

Статистический анализ плотности хребтов и долин по цифровой модели рельефа

А.А. Златопольский

*Институт космических исследований РАН, Москва, 117997, Россия
E-mail: aazlat@gmail.com*

В статье представлены результаты анализа хребтов и долин, которые обнаруживаются в цифровых моделях рельефа (ЦМР). С помощью технологии LESSA (Lineament Extraction and Stripe Statistical Analysis) автоматически выявлялись оси хребтов и долин и рассчитывалась их плотность в большом окне. По данным SRTM проанализировано 12 разнообразных участков земной поверхности, а масштаб выявляемых элементов рельефа изменялся в 16 раз. Оказалось, что плотность, измеренная внутри одной территории, изменяется в не очень большом диапазоне, а значение середины диапазона оказалось очень близким как для разных территорий, так и для разных масштабов анализа. Конкретное значение этой постоянной величины определяется эффективным разрешением инструмента анализа ЦМР, т.е. тем, насколько мелкие и близкие друг к другу объекты позволяет обнаружить используемый инструмент поиска хребтов и долин. Для технологии LESSA это значение средней плотности близко к 0,2. Обнаруженный факт позволяет, в частности, так выбирать размер окна для статистического исследования хребтов и долин, что в него заведомо будет попадать достаточно объектов для достоверных измерений.

Ключевые слова: цифровая модель рельефа, хребты и долины, плотность

*Одобрена к печати: 25.05.2016
DOI: 10.21046/2070-7401-2016-13-3-209-216*

Введение

Анализируя по цифровой модели (ЦМР) статистические свойства элементов рельефа (долин и хребтов), необходимо учитывать число этих элементов в окне расчета, или, другими словами, плотность этих элементов. Эту плотность будем определять отношением суммарной длины найденных хребтов и долин к площади окна. И длина, и площадь измеряется в пикселях. Эту величину необходимо знать хотя бы для того, чтобы оценивать достоверность и статистическую обоснованность получаемых результатов. А кроме того, эта величина используется в качестве самостоятельного признака при геологическом анализе территорий. В ходе практических исследований различных территорий обратил на себя внимание тот факт, что плотность, хотя и изменяется внутри исследуемой области, но в целом мало меняется при переходе от области к области и от одного масштаба исследования к другому. Этот факт был коротко отмечен в предыдущей работе (Златопольский, 2015), и здесь мы уточним как количественное описание, так и природу этой закономерности.

Отметим следующие существенные моменты предлагаемого анализа. Во-первых, как сказано выше, рельеф задан в растровом виде (ЦМР). Во-вторых, элементы рельефа выявляются локально, по их геометрии, как складка местности, а не путем последовательного прослеживания, как часто определяют линии дренажной сети. И в третьих, элементы рельефа ищутся автоматически, что позволяет объективно, в едином ключе измерять и сравнивать характеристики элементов разных территорий, разного размера (масштаба, ранга). В этом исследовании использована технология LESSA (Златопольский, 2011),

которая учитывает особенности работы на растре – связность (толщину) получаемых линий, их расположение относительно растра (например, хребет в 7 пикселей при расположении вдоль растра может занимать только 5 пикселей, если он расположен по диагонали к растру) и т.п. Для каждой точки ЦМР рассматривается ее небольшая окрестность (10 пикселей в диаметре) и проверяется, нет ли в этой окрестности двух склонов противоположной ориентации, которые сходятся в этой точке. Для того, чтобы по возможности снизить субъективность результатов, параметры этой процедуры постоянны и учитываются элементы рельефа с любой крутизной склона. Отметим, что тестовые проверки показали нецелесообразность работы только с долинами – данных анализа при этом вдвое меньше, а результаты, по сути, те же.

Далее мы рассмотрим, как плотность элементов рельефа меняется внутри исследуемых областей, при переходе от области к области, как она меняется при изменении размера (масштаба) выявляемых объектов, и постараемся дать объяснение полученным результатам.

Материалы и методы

Исследование проведено на цифровых моделях рельефа (SRTM), которые получены в открытом доступе программой Global Mapper с разрешением 0,065 км/пикс. Использовано 12 областей размером около 400000 кв. км, расположенных от Испании до востока Китая и от Индии до центральной России, с совершенно различным горным и равнинным рельефом (*табл. 1*). Не включались области с заведомо бедным «пустынным» рельефом. Из анализа исключались участки без данных (например, водная поверхность).

Анализ осуществлялся следующим образом. В области выявлялись пиксели, относящиеся к осям долин и хребтов, и в скользящем окне рассчитывалась их плотность. Для того чтобы заменить словосочетание «хребты и долины» одним словом, будем далее использовать слово «линформа» (сокращение от «линейная форма» рельефа). Для того чтобы выявлять не случайные детали, а общие закономерности, использовано большое круглое окно диаметром 1440 пикселей или 94 км на местности.

Нас интересовало то, как значения плотности меняются внутри области и как этот диапазон изменяется от области к области. Для этого мы измеряли диапазон значений плотности для каждой области, при определении которого был отброшен 1% нехарактерных крайних значений (квантильный диапазон). По минимальному и максимальному значению диапазона находится его абсолютная величина, D , центр (середина) диапазона, C , и относительная величина диапазона (в процентах) $R = D/C \times 100\%$. Тесты показали, что значения C очень близки к плотности по всей области (превышают ее на 3–6%).

Измерив эти показатели для различных областей и для объектов разного масштаба, мы сможем определить, как сильно изменяется диапазон плотности, как изменяется центр этого диапазона и есть ли закономерности в этих изменениях.

Таблица 1. Исследованные области и статистика плотности хребтов и долин.

№	Площадь (км)	Размер (в пиксе)	Расположение		D	C	R %
1	685×625 км	10505 × 9585	37,5÷43 с.ш.	0,75÷8,5 з.д.	0,0516	0,2049	25,2
2	627×672 км	9631 × 10329	30÷36 с.ш.	36,5÷43 в.д.	0,0719	0,2313	31,1
3	527×627 км	8097 × 9635	43,2÷48,7 с.ш.	38,5÷45 в.д.	0,1209	0,1957	62,0
4	526×792 км	8060 × 12145	41÷48 с.ш.	21,5÷27,75 в.д.	0,0849	0,2134	39,8
5	617×682 км	9486 × 10496	48÷57 с.ш.	52÷58 в.д.	0,0624	0,2042	30,6
6	595×622 км	9140 × 9555	46,8÷52,2 с.ш.	68÷75,8 в.д.	0,0505	0,2275	22,2
7	645×676 км	9914 × 10388	34÷40 с.ш.	70÷77 в.д.	0,1141	0,2112	54,0
8	632×670 км	9703 × 10291	19÷25 с.ш.	79÷85 в.д.	0,0644	0,2261	28,5
9	536×783 км	8216 × 12010	29÷36 с.ш.	83÷88,5 в.д.	0,0517	0,1851	35,9
10	619×568 км	9513 × 8719	46÷51 с.ш.	96÷104 в.д.	0,0699	0,1784	39,2
11	586×682 км	9000 × 10480	29÷35 с.ш.	105÷111 в.д.	0,0674	0,2069	32,5
12	651×651 км	10000 × 10000	48÷56 с.ш.	111÷123 в.д.	0,0994	0,1695	58,6

Плотность элементов рельефа для различных территорий

Начнем с самой общей оценки изменчивости плотности. Минимальное значение нижней границы диапазона плотности по всем 12 областям оказалось 0,1198 (область 12), а максимальное значение верхней границы – 0,2683 (область 7), т.е. разброс плотностей 0,1485, или в 2,24 раза. Диапазоны плотностей по отдельным областям, D, различаются существенно – от 0,0505 до 0,1209, а средняя по областям величина диапазона – 0,0758. При этом для разных областей относительный разброс, R, меняется так же существенно, как и абсолютные значения диапазонов: от 22,2% до 62% при среднем по всем областям значении $\bar{R} = 37,6\%$.

Это связано с тем, что центры диапазонов, C, изменяются гораздо менее существенно, от 0,1695 до 0,2313, т.е. в 1,36 раз, что почти в 2 раза меньше, чем разброс собственно плотности. Среднее по областям значение центра диапазона $\bar{C} = 0,2045$, и разброс значений C относительно \bar{C} невелик и составляет 30%.

В расчетах плотности использовано очень большое окно, но и в меньших окнах описанная ситуация сохраняется. Для примера, в области 12 с минимальным значением C проведены расчеты – плотности, диапазона (D и C) – в окнах гораздо меньшего размера. При последовательном уменьшении окна по площади в 4, 16, 64, 256 раз значение C меняется, не монотонно, в небольшом диапазоне 0,1692–0,1758. А диапазон, D, монотонно растет, и для минимального окна (диаметр 88 пикселей) он в 1,62 раза больше исходного диапазона. Рост диапазона может быть обусловлен как ростом числа измерений (числа окон) в сотни раз, так и появлением нехарактерно больших и малых значений плотности в меньших окнах, хотя и в минимальном окне (в соответствии с найденной плотностью) – более 1000 пикселей элементов рельефа, что достаточно для достоверной статистики.

Повторив измерения в уменьшающихся окнах для области 2 с максимальной плотностью, получаем схожие результаты – значения C меняются в интервале 0,21–0,2312, при монотонном росте диапазонов (в 1,59 раз для минимального окна). В итоге, можно отметить, что результаты, приведенные в начале параграфа, не обусловлены конкретным размером окна анализа.

Таким образом, значения плотности внутри одной области могут изменяться более чем в 2 раза, но плотность всех областей оказывается почти в том же круге значений. Дело в том, что центры всех диапазонов находятся недалеко ($\pm 15\%$) от значения 0,2 и все диапазоны существенно перекрываются, так что максимальный разброс плотности в одной области – 0,1209, а разброс по всем областям – 0,1485.

Плотность элементов рельефа для различных масштабов.

Рассмотрим, как статистика плотности меняется при смене масштаба анализа (величины разрешения). Для каждой области проводилось 9 шагов измерений с масштабом анализа, различающимся в $\sqrt{2}$ раз. Для того чтобы при смене масштаба определять плотность в одних и тех же окнах на местности, размер окна в пикселях уменьшался соответственно разрешению. Напомним, что, как мы видели выше, нельзя ожидать изменения C просто из-за уменьшения размера окон в пикселях. Первая редукция ЦМР в $\sqrt{2}$ раз выполнялась бикубическим алгоритмом, а дальше достаточно было попеременно сжимать полученную пару ЦМР в 2 раза (усредняя каждую ячейку 2 x 2 пикселя).

Оказалось, что C изменяется при разных масштабах анализа не существеннее, чем при смене областей: минимальное значение – 0,1695 (область 12, масштабный шаг 1), а максимальное – 0,2332 (область 4, шаг 5), т.е. в 1,35 раза. А среднее значение C по всем масштабам и областям – 0,2056, т.е. почти такое же, как среднее по областям на шаге 1 – 0,2045. Не обнаружилось закономерности в изменении среднего, \bar{C} , при смене масштаба (размера) анализируемых элементов рельефа. Оказалось, что изменения C совсем не велики (разброс 0,0127) и они не монотонны, что приводит к разному соотношению плотности по областям при разных масштабах. Если расставить области по росту величины C , то на шаге 1 первые 4 области идут по порядку – 12, 10, 9, 3, а на шаге 6 порядок меняется на 10, 9, 3, 7.

А вот относительный разброс плотностей, R , различается при разных масштабах анализа существенно больше, чем по областям – от 11% до 74%. Средняя величина по 12 областям, \bar{R} , падает от 37,6% до 25,5%, при среднем значении по всем областям и масштабам – 31,6%. По отдельным областям изменения R при смене масштаба анализа происходят различным образом (*рис. 1*). Хотя у большей части областей при крупных масштабах эта величина падает, а не растет, как отмечалось выше, при уменьшении размера окна в одном масштабе. Возможно, дело в том, что в эксперименте с одним масштабом резко растет число измерений, а здесь оно сохраняется.

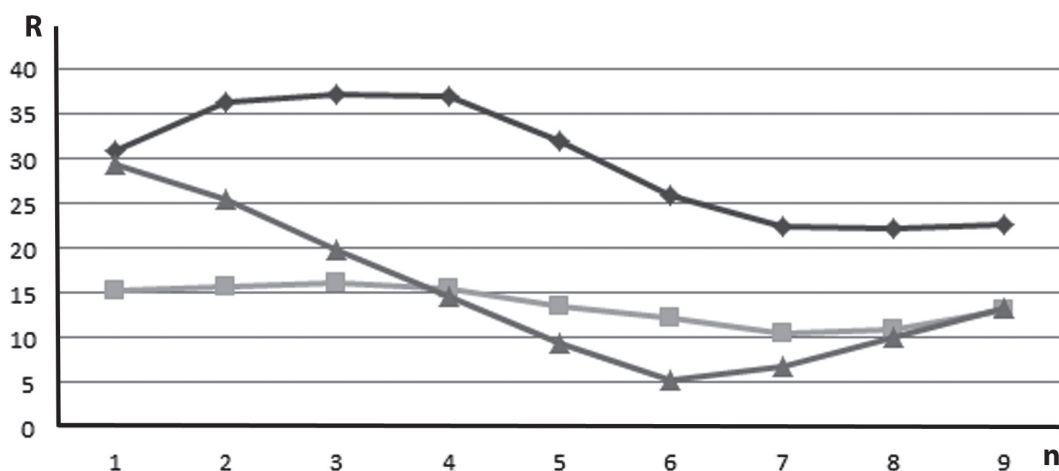


Рис. 1 Относительный разброс плотности, R, (в процентах) в зависимости от масштаба анализа (номер масштабного шага, n) для областей - 3 (ромбы), 5 (квадраты) и 12 (треугольники)

В итоге общее изменение плотности и по разным областям, и по элементам рельефа разного масштаба оказалось от 0,1152 (область 3, шаг 3) до 0,2735 (область 9, шаг 8), т.е. изменяется почти так же, как и по разным областям в исходном масштабе. В целом значения плотности меняются заметно, но разброс не превышает $\pm 38\%$ от среднего по областям и масштабам значения (0,2056). При этом плотность внутри отдельных областей в среднем меняется в 2 раза меньше, а именно $\pm 15\%$ относительно С. А сами значения, С, близки для разных областей (0,2056 $\pm 15\%$) и мало изменяются при смене масштабов.

Почему мало изменяется средняя плотность элементов рельефа

Напомним, что все эксперименты производились на единообразных данных одного качества, с достаточно хорошо представленными мелкими деталями. Обнаружено, что плотность точек линформ в рельефе при достаточно большом окне, хотя и меняется существенно (больше, чем в 2 раза), но в среднем для разных областей и масштабов анализа изменяется в районе одного и того же значения – 0,2. Обратим внимание, что описанная методика направлена на поиск всех точек линформ, вне зависимости от амплитуды изменений (глубины вреза). Таким образом, получается, что если выявлять всё похожее на линформу, то таких объектов будет одинаково много и для разных областей, и для разных масштабов.

Обсудим, почему плотность в среднем меняется мало для объектов разного масштаба, а затем – почему ее величина изменяется именно около этого значения. Пусть мы проанализировали ЦМР в окне и перешли к более детальному разрешению. При этом проявляются пространственные гармоники более высокой частоты. Возможно, они более слабые, но они есть, а мы выявляем всё. В результате, детектируются не только более мелкие детали прежних линформ, но и много новых. Если бы при росте детализации в К раз сохранялись только прежние линформы, то их длина в пикселях выросла бы в К раз (или немного больше), а значит плотность упала бы в К раз. Но дело в том, что и расстояние между преж-

ними линформами выросло в K раз и в этих местах обнаруживается много новых более мелких объектов, которых соответственно в K раз больше, чем прежних, так что плотность не падает. Эта картина описывает ситуацию в среднем, а в конкретных случаях плотность в окне, конечно, может и изменяться.

Теперь разберемся в том, почему значения плотности колеблются вокруг значения 0,2, а, скажем, не около 0,05 или 0,4. Локальный алгоритм поиска линформ опирается на некоторую окрестность точки ЦМР, в которой ищутся склоны. В алгоритме заложены правила и параметры, по которым принимается решение о том, есть ли в этой окрестности склоны и линформы. В результате, алгоритм и его параметры определяют размер объектов, которые могут быть обнаружены, а более мелкие объекты напрямую не детектируются (об эффективном пространственном разрешении инструмента см. (Златопольский, 2014)).

В использованном алгоритме анализируется окрестность размером 10 пикселей в диаметре, и, несомненно, обнаруживаются линформы длиной больше 9 пикселей и со склонами шириной больше 4 пикселей. Как правило, на результат поиска более мелких деталей влияет, главным образом, именно ширина склонов или расстояние между соседними однонаправленными объектами (длина хребтов и долин обычно достаточна). При ширине склона в 4 пикселя линформы обнаружатся хорошо, а при ширине в 1 пиксель – не обнаружатся совсем.

Проясним ситуацию с плотностью на искусственном примере, который легко себе представить. Пусть мы имеем ЦМР, все строки которой одинаковы, а вдоль строки высота меняется по синусоиде с периодом в K пикселей, т.е. имеем параллельные хребты и долины меридионального направления. Если $K=20$, то данный алгоритм найдет все оси хребтов и долин, т.е. получится 2 точки линформ на период и плотность этих точек 0,1. При $K=10$ также выделятся все точки линформ, с промежутком между ними в 4 пикселя и плотностью 0,2. При $K=6$ такие идеальные линформы также найдутся с промежутком между ними в 2 пикселя и плотностью 0,33. Но при $K=2$ – ничего не выделится.

Из этого примера понятно, что хотя в идеале используемым алгоритмом могут детектироваться и более мелкие, более частые линформы, дающие плотность 0,33 и меньше, но в реальных ЦМР обнаруживаются в основном несколько более крупные детали, дающие плотность около 0,2.

Вернемся к нашему мысленному эксперименту, чтобы получить еще одно косвенное подтверждение предложенному пониманию ситуации. После детектирования хребтов и долин появляются пиксели, относящиеся к линформам, а остальные пиксели назовем «пустыми». При синусоиде с периодом $K=6$ получаем линформы, которые располагаются с промежутком в 2 «пустых» пикселя, а значит, все «пустые» пиксели лежат на расстоянии в 1 пиксель от линформ. Если период $K=8$, то между линформами – 3 «пустых» пикселя и 66,6% из них лежат на расстоянии в 1 пиксель от линформ, а 33,3% – на расстоянии в 2 пикселя. Для $K=10$ аналогично имеем соотношение этих расстояний – 50% и 50%, а для $K=12$ – 40%, 40% и 20%. Ситуация понятна – чем дальше друг от друга расположены линформы, тем больше «пустых» пикселей, которые удалены от них все дальше. Эту удален-

ность от линформ мы можем практически измерить, чтобы через нее оценить расстояние между линформами. Нужно учитывать, что на растре измерение таких небольших расстояний происходит с существенной погрешностью, но общей картине можно доверять. Для области 5, масштаб 1 оказалось, что 63% «пустых» пикселей находятся на расстоянии 1 от линформ, 29% – на расстоянии 2, 6% – на расстоянии 3 и 2% – на расстоянии 4 и больше. Для области 12, масштаб 1, где средняя плотность самая маленькая, соответствующие цифры – 51%, 31%, 10,5%, 7,5%. Больше всего эти цифры соответствуют искусственному примеру с периодом около 10, который и дает плотность 0,2. Если вывести на дисплей и рассмотреть результаты поиска осей элементов рельефа реальных ЦМР, то между ними обычно окажется 3–5 пикселей, хотя бывает и 1–2.

Выводы

В статье приведены результаты статистического анализа плотности элементов рельефа, которая измерена на качественной ЦМР в большом скользящем окне. Для одной территории плотность изменяется в не особенно большом диапазоне вокруг некоторого значения, которое оказалось очень близким и для разных географических областей, и для разных масштабов анализа. В итоге обнаружено значение плотности, вокруг которого в некотором диапазоне ($\pm 40\%$) находятся значения плотности для разных территорий и разных масштабов анализа. Этот факт можно учитывать при статистическом анализе элементов рельефа. Например, определив указанную общую среднюю величину, можно выбирать размер окна статистического анализа, в котором, как правило, будет достаточно данных для достоверного измерения.

Еще один вывод состоит в том, что при тематическом анализе плотности хребтов и долин в первую очередь важны относительные особенности – аномалии плотности, границы областей разной плотности и т.п., а вот абсолютные значения плотности могут пригодиться меньше, так как сходны у разнородных областей.

Общая средняя величина плотности зависит от алгоритма поиска элементов рельефа, от его детальности (эффективного разрешения). В технологии LESSA оси хребтов и долин выявляются примерно на расстоянии 4 пикселя друг от друга, что и дает среднюю плотность около 0,2. Для сопоставления результатов было бы интересно, если бы пользователи других систем поиска элементов рельефа измерили и опубликовали получаемую ими среднюю плотность.

Некоторые предварительные исследования позволяют предположить, что аккуратное введение ограничений на поиск элементов рельефа (отбраковка) не изменит обнаруженную ситуацию, а приведет только к изменению значения средней плотности. Например, если в LESSA учитывать только долины, то сохранится постоянство средней плотности (около значения 0,1). Но если ввести ограничение на крутизну склонов, то плотность будет отражать уже не столько свойства рисунка рельефа, сколько дисперсию высот.

Вполне возможно, что похожая ситуация – мало изменяющаяся средняя плотность выявляемых особенностей – существует и при анализе локальными детекторами изображений, например, при поиске границ.

Литература

1. Златопольский А.А. Новые возможности технологии LESSA и анализ цифровой модели рельефа. Методический аспект // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 3. С. 38–46.
2. Златопольский А.А. Эффективное пространственное разрешение данных и инструмента. На примере мультимасштабного анализа ЦМР // Современные проблемы дистанционного зондирования Земли из космоса. 2014. Т. 11. № 2. С. 18–25.
3. Златопольский А.А. Мультимасштабный анализ цифровой модели рельефа. Экспериментальные закономерности // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 3. С. 21–29.

Statistical analysis of the ridges and valleys density via digital terrain model

A.A. Zlatopolsky

*Space Research Institute RAS, Moscow 117997, Russia
E-mail: aazlat@gmail.com*

The paper sets out results of ridge and valley analysis detected in the digital terrain models (DTM). Using LESSA (Lineament Extraction and Stripe Statistical Analysis) software axes of ridges and valleys were determined, and their density was calculated within in a big window. SRTM data were used and the surfaces of 12 very distinct regions of the Earth were analyzed and the size of the detected relief elements changed 16-fold. It was found that the density calculated for one region can change not in a big range inside the region, and the middle of that range is very similar for different regions and different analysis scales. The particular value of this mean density is determined by the effective resolution of the DTM analysis tool, i.e. how much small and close to each other objects could be detected by a given ridge and valley detecting tool. For the LESSA technology, this mean density value is approximately 0,2. This finding can help, for instance, predetermine the size of the window that contains enough objects for the reliable statistical analysis of valleys and ridges.

Keywords: digital terrain model, ridges and valleys, density

*Accepted: 25.05.2016
DOI: 10.21046/2070-7401-2016-13-3-209-216*

References

1. Zlatopol'sky A.A. Novye vozmozhnosti tekhnologii LESSA i analiz tsifrovoi modeli rel'efa. Metodicheskii aspekt (New LESSA technology resources and digital terrain model analysis. Methodology), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 3, pp. 38–46.
2. Zlatopol'sky A.A., Effektivnoe prostranstvennoe razreshenie dannykh i instrumenta. Na primere mul'timasshtabnogo analiza TsMR (Effective data and instrument spatial resolution. By the example of the multiscale DTM analysis), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2014, Vol. 11, No. 2, pp. 18–25.
3. Zlatopol'sky A.A., Mul'timasshtabnyi analiz tsifrovoi modeli rel'efa. Eksperimental'nye zakonomernosti (Multiscale digital terrain map analysis. Experimental regularities), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 3, pp. 21–29.