

УДК 551.4.012:550.81(597)

СПЕКТРАЛЬНЫЙ АНАЛИЗ РЕЛЬЕФА С ПОСТРОЕНИЕМ НЕЙРОННОЙ СЕТИ ДЛЯ РЕШЕНИЯ ПОИСКОВЫХ ЗАДАЧ НА ПРИМЕРЕ ГОРНОГО МАССИВА ЛУК-ТЬЕН (СЕВЕРНЫЙ ВЬЕТНАМ)[#]

© 2023 г. И. С. Сергеев^{1,*}, К. А. Кукса¹, А. Б. Глебова¹

¹Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

*E-mail: igorsergeev.spb@gmail.com

Поступила в редакцию 25.06.2021 г.

После доработки 07.03.2022 г.

Принята к публикации 22.12.2022 г.

Территория исследования расположена на севере Вьетнама в провинции Йенбай и представляет собой крупный ($14.5 \times 6.5 \times 0.8$ км) структурно-денудационный останец на периферии сильного расчлененного низкогогорья Кон Вой, а также склоны и днища прилегающих речных долин. Для территории известны проявления камнесамоцветной минерализации в виде жильных образований в толщах мраморов. Район относительно труднодоступен для полевых изысканий, поэтому для предварительной оптимизации проведения геолого-поисковых работ стояла задача на основе анализа имеющейся геолого-геоморфологической информации получить данные о возможной локализации участков полезной минерализации. Для этого методом дискретного преобразования Фурье был рассчитан амплитудный спектр расчленения рельефа для участков, связанных с жильными геологическими образованиями в приповерхностной части мраморных толщ. Бинарная классификация (на потенциальные участки с полезной минерализацией и без нее) полученных числовых показателей амплитуд высот, отвечающих гармоническим колебаниям разных пространственных частот, осуществлена с помощью простой нейронной сети – двухслойного персептрона. Расчетный алгоритм был реализован на языке Python. Применение данной методики позволило выполнить прогноз на рубиново-шпинельную минерализацию в коренном залегании на изучаемую площадь более 200 км^2 . Полевыми исследованиями в 2019 г. выполнена заверка прогнозных данных, заключающаяся в минералогическом и геохимическом опробовании доступной части спрогнозированных точек. Получена оценка прогнозной силы использованной методики: каждый третий (~35%) спрогнозированный нейронной сетью участок фактически содержит коренные источники рубинов и шпинелей на рассмотренной территории.

Ключевые слова: двумерное спектральное преобразование рельефа, поисковая геоморфология, морфометрические методы, машинное обучение, ГИС, ЦМР

DOI: 10.31857/S2949178923030106, **EDN:** WDVKDT

ВВЕДЕНИЕ

Комплекс геологических методов поиска и разведки месторождений полезных ископаемых использует, кроме прочего, геоморфологические знания о конформности (дисконформности) рельефа геологическому субстрату (Лопатин и др., 2012). Это позволяет решать обратную задачу – по характеру рельефа и его конкретным индикаторам (в первую очередь, морфологическим или даже исключительно морфометрическим) судить о

вероятной локализации месторождений. Общая схема такого подхода – т.н. “предиктивного моделирования” – обычно включает в себя этапы:

1) сбор информации о морфологии рельефа в виде набора индикаторов – различных морфометрических переменных, в том числе производных от абсолютной высоты и ее пространственного распределения. Все многообразие форм земной поверхности выражается в огромном числе комбинаций значений разных морфометрических индикаторов, прогнозная сила которых на этом этапе еще не определена;

2) классификация участков земной поверхности на основе экспертных данных о проявлениях полезной минерализации (фактически, инвентаризации известных месторождений) и всего набора ее потенциальных морфометрических предик-

[#] Ссылка для цитирования: Сергеев И.С., Кукса К.А., Глебова А.Б. (2023). Спектральный анализ рельефа с построением нейронной сети для решения поисковых задач на примере горного массива Лук-Тьен (Северный Вьетнам) // Геоморфология и палеогеография. Т. 54. № 3. С. 138–149. <https://doi.org/10.31857/S2949178923030106>; <https://elibrary.ru/WDVKDT>

торов. Здесь, во-первых, оценивается, какие переменные и их конкретные числовые диапазоны являются наиболее характерными для искомых участков земной поверхности, а во-вторых, производится экстраполяция выявленных закономерностей на всю интересующую площадь. Таким образом, превентивно выявляются новые потенциальные участки для полевой разведки, за счет заведомо бесперспективных (по геоморфологическим критериям) участков сокращаются временные и финансовые затраты.

Так как источником информации о морфологии земной поверхности чаще всего служат цифровые модели рельефа (ЦМР) среднего и высокого разрешения, то данный подход к прогнозу месторождений реализуется, как правило, инструментарием компьютерных информационных систем.

Современные информационные системы содержат самые различные алгоритмы машинного обучения: от простых статистических байесовских классификаторов до сложных ресурсоемких алгоритмов нейронных сетей (Плас, 2018; Мюллер, Гвидо, 2018). Все это позволяет перенести в сферу искусственного интеллекта первичный анализ данных о рельефе, а результаты подобного анализа контролирует эксперт как на стадии обучения нейронной сети, так и на стадии верификации результатов. Исходные данные для осуществления подобных исследований — обычно, ЦМР Aster GDEM, SRTM и др. с шагом регулярной сетки высот от 30 до 90 м. Такое пространственное разрешение исходных данных позволяет создавать прогнозные модели с точностью локализации объектов до первых сотен метров, а в некоторых случаях и до десятков метров.

Вторым необходимым источником данных для решения прогнозных рекогносцировочных задач является информация об уже известных объектах как вторичной, так и первичной минерализации. Первичные источники представлены геологическими телами, выходящими на дневную поверхность и “тяготеющими” к определенным формам рельефа с резкими границами и высокими градиентами морфометрических параметров. В то же время вторичные источники, образованные экзодинамическими процессами, приурочены к более спокойному рельефу речных долин и межгорных котловин, в том числе к основанию пологих склонов, к речным и морским террасам (Лоскутов, 1999).

Выражение экспонированных поверхностей геологических тел с полезной минерализацией в виде конкретных форм и элементов рельефа (например, структурно-денудационных уступов поверхностей отпрепарированных даек), а также приуроченность очагов вторичной минерализации к другому набору форм (различного рода топографические ловушки в долинной сети) — это

основание для разработки методики использования морфометрических критериев для прогноза размещения месторождений. Входными признаками для подобных алгоритмов могут быть, например, амплитуды двумерного спектра рельефа. Спектральный (гармонический) анализа рельефа применяется в геоморфологии с 60-х гг. прошлого века (Хромченко, 1967; Pike, Rozema, 1975), несмотря на определенные трудности геоморфологической интерпретации результатов спектрального разложения (Харченко, 2017; Харченко, Казаков, 2018). Достоинства гармонического разложения рельефа связаны с тем, что двухмерный амплитудно-частотный спектр рельефа позволяет автоматизировать выявление участков с однородным рисунком топографического расчленения, зачастую отвечающим схожим структурам земной коры, — этот метод позволяет с высокой информативностью характеризовать двухмерный “сигнал” (в нашем случае — ЦМР) не столько в точке, сколько в выбранной ее окрестности (Дэвис, 1990; Сергиенко, 2002).

В настоящей работе представлена методика поиска участков территории с заданными морфометрическими характеристиками рельефа для задач предсказания выходов жильных образований камнесамоцветной минерализации (рубины и шпинель), основанная на двумерном гармоническом разложении ЦМР. Показаны результаты апробации методики для прогноза положения потенциальных проявлений полезной минерализации в коренном залегании в пределах одного из районов п-ова Индокитай.

ОБЪЕКТЫ ИССЛЕДОВАНИЯ

Исследуемый район расположен в пределах северного Вьетнама (координаты центра исследуемого полигона — $22^{\circ}2'18.6''$ в.д. $104^{\circ}48'7.2''$ с.ш.). Территория представляет собой низкогорье (Кон Voi или Слоновьи горы) с максимальными отметками, едва превышающими 1000 м, а фоновыми значениями высот вершинных поверхностей — 400–700 м, она разбита серией крупных разломов, ориентированных в направлении СЗ–ЮВ, эти разломы, будучи освоены речной сетью, определяют общий характер рисунка рельефа участка (Geological and mineral..., 2000). Крупнейший из разломов в регионе — планетарный разлом Красной реки (или р. Хонгха) — протягивается на 1500 км от юго-восточной окраины Тибета до Тонкинского залива и задает преобладающее направление горных кряжей и разделяющих их долин. Рассматриваемый участок расположен от этого разлома в ~50 км к северо-востоку, и очерчивается с двух сторон разломными зонами более низкого иерархического уровня. Объект нашего интереса — отдельный кряж (рис. 1), размерами 14.5×6.5 км и относительным превышением высшей точки

МЕТОДЫ

поверхности над днищами окружающих долин порядка 800 м, в публикациях носит название Large Rock (Pham et al., 2017) или Большая Скала, расположен непосредственно к северу от оз. Тха-Ба. Максимальная абсолютная высота в пределах кряжа ~860 м (рис. 2, (а)). Окружающие его долины находятся в интервале высот 50–150 м, при ширине по бровкам до 2 км.

Территория сложена мраморами графитизированными (преобладают), гнейсами, гранитами возрастом поздний протерозой – ранний кембрий (рис. 2, (б)), а разломная сеть (к элементам которой и приурочены месторождения самоцветов) обусловлена новейшими ранненеогеновыми тектоническими движениями, спровоцированными приращением индостанского блока к южно-китайскому и современным развитием юго-восточного сегмента Альпийского-Гималайского орогенного пояса.

Поверхность рассматриваемого кряжа разбита системами малых долин с характерным СЗ-ЮВ и С-Ю простиранием и амплитудами высот в несколько десятков метров. Крутизна бортов кряжа значительна и составляет 10°–29° – для юго-западных, и 8°–19° – для северо-восточных склонов. На высотах 400–500 м наблюдаются пологие, относительно неглубокие расчлененные остаточные поверхности выравнивания. Для территории характерны карстовые процессы и явления. Территория получает значительное увлажнение – ~2000 мм/год, что в совокупности с выходом к поверхности карстующихся пород обуславливает проявления как покрытого, так и поверхностного карста. Склоновые процессы замедлены сомкнутым растительным покровом и наблюдаются только на редких открытых участках с крутизной более 35°, либо оказываются активированы антропогенной деятельностью. Наносы со склонов малых долин поступают в русловую сеть на поверхности кряжа и затем выносятся в более крупные долины СЗ-ЮВ простирания, дренируемые реками Бьетх, Тюнг, Тяй и др.

Характерные особенности рельефа, перечисленные ранее, по всей видимости определяются наличием мраморов на исследуемой территории с многочисленными интрузиями основного и кислого состава и тектоническими разломами (Khoi et al., 2016). Участки территории, подстилаемые мраморами, в значительной степени подвержены процессам растворения, поэтому преобладают разномасштабные формы карстового рельефа. При этом карстовые депрессии зачастую заполнены нерастворимым рыхлым высвобожденным материалом. Но локальные мощности рыхлых отложений здесь невелики – от 0.5 до 1.5 м (Long et al., 2013).

При подготовке полевых геологических поисковых работ была составлена серия вспомогательных карт-схем района в ГИС по данным дистанционного зондирования. Морфометрический анализ рельефа осуществлялся по данным глобальной модели рельефа Aster разрешением 30 м. Построение линий тока с использованием модуля гидрологии ПО ArcGIS позволило проследить основные направления выноса рыхлого вещества и, как результат, ареалы развития аллювиальных толщ с проявлениями полезной минерализации. Также выделены участки интенсивного карста по аномалиям профильной кривизны рельефа. Построенные карты были использованы для анализа морфологических особенностей территории. Была предпринята попытка применения спектрального анализа рельефа для выявления перспективных участков геологического маршрутного опробования. В работе, кроме прочего, была использована информация о существующих местонахождениях рубинов и шпинели по (Long et al, 2013). Следует отметить, что известные разработки россыпей рубинов и шпинели приурочены к долинам рек в среднем и нижнем течении и к зонам сгущения проявлений карстовых процессов. Учет этих данных и имеющийся опыт в применении спектрального анализа рельефа позволил создать методику для прогноза камнесамоцветной минерализации на рассматриваемой территории как часть геолого-геоморфологических поисковых мероприятий.

Описываемая методика включила в себя расчет гармонических составляющих рельефа земной поверхности и создания нейронной сети, где в качестве входных признаков были использованы амплитуды двумерного спектра рельефа. В качестве входных данных нейронов было использовано 40 спектральных признаков (амплитуд отдельных гармоник в спектре поля высот), полученных скользящей палеткой 240 × 240 м (т.е. 8 × 8 ячеек ЦМР) по общей площади участка 272 км². Исходные данные о рельефе представлены регулярной сеткой с 302088 высотных значений с шагом 30 м.

Основу представляемой методики составили два хорошо проработанных научных метода обработки данных: 1-й – это дискретное разложение Фурье и 2-й – машинное обучение с учителем нейронной сети. Краткое описание этих методов, объединенных нами в общей методике поиска участков поверхности, наиболее схожих по спектральным характеристикам рельефа с обучающей выборкой (включающей фрагменты ЦМР для участков с достоверно установленными ранее месторождениями), приведено ниже.

Метод спектрального анализа основывается на следующих положениях. Любой сигнал можно разложить на ряд гармонических функций (ряд



Рис. 1. Вверху – положение района работ в пределах Юго-Восточной Азии показано красной точкой, внизу – гипсометрическая схема района работ.

Fig. 1. Up – interest region location on the map of South-East Asia is shown with the red point, down – hypsometric scheme of the interest region.

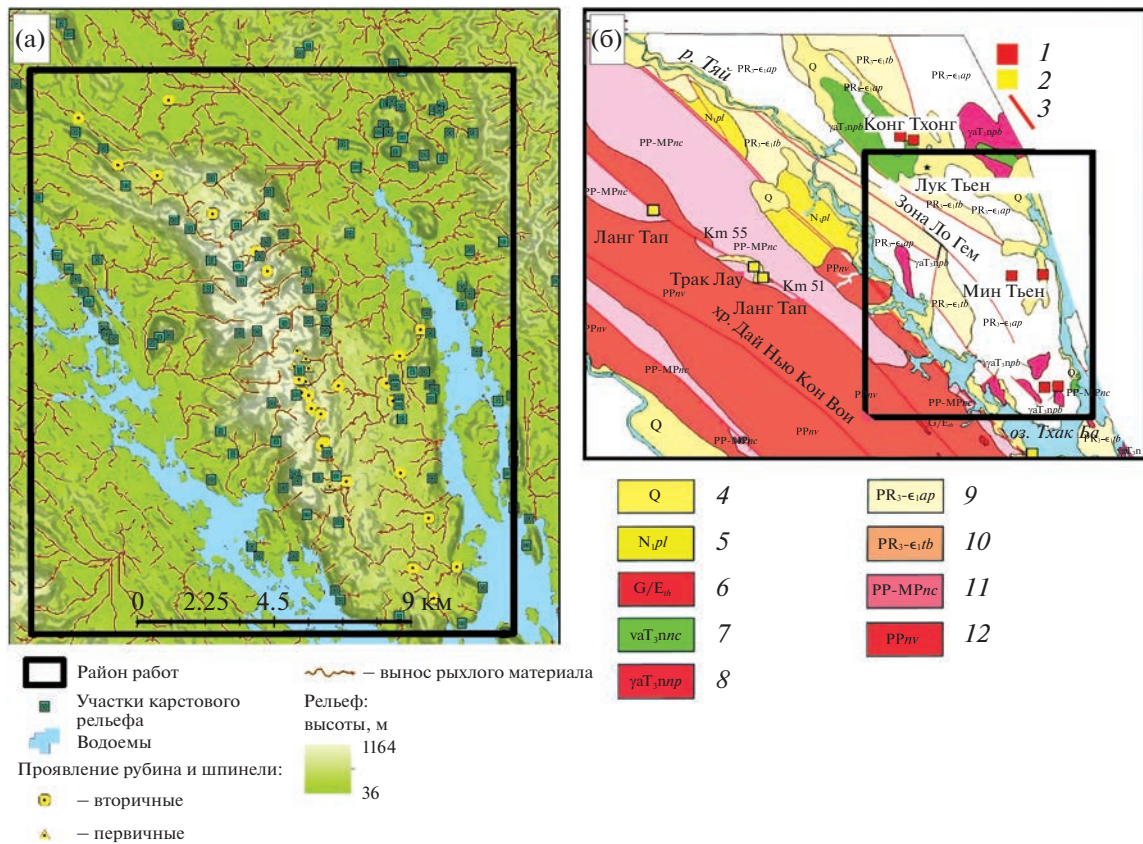


Рис. 2. Рельеф, геологическое строение района исследования и инвентаризация месторождений рубина и шпинели: (а) — рельеф района и позиции проявлений самоцветов; (б) — схематическая геологическая карта. *Проявления камне-самоцветной минерализации:* 1 — в мраморах, 2 — в гнейсах; 3 — разломы. *Геологическое строение:* 4 — четвертичные отложения, 5 — неогеновая формация Фанг-Луонг, 6 — комплекс Тан Хуонг (граниты, граносиениты, пегматиты), 7 — комплекс Нуи Чуа (габбро, габбродиориты, диориты), 8 — комплекс Фиа Биок (биотитовые и пегматитовые граниты), 9 — комплекс Ан Фу (графитизированные мраморы, слюды), 10 — комплекс Тхак Ба (кварцевые слюды, эпидот-биотитовые сланцы, мраморы), 11 — комплекс Нгои Тчи (кварц-биотит-силлиманитовые сланцы, амфиболиты, кварциты, мраморы), 12 — комплекс Нью Вои (плагиногнейсы, гнейсы, амфиболиты, мраморы, сланцы).

Fig. 2. Topography, geological settings and inventory of ruby and spinel deposits: (a) — interest region terrain and positions of gems signs, (b) — schematic geological map. *The signs of the gems:* 1 — in the marbles, 2 — in the gneisses; 3 — the faults. *Geological division:* 4 — quaternary deposits, 5 — Neogene Fang-Luong formation, 6 — Tan Huong formation (granite, granosyenite, pegmatite), 7 — Nui Chua formation (gabbro, gabbro-diorite, diorite), 8 — Phia Biok formation (biotite and pegmatite granite), 9 — An Phue formation (graphitized marble, mica), 10 — Tkhak Ba (quartz mica, epidote-biotite shale, marble), 11 — Ngoi Tchi formation (quartz-biotite-sillimanite shale, amphibolite, quartzite, marble), 12 — Niu Voi formation (plagiogneiss, gneiss, amphibolite, marble, shale).

Фурье) таким образом, что при суммировании функций этого ряда получается исходный сигнал (Davis, Choinacki, 2017). На основе этого имеется возможность анализировать разночастотные составляющие сигнала, которые представлены набором гармонических функций — спектром. В практике обработки ЦМР используется вариант дискретного преобразования Фурье (Сергиенко, 2002).

Спектр несет информацию о распределении “энергии” (в нашем случае — амплитуды высот) в сигнале между его разночастотными периодическими составляющими — гармониками, применительно к рельефу возможно оценить преобла-

дающие линейные размеры периодов расчленения изучаемого участка земной поверхности.

Двумерное дискретное преобразование Фурье, выполняемое одновременно по осям X и Y , позволяет выявить параметры всех возможных периодических составляющие не профиля, а поверхности в пределах заданной точности (дискретности) ЦМР (Дэвис, 1990). Основным результатом преобразования в спектр исходного участка поверхности является набор гармонических составляющих, каждая из которых имеет различную амплитуду и фазовое смещение относительно нулевого отсчета. Одним из ограничений разложения в спектр является то, что в спектре отсутствует информация о координатах конкретной точки исход-

ного участка, так как спектр — это набор волновых изменений, затрагивающих все рассматриваемое пространство (т.н. глобальное спектральное преобразование). Для преодоления этого неудобства на больших площадях и при избыточной информации о рельефе следует применять скользящую палетку спектрального анализа небольшого размера.

Обычно в практике работы с ЦМР (например, при вычислении крутизны или видов кривизны) используется скользящая палетка с минимальным размером — 3×3 ячейки, т.е. рассматриваются только ближайшие соседи центральной ячейки. Для спектральной палетки был выбран квадрат со стороной $n = 8$ ячеек, т.е. преобразованию Фурье подвергались последовательно фрагменты исходной ЦМР по 8×8 ячеек или 240×240 м, что примерно отвечает среднему размеру участка со следами самоцветной минерализации для известных месторождений. Оптимизация результата вычисления спектральной палетки в программной среде Python была достигнута использованием функции “fft2”, которая вычисляет двумерное дискретное преобразование Фурье исходных данных, представленных действительными числами (в нашем случае — абс. высотами). В результате вычисления n^2 значений высот исходной палетки получаем $n(n/2 + 1)$ — всего 40 амплитуд спектра или спектральных признаков, которые записываются как вложенный массив в одну ячейку. При следующем шаге скольжения на одну ячейку входные данные функции обновляются, выполняется расчет спектральных признаков — результат записывается в соответствующую ячейку результирующего массива спектральных признаков. Таким образом, выходные данные после прохода вычислительной палеткой всех исходных значений высот размером Y и X представляют собой массив данных в ГИС с размерностью: $((Y-8) \cdot (X-8)) \times (n(n/2 + 1))$ или $(293304) \times (40)$. Такое представление удобно для дальнейшей обработки результата средствами библиотек программного кода Scikit-learn для Python при создании алгоритма машинного обучения (Мюллер, Гвидо, 2018).

Термин “машинное обучение” широко применяется в информационных сетях и системах в прикладных аспектах решения научных задач. Оно нацелено на извлечение полезной информации из данных, как правило, глубоких закономерностей, подтверждаемых эмпирически, но не наблюдаемых в массиве информации “невооруженным взглядом”. В своей основе решаемая нами в данной статье задача — классификация, она успешно решается на основе контролируемого обучения модели или “обучения с учителем”. Созданная в работе математическая модель была построена на основе обучающих данных амплитуд

двумерного спектра участков рельефа, имеющих жильные проявления полезной минерализации, и участков, не имеющих таковой. Это означает следующее: на основе известных данных (Long et al., 2013) были заданы два класса объектов распознавания: первый класс — это участки поверхности, содержащие коренные месторождения шпинели и рубинов — 7 объектов; второй класс — участки поверхности нулевого класса, не содержащие проявлений коренных источников в количестве 25 объектов. При этом каждый объект был представлен четырьмя десятками спектральных признаков формы его поверхности. Все это было оформлено в виде матрицы размерностью $(32) \times (40)$. Данный массив обрабатывался программным кодом библиотеки Scikit-learn (Мюллер, Гвидо, 2018). Программный модуль разбивает входной массив на два: на одном — происходит обучение нейронной сети, на второй части массива делается прогноз и полученный результат классификации сравнивается с известной принадлежностью участка к одному из классов. Сам алгоритм обучения строится на основе известного представления искусственного нейрона — персептрона, как множественной линейной регрессии.

Сущность его состоит в итеративной подгонке коэффициентов уравнения линейной регрессии при независимых переменных (амплитудах гармонических колебаний высот, отвечающих разным пространственным частотам) таким образом, чтобы зависимая переменная была линейным отражением меры (например, вероятности или др.) отнесения классифицируемых объектов к одному из двух классов — скажем, перспективному на самоцветы участку и не перспективному. Иными словами, алгоритм пытается определить, какие морфометрические (а именно спектральные) характеристики рельефа наилучшим образом соответствуют в пространстве участкам обоих классов, какие характеристики более значимы (обладают большим предсказательным потенциалом и большим коэффициентом в уравнении), а какие — менее. Использование двухслойной нейронной сети позволяет проследить не только прямую корреляцию в пространстве каждой отдельной спектральной характеристики рельефа (амплитуды гармоник) и полезной минерализации, но и линейных комбинаций амплитуд, соответствующих разным частотам на местности. В геоморфологическом смысле это значит, что потенциальными признаками для прогноза были, например, соотношения линейных размеров долин разного порядка и других разномасштабных, наложенных друг на друга форм рельефа.

Аналитически нейронная сеть представлена следующим набором выражений (Мюллер, Гвидо, 2018):

Первый скрытый слой:

$$h1[0] = w[0.0]x[0] + w[1.0]x[1] + w[2.0]x[2] + \dots + w[k.0]x[k]$$

$$h1[1] = w[0.1]x[0] + w[1.1]x[1] + w[2.1]x[2] + \dots + w[k.1]x[k]$$

$$h1[2] = w[0.2]x[0] + w[1.2]x[1] + w[2.2]x[2] + \dots + w[k.2]x[k]$$

...

$$h1[N] = w[0.N]x[0] + w[1.N]x[1] + w[2.N]x[2] + \dots + w[k.N]x[k]$$

Второй скрытый слой:

$$h2[0] = u[0.0]h1[0] + u[1.0]h1[1] + u[2.0]h1[2] + \dots + u[N.0]h1[N]$$

$$h2[1] = u[0.1]h1[0] + u[1.1]h1[1] + u[2.1]h1[2] + \dots + u[N.1]h1[N]$$

$$h2[2] = u[0.2]h1[0] + u[1.2]h1[1] + u[2.2]h1[2] + \dots + u[N.2]h1[N]$$

...

$$h2[N] = u[0.N]h1[0] + u[1.N]h1[1] + u[2.N]h1[2] + \dots + u[N.N]h1[N]$$

Выход:

$$\bar{y} = v[0]h2[0] + v[1]h2[1] + v[2]h2[2] + \dots + v[N]h2[N],$$

где w , u , v — весовые коэффициенты, $h1$, $h2$ — персептроны первого и второго слоя; x — входные значения амплитуд признаков; N — количество персептронов первого и второго слоя персептронов, равное 200; k — количество входных значений.

Данный алгоритм на выходе дает значение $\bar{y}$, которое является контролирующим, — при выполнении множества итераций весовые коэффициенты, входящие в уравнения слоев $h1(\dots)$, $h2(\dots)$ и выхода $\bar{y}(\dots)$, меняются для достижения минимума $\bar{y}$. После того как функция $\bar{y}$ приняла минимум, сеть весовых коэффициентов фиксируется, обучение модели заканчивается, а выходное значение принимается соответствующим классу входного объекта и считается, что нейронная сеть обучена. Одним из результатов работы алгоритма является статистическая оценка точности прогноза (ассигасу) — мера успешности обучения нейронной сети.

На следующем этапе полученные регрессионные уравнения (собственно, персептрон) применили к рассчитанным для всей площади спектральным признакам рельефа, в результате получили массив из 293 304 строк и двух колонок, где в первой колонке записан прогнозируемый класс, а во второй записана вероятность прогнозируемого класса. Таким образом, можно не только говорить о наиболее вероятном классе для любого участка на территории, но и о надежности отнесения его к выбранному классу (например, можно отсеять спорные участки и выбрать лишь наиболее надежные прогнозы). Каждой строке массива соответствуют географические координаты X и Y , что дает возможность результаты прогноза перенести в ГИС и представить в виде карты (рис. 3). В результате для изучаемого района получено 87 прогнозных участков. Для анализа прогнозной силы предлагаемой методики использованы только те результаты опробования, которые получены в ходе полевых исследований.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Возможность успешного прогноза расположения участков поверхности, содержащей коренные источники камнесамоцветной минерализации, основана на том факте, что часто встречаемые проявления жильных геологических тел в рельефе земной поверхности образуют выпуклые формы с относительно резкими границами (Лоскутов, 1999). Это связано с избирательной денудацией вмещающих пород, как правило, они менее стойкие к экзогенным процессам. В пределах изучаемой территории породами, вмещающими полезную минерализацию, являются мраморы позднепротерозойского и раннекембрийского возраста, которые неустойчивы к химическому выветриванию в условиях гумидного климата. В отличие от магнезиальных и алюмосиликатных минералов, формирующие жилы и гнезда с рубинами и шпинелью. На момент проведения полевых исследований на территории массива Лук-Тьен было известно по литературным источникам 32 проявления полезной минерализации, из которых 7 являются коренными (Long et al., 2013). Эти данные явились основой обучения нейронной сети. При этом в качестве предикторов применены 40 признаков (амплитуд двумерного спектра) поверхности для характеристики морфологии рельефа, отвечающей бинарной классификации тестовых участков — содержание тела с полезным компонентом или же не содержащие. Средняя оценка качества обучения или точности прогноза (ассигасу) по результатам нескольких циклов обучения тестовой выборки составила 0.7–0.75 при максимально возможной, равной 1. После более тщательного подбора рекомендуемых параметров описанного выше классифика-

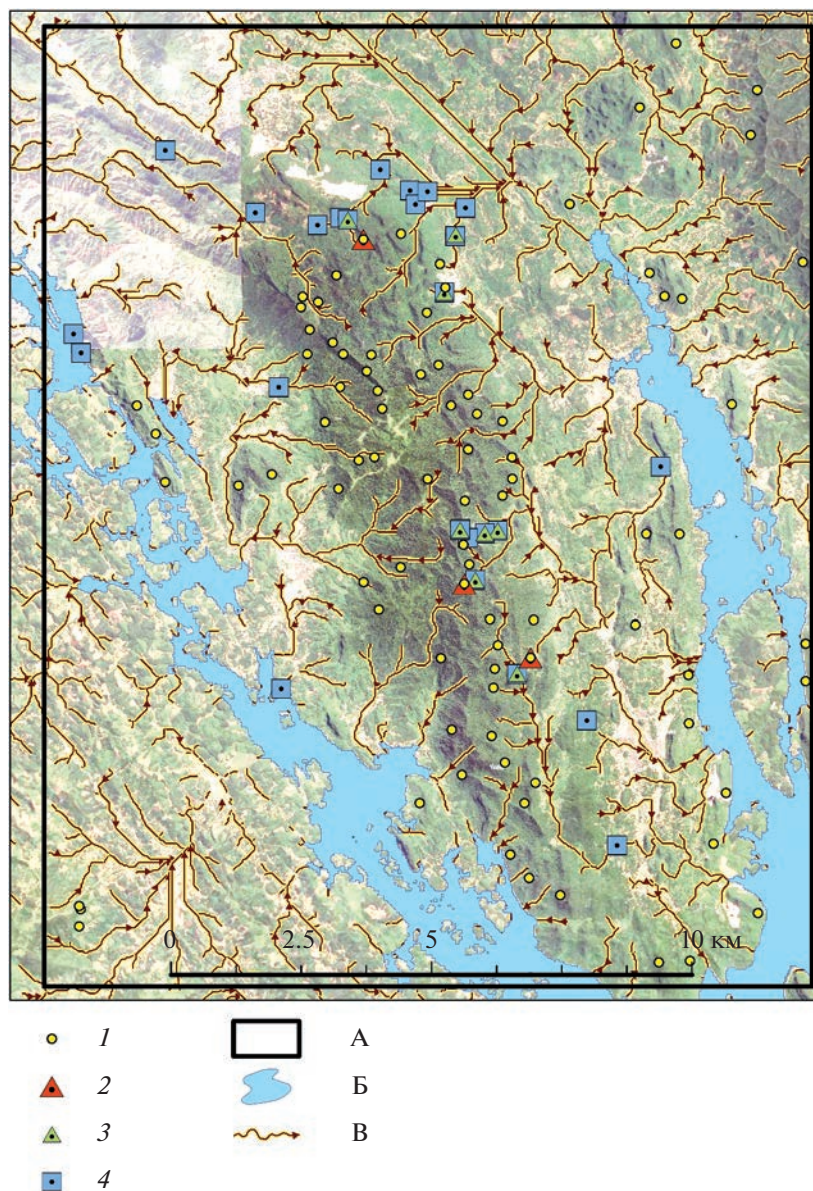


Рис. 3. Результаты прогноза коренных источников камнесамоцветной минерализации: 1 – перспективные участки, спрогнозированные нейронной сетью (87 шт.), 2 – участки, подтвержденные прямыми признаками (3 шт.), 3 – участки, подтвержденные косвенными признаками (8 шт.), 4 – точки опробования с признаками полезной минерализации (31 шт.). А – район работ, Б – водные объекты, В – направления потоков наносов.

Fig. 3. The results of the prediction of primary sources of gemstone mineralization: 1 – promising areas predicted by a neural network (87 units), 2 – areas confirmed by direct signs (3 units), 3 – areas confirmed by indirect evidences (8 units), 4 – sampling points with signs of useful mineralization (31 units). A – interest area, Б – water surface, В – sediment flows ways.

тора на основе многослойной нейронной сети (Мюллер, Гвидо, 2018) оценка обучения была улучшена до 0.83. Полученные данные были вынесены на карту-схему (на рис. 3 – обозначены под номером 1) и требовали проверки, которая и была выполнена в ноябре 2019 г. в ходе проведения полевых исследований.

В процессе проведения полевых работ мы соотнесли известные и вновь обнаруженные нами проявления полезной минерализации, как в пря-

мых признаках, так и во вторичных (в присклоновых осыпных ореолах рассеяния, в подвижных коллювиально-делювиальных толщах, в суффозионных отложениях) с прогнозными участками, расположенными в той же локализации. При этом получено подтверждение для трех площадок по прямым признакам – полезная минерализация непосредственно выделена из коренного залегания из жил в мраморах (на рис. 3 – под номером 2). Косвенные признаки позволили выделить бли-

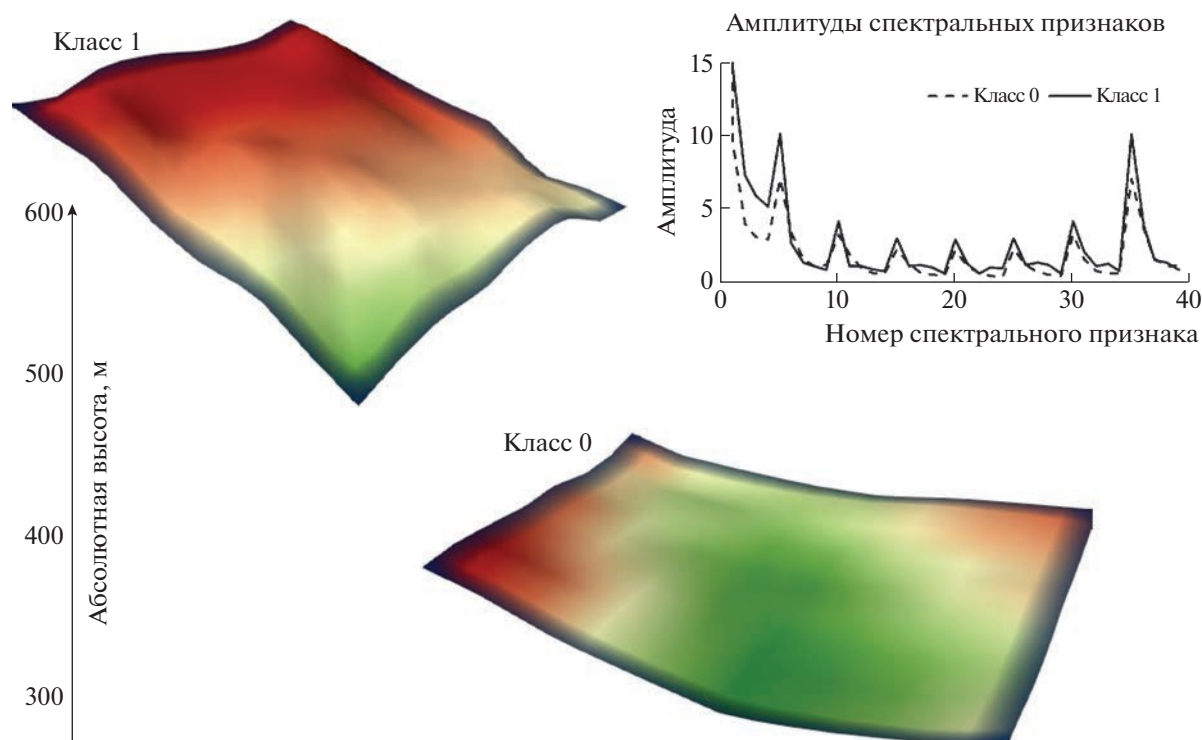


Рис. 4. Амплитуды спектральных признаков рельефа выделенных классов (справа сверху) и модели соответствующего обобщенного рельефа: класс 1 — генерализованный облик участков, содержащих коренные источники полезной минерализации (слева сверху); класс 0 — генерализованный облик участков, не содержащих коренных источников (справа внизу).

Fig. 4. Magnitudes of the spectral features of the land surface for selected classes (top right) and models of their generalized terrain image: class 1 — terrain of areas containing primary sources of gems (top left); class 0 — terrain of areas that do not contain primary gems sources (bottom right).

жайшие прогнозные участки как высоковероятные объекты первичной минерализации и они, *также*, вынесены на прогнозную карту (на рис. 3 — под номером 3). Путем использования показанной выше методики было спрогнозировано 87 площадок, потенциально содержащих коренные источники полезной минерализации. В процессе работ был взят ряд геологических проб. По данным шлихового и геохимического опробования положительные результаты на полезную минерализацию получены в 31 точке наблюдения (на рис. 3 — под номером 4). Из них 11 точек расположены в той же локализации, что и прогнозные участки. Оставшиеся 20 точек опробования так же имеют положительные результаты на полезную минерализацию, но прогнозная модель не предсказала эти участки (рис. 3). Это дает основания считать 20 объектов не подтвержденными. В статистических терминах это называется ложно отрицательный прогноз. Тогда как ложно положительных результатов не выявлено — в 11 точках прогноза получено 11 положительных проб. Таким образом, прогнозная сила (качество модели, или recall) рассмотренной методики составила $(11)/(11 + 20)$ или 0.35. К сожалению, не были

охвачены наблюдения оставшиеся 76 спрогнозированных участков в силу труднодоступности. Поэтому мы не можем сказать с какой-либо долей уверенности о наличии в их локализации коренных источников. Тем не менее полученный результат подтвержденной части прогноза, равный 35%, указывает, что примерно каждый третий спрогнозированный участок по вышеприведенной методике содержит полезную минерализацию в коренном заложении. Таким образом, в процессе полевых наблюдений удалось подтвердить 11 спрогнозированных объектов коренного залегания полезной минерализации.

При обучении нейронной сети, путем сравнения обобщенных спектров, отвечающих классам 1 (полезный компонент) и 0 (его отсутствие), были выявлены характерные черты рельефа, присущие участкам коренных источников рубино-шпинельной минерализации. Различие амплитуд спектра для этих классов показано на рис. 4 справа сверху на графике. По горизонтальной оси отложены порядковые номера гармоник двумерного спектра (спектральных признаков), а по вертикальной оси — амплитуда. Как видно на графике, классы уверенно различимы по амплитуде. Осо-

бенно это заметно для гармоник под номерами 0–5, 16–19, 24–25, 28–30, 33–35. Следует учесть, что двумерный гармонический ряд обычно представляется как таблица, но в данном случае, для отображения на плоском графике спектральная матрица развернута по одной оси. Даже небольшие относительные различия амплитуд стандартного набора гармоник (особенно для низкочастотных составляющих) от участка к участку указывают на разный рельеф этих площадок. Это также видно по восстановленному рельефу путем обратного преобразования Фурье из обобщенных спектральных признаков (рис. 4): вверху слева участки рельефа с крутыми склонами в пригребневых частях водоразделов на высотах 500–600 м с сильно трещиноватой поверхностью положительной кривизны с расчленением около 200 м соответствуют выходам полезной минерализации в коренном залегании. И наоборот, рельеф оснований речных долин в межгорных впадинах с плавными характеристиками отрицательной кривизны на высотах 300–350 м с расчленением не более 150 м не является прямым признаком первичной минерализации – характерный облик рельефа таких участков показан на рис. 4 внизу справа. Таким образом, с применением машинного обучения удалось формализовать процедуру поиска по объективным спектральным критериям рельефа для территории исследований площадью более 200 км².

ЗАКЛЮЧЕНИЕ

В ходе подготовки и проведения геолого-геоморфологических поисковых работ в северном Вьетнаме была апробирована методика прогноза участков с рубиново-шпинельной минерализацией на основе морфометрических данных. Метод основан на разложении ЦМР участка на гармонические составляющие с усовершенствованием алгоритмов, предложенных ранее (Сергеев и др., 2020). Были проанализированы спектральные характеристики рельефа участков, содержащих проявления первичной и вторичной полезной минерализации, известные по литературным источникам. В качестве метода для бинарной классификации территории были выбрана нейронная сеть, а именно двухслойный персептрон, достигнутая точность созданной модели довольно высока – 0.83, однако эта оценка получена при приложении модели к тем же данным, на которых она была обучена. Независимая верификация по результатам полевых работ, шлихового и геохимического опробования показала прогнозную силу модели 0.35, что, однако, также может считаться высоким показателем, с учетом предсказания редкого (т.е. проявляющегося на малой доле площади территории) явления.

Высокая прогнозная сила, однако, налагает и некоторые ограничения в виде переобучения мо-

дели – для другой территории данная обученная нейронная сеть не даст столь же высокого и при этом достоверного результата. Тем не менее намечаются и пути решения этой проблемы: для новой территории необходимы фактические данные о принадлежности хотя бы нескольких участков к обоим (прогнозируемый класс – месторождения, фоновый класс – остальная территория) классам, включение этих участков в выборку для обучения модели. Более того, количество независимых переменных в модели желательно расширить категориальными данными: возрастом и составом горных пород, информацией об их трещиноватости, тектонических нарушениях. Однако если данные свойства субстрата находят хотя бы частичное выражение в морфологии рельефа – они могут быть учтены в качестве предикторов косвенно, через морфометрические индикаторы. Тем не менее в случае включения новых переменных в модель: нейронная сеть заново обучается и строится прогноз на новую территорию.

Более того, если какой-либо другой район является схожим с исходной территорией с геолого-геоморфологических позиций, то модель может с высокой долей уверенности экстраполироваться и на новые территории-аналоги без дополнительной корректировки. Преимущества такого подхода очевидны: первичный трудоемкий анализ больших массивов исходных данных сводится в машинную область, а статистически значимые результаты этого анализа передаются экспертам и проверяются полевыми наблюдениями. Таким образом, методика поиска участков с заданными особенностями рельефа на основе спектрального анализа может быть использована как в самостоятельном структурно-геоморфологическом анализе, так и для задач прогнозирования в комплексных поисковых геолого-геоморфологических исследованиях.

БЛАГОДАРНОСТИ

Авторы выражают огромную признательность П.Б. Соколову за оказанную помощь в организации и проведении полевых работ. Отдельная благодарность господину Г.А. Гуссиасу (G.A. Gussias) за помощь в проведении экспедиционных маршрутов.

СПИСОК ЛИТЕРАТУРЫ

- Дэвис Дж. С. (1990). Статистический анализ данных в геологии. Кн. 2. М.: Недра. С. 427.
- Лопатин Д.В., Шавель Н.И., Нефёдов Д.А. (2012). Структурная и поисковая геоморфология. М.: Академия. С. 240.
- Лоскутов Ю.И. (1999). Геоморфологические основы поисков месторождений твердых полезных ископаемых. Автореф. дис. ... докт. геогр. наук. Новосибирск: СНИИГГиМС. С. 39.

- Мюллер А., Гвидо С. (2018). Введение в машинное обучение с помощью PYTHON. Руководство для специалистов по работе с данными. М.—СПб.—Киев: Диалектика. С. 471.
- Плас Дж. В. (2018). Python для сложных задач: наука о данных и машинное обучение. СПб.: Питер. С. 576.
- Сергеев И.С., Егоров И.В., Глебова А.Б. (2020). Спектральный анализ рельефа для решения прогнозно-поисковых задач на примере рифтовой зоны Срединно-Атлантического хребта // Геоморфология. № 4. С. 34–44.
<https://doi.org/10.31857/S0435428120040094>
- Сергиенко А.Б. (2002). Цифровая обработка сигналов. СПб.: Питер. С. 608.
- Харченко С.В. (2017). К вопросу о применении гармонического анализа при количественной характеристике рельефа // Геоморфология. № 2. С. 14–24.
<https://doi.org/10.15356/0435-4281-2017-2-14-24>
- Харченко С.В., Казаков С.Г. (2018). Спектральные характеристики рельефа суши в задачах его автоматизированной классификации (на примере Южной Америки) // Вестник Московского государственного областного университета. Серия: Естественные науки. № 4. С. 39–49.
<https://doi.org/10.18384/2310-7189-2018-4-39-49>
- Хромченко А.И. (1967). Кривая пересеченности и корреляционная функция рельефа // Рельеф Земли и математика. М.: Мысль. С. 50–71.
- Davis J.D., Chojnacki J.D. (2017). Two-dimensional discrete Fourier transform analysis of karst and coral reef morphologies // Transactions in GIS. Vol. 21. No. 3. P. 521–545.
<https://doi.org/10.1111/tgis.12277>
- Geological and mineral resources map of Vjetnam. (2000). 1: 200 000. Bæc Quang. F-48-X. Publication permit No.212/CXB. Published and copyright by department of geology and minerals of Vjetnam. Hµ NËi.
- Khoi N.N., Hauzenberger Ch.A., Tuan D.A. et al. (2016). Mineralogy and petrology of gneiss hosted corundum deposits from the Day Nui Con Voi metamorphic range, Ailao Shan—Red River shear zone (North Vietnam) // N. Jb. Miner. Abh. (J. Min. Geochem.). Vol. 193/2. P. 161–181.
<https://doi.org/10.1127/njma/2016/0300>
- Long Pham V., Pardieu V., Giuliani G. (2013). Update of gemstone mining in Luc Yen, Vietnam // Gems & Gemology. Vol. 49. No. 4. P. 31–46.
<https://doi.org/10.5741/GEMS.49.4.233>
- Pham B.T., Prakash I., Bui D.T. (2018). Spatial prediction of landslides using a hybrid machine learning approach based on random subspace and classification and regression trees // Geomorphology. Vol. 303. P. 256–270.
<https://doi.org/10.1016/j.geomorph.2017.12.008>
- Pike R.J., Rozema W.J. (1975). Spectral analysis of landforms // Annals of the Assoc. of American Geographers. Vol. 65. No. 4. P. 499–516.
<https://doi.org/10.1111/j.1467-8306.1975.tb01058.x>

SPECTRAL ANALYSIS OF LAND SURFACE WITH THE CONSTRUCTION OF A NEURAL NETWORK FOR GEMS SEARCH ON THE EXAMPLE OF THE LUK TIEN MOUNTAIN RANGE (NORTHERN VIETNAM)¹

I. S. Sergeev^{a, #}, K. A. Kuksa^a, and A. B. Glebova^a

^aSt. Petersburg State University, St. Petersburg, Russia

[#]E-mail: igorsergeev.spb@gmail.com

The study area is located in the north of Vietnam in the province of Yen Bai and it is a large ($14.5 \times 6.5 \times 0.8$ km) structural and denudational butte on the periphery of high-dissected low mountains Con Voi, and they are also slopes and bottoms of the neighbor rivers valleys. There are a lot of gemstone outcrops on the territory related with the vein formations in the strata of marbles. The area is relatively difficult to access for geological fieldworks. Therefore, in order to organize and conduct field geological prospecting work, the task was to obtain preliminary data on the possible localization of useful mineralization areas based on the analysis of available geological and geomorphological information. For the task, the spectral regularities of the land surface dissection spatially associated with veined geological formations in the near-surface part of the marble strata were studied, we used the discrete Fourier transform for this. The binary classification (for classes of potentially useful and useless areas) of the elevation amplitudes according with different spatial frequency of topographic dissection was provided with the simple neural network – two-layer perceptron. This algorithm is implemented on the basis of the scientific analysis libraries of the Python. The application of this technique made it possible to carry out a prediction for ruby-spinel mineralization in bedrock over a study area of more than 200 km². Fieldworks in 2019 verified the predicted data by the ways of mineralogical and geochemical testing of the accessible part of the predicted points. An average estimate of the predictive strength of the method used was obtained as 35% – every third site predicted by the neural network actually contains the primary sources of rubies and spinels in the territory under consideration.

¹ For citation: Sergeev I.S., Kuksa K.A., Glebova A.B. (2023). Spectral analysis of land surface with the construction of a neural network for gems search on the example of the Luk Tien mountain range (Northern Vietnam). *Geomorfologiya i Paleogeografiya*. Vol. 54. No. 3. P. 138–149 (in Russian). <https://doi.org/10.31857/S2949178923030106>; <https://elibrary.ru/WDVKDT>

Keywords: 2D spectral terrain decomposition, search geomorphology, morphometric methods, machine learning, GIS, DEM

ACKNOWLEDGMENTS

The authors express their deep gratitude to P.B. Sokolov for his assistance in organizing and conducting field work. Special thanks to Mr. G.A. Gussias for help in organizing expedition routes.

REFERENCES

- Davis J.D., Chojnacki J.D. (2017). Two-dimensional discrete Fourier transform analysis of karst and coral reef morphologies. *Transactions in GIS*. Vol. 21. No. 3. P. 521–545.
<https://doi.org/10.1111/tgis.12277>
- Djevis Dzh. S. (1990). *Statisticheskii analiz dannykh v geologii. Kn. 2* (Statistical analysis of data in geology. Vol. 20. Moscow: Nedra (Publ.). 427 p. (in Russ.))
- Geological and mineral resources map of Vjetnam. (2000). 1: 200000. Bæc Quang. F-48-X. Publication permit No.212/CXB. Published and copyright by department of geology and minerals of Vjetnam. Hµ NËi.
- Hromchenko A.I. (1967). *Krivaya peresechennosti i korrelyatsionnaya funktsiya rel'efa* (Dissection Curve and the Correlation Function of the Earth Surface). Relief of the Earth and Mathematics. Moscow: Mysl' (Publ.). P. 50–71. (in Russ.)
- Kharchenko S.V. (2017). On the question of the application of harmonic analysis in the quantitative characterization of the relief. *Geomorfologiya*. No. 2. P. 14–24. (in Russ.).
<https://doi.org/10.15356/0435-4281-2017-2-14-24>
- Kharchenko S.V., Kazakov S.G. (2018). The spectral landform signatures for automatic terrain classification (on the example of South America). *Bulletin of the Moscow State Regional University. Series: Natural Sciences*. No. 4. P. 39–49. (in Russ.).
<https://doi.org/10.18384/2310-7189-2018-4-39-49>
- Khoi N.N., Hauzenberger Ch.A., Tuan D.A. et al. (2016.). Mineralogy and petrology of gneiss hosted corundum deposits from the Day Nui Con Voi metamorphic range, Ailao Shan–Red River shear zone (North Vietnam). *N. Jb. Miner. Abh. (J. Min. Geo-chem.)*. Vol. 193/2. P. 161–181.
<https://doi.org/10.1127/njma/2016/0300>
- Long Pham V., Pardieu V., Giuliani G. (2013). Update of gemstone mining in Luc Yen, Vietnam. *Gems & Gemology*. Vol. 49. No. 4. P. 31–46.
<https://doi.org/10.5741/GEMS.49.4.233>
- Lopatin D.V., Shavel' N.I., Nefjodov D.A. (2012). *Strukturnaya i poiskovaya geomorfologiya* (Structural and search geomorphology). Moscow: Academy (Publ.). 240 p. (in Russ.)
- Loskutov Ju.I. (1999). *Geomorfologicheskie osnovy poiskov mestorozhdenii tverdykh poleznykh iskopaemykh* (Geomorphological foundations of the search for deposits of solid minerals). PhD thesis. Novosibirsk: SNIIGGiMS (Publ.). 39 p. (in Russ.)
- Mjuller A., Gvido S. (2018). *Vvedenie v mashinnoe obuchenie s pomoshh'yu PYTHON. Rukovodstvo dlya spetsialistov po rabote s dannymi* (Introduction to machine learning with PYTHON. A guide for data scientists). Moscow–Sankt-Peterburg–Kyiv: Dialectics (Publ.). 471 p. (in Russ.)
- Pham B.T., Prakash I., Bui D.T. (2018). Spatial prediction of landslides using a hybrid machine learning approach based on random subspace and classification and regression trees. *Geomorphology*. Vol. 303. P. 256–270.
<https://doi.org/10.1016/j.geomorph.2017.12.008>
- Pike R.J., Rozema W.J. (1975). Spectral analysis of landforms. *Annals of the Assoc. of American Geographers*. Vol. 65. No. 4. P. 499–516.
<http://dx.doi.org/10.1111/j.1467-8306.1975.tb01058.x>
- Plas Dzh.V. (2018). *Python dlja slozhnykh zadach: nauka o dannykh i mashinnoe obuchenie* (Python for Complex Problems: Data Science and Machine Learning). Sankt-Peterburg: Piter (Publ.). 576 p. (in Russ.)
- Sergeev I.S., Egorov I.V., Glebova A.B. (2020). Spectral analysis of ocean bottom for geological prospecting (on the example of the rift zone of the Mid-Atlantic ridge). *Geomorfologiya*. No. 4. P. 34–44. (in Russ.).
<https://doi.org/10.31857/S0435428120040094>
- Sergienko A.B. (2002). *Tsifrovaya obrabotka signalov* (Digital signal processing). Sankt-Peterburg: Piter (Publ.). 608 p. (in Russ.).