

УДК 551.435.627: 528.8: 004.8

КОМПЛЕКСНАЯ ОЦЕНКА ОПОЛЗНЕВОЙ ОПАСНОСТИ НА ТЕРРИТОРИИ КЫРГЫЗСТАНА С ИСПОЛЬЗОВАНИЕМ СПУТНИКОВЫХ ДАННЫХ И МАШИННОГО ОБУЧЕНИЯ

Самсакова А.К., Орозобекова А.К.

КГТУ им. И. Раззакова

В данной работе рассматривается применение геоинформационных систем (ГИС) и методов машинного обучения для картографирования оползневой опасности в южных регионах Кыргызстана. В качестве исходных данных использованы цифровая модель рельефа SRTM и статистические сведения о ранее зарегистрированных оползнях. На основе анализа морфометрических характеристик рельефа, включая уклон, экспозицию, кривизну и топографический индекс влажности (TWI), сформирован набор признаков для обучения модели. Для оценки оползневой опасности применён алгоритм Random Forest. В результате построена карта вероятности возникновения оползней, позволяющая выявить наиболее уязвимые и потенциально опасные участки исследуемой территории.

Ключевые слова: картографирование оползневой опасности, геоинформационные системы, машинное обучение, Random Forest, SRTM, цифровая модель рельефа, топографический индекс влажности (TWI)

COMPREHENSIVE LANDSLIDE HAZARD ASSESSMENT IN KYRGYZSTAN BASED ON SATELLITE DATA AND MACHINE LEARNING APPROACHES

Samsakova A.K., Orozbekova A.K.

KSTU named after. I. Razzakova

This study examines the application of geographic information systems (GIS) and machine learning methods for landslide susceptibility mapping in the southern regions of Kyrgyzstan. The analysis is based on the SRTM digital

elevation model and statistical data on previously recorded landslides. Morphometric terrain characteristics, including slope, aspect, curvature, and the Topographic Wetness Index (TWI), were derived and used as input features for model training. A Random Forest algorithm was employed to assess landslide susceptibility. As a result, a landslide probability map was generated, enabling the identification of the most vulnerable and potentially hazardous areas within the study region. The results confirm the effectiveness of integrating geospatial analysis and machine learning techniques for regional-scale natural hazard assessment.

Keywords: landslide susceptibility mapping, geographic information systems (GIS), machine learning, Random Forest, SRTM, digital elevation model (DEM), Topographic Wetness Index (TWI)

КЫРГЫЗСТАНДАГЫ ЖЕР КӨЧКҮСҮ КОРКУНУЧУН КОМПЛЕКСТҮҮ БААЛОО СПУТНИК МААЛЫМАТТАРЫН ЖАНА МАШИНА ОКУТУУНУ КОЛДОНУУ МЕНЕН

Самсакова А.К., Орозобекова А.К.

И. Раззаков атындагы КМТУ

Бул макалада Кыргызстандын түштүк аймактарында жер көчкү коркунучун картага түшүрүү үчүн географиялык маалымат системаларын (ГИС) жана машиналык окутуу ыкмаларын колдонуу каралат. Колдонулган баштапкы маалыматтар санариптик бийиктик модели (SRTM) жана мурда катталган жер көчкүлөр боюнча статистикалык маалыматтар болгон. Жердин жантайыңкылыгы, аспекти, ийрилиги жана топографиялык нымдуулук индекси (TWI) сыяктуу морфометриялык мүнөздөмөлөрүн талдоонун негизинде моделдик окутуу үчүн бир катар өзгөчөлүктөр иштелип чыккан. Жер көчкү коркунучун баалоо үчүн Random Forest алгоритми колдонулган. Натыйжада, изилдөө аймагынын эң аялуу жана потенциалдуу кооптуу аймактарын аныктоого мүмкүндүк берген жер көчкү ыктымалдуулугу картасы түзүлгөн.

Баштапкы сөздөр: жер көчкү коркунучун картага түшүрүү, географиялык маалымат системалары, машиналык окутуу, Random Forest, SRTM, санариптик бийиктик модели, топографиялык нымдуулук индекси

Оползни относятся к числу наиболее опасных природных процессов, регулярно происходящих на территории Кыргызстана. Из-за горного рельефа, высокой сейсмичности и неустойчивых пород страна особенно

подвержена этим явлениям. По данным Министерства чрезвычайных ситуаций, на сегодняшний день на территории республики выявлено более 4554 оползней, из которых около 98% сосредоточено в южных регионах (Ошская, Джалал-Абадская, Баткенская области). С 1991 по 2022 год зафиксировано 593 оползневых процесса, унесших жизни 275 человек, разрушивших тысячи домов и угрожающих инфраструктуре.

В настоящее время оценки риска опасных природных процессов и явлений выполняются несколькими способами. К первой группе методов оценки риска можно отнести, те которые практически подменяют понятие риска понятием опасность. Риск есть вероятность нежелательных последствий, а опасность – потенциальная угроза. Опасность, как правило, качественная характеристика, полученная различными способами. Среди них можно отметить экспертную оценку, широко распространенную балльную оценку. К другой группе методов оценки риска относятся те из них, которые оценивают вероятность последствий. Они, как правило, основаны на теоретических и статистических исследованиях [1].

Традиционные методы мониторинга, основанные на полевых наблюдениях, отличаются высокой точностью, но охватывают лишь ограниченные участки, поэтому в последние годы всё большее внимание уделяется применению спутниковых данных и методов машинного обучения, которые позволяют оперативно выявлять потенциально опасные зоны и анализировать динамику изменений на больших территориях.

Целью данной работы является разработка и апробация подхода к комплексной оценке оползневой опасности на территории Кыргызстана с использованием данных дистанционного зондирования Земли и машинного обучения. Такой подход поможет повысить точность прогнозирования, улучшить управление рисками и укрепить систему предупреждения чрезвычайных ситуаций.

Исходные данные

Разработка структурной схемы системы анализа данных и визуализации оползнеопасных участков Кыргызской Республики, позволяющая визуализировать и анализировать данные об оползнеопасных регионах, требует объединения разных методов решения поставленной задачи, а также комплексного подхода, так как предполагает обработку пространственных данных, расчётов факторов рельефа, построения модели машинного обучения и представления результатов пользователям в понятной форме. Поэтому первым этапом практической

реализации была разработка структурной схемы системы, позволяющая показать её основные модули, последовательность обработки данных и взаимосвязь между полученными модулями.

Структурная схема показывает общий процесс исследования в виде алгоритма, начиная от подготовки исходных данных, заканчивая отображением полученных карт вероятностей и рисков оползневой опасности в web-интерфейсе. Благодаря этому, можно увидеть целостность разработки, при этом определить назначение каждого модуля полученной системы.

Для проведения исследования использовались пространственные и статистические данные. Комплексный подход к формированию исходной базы позволил учесть основные факторы, влияющие на формирование оползневых процессов. В качестве основного источника топографической информации была использована цифровая модель рельефа SRTM с пространственным разрешением 30 метров, обеспечивающая достаточную детализацию для регионального анализа. На её основе были рассчитаны ключевые морфометрические параметры рельефа, включая уклон, экспозицию и кривизну поверхности.

Дополнительно был вычислен топографический индекс влажности (Topographic Wetness Index, TWI), характеризующий пространственное распределение потенциального накопления влаги и играющий важную роль в формировании условий оползнеобразования.

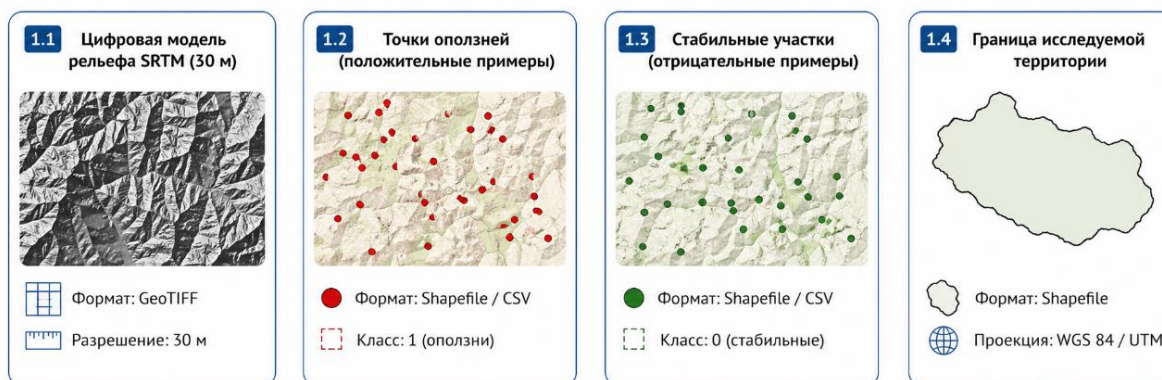


Рис.1. Исходные пространственные данные, используемые в системе.

В качестве обучающих и валидационных данных использовались сведения о ранее зарегистрированных оползнях, полученные из доступных отчётов МЧС КР. Эти данные позволили сформировать выборку для обучения модели и последующего анализа.

Таким образом, использование совокупности указанных источников обеспечило репрезентативность исходных данных и позволило учесть

ключевые природные факторы, влияющие на устойчивость склонов и развитие оползневых процессов.

Методика исследования.

Методика исследования основана на интеграции геоинформационного анализа и методов машинного обучения для оценки оползневой опасности. Общая схема включала расчет факторов рельефа, формирование обучающей выборки и построение прогностической модели.

Анализ выполнялся в среде **Python** с использованием библиотек **rasterio**, **numpy** и **scipy**, обеспечивающих обработку растровых данных и вычисление производных параметров. На основе цифровой модели рельефа (DEM) были рассчитаны ключевые критерии, оказывающие влияние на развитие оползневых процессов:

Уклон поверхности (slope) рассчитывался как величина градиента высоты (1)

$$S = \arctan \left(\sqrt{\left(\frac{dz}{dx}\right)^2 + \left(\frac{dz}{dy}\right)^2} \right) \quad (1)$$

где z – значения высоты, $\left(\frac{dz}{dx}\right)$ и $\left(\frac{dz}{dy}\right)$ – частные производные по координатам.

Экспозиция склонов (aspect) определялась как направление максимального уклона (2)

$$A = \arctan \left(\frac{\frac{dz}{dy}}{-\frac{dz}{dx}} \right) \quad (2)$$

Экспозиция показывает ориентацию склонов относительно сторон света.

Кривизна поверхности (curvature) характеризует форму рельефа и рассчитывается как вторая производная высоты (3)

$$C = \frac{d^2z}{dx^2} + \frac{d^2z}{dy^2} \quad (3)$$

Данный показатель отражает выпуклость или вогнутость поверхности.

Топографический индекс влажности (TWI) вычислялся по формуле (4), данный критерий отражает потенциальное накопление влаги и является важным фактором в процессах оползнеобразования.

$$TWI = \ln \left(\frac{A_s}{\tan S} \right) \quad (4)$$

где A_s – удельная площадь водосбора, S – уклон поверхности.

Полученные параметры были приведены к единому пространственному разрешению и использованы в качестве входных признаков для последующего моделирования.

Для построения модели была сформирована выборка, включающая точки с известными случаями оползней, а также точки, соответствующие стабильным участкам территории. В качестве положительных примеров использовались координаты зарегистрированных оползней, тогда как отрицательные примеры формировались путем генерации случайных точек вне зон проявления оползневых процессов. Для каждой точки были извлечены значения рассчитанных факторов рельефа, что позволило сформировать матрицу признаков. Такой подход обеспечил представление как опасных, так и устойчивых участков территории.

В качестве алгоритма машинного обучения использовалась модель Random Forest (RandomForestClassifier), представляющая собой группу деревьев решений. Построение модели выполнялось на сформированной выборке с разделением на три класса оползневой опасности: 0 – низкий риск; 1 – средний риск; 2 – высокий риск. Параметры модели подбирались таким образом, чтобы обеспечить баланс между обобщающей способностью и устойчивостью к переобучению. В процессе моделирования учитывались нелинейные зависимости между факторами рельефа и вероятностью возникновения оползней.

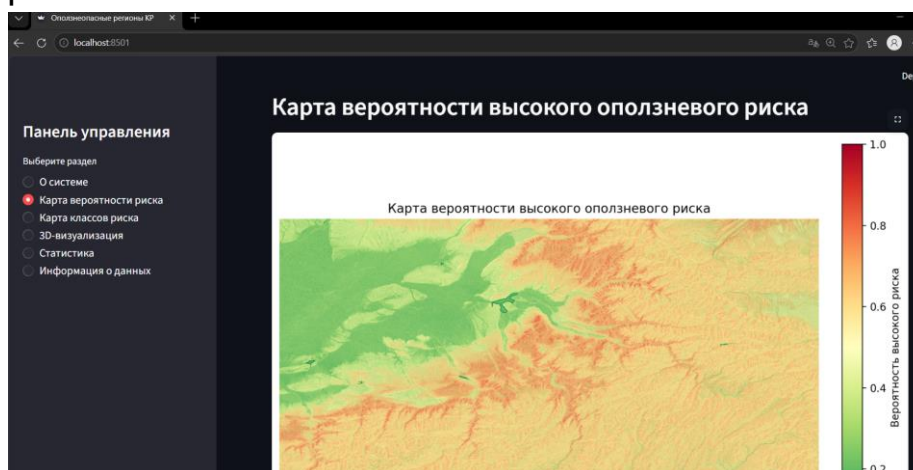


Рис.2. Карта вероятности оползневого риска

Итоговая модель была применена к растровым данным, что позволило получить пространственное распределение оползневой опасности в виде карт классификации и вероятностей.

Результаты.

В результате применения разработанного подхода была построена модель оценки оползневой опасности на основе алгоритма Random Forest.



Рис 3. График распределения значений вероятности риска.

По итогам тестирования модель продемонстрировала среднюю точность классификации на уровне около 0.85, что свидетельствует о её высокой предсказательной способности при решении задачи пространственного анализа. С использованием обученной модели можно выполнить прогнозирование и на других территориях Кыргызстана, при наличии цифровой модели рельефа (DEM). В результате были получены два основных выходных продукта: растр классификации (hazard_pred.tif), отражающий распределение территории по классам оползневой опасности; растр вероятности высокого риска (hazard_proba_high.tif), содержащий значения вероятности возникновения оползней. Анализ пространственного распределения полученных результатов показал, что зоны с высокой вероятностью возникновения оползней преимущественно приурочены к участкам с крутыми склонами, значительными значениями кривизны и повышенным уровнем увлажнения, что соответствует известным закономерностям развития оползневых процессов. Визуализация результатов была выполнена в среде QGIS с использованием псевдоцветового отображения (Singleband Pseudocolor) (рис.1), где значения вероятности интерпретируются в виде градиента от зелёного (низкий риск) к красному (высокий риск). Полученные карты позволяют выявить наиболее уязвимые участки территории и могут служить основой для последующего анализа, мониторинга и принятия решений.

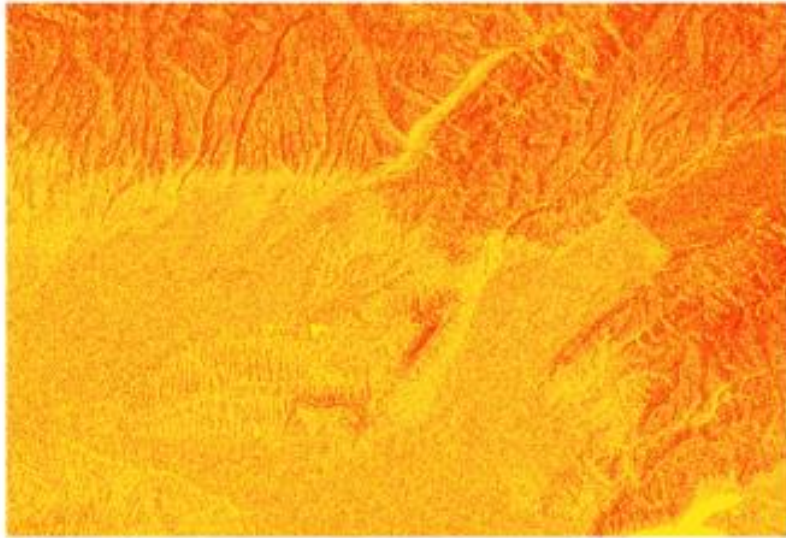


Рис.1. Визуализация полученного растра вероятности высоко риска (hazard_proba_high.tif) на примере территории Узгенского района Ошской области.

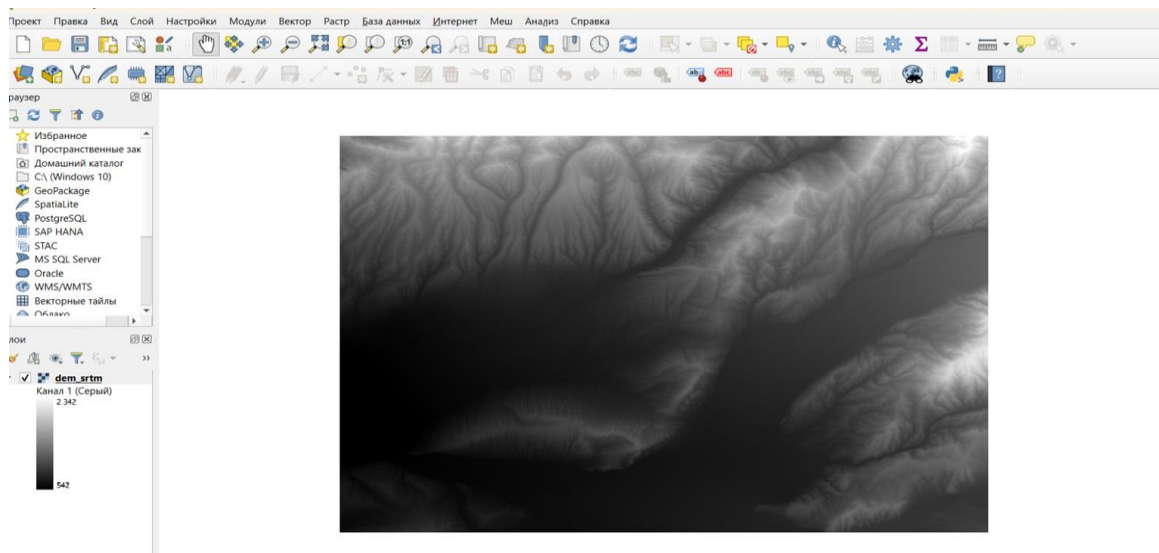
И так, создан алгоритм, позволяющий выполнять оценку оползневой опасности на основе цифровых моделей рельефа с использованием методов машинного обучения.

Для этого были решены следующие задачи:

- сбор и подготовка пространственных данных (DEM, координаты населённых пунктов, оползневые события);
- расчет морфометрических показателей рельефа: уклон, экспозиция, кривизна, индекс влажности (TWI);
- формирование обучающей выборки для многоклассовой классификации уровней риска (0 — нет риска, 1 — риск, 2 — сильный риск);
- обучение модели машинного обучения и визуализация результатов в QGIS.

Исходными данными были:

1. Цифровая модель рельефа (DEM SRTM 30 м) — загружена через OpenTopography API;
2. Координаты населённых пунктов Узгенского района Ошской области в формате CSV;
3. Модель рельефа: dem_srtm.tif;
4. Производные факторы: elev.tif, slope.tif, aspect.tif, curvature.tif, twi.tif;
5. Метка риска (label): 0 — нет, 1 — риск, 2 — сильный риск.



Реализована полная цепочка оценки оползневой опасности: от загрузки данных и расчета факторов до построения прогнозной карты.

Модель позволяет проводить оценку риска на новых территориях при наличии DEM и расширяться за счёт добавления новых факторов (растительность, осадки, литология).

Выводы: В ходе проведённого исследования был разработан подход к оценке оползневой опасности на основе геоинформационных технологий и методов машинного обучения. Использование цифровой модели рельефа SRTM и производных морфометрических параметров позволило сформировать информативный набор признаков, отражающих основные факторы, влияющие на устойчивость склонов. Применение алгоритма Random Forest обеспечило построение модели, способной учитывать сложные нелинейные зависимости между характеристиками рельефа и вероятностью возникновения оползней.

Полученные результаты показали высокую эффективность предложенного подхода при решении задачи пространственного анализа оползневой опасности. По итогам исследований была построена карта вероятности возникновения оползней, позволяющая выявить наиболее уязвимые участки исследуемой территории, которая может быть использована для мониторинга природных рисков, планирования хозяйственной деятельности и принятия решений. Но несмотря на достигнутые результаты, следует учитывать ограничения исследования, связанные с использованием данных с ограниченным пространственным разрешением и неполнотой информации о зарегистрированных оползнях.

В дальнейшем перспективным направлением является использование более детализированных данных и расширение набора факторов, влияющих на развитие оползневых процессов.

ЛИТЕРАТУРА

1. Breiman L. Random Forests Breiman L. Random Forests // *Machine Learning*. - 2001. - Vol. 45, № 1. - С. 5–32.
2. Moore I.D. Digital terrain modelling review Moore I.D., Grayson R.B., Ladson A.R. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications // *Hydrological Processes*. - 1991. - Vol. 5. - С. 3–30.
3. Reichenbach P. Landslide susceptibility models review Reichenbach P., Rossi M., Malamud B.D., Mihir M., Guzzetti F. A review of statistically-based landslide susceptibility models // *Earth-Science Reviews*. - 2018. - Vol. 180. - С. 60–91.
4. Tarboton D.G. Flow direction method Tarboton D.G. A new method for the determination of flow directions and upslope areas in grid digital elevation models // *Water Resources Research*. - 1997. - Vol. 33, № 2. - С. 309–319.