Watersheds in disordered media

Nuno A. M. Araújo1*, K. Julian Schrenk2†, Hans J. Herrmann3,4† and José S. Andrade Jr.3,4†

1 Departamento de Física, Centro de Física Teórica e Computacional, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal
2 Department of Chemistry, University of Cambridge, Cambridge, UK
3 Departamento de Física, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil
4 Computational Physics, Institute for Building Materials, ETH Zurich, Zurich, Switzerland

1. INTRODUCTION

Although both start in the mountains of Switzerland, the Rhine and Rhone rivers diverge while flowing toward different seas. While the Rhine empties into the North Sea, the Rhone drains into the Mediterranean. Similarly, all over the world one finds rivers with close sources but distant mouths. For example, the Colorado and the Rio Grande even open toward different oceans. When looking at a landscape, how to identify the regions draining toward one side or the other? When rain falls or snow melts on a landscape, the dynamics of the surface water is determined by the topography of the landscape. Water flows downhill and over-passes small mounds by forming lakes which eventually overspill. A drainage basin is then defined as the region where water flows toward the same outlet. Its shape and extension strongly depend on the topography. The line separating two adjacent basins is the watershed line, which typically wanders along the mountain crests [1].

Since watersheds provide information about the dynamics of surface water, they play a fundamental role in water management [2–4], landslides [5–8], and flood prevention [8–10]. Watersheds are also of relevance in the political context they have been used to demarcate borders between countries such as the one between Argentina and Chile [11]. Thus, the understanding of their statistical properties and resilience to changes in the topography of the landscape are two important questions that we will review here. In general, for every landscape, several outlets can be defined, each one with a corresponding drainage basin. Sets of small drainage basins eventually drain toward the same outlet forming a even larger basin. Such hierarchy results in a larger number of watersheds. However, without loss of generality, the study of watershed lines can be simplified by splitting the landscape into only two large basins, each one draining toward opposite boundaries. Figure 1 shows how this watershed can be identified by flooding the landscape from the valleys (see Supplemental video). Two sinks are initially defined (the lower-left and upper-right boundaries in the example). While flooding the landscape, each time two lakes (A and B) draining toward opposite sinks are about to connect, one imposes a physical barrier between the two. In the end, the watershed line is the line formed by the barriers that separate these two lakes.

The concept of watersheds is also of interest in other fields like, e.g., in medical image processing [12]. There, computed tomography scans need to be segmented to identify different tissues. The pictures are discretized into pixels and a number is assigned to each one of them according to the intensity. The segmentation procedure consists in clustering neighboring pixels following the order of increasing intensity gradient, splitting the image into different parts (tissues). The equivalent to the watershed line corresponds to the line separating two different tissues [13, 14].

It was recently shown that watersheds can be described in the context of percolation theory in terms of bridges and cutting bond models [15], with numerical evidence that they are

Keywords: watersheds, landscapes, correlations, percolation, perturbations

What is the best way to divide a rugged landscape? Since ancient times, watersheds separating adjacent water systems that flow, for example, toward different seas, have been used to delimit boundaries. Interestingly, serious and even tense border disputes between countries have relied on the subtle geometrical properties of these tortuous lines. For instance, slight and even anthropogenic modifications of landscapes can produce large changes in a watershed, and the effects can be highly nonlocal. Although the watershed concept arises naturally in geomorphology, where it plays a fundamental role in water management, landslide, and flood prevention, it also has important applications in seemingly unrelated fields such as image processing and medicine. Despite the far-reaching consequences of the scaling properties on watershed-related hydrological and political issues, it was only recently that a more profound and revealing connection has been disclosed between the concept of watershed and statistical physics of disordered systems. This review initially surveys the origin and definition of a watershed line in a geomorphological framework to subsequently introduce its basic geometrical and physical properties. Results on statistical properties of watersheds obtained from artificial model landscapes generated with long-range correlations are presented and shown to be in good qualitative and quantitative agreement with real landscapes.

Editorial information

Reviewed by:
Anna Carbone, Politecnico di Torino, Italy
Roberto F. S. Andrade, Universidade Federal do Bahia - Instituto de Física, Brazil

*Correspondence:
Nuno A. M. Araújo, Departamento de Física, Centro de Física Teórica e Computacional, Faculdade de Ciências, Universidade de Lisboa, Avenida Professor Gama Pinto 2, P-1649-003 Lisboa, Portugal
e-mail: nnmarujo@fc.ul.pt
†These authors have contributed equally to this work.

Reviewed by:
Anna Carbone, Politecnico di Torino, Italy
Roberto F. S. Andrade, Universidade Federal do Bahia - Instituto de Física, Brazil

*Correspondence:
Nuno A. M. Araújo, Departamento de Física, Centro de Física Teórica e Computacional, Faculdade de Ciências, Universidade de Lisboa, Avenida Professor Gama Pinto 2, P-1649-003 Lisboa, Portugal
e-mail: nnmarujo@fc.ul.pt
†These authors have contributed equally to this work.
FIGURE 1 | Watershed dividing a landscape into two parts. The landscape (A) is flooded from the valleys such that all regions lower than a certain height are covered with water (B). As the water level rises, lakes merge under the constraint that no lake can connect two predefined opposite boundaries of the landscape (C). In this example, we have taken the lower-left and the upper-right boundaries. Thus, a watershed line emerges separating the two final lakes (D). These lakes (A and B) are the drainage basins of this landscape.

Schramm-Loewer Evolution (SLE) curves in the continuum limit [16]. This association explains why watersheds on uncorrelated landscapes as well as other statistical physics models, such as, optimal path cracks [17–19], fuse networks [20], and loopless percolation [15], belong to the same universality class of optimal paths in strongly disordered media.

2. THE LANDSCAPE

The motion of surface water is defined by the topology of the landscape, usually characterized by the spatial distribution of heights. Although this distribution is a continuous function for real landscapes, it is typically coarse-grained and represented as a digital elevation map (DEM) of regular cells (e.g., a square lattice of sites or bonds) to which average heights can be associated. This process is exemplarily shown in Figures 2A–C. In fact, modern procedures of analyzing real landscapes numerically process Grayscale Digital Images where the gray intensity of each pixel is transformed into a height, resulting in a DEM. As such, discretized maps have been useful to delimit spatial boundaries in a wide range of problems, from tracing watersheds and river networks in landscapes [21–25] to the identification of cancerous cells in human tissues [13, 26], and the study of spatial competition in multispecies ecosystems [27, 28].

To study watersheds theoretically one can generate artificial landscapes. Starting with a regular lattice, a numerical value is randomly assigned to each element, corresponding to its average height. If these values are spatially correlated, a correlated artificial landscape is obtained. Otherwise, the landscape is called an uncorrelated artificial landscape. Natural landscapes are characterized by long-range correlations [29]. In Section 7 we will review how to generate correlated artificial landscapes and their main properties.

The DEM can be further simplified by mapping them onto ranked surfaces, where every element (site or bond) has a unique rank associated with its corresponding value [15]. A ranked surface can be defined in the following way. Given a two-dimensional discretized map of size $L \times L$, one generates a list containing the heights of its elements (sites or bonds) in crescent order, and then replaces the numerical values in the original map by their corresponding (integer) ranks. As depicted in Figure 2D, the result is a ranked surface.

3. HOW TO DETERMINE THE WATERSHED

Traditional cartographic methods for basin delineation have relied on manual estimation from iso-elevation lines. Instead, modern procedures are now based upon automatic processing of images as the ones obtained from satellites [30, 31]. These images are typically coarse-grained into discretized maps with each cell $h_i$ having a height $h_i$ as explained in Section 2. A recently proposed algorithm to identify watersheds that became rather popular consists in the flooding procedure described in Figure 3 [31], considering two sinks such as, for example, two opposite boundaries. Cells in the discretized map are ranked according to their height, which leads to a ranked surface, and are sequentially occupied, following the rank, from the lowest to the highest. Neighboring occupied cells are then connected and considered part of the same drainage basin, except when during this process their connection would promote the agglomeration of the two basins draining toward different sinks. In this case, their connection is avoided, since they should belong to different basins. The edge between...
Araújo et al. Watersheds in disordered media

FIGURE 3 | Flooding algorithm to determine the watershed. Two sinks are considered: the top row and the bottom one. Cells in the discretized map are occupied following the rank, from the lowest to the highest. Neighboring occupied cells are considered connected and part of the same basin (gray cells). The basins connected to the top row (blue) correspond to the ones draining toward the top sink. Similarly, the ones connected to the bottom row (green) drain toward the bottom sink (upper-left panel). If two neighbors of the next cell in the rank belong to basins of different sinks, the edges in contact with the cells are marked as elements of the watershed (red-thick edges). The process proceeds iteratively until all cells are flooded (bottom panels). At the end (bottom-right panel), the watershed splits the discretized map into two basins (blue and green).

them is part of the watershed and, at the end, the set of such edges forms one single watershed line that splits the landscape into two drainage basins (see Figure 3).

The flooding procedure implies visiting every cell in the discretized map. Fehr and co-workers [32] have devised a more efficient identification algorithm where only a fractal subset needs to be visited. The algorithm is based on Invasion Percolation (IP) [33] and consists in the following procedure. One initially defines two sinks (for example, the bottom and top rows of the discretized map) and considers that the (non-invaded) cell with the lowest height on the perimeter of the (already) invaded region is the next cell to be invaded. If when starting from one cell a sink is invaded, the initial cell and the entire invaded region is considered to belong to the catchment basin of that sink. Consider that the invasion is initially started from the cell in the bottom-right corner. Since this cell is part of the basin of the bottom sink, the invaded region has only one cell. One proceeds with a new invasion from the next cell upwards (cell 35 in Figure 4) and a new invasion cluster is grown until a sink is reached. Sequentially, all other cells are considered in the same way. The watershed is then identified as the line splitting the landscape into two catchment basins. The efficiency is significantly improved if, once the first cell-edge belonging to the watershed has been identified, the exploration continues from the cells in the neighborhood of this edge (see Figure 4). Thus, instead of visiting all cells, only a subset of \( N_{\exp} \) cells needs to be explored. For uncorrelated random landscapes, the size of this subset scales with the linear size of the DEM \( L \) as \( N_{\exp} \sim L^D \), with \( D = 1.8 \pm 0.1 \) [32].

4. Fractal Dimension

Breyer and Snow [35] have studied 12 basins in the United States and concluded that their watershed lines are self-similar objects [36, 37]. A self-similar structure is characterized by its fractal dimension \( d_f \), which is defined here through the scaling of the number \( M \) of edges belonging to the watershed with the linear size \( L \) of the discretized map,

\[
M \sim L^{d_f}.
\]

They obtained fractal dimensions in the range 1.05 – 1.12. Fehr and co-workers [32], with the method described in Section 3 confirmed this self-similar behavior over more than three orders of magnitude and measured the fractal dimension for watersheds in several real landscapes, as summarized in Table 1.
Table 1 | Fractal dimension of watersheds for natural landscapes obtained from satellite imagery [30] in [34].

<table>
<thead>
<tr>
<th>Landscape</th>
<th>$d_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpes</td>
<td>1.10</td>
</tr>
<tr>
<td>Europe</td>
<td>1.10</td>
</tr>
<tr>
<td>Rocky Mountains</td>
<td>1.11</td>
</tr>
<tr>
<td>Himalayas</td>
<td>1.11</td>
</tr>
<tr>
<td>Kongo</td>
<td>1.11</td>
</tr>
<tr>
<td>Andes</td>
<td>1.12</td>
</tr>
<tr>
<td>Appalachians</td>
<td>1.12</td>
</tr>
<tr>
<td>Brazil</td>
<td>1.12</td>
</tr>
<tr>
<td>Germany</td>
<td>1.14</td>
</tr>
<tr>
<td>Big Lakes</td>
<td>1.15</td>
</tr>
</tbody>
</table>

The values for the Alps and Himalayas are from Fehr et al. [32]. The error bars are of the order of 2%.

For uncorrelated artificial landscapes the watershed fractal dimension has been estimated to be $d_f = 1.2168 \pm 0.0005$ [32, 38, 39]. This value has drawn considerable attention since it was also found in several other physical models such as optimal paths in strong disorder and optimal path cracks [17–19, 40, 41], bridge percolation [15, 38, 42], and the surface of explosive percolation clusters [43, 44]. The conjecture that all these models might belong to the same universality class has opened a broad range of possible implications and applications of the properties of watersheds. As discussed in Schrenk et al. [15], the relation between most of such models can be established when they are described within the framework of ranked surfaces (see also Section 2).

5. WATERSHEDS IN THREE AND HIGHER DIMENSIONS

Up to now, we have focused on the watershed line that divides the landscape into drainage basins for water on the surface. However, in reality, water also penetrates the soil and flows underground. Thus, the concept of watersheds and ranked surfaces can be extended to three dimensions. The soil can be described as a porous medium consisting of a network of pores connected through channels. When a fluid penetrates through this medium, a threshold pressure $p_k$ can be defined for each channel $k$ such that the channel can only be invaded when $p \geq p_k$, where $p$ is the fluid pressure. In general, a channel $k$ is closed when $p_k > p$, and open otherwise. This system can be mapped into a three dimensional ranked volume, where the lattice elements are the channels and the rank is defined by the increasing order of the threshold pressure. The sequence in the rank corresponds to the order of channel openings when the fluid pressure is quasistatically raised from zero. Analogously to the ranked surfaces, one can split the space in two regions, draining toward opposite boundaries.

The watershed in three dimensions is a surface of fractal dimension $d_f = 2.487 \pm 0.003$ [39]. An example for a simple-cubic ranked volume is shown in Figure 5. In general, for lattices of size $L^d$, where $d$ is the spatial dimension, the watershed blocks connectivity from one side to the other. Thus, the watershed fractal dimension must follow $d - 1 \leq d_f \leq d$, i.e., with increasing dimension $d_f$ also increases [15].

6. IMPACT OF PERTURBATION ON WATERSHEDS

The stability of watersheds is also a subject of interest. For example, changes in the watershed might affect the sediment supply of rivers [45]. Also, the understanding of the temporal evolution of drainage networks provides valuable insight into the biodiversity between basins [46]. Geographers and geomorphologists have found that the evolution of watersheds is typically driven by local changes of the landscape [47]. These events can be triggered by various mechanisms like erosion [46, 48, 49], natural damming [8], tectonic motion [50–52], as well as volcanic activity [53]. Although rare, these local events can have a huge impact on the hydrological system [47–49, 54]. For example, it was shown that a local height change of less than two meters at a location close to the Kashabowie Provincial Park, some kilometers North of the US-Canadian border, can trigger a displacement in the watershed such that the area enclosed by the original and the new watersheds is about 3730 km² [55].

The stability of watersheds is also relevant in the political context. For example, Chile and Argentina share a common border with more than 5000 kilometers, which was the source of a long dispute between these countries [11]. A treaty established this border as being the watershed between the Atlantic and Pacific Oceans for several segments. In 1902, the Argentinian Francisco Moreno contributed significantly to elucidate the technical basis for dispute. He proved that during the quaternary glaciations, the watershed line changed. In particular, several Patagonian lakes currently draining to the Pacific Ocean were in fact originally part of the Atlantic Ocean basin. Consequently, he argued, instead of belonging to Chile they should be awarded to Argentina.

Douglas and Schmeeckle [45] have performed fifteen table top experiments to study the mechanisms of drainage rearrangements. In spite of being diverse, the mechanisms triggering the
evolution of watersheds are all modifications of the topography [8, 46, 50]. In that perspective, the effect of such events can be investigated by applying small local perturbations to natural and artificial landscapes and analyzing the changes in the watershed [34, 55]. Specifically, one starts with a discretized landscape and computes the original watershed. A local event is then induced by changing the height of a site \( k \), \( h_k \rightarrow h_k + \Delta \), where \( \Delta \) is the perturbation strength, and the new watershed is identified. The impact of perturbations on watersheds in two and three dimensions will be discussed in the following.

### 6.1. IMPACT OF PERTURBATIONS IN TWO DIMENSIONS

Fehr and co-workers [55] have normalized the perturbation strength by the height difference between the highest and lowest height of the landscape. For each landscape, they have sequentially perturbed every site of the corresponding discretized map, such that each perturbed landscape differs from the original only in one single site. The effect of those perturbations affecting the watershed has been quantified by the properties of the region enclosed by the original and the new watersheds. For that region, they have measured its area, corresponding to the number of sites, \( N_s \), in the discretized map and the distance \( R \) between its original and new outlets, defined as the points where water escapes from this region. Note that, there are only two outlets involved in this procedure: one related to the original landscape and the other to the new one. Since \( \Delta \) is strictly positive, the outlet in the new landscape corresponds always to the perturbed site \( k \).

Scale-free behavior has been found for the distribution \( P(N_s) \) of the number of enclosed sites \( N_s \), the probability distribution \( P(R) \) of the distance \( R \) between outlets, and the dependence of the average \( \langle N_s \rangle \) on \( R \) [34, 55]. Specifically,

\[
P(N_s) \sim N_s^{-\beta}, \tag{2a}
\]
\[
P(R) \sim R^{-\rho}, \tag{2b}
\]
\[
\langle N_s \rangle \sim R^{\sigma}, \tag{2c}
\]

where the measured exponents are summarized in Table 2. The power-law decay (2b) and the relation (2c) imply that a local perturbation can have a large impact on the shape of the watershed even at very large distances, hence having a non-local effect. Additionally, the analysis of the fraction of perturbed sites affecting the watershed revealed a power-law scaling with the strength \( \Delta \). This finding supports the conclusion that changes in the watershed can be even triggered by anthropological small perturbations [55].

For the region enclosed by the original and the new watersheds, an invasion percolation (IP) cluster can be obtained by imposing a pressure drop between the outlet in the new watershed and the one in the original one, always invading along the steepest descent of the entire cluster perimeter. The size distribution \( P(M_{IP}|R) \) of these clusters, for each fixed distance \( R \) between outlets, has been shown to scale as,

\[
P(M_{IP}|R) \sim M_{IP}^{-(1+\alpha^*)}, \tag{3}
\]

where \( \alpha^* \approx 1.39 \). This exponent corresponds to the one found for the size distribution of point-to-point IP-clusters [57]. In that process, invasion clusters are obtained in a random medium by invading from one point to another at a certain distance. By contrast, in the watershed case the invasion is always started from the outlet on the new watershed. This difference between starting at any point or in the outlet justifies the additional factor of unity in the scaling [Equation (3)], since the probabilities need to be rescaled by the size of the IP-cluster [55].

### 6.2. IMPACT OF PERTURBATIONS IN THREE DIMENSIONS

The impact of perturbations on watersheds has been also analyzed in three dimensions [34]. Similar to 2D, a perturbation is induced by changing the value of a site \( k \), \( h_k \rightarrow h_k + \Delta \), where \( \Delta \) is the perturbation strength. The number of sites \( N_s \) enclosed by the original and new watersheds corresponds to a volume. Power-law scaling in terms of Equations (2a)–(2c) has also been observed with the exponents summarized in Table 2. The value of \( \alpha^* \) is similar to the one found in two dimensions. According to Lee [58], the size distribution of the point-to-point IP-cluster is independent on the dimensionality of the system. Therefore, the numerical agreement between \( \alpha^* \) for different spatial dimensions supports the relation with invasion percolation.

### 7. WATERSHEDS ON LONG-RANGE CORRELATED LANDSCAPES

Results discussed heretofore were obtained on random uncorrelated landscapes. However real landscapes are characterized by spatial long-range correlated height distributions. Numerically, such distributions can be generated from fractional Brownian motion (fBm) [36, 59], using the Fourier filtering method [19, 55, 60–70]. This method allows to control the nature and the strength of correlations, which are characterized by the Hurst exponent \( H \). The uncorrelated distribution of heights is solely obtained for \( H = 0 \), and found good quantitative agreement with the exponents obtained for natural landscapes, typically with \( 0.3 < H < 0.5 \), which is the known range of Hurst exponents for real landscapes on length scales.

<table>
<thead>
<tr>
<th>Exponent</th>
<th>Uncorrelated landscapes</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d )</td>
<td>2</td>
</tr>
<tr>
<td>( d_f )</td>
<td>2.168 ± 0.0005</td>
</tr>
<tr>
<td>( \beta )</td>
<td>1.16 ± 0.03</td>
</tr>
<tr>
<td>( \rho )</td>
<td>2.21 ± 0.01</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>2</td>
</tr>
<tr>
<td>( \alpha^* )</td>
<td>1.39 ± 0.03</td>
</tr>
</tbody>
</table>

Table 2: Exponents for the impact of perturbations on watersheds, calculated for uncorrelated artificial landscapes in two and three dimensions, from [34, 39, 55, 56].
larger than 1 km (see Pastor-Satorras and Rothman 29 and references therein). They also obtained $\alpha$, $\beta$ and $\rho$ for several values of $H$, observing that both $\beta$ and $\rho$ increase with $H$, finding also good quantitative agreement. Thus, their model provides a complete quantitative description of the effects observed on natural landscapes.

8. FINAL REMARKS
Here we solely discussed cases with one watershed, but the same theoretical framework can be straightforwardly extended to tackle other space partition problems. For example, the identification of the entire set of watersheds of a landscape with multiple outlets helps identifying the catchment areas contributing to each river or reservoir [72]. A systematic study of disordered media with multiple outlets is still missing. Examples of open questions are: How does the distribution of catchment basins or the number of triplets (points where two watersheds meet) depend on the number of outlets? And, how does the statistics of perturbations change in the presence of triplets? The division of a volume into several parts is also a problem of practical interest in the extraction of resources from porous soils [73]. Studies of such systems have mainly considered uncorrelated disordered media. The role of long-range correlation there is still an open problem.

ACKNOWLEDGMENTS
We acknowledge financial support from the European Research Council (ERC) Advanced Grant 319968-FlowCCS, the Brazilian Agencies CNPq, CAPES, FUNCAP and FINEP, the FUNCAP/CNPq Pronex grant, the National Institute of Science and Technology for Complex Systems in Brazil, the Portuguese Foundation for Science and Technology (FCT) under contracts no. IF/00255/2013, PEst-OE/FIS/UI0618/2014, and EXCL/FIS-NAN/0083/2012, and the Swiss National Science Foundation under Grant No. P2EPZP-152188.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/Journal/10.3389/fphy.2015.00005/abstract

REFERENCES