

Выводы

Обобщая полученный опыт, можно выделить три основных результата, представляющих практический интерес для разработчиков аналогичных систем.

Применение пространственных индексов является необходимым, но недостаточным условием обеспечения производительности. Использование GIST-индекса для зональной статистики снизило долю отказов с 18% до 3%, что является улучшением, однако не соответствует требованиям. Нулевое значение показателя было достигнуто только после добавления кэширования. Распространённая в литературе по геоинформационным системам рекомендация применения GIST-индекса как универсального решения не подтверждается полученными результатами.

Распределение вычислений между серверной и клиентской сторонами должно определяться результатами измерений, а не априорными предположениями. Исходное предположение о целесообразности серверного построения тепловой карты ввиду более высоких аппаратных ресурсов сервера не подтвердилось: расчёт средствами WebGL на стороне клиента обеспечил более высокую производительность и дополнительное преимущество - пересчёт при изменении масштаба без обращения к серверу.

Разнородность систем координат является характерной особенностью пространственных данных постсоветских стран, а не исключением. Хранение данных в исходных системах координат с перепроецированием на этапе запроса оказалось предпочтительнее предварительной конвертации: к системе было подключено семь новых источников данных без внесения изменений в схему базы данных.

ЛИТЕРАТУРА

1. Атлас селе-паводковой опасности на территории Кыргызской Республики / Департамент мониторинга, прогнозирования чрезвычайных ситуаций при МЧС КР. - Бишкек: МЧС КР, 2025. - 164 с. - ISBN 978-9967-9581-2-8.
2. Hungr O., Leroueil S., Picarelli L. The Varnes classification of landslide types, an update // Landslides. - 2014. - Vol. 11, № 2. - P. 167–194.
3. Северский И.В., Благовещенский В.П. Оценка лавинной и селевой опасности горных территорий. - Алма-Ата: Наука, 1983. - 220 с.
4. Obe R.O., Hsu L.S. PostGIS in Action. 3rd ed. - Shelter Island: Manning Publications, 2021. - 600 p.

5. Kirschbaum D., Stanley T. Satellite-based assessment of rainfall-triggered landslide hazard for situational awareness // *Earth's Future*. - 2018. - Vol. 6, № 3. - P. 505–523.
6. МЧС КР. Отчёт о состоянии природных рисков на территории КР. - Бишкек: МЧС КР, 2022. - 87 с.
7. Google Earth Engine [Электронный ресурс]. - URL: <https://developers.google.com/earth-engine> (дата обращения: 10.04.2026).
8. Martha T.R., Kerle N., van Westen C.J. Segment optimization for knowledge-based landslide detection // *IEEE TGRS*. - 2011. - Vol. 49, № 12. - P. 4928–4943.