A universal approach for drainage basins

Erneson A. Oliveira1,2,3, Rilder S. Pires3, Rubens S. Oliveira3, Vasco Furtado1, Hans J. Herrmann3,4,5 & José S. Andrade Jr.3

Drainage basins are essential to Geohydrology and Biodiversity. Defining those regions in a simple, robust and efficient way is a constant challenge in Earth Science. Here, we introduce a model to delineate multiple drainage basins through an extension of the Invasion Percolation-Based Algorithm (IPBA). In order to prove the potential of our approach, we apply it to real and artificial datasets. We observe that the perimeter and area distributions of basins and anti-basins display long tails extending over several orders of magnitude and following approximately power-law behaviors. Moreover, the exponents of these power laws depend on spatial correlations and are invariant under the landscape orientation, not only for terrestrial, but lunar and martian landscapes. The terrestrial and martian results are statistically identical, which suggests that a hypothetical martian river would present similarity to the terrestrial rivers. Finally, we propose a theoretical value for the Hack’s exponent based on the fractal dimension of watersheds, \( \gamma = D/2 \). We measure \( \gamma = 0.54 \pm 0.01 \) for Earth, which is close to our estimation of \( \gamma \approx 0.55 \). Our study suggests that Hack’s law can have its origin purely in the maximum and minimum lines of the landscapes.

Drainage basins play a fundamental role in the hydrologic cycle, which make them essential to diversity and maintenance of Life on Earth1. The longstanding problem of characterising drainage basins has drawn much attention due to its importance in a variety of environmental issues, such as water management2–4, landslide and flood prevention5–9, and aquatic dead zones10–12. In this context, drainage basins, or simply basins, are all land areas sloping toward a single outlet, e.g. a river mouth or points of higher infiltration or evaporation rates. They are outlined by abstract boundary lines, called topographic divides or watersheds. The concept of watersheds appears in many other seemingly unrelated areas like percolation theory13,14, image segmentation and medicine15–18, and even international borders19,20.

Watersheds are recognized as fractals in several cases21–24, exhibiting self-similarity and a well defined fractal dimension. There are several objects that share the same fractal dimension of watersheds on uncorrelated random substrates (viz. \( D \approx 1.22 \)), such as optimal paths under strong disorder25–28, minimum spanning trees on random networks29, backbones of the optimal path crack30–32, and bridge bonds on ranked surfaces33. All these loopless paths belong to the same universality class as watersheds. Furthermore, it is also known that watersheds are Schramm-Loewner evolution curves34 and that it is possible to define hydrological watersheds35, where the infiltration process in the soil is taken into account.

The availability of digital elevation models (DEMs) allowed the development of modern techniques to automatically delineate watersheds. Nowadays, these methods are used in the standard Geographic Information System (GIS) software and are well established as a fundamental tool in Geoscience36. The basic idea behind these methods is the calculation of the flow directions, which can be defined in several ways depending on the assumptions on the flow and number of neighbouring cells in the grid37–44. Simultaneously to the development of these methods, the concept of watersheds were also introduced in the context of contour delineation45 and became a widely used tool for image processing. The key difference among these methods is that those adopted in image processing do not use channels to define watershed lines. Instead, they simply search the landscape for the highest lines that divide it. In this context, even the most modern computational models are not able to get all watersheds of the globe at once, they are usually limited to specific regions. As compared to all these methods, methods based

1Programa de Pós Graduação em Informática Aplicada, Universidade de Fortaleza, 60811-905, Fortaleza, Ceará, Brazil. 2Mestrado Profissional em Ciências da Cidade, Universidade de Fortaleza, 60811-905, Fortaleza, Ceará, Brazil. 3Departamento de Física, Universidade Federal do Ceará, Campus do Pici, 60451-970, Fortaleza, Ceará, Brazil. 4PMMH, ESPCI, 7 quai St Bernard, 75005, Paris, France. 5ETH Zürich, Computational Physics for Engineering Materials, Institute for Building Materials, Wolfgang-Pauli-Strasse 27, Hit, CH-8093, Zürich, Switzerland. Correspondence and requests for materials should be addressed to E.A.O. (email: erneson@eaoliveira.com)
on invasion percolation as ours do not need to explore the entire space but only a fractal subset of it. Therefore, our method becomes more and more efficient the larger the landscape.

Here, we propose a simple model to fully delineate multiple drainage basins for any given landscape based on the traditional Invasion Percolation (IP) model46, which is known to be a Self-Organized Criticality (SOC) model47, i.e. the IP model does not need a tuning parameter to evolve and identify the watershed. Our method allows us to study the morphological segmentation of global maxima and minima in image processing, even if the landscape stands for the gray scale of a brain Magnetic Resonance Imaging (MRI) or the heights of a celestial body with or without river channels. The novelty of our approach is to characterize all basins from a single height dataset through the definition of a reference (sea) level, i.e. our approach is free of parameter tuning.

**The Model**

In 2009, Fehr et al.22,23 introduced a model, called Invasion Percolation-Based Algorithm (IPBA), in order to extract watersheds from landscapes. The IPBA was proposed for a regular square lattice of size $L \times L$ with fixed boundary conditions in the vertical direction and periodic boundary conditions in the horizontal direction, where the height of each site $i$ was represented by $h_i$. It was also defined that the upper and lower lines of the lattice represent the sinks of two basins, e.g. one at the North (N) and other at the South (S), respectively. In this context, the following rule for the identification of the basins was proposed: For each site $i$, one applies the IP model, defining that the basin ($N$ or $S$) to which the site $i$ belongs is the one that the IP invaded cluster reaches first. Thus, all sites of the lattice belong to one of the two basins and the interface line between them defines the watershed. To improve the computational performance of the task of finding the interface line, an efficient sweeping strategy was also introduced: (i) Initially, the sites are chosen along a straight line that connects the sinks. Therefore, when the IP processes from two neighbouring sites evolve to different sinks, a segment of the watershed lies between them. (ii) From then on, the sites are chosen only in the neighbourhood of the already known watershed segments in order to reveal more segments of the watershed, resulting at the end in the complete watershed. Fehr et al.22 also showed that the IPBA follows the same dynamics as the Vincent-Soille algorithm15. The Vincent-Soille algorithm has a direct interpretation, which is basically a flooding process, although it is computationally inefficient. The main advantage of the IPBA is its computational performance, the IPBA presents a sublinear complexity time since it only explores a fractal subset of space of fractal dimension $91/48^{22}$ and this characteristic allows us to perform a global analysis to the drainage basins.

Our aim is to define a robust mathematical model for the delineation of multiple drainage basins through an extension of the IPBA. Suppose a regular rectangular lattice $L_x \times L_y$, where the height of each site $i$ is $h_i$, analogous to the original model. We introduce a height threshold $h^*$ such that, if $h_i > h^*$, then the $i$th site belongs to a cluster, which we call height cluster, composed by all connected sites with height above that threshold. Otherwise, the $i$th site does not belong to any cluster. As explained in the following section, we adopted $h^* = 0$ throughout this study, which for Earth corresponds to sea level. For this particular choice, the height clusters define continents and islands on Earth, as shown in Fig. 1A. Here, the sinks $S_k$ ($k = 1, 2, \ldots, N_b$) are all the $N_b$ border sites of the height clusters, i.e. the sea shore on Earth. Consequently, we know a priori that they define $N_b$ drainage basins separated by several interface lines, but their specific sizes and shapes need to be determined. Similarly to the ideas proposed by Fehr et al.22, we define the following rule to identify basins present in the height clusters: The IP model is applied for each site $i$ defining that the basin ($S_b$) at which the site $i$ belongs is the one that the IP invaded

**Figure 1.** Extension of the Invasion Percolation-Based Algorithm (IPBA). We use a Digital Elevation Model (DEM) from the Island of Hawaii’s region to show the three steps of our approach. (A) Single height cluster composing by all sites with $h_i > h^*$, where $h^* = 0$ (sea level) throughout the study. The landscape colour is arranged from dark purple (small heights) to light yellow (large heights) in linear scale. We represent in different colours all sinks, namely, the sites that are in the height cluster border, in this case, the sea shore. Finally, the Invasion Percolation (IP) cluster of site $i$ is shown in black. The IP process starts at the site $i$ and finishes at the sink $S_b$ (B) All drainage basins identified with our algorithm. The basins have the colours of their sinks and the black lines stand for the watershed network. (C) All drainage anti-basins from the largest drainage basin. The anti-basins are also represented by several colours and the black lines, in this case, stand for the anti-watershed network.
cluster reaches first (see Fig. 1A). As depicted in Fig. 1B, the set of interface lines forms the watershed network that separates all basins in the height clusters. We also use a strategy analogous to the original IPBA to improve the performance of finding the watershed network. Here, the sweeping occurs in each basin $S_i$ as follows: (i) The sink $S_i$ defines the ends (initial and final segments) of its yet not identified watershed. (ii) For each basin, the sweeping occurs only at sites neighboring the already known watershed segments in order to reveal the missing ones. In other words, we scanned the sites along the watershed inner perimeter neighbourhood of each basin. This strategy drastically reduces the number of times that we need to apply the IP algorithm for the identification of the watershed network. (iii) Optionally, a simple burning algorithm can be applied to each basin in order to evaluate its area.47

Actually, we consider two versions of our algorithm along the study: A version with traditional periodic boundary conditions in horizontal direction and unconventional periodic boundary conditions in vertical direction for real landscapes, and another version with fixed boundary conditions on both directions for artificial landscapes. The unconventional periodic boundary conditions, adopted for real landscapes, are defined by imposing that each site in the top (bottom) row be neighbour of every other site in the top (bottom) row. These boundary conditions represent a mapping of a sphere into a lattice. They are needed because we performed all measures on the entire globe for real landscapes.

A natural extension of the watershed concept is the definition of its reciprocal line, called anti-watershed. The anti-watersheds are composed by lines of minimal heights, in contrast to the watersheds, defined in terms of lines of maximal heights. Therefore, if the basins can (grossly) be understood as “cavities” in a surface, then the anti-basins can also be understood as “humps” on that same (inverted) surface. We emphasize that anti-watersheds do not necessarily represent river channels, since a river channel (or a channel) is a “clearly defined watercourse (”natural or man-made channel through which water may flow”) which periodically or continuously contains moving water”, or is a “watercourse forming a connecting link between two water bodies”, or even the “deepest portion of a watercourse, in which the main stream flows”,48 while a anti-watershed line may never contain any flowing water. Given an upside-down landscape, the same approach for watersheds can be considered to define the anti-watersheds. In this case, the watershed network represents the anti-watershed network in the original landscape. Furthermore, we can also define an anti-watershed network within a drainage basin (see Fig. 1C). Here, the lines of minimal heights represent the most deeper rivers and their tributaries in a lot of situations.

We highlight that our model just takes into account the drainage basins that flow out to the oceans, i.e. we are ignoring the endorheic drainage basins (or simply endorheic basins), which are those basins that flow out to any place other than the oceans, e.g. lakes or swamps, their amounts of water being balanced by infiltration or evaporation.49 Nonetheless, we could also include such basins in our model introducing additional sinks $S'_k$ (where $k = 1, 2, \ldots, N_e$ and $N_e$ is the number of endorheic basins), located at the points of higher evaporation rates, for example. In this context, the total number of drainage basins ($N_e + N_s$) would depend on the spatial resolution of the data and the information available about the endorheic basins. Our code is available at https://github.com/erneson.

Results

We applied the IPBA model to real and artificial landscapes in order to study the statistical properties of the drainage basins on Earth, Moon, and Mars. For real landscapes, we use three different DEMs throughout this study: the General Bathymetric Chart of the Oceans (GEBCO)50, the Lunar Orbiter Laser Altimeter (LOLA)51, and the Mars Orbiter Laser Altimeter (MOLA)52. Such datasets consist of the map of heights for the Earth, Moon and Mars. Moreover, we obtained the artificial landscapes through the fractional Brownian motion (fBm)53. We chose this method because we are interested in generating landscapes with tunable spatial long-range correlations since it is known that such correlations change the statistical properties of watersheds.24 We show all landscapes in Figs 2A–C and 3A–D. We considered two scenarios for GEBCO, LOLA, and MOLA datasets: the original and the upside-down landscape orientations. We emphasize that we chose the GEBCO dataset to represent the terrestrial landscapes because it includes altimetric and bathymetric heights, which makes it more similar to LOLA and MOLA datasets. In other words, we used the GEBCO dataset in order to make a general and uni-form comparison between Earth and two celestial bodies that do not have oceans (Moon and Mars). To perform a global analysis, the resolutions of the real datasets were decreased by a factor of $r = 8$, where $r$ is the resolution factor. Thus, each tile of $8 \times 8$ sites was replaced by a single site with a value given by the mean of all 64 original sites. The sizes of the used lattices were $5400 \times 23040$ for GEBCO (original $43200 \times 21600$), $5760 \times 2880$ for LOLA (original $46080 \times 23040$), and $5760 \times 2880$ for MOLA (original $46080 \times 23040$). For the fBm landscapes, we considered only the original orientation due to the natural symmetry of the Gaussian distribution used in the fBm. In this case, we averaged our simulations over 10 samples of lattices of $4096 \times 4096$ for the traditional range of the Hurst exponent $H$ ($0 \leq H \leq 1$). The following subsections show the corresponding results.

Real landscapes. Here, we show the main results of our approach applied to real landscapes. In Fig. 2D,G, we show the results of our model applied to the original and upside-down landscapes of the Earth. In Fig. 2E,H,F and I, we use the height threshold $h^* = 0$ (hypothetical sea level) and perform the same analysis used for Earth’s landscape to obtain the basins and anti-basins for the Moon and Mars. We note that the basins and anti-basins look similar in the terrestrial and martian landscapes, while in the lunar landscape, they are affected by the impact basins, i.e. craters originated from the impact of asteroids. This similarity is quantified here through the statistical distributions of perimeters and areas of the basins and anti-basins for Earth, Moon and Mars. We included the lunar analysis as a counterpoint to show that the observed similarity between the terrestrial and Martian results is indeed genuine. The log-log plots of all these distributions shown in Fig. 4A–B clearly indicate the presence of long tails extending over several orders of magnitude. Moreover, these tails approximately follow power-law
behaviors, $P(s) \propto s^\alpha$ and $P(A) \propto A^\beta$, for the perimeters and areas, respectively. By performing Ordinary Least Square (OLS) fits to the corresponding distributions, we obtained estimates for the power-law exponents $\alpha$ and $\beta$ that are summarized in Table 1. Similar results were found using a Maximum Likelihood Estimator (MLE) as shown in the Supplementary Information (SI).

Figure 2. Landscape, basins and anti-basins for Earth, Moon, and Mars. (A) General Bathymetric Chart of the Oceans (GEBCO)\textsuperscript{50}. (B) Lunar Orbiter Laser Altimeter (LOLA)\textsuperscript{51}. (C) Mars Orbiter Laser Altimeter (MOLA)\textsuperscript{52}. In all three cases, the heights are in units of kilometre (km) and represented in linear scale. (D–F) Drainage basins for heights above $h^*$ of the GEBCO, LOLA, and MOLA landscapes. (G–I) Drainage anti-basins for heights below $h^*$ of the GEBCO, LOLA, and MOLA landscapes. Note that $h^*$ = 0 (sea level for Earth and hypothetical sea level for Moon and Mars) throughout the study. Here, we use the following abbreviations: Terrestrial Basins (TB), Terrestrial Anti-basins (TA), Lunar Basins (LB), Lunar Anti-basins (LA), Martian Basins (MB), and Martian Anti-basins (MA). Sites below the height threshold $h^*$ are shown in grey.

Figure 3. Artificial landscapes and their respective basins. Fractional Brownian motion (fBm) landscapes generated by the Fourier filtering method (Ffm) for four regimes. A typical uncorrelated landscape ($H = -1$) is shown in (A), a landscape generated with negative correlations ($H = 0.2$) in (B), a Brownian motion landscape ($H = 0.5$) in (C), and a landscape with positive correlations ($H = 0.8$) in (D). All fBm landscapes shown share the same size $L = 1024$ and the same random seed. The heights are represented in linear scale, where lighter (darker) colours stand for higher (lower) heights. The basins defined by our approach corresponding to the fBm landscapes shown in (A–D) are shown in (E–H), respectively. Sites below the height threshold $h^*$ are shown in grey.
We found that the terrestrial and martian results are statistically identical, which suggests the surfaces of both planets underwent a similar formation history and that a hypothetical martian river would present some level of similarity to the terrestrial rivers, since both landscapes share the same statistics for watershed (maxima) and anti-watershed (minima) lines. There is already evidence that the Martian channels were formed by a series of fluvial and non-fluvial processes. Furthermore, we found evidence that the obtained perimeter and area exponents are independent of the resolution factor. It is known that the effects of finite size are attenuated as the system size increases. We also obtained similar results for other values of $h^*$ on Earth, Moon and Mars.

In Fig. 5, we show the Amazon basin and its associated anti-basins defined by our algorithm on GEBCO dataset at the original resolution. In this special case, we are removing all internal sinks from the South American continent, i.e. we are allowing the existence of sites with negative heights within South America. We emphasise that several rivers (the deeper ones) follow the anti-watershed lines. The black lines in Fig. 5 stand for the anti-watershed network, which shows an impressive similarity with the Amazon river network (See SI for further comparisons). This similarity allows us to obtain the length of the longest river in a basin by approximating it by the length of the longest path from the point of minimal height (the mouth of the river) of the largest tree on the anti-watershed network. We defined the height of each point of the anti-watershed lines as the mean of the heights of its neighbouring sites. Therefore, we are ignoring any issues related to the initiation point of river channels.

**Table 1.** Exponents $\alpha$ and $\beta$ for real landscapes. These exponents were determined through the Ordinary Least Square (OLS) fits to the corresponding distributions. Similar results were found using a Maximum Likelihood Estimator (MLE) (See SI).

| Cases               | Abbreviation | $\alpha$ ± | $\beta$ ± |
|---------------------|--------------|------------|-----------|
| Terrestrial Basins  | TB           | $2.30 \pm 0.03$ | $1.74 \pm 0.02$ |
| Terrestrial Anti-basins | TA       | $2.26 \pm 0.03$ | $1.72 \pm 0.02$ |
| Lunar Basins        | LB           | $2.70 \pm 0.02$ | $1.93 \pm 0.01$ |
| Lunar Anti-basins   | LA           | $2.64 \pm 0.04$ | $1.89 \pm 0.02$ |
| Martian Basins      | MB           | $2.33 \pm 0.04$ | $1.75 \pm 0.02$ |
| Martian Anti-basins | MA           | $2.26 \pm 0.02$ | $1.73 \pm 0.02$ |

Figure 4. Log-log plots of the perimeter and area distributions for basins and anti-basins on real and artificial landscapes. (A) The perimeter distributions for basins and anti-basins on Earth, Moon and Mars. In both, we use the following abbreviations: Terrestrial Basins (TB), Terrestrial Anti-basins (TA), Lunar Basins (LB), Lunar Anti-basins (LA), Martian Basins (MB), and Martian Anti-basins (MA). We show all exponents for real landscapes in Table 1. (C) The perimeter distributions for several values of the Hurst exponent $H$. (D) The surface area distributions for several values of $H$. All exponents are calculated through the Ordinary Least Square (OLS) fits. We also obtain similar exponents using a Maximum Likelihood Estimator (MLE) (See SI).
As a perspective for future work, quantitative analyses can be performed comparing the anti-watershed and river networks, e.g., see the proposed method by Grieve et al.58. We also performed the verification of Hack’s law59 for the entire planet. This scaling law establishes the relation between the areas ($A$) of the basins and the maximum lengths ($\ell$) of their rivers, i.e.,

$$\ell \propto A^\gamma,$$  \hspace{1cm} (1)

where $\gamma$ is known as the Hack exponent. In Fig. 6, we show Hack’s law for the original orientation of the terrestrial landscape considering only the basins with area greater than $A^* = 100$ km$^2$. We chose an area threshold $A^*$ to ensure that the analyzed basins were not too much affected by the resolution of the dataset. We found the Hack exponent $\gamma = 0.54 \pm 0.01$ with coefficient of determination $R^2 = 0.87$, which is very close to our theoretical value of $\gamma \approx 0.55$ for Earth.

**Artificial landscapes.** We applied the extension of the IPBA to artificial landscapes. In Fig. 3E–H, we show all basins defined by our model for the same samples presented in Fig. 3A–D. As shown in Fig. 4C,D, the perimeter and area distributions obtained from these landscapes are systematically affected by the presence of spatial correlations, quantified here in terms of the parameter $H$. Moreover, except for the case of $H = -1$, all distributions generated from artificial landscapes can be approximately described by power laws, $P(s) \propto s^{-\alpha}$ for perimeters, and $P(A) \propto A^{-\beta}$ for areas. For each value of $H$, we averaged both distributions for all samples. The insets of the Fig. 4C,D show that $\alpha$ and $\beta$ decrease with the Hurst exponent $H$. The exponents $\alpha$ and $\beta$ range between 3.36
and 1.93 and between 2.39 and 1.59, respectively. In the uncorrelated case ($H = -1$), however, we obtained less than one order of magnitude for $\ell$ and $A$ precluding the same kind of analysis.

In Fig. 6, we show Hack's law obtained for $H = 0.7$ (a close value of $H$ is usually obtained for real landscapes), considering only the basins with area above $A^* = 1024$. Such result led us to the following conjecture: Let the basin area be $A$, the longest anti-watershed line be $\ell$, and assuming that the anti-watershed lines are indeed watershed lines of the upside-down landscapes, it is known that $\ell \propto L^D$, where $L$ is the linear length of the system and $D$ is the fractal dimension of the watershed lines. Since $A \propto L^2$, we have:

$$L^D \propto \ell \propto A^* \propto L^2,$$

(2)

which gives $\gamma = D/2$. In other words, the Hack exponent depends on the fractal dimension of the anti-watershed lines. Fehr et al. showed that the fractal dimension of the watersheds decreases with the Hurst exponent $H$, similarly to the Optimal Path Cracks, and the coastlines on correlated landscapes. The fractal dimension of the watershed lines ranges between $D = 1.0$ and $D = 1.22$, what gives to the Hack exponent a corresponding range from $\gamma = 0.5$ to $\gamma = 0.61$. For real landscapes, the fractal dimensions of the watershed lines are around 1.10 ($D = 1.10$ for the Alps, $D = 1.11$ for the Himalaya, and $D = 1.12$ for the Andes). Therefore, our expected Hack exponent for Earth should be $\gamma \approx 0.55$, a value which is in good agreement with the result shown in Fig. 6 ($\gamma = 0.521 \pm 0.002$ with coefficient of determination $R^2 = 0.94$ for artificial landscapes). This result suggests that Hack's law, often observed for river networks, is an intrinsic effect of topography, i.e. it depends, in essence, on the watershed and anti-watershed lines. In other words, Hack's law may have a purely geometrical origin and does not depend on physical laws governing the water flow on a surface.

**Methods**

**Real landscapes.** We use three different DEM datasets throughout this study. The first dataset is the General Bathymetric Chart of the Oceans (GEBCO) consisting of altimetric and bathimetric heights, i.e. the heights above and below the sea level, around the Earth globe. The resolution of this dataset is 30 arc–seconds (30/3600 decimal degrees) in both coordinates, equivalent to a square lattice with edge length of 0.926 kilometers ($\text{km}$) at the Equator line. The other two are the Lunar Orbiter Laser Altimeter (LOLA) and Mars Orbiter Laser Altimeter (MOLA) consisting of the map of heights for the Moon and Mars, respectively. The LOLA resolution is about 0.118 km, while MOLA resolution is 0.463 km, both in relation to their corresponding “Equator”. We show the three datasets in Fig. 2A–C. In addition, we point out that all datasets are freely available and are friendly ready-to-use, i.e. all technical preprocessing steps were already performed by the GEBCO, LOLA, and MOLA research teams. We do not perform any additional preprocessing (such as any hydrologic correction or void filling), except the addition of a tiny noise in the heights in both coordinates, equivalent to a square lattice with edge length of 0.926 km.

We also emphasise that all three datasets are available in the Geographic Coordinate System (GCS), more precisely, in latitude and longitude grids in the image format TIFF, i.e. they are mapped on spheres, the terrestrial (with radius $R_{\text{Earth}} = 6378.137 \text{ km}$), the lunar (with radius $R_{\text{Mars}} = 1737.4 \text{ km}$), and the martian (with radius $R_{\text{Mars}} = 3396.19 \text{ km}$). For GEBCO, the reference surface (the zero height) is defined by the terrestrial geoid. The geoid is the natural shape that a static fluid would present due to the gravitational potential of its celestial body. On Earth, the oceans could be considered static and, consequently, they are well approximated by such a surface. For LOLA and MOLA, the geoid concept is generalised by the gravitational equipotential surface with the mean lunar and martian radius at the Equator, respectively, defining hypothetical sea levels. We adopt the height threshold $h^* = 0$ for all landscapes in order to make a general comparative analysis.

Here, we perform the calculation of the area of each site by the composition of two spherical triangles (the site areas for artificial landscapes have no unit of measure and are all unitary). The area of a spherical triangle with edges $a$, $b$ and $c$ is given by:

$$A = 4R_i^2 \tan \left[ \tan \left( \frac{a}{2} \right) \tan \left( \frac{b}{2} \right) \tan \left( \frac{c}{2} \right) \right]^{1/2},$$

(3)
where \( s = (a/R_k + b/R_i + c/R_j)/2 \), \( s_i = s - a/R_k \), \( s_j = s - b/R_i \), and \( s_\ell = s - c/R_j \). In this formalism, \( R_k \) is the sphere radius, where, in our case, \( k = \{ \text{earth, moon, mars} \} \), and the edge lengths are calculated by the great circle (geodesic) distance between two points \( i \) and \( j \) on the sphere surface given by the Haversine formula\(^\text{iv}\):

\[
d_{ij} = 2R \sin \left( \frac{\sin \phi_i \sin \phi_j + \cos \phi_i \cos \phi_j \sin \Delta \lambda}{2} \right),
\]

where \( \Delta \phi = \phi_i - \phi_j \) and \( \Delta \lambda = \lambda_i - \lambda_j \). The values of \( \lambda_i (\lambda_j) \) and \( \phi_i (\phi_j) \), measured in radians, are the longitude and latitude, respectively, of the point \( i (j) \). Therefore, we are able to define the site areas, and, consequently, obtain the total area of a given basin, since each basin is composed by a set of sites.

**Artificial landscapes.** We obtained artificial landscapes through the fractional Brownian motion (fBm)\(^\text{v}\) in order to study the watershed and anti-watershed networks. One of the most established method to generate a fBm is the so-called Fourier filtering method (Ffm)\(^\text{vi}\). The basic idea of the Ffm is to define random Fourier coefficients in the reciprocal space, distributed according to the following power-law spectral density:

\[
S(f_1, f_2, \ldots, f_d) = \left( \frac{1}{\sum_{i=1}^{d} f_i^2} \right)^w,
\]

where \( f_i \) is the frequency of the dimension \( i \), \( d \) is the topological dimension, and \( w \) is the spectral exponent. Subsequently, the inverse Fourier transform is applied to generate a correlated distribution in the real space. In our case \( d = 2 \), the correlated distribution is a landscape. Each landscape is characterised by an exponent \( H \), called Hurst exponent, related to the spectral exponent by \( w = 2H + d = 2H + 2 \). Four cases can be distinguished: (i) For \( H = -1 \), the uncorrelated landscape (see Fig. 3A). (ii) For \( 0 < H < 1/2 \), the landscape has a negative correlation, i.e. the increments are anticorrelated (see Fig. 3B). (iii) For \( H = 1/2 \), the landscape is correlated, but the increments are uncorrelated (see Fig. 3C), which is the case of the classical Brownian motion\(^\text{ix}\). (iv) Finally, for \( 1/2 < H < 1 \), the landscape has a positive correlation, i.e. the increments are correlated (see Fig. 3D).

**Data Availability**

All data used in this manuscript are free available. Please, check the references related to GEBCO, LOLA, and MOLA datasets.

**References**

1. Fetter, C.W. *Applied Hydrogeology* (Pearson Education Limited, 2014).
2. Vörösmarty, C. J., Federer, C. A. & Schloss, A. L. Potential evaporation functions compared on US watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modelling. *J. Hydrol.* 207, 147–169, https://doi.org/10.1016/S0022-1694(98)00109-7 (1998).
3. Knecht, C. L., Trump, W. ben Avraham, D. & Ziff, R. M. Retention capacity of random surfaces. *Phys. Rev. Lett.* 108, 045703, https://doi.org/10.1103/PhysRevLett.108.045703 (2012).
4. Brooks, K. N., Folliotii, P. F & Magner, J. A. *Hydrology and the management of watersheds* (Wiley-Blackwell, 2012).
5. Dhakal, A. S. & Sidle, R. C. Distributed simulations of landslides for different rainfall conditions. *Hydrol. Process.* 18, 757–776, https://doi.org/10.1002/hyp.1365 (2004).
6. Pradhan, B., Singh, R. P & Buchroithner, M. F. Estimation of stress and its use in evaluation of landslide prone regions using remote sensing data. *Adv. Space Res.* 37, 698–709, https://doi.org/10.1016/j.asr.2005.03.137 (2006).
7. Lazzari, M., Geraldi, E., Lapenna, V. & Loperte, A. Natural hazards vs human impact: an integrated methodological approach in geomorphological risk assessment on the Tursi historical site, Southern Italy. *Landslides* 3, 275–287, https://doi.org/10.1007/s10346-006-0055-y (2006).
8. Lee, K. T. & Lin, Y. T. Flow analysis of landslide dammed lake watersheds: a case study. *J. Am. Water Resour. Assoc.* 42, 1615–1628, https://doi.org/10.1111/j.1752-1688.2006.tb06024.x (2006).
9. Yang, D., Zhao, Y., Armstrong, R., Robinson, D. & Brodzik, M. J. Streamflow response to seasonal snow cover mass changes over large Siberian watersheds. *J. Geophys. Res. Earth Surf.* 112, F02S22, https://doi.org/10.1029/2006JF000518 (2007).
10. Diaz, R. J. & Rosenberg, R. Spreading Dead Zones and Consequences for Marine Ecosystems. *Science* 321, 925–939, https://doi.org/10.1126/science.1156401 (2008).
11. Breitburg, D. et al. Declining oxygen in the open global ocean and coastal waters. *Science* 359, eaam7240, https://doi.org/10.1126/science.aam7240 (2018).
12. Goudie, A. S. *The Human Impact on the Natural Environment: Past, Present, and Future* (Wiley-Blackwell, 2018).
13. Araujo, N. A. M., Grassberger, P., Kahng, B., Schrenk, K. J. & Ziff, R. M. Recent advances and open challenges in percolation. *Eur. Phys. J. Spec. Top.* 223, 2367–2381, https://doi.org/10.1140/epst/2014-02266-y (2014).
14. Saberi, A. A. Recent advances in percolation theory and its applications. *Phys. Rep.* 578, 1–32, https://doi.org/10.1016/j.physrep.2015.03.003 (2015).
15. Vincent, L. & Solle, P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.* 13, 583–598, https://doi.org/10.1109/34.87344 (1991).
16. Grau, V., Mewes, A. U. J., Alcaniz, M., Kikinis, R. & Warfield, S. K. Improved watershed transform for medical image segmentation using prior information. *IEEE Trans. Med. Imaging* 23, 447–458, https://doi.org/10.1109/TMI.2004.824224 (2004).
17. Ng, H. P., Ong, S. H., Foong, K. W. C., Goh, P. S. & Nowinski, W. L. Medical image segmentation using k-means clustering and improved watershed algorithm. *Proc. IEEE Southwest Symp. Image Anal. Interpert.* 61–65 https://doi.org/10.1109/SSIAM.2006.1633722 (2006).
18. Cousty, J., Bertrand, G., Najman, L. & Couprie, M. Watershed cuts: Thinnings, shortest path forests, and topological watersheds. *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 925–939, https://doi.org/10.1109/TPAMI.2009.71 (2010).
19. United Nations (1902) Reports of international arbitral awards. Available at http://legal.un.org/rria/cases/vol_IX/37-49.pdf Accessed November 13, 2018.
20. Eidgenössisches Departement für auswärtige Angelegenheiten (2009) Verlauf der grenze zwischen schweiz und italien. Available at https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-28489.html Accessed November 13, 2018.
Acknowledgements
We gratefully acknowledge CNPq, CAPES, FUNCAP and the National Institute of Science and Technology for Complex Systems in Brazil for financial support.

Author Contributions
E.A.O., H.J.H. and J.S.A. designed research; E.A.O., R.S.P. and R.S.O. performed research; E.A.O., R.S.P. and J.S.A. analyzed data; and E.A.O., R.S.P., R.S.O., V.F., H.J.H., and J.S.A. wrote the paper. All authors reviewed the manuscript.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-46165-0.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019