RESEARCH ARTICLE
doi:10.59236/geomorphica.v2i1.51

GEOMORPHICA

Methods for quantifying spatial and temporal variation in landscape evolution model outputs

Vivian Carmello Grom *[1], **Adam Matthew Forte** [1]

[1]Department of Geology Geophysics, Louisiana State University, Baton Rouge, United States of America

Author contributions: *Conceptualization*: V.C. Grom, A.M. Forte. *Funding acquisition*: V.C. Grom, A.M. Forte. *Software*: V.C. Grom. *Original draft*: V.C. Grom. *Review & editing*: A.M. Forte.

Abstract The increasing complexity of landscape evolution models (LEMs) has expanded their application to diverse problems, yet methods for quantitatively analyzing model outputs remain underexplored. Here we propose using Shannon's entropy, Moran's I, and Geary's C to enhance the analysis of LEM outputs, offering insights into both intra- and inter-simulation dynamics and differences. These metrics enable the quantitative and visual identification of dissimilarities and variability across simulations. As case studies, we applied these metrics to three LEM experiments: steady uplift (Experiment 1), periodic alternating uplift (Experiment 2), spatially variable uplift (Experiment 3), and also an analysis of a repeat digital elevation model series (Experiment 4). Incorporating these metrics into LEMs as a comparison method provides a systematic framework for evaluating differences in information content, assessing spatial consistency, and identifying divergence among simulations. While our focus is on LEM outputs, as we demonstrate with Experiment 4, these methods are adaptable to any matrix-based data, enabling comprehensive model evaluations and facilitating improved decision-making in topographic analysis and landscape modeling studies.

Received:
February 05, 2025
Accepted:
January 11, 2026
Published:
March 23, 2026

Editor-in-Chief:
Alice Lefebvre
Associate Editors:
Fabiano Pupim
Roberto Fernández
Copy Editor:
Olivier Gourgue
Production Editor:
Leo Guerrero

Reviewers:
Anonymous

1 Introduction

Earth surface models are powerful tools for exploring how landscapes evolve over time, but comparing their outputs is far from straightforward. Subtle differences in model behavior and outcomes can be difficult to quantify, yet capturing them is essential for drawing meaningful conclusions. This challenge is especially pressing when models must be compared both against each other and against real-world observations. The difficulty increases when dealing with multidimensional models, where a unified approach is needed to evaluate not only the simulated topography but also associated variables such as soil depth, in both space and time.

Previous efforts have explored various approaches to address this need. For instance, statistical metrics such as root mean square error (RMSE) and mean absolute error (MAE) are commonly used to assess the accuracy of model predictions relative to observed data, but these metrics often fail to capture spatial variability within model outputs (Willmott and Matsuura, 2005; Chai and Draxler, 2014), and also do not factor in details of drainage network that are known to be sensitive to initial conditions (Ijjász-Vásquez et al., 1992; Tucker and Hancock, 2010). Alternatively, spatial analysis techniques, such as landscape metrics and patch-based comparisons, have been applied to quantify differences in landscape structure and patterns (Turner and Gardner, 2015). Additionally, hydrological models often incorporate goodness-of-fit measures like Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) and Kling-Gupta efficiency (Gupta et al., 2009) to evaluate performance, but these tend to focus on temporal dynamics rather than spatial heterogeneity.

While effective in identifying broad patterns, such approaches primarily focus on comparing models to real-world observations and may not be well suited for examining variation within models. They often lack the quantitative rigor needed to capture nuanced differences, particularly when comparisons involve multiple, spatially variable aspects of model outputs. Spatial statistics metrics such as Moran's I and Geary's C, commonly applied in other disciplines but not widely considered in landscape evolution modeling, offer tools for quantifying spatial correlation and variability, providing a more robust framework for comparing model outputs in a spatially explicit manner (Cliff and Ord, 1970; Upton et al., 1994; Kissling and Carl, 2008).

Complementing these measures, entropy-based metrics provide a way to assess landscape complexity, heterogeneity, and evolution. The use of entropy to quantify the complexity of spatial patterns is well established in landscape ecology, often using Boltzmann and Shannon's entropy measures (Joshi et al., 2006; Quijano and

*Corresponding author: vgrom1@lsu.edu

Lin, 2014; Vranken et al., 2015; Cushman, 2018, 2021; Gao and Li, 2019; Liu and Zhao, 2022; Mattos et al., 2022). In geomorphology, particularly in erosional landscapes, entropy has been proved useful for understanding the arrangement of rivers, where prior work suggests that river networks evolve to minimize and uniformly distribute entropy production, energy expenditure (Rinaldo et al., 1992; Rodríguez-Iturbe et al., 1992), and power expenditure (Leopold and Langbein, 1962; Cherkauer, 1973; Molnár and Ramírez, 1998; Chembolu and Dutta, 2018). However, despite these established applications, entropy has not previously been applied as a quantitative tool for evaluating landscape evolution model outputs, making its integration with spatial autocorrelation metrics a novel approach for assessing model behavior and landscape complexity.

Therefore, we provide a proof-of-concept exploration of the use, and potential utility, of using these methods for evaluating intra- and inter-earth surface process model variation. Specifically, we assess and compare model outputs with an approach that combines entropy and spatial correlation metrics to provide a comprehensive and quantitative analysis of landscape dynamics. By capturing both the complexity and spatial variability of landscapes, this approach provides a novel and efficient way to assess differences across suites of models, enabling deeper insights into their performance and the processes they simulate.

2 Background

In the section that follows, we review the underpinnings of the metrics we employ. We first consider Shannon's entropy, which is a concept from information theory that allows the measurement of the uncertainty or information content of a random variable. Subsequently, we consider Moran's I and Geary's C, which are tests for spatial autocorrelation. While the former uses the concept of standardized spatial covariance to calculate its index, the latter relies on the sum of squared differences. Ultimately, as we demonstrate, these metrics can offer valuable insights into the divergence or convergence between simulation results.

2.1 Shannon's entropy

Shannon's entropy, named after Claude Shannon, is a foundational concept in information theory that quantifies the uncertainty or randomness associated with a random variable. It measures the average amount of information or "surprise" that one would expect when observing the outcome of an event (Shannon, 1948; Karaca and Moonis, 2022). Essentially, Shannon's entropy provides a way to understand how much unpredictability exists in a set of possible outcomes.

The formula for Shannon's entropy (H) for a discrete random variable X with probability mass function $p(x)$ is given by Equation 1, where x_i represents each possible outcome of the random variable, $p(x_i)$ is the probability of occurrence of x_i. The logarithm is usually taken to the base 2, so the resulting entropy is measured in bits:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\,\log_n(p(x_i)) \tag{1}$$

Shannon developed his concept of entropy by studying the problem of transmitting information accurately over noisy communication channels. The primary challenge was how to convey messages reliably when interference or noise could distort the signals being sent. To address this, Shannon's work focused on determining the minimum amount of redundancy that must be included in a message to ensure that it could be understood correctly by the receiver, even if parts of the message were lost or altered during transmission. This redundancy helps preserve the integrity of the information, making it possible to reconstruct the original message despite any degradation that might occur.

Shannon's idea of using redundancy to combat uncertainty in communication has a parallel in Boltzmann's work in thermodynamics (Gao and Li, 2019). Boltzmann's theory, particularly his interpretation of the second law of thermodynamics, explains how physical systems tend to evolve toward states of higher entropy or disorder. In thermodynamics, a macrostate refers to the overall state of a system, characterized by macroscopic properties like temperature and pressure. Each macrostate can correspond to numerous microstates, which are the different possible configurations of the system's individual particles. According to Boltzmann, systems naturally evolve toward the macrostate that can be derived by the largest number of microstates because this state is statistically the most likely.

The connection between Shannon's information theory and Boltzmann's thermodynamics lies in their shared focus on entropy, a measure of uncertainty or disorder. In both fields, entropy reflects the tendency of systems – whether they are messages traveling through a noisy channel or particles in a physical system – to move toward states of greater uncertainty or randomness. This convergence of concepts provides a powerful framework for understanding the complexity and variability in landscape patterns. Just as information theory uses entropy to measure uncertainty in communication, landscape ecology and modeling can use entropy to quantify the unpredictability and complexity of spatial patterns (MacArthur, 1955; Margalef, 1958; Ricotta, 2000; Ulanowicz, 2001; Vranken et al., 2015). Both disciplines offer insights into how systems, whether informational or physical, evolve toward states of higher entropy, providing tools for analyzing the behavior of complex systems.

While entropy is often associated with chaos, in this context it measures the uncertainty or vagueness in a system. For example, considering the scenario of tossing a fair coin, if we toss the coin once, there are two equally likely outcomes: heads or tails. The entropy here is at its maximum because each outcome is equally probable, meaning there is a high level of uncertainty or "information content" in predicting the result of a single toss. However, if we toss a weighted coin that always lands on heads, the entropy is zero because there is no uncertainty in the outcome.

2.2 Moran's I

Developed by Patrick Alfred Pierce Moran to understand stochastic phenomena distributed in two or more dimensions, like the distribution of soil fertility over a field, or the relations between the velocities at different points in a turbulent fluid, Moran's Index is one of the oldest statistics used to explore spatial autocorrelation, using both feature locations and feature values simultaneously (Moran, 1948; Upton et al., 1994; Mathur, 2015; Boots and Getls, 2020). The index is produced by standardizing the spatial autocovariance by the variance of the data. In Equation 2, $\bar{x}$ is the mean of $x_{(i,j)}$, $w_{(i,j)}$, is the spatial weight between features i and j, N is the total number of features, and S_0 is the aggregate of all spatial weights.

$$I = \frac{N}{S_0} * \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{(i,j)}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2}; \quad j \neq i \tag{2}$$

The resulting Moran's I statistic ranges from -1 to 1. Negative values suggest a negative spatial correlation, indicating that neighboring features tend to have dissimilar values. Values near 0 indicate no spatial correlation, suggesting a random spatial distribution. Positive values indicate a positive spatial correlation, implying that neighboring features tend to have similar values.

Moran's I finds extensive application in various fields such as geography (Chung and Noguchi, 1998; Dejene and Muleta, 2020), urban planning (Fan and Myint, 2014), ecology (Koenig and Knops, 1998; Kuuluvainen et al., 1998; Sokal et al., 1998; Mathur, 2015), and epidemiology (Walter, 1993; Shrestha et al., 2022). It helps researchers and practitioners understand spatial patterns, identify clusters, and make informed decisions about resource allocation, policy formulation, and spatial planning. In summary, Moran's I provide a valuable tool for exploring spatial relationships and patterns within dataset. The p-value in the context of Moran's I assesses the significance of the observed spatial autocorrelation. A low p-value indicates that the observed pattern is unlikely to have occurred by chance.

2.3 Geary's C

Geary's C is a statistic used to measure spatial autocorrelation by quantifying the covariation between pairs of data values. Developed by Frank Brian Geary, it utilizes the sum of squared differences between pairs of data values to assess spatial relationships(Geary, 1954).

The calculation of Geary's Contiguity ratio "c" involves comparing the observed sum of squared differences to the expected sum under the assumption of spatial randomness. The calculation is similar to Moran's I, but instead of being calculated from the deviations from the mean, here the index comes from the values from the variables (Mathur, 2015). In Equation 3, $\bar{x}$ is the mean of x_i, $w_{(i,j)}$ is the spatial weight between feature i and j, N is the total number of features, and S_0 is the aggregate of all spatial weights.

$$C = \frac{(N-1)}{2S_0} * \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{(i,j)}(x_i - x_j)^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2}; \quad j \neq i \tag{3}$$

The expected value for Geary's C is between 0 and 2. Values lower than 1 indicate positive spatial autocorrelation, suggesting that nearby data values tend to be more similar than expected under spatial randomness. Conversely, values larger than 1 indicate negative spatial autocorrelation, implying that nearby data values tend to be more dissimilar than expected. Values close to 1 indicate no autocorrelation, or randomness. Compared to Moran's I, then, the interpretation is basically the opposite (Mathur, 2015).

Geary's C is commonly employed in spatial analysis, geography, and ecology (Anselin, 2019; Lin, 2023) to assess spatial patterns, similarly to Moran's I. In summary, Geary's C serves as a valuable tool for quantifying spatial autocorrelation, providing insights into the degree to which neighboring data values exhibit similarity or dissimilarity within a spatial context.

3 Implementation

In landscape evolution modeling, we propose here the use of Shannon's entropy to identify which areas of the model output have the highest variability or contrast, providing a measure of the "information" or "surprise" associated with different landscape configurations. This is useful because it allows researchers to pinpoint regions where models differ significantly, offering insights into how different processes might be influencing landscape changes. Understanding these differences is crucial for improving model accuracy and for gaining a more comprehensive understanding of the natural processes shaping our environment.

The key idea is that events with lower probabilities contribute more to the entropy, and events with higher probabilities contribute less. If the probability distribution is more spread out, indicating higher uncertainty, the entropy will be higher. Conversely, if the probability distribution is more concentrated, indicating lower uncertainty, the entropy will be lower. In our use case, the term "probability" is used as a conceptual framework to describe the relative likelihood or frequency of particular landscape patterns or states, rather than as a formal calculation of the probability of specific events occurring. Unlike traditional probability models that predict the occurrence of discrete events, our use of probability focuses on characterizing the distribution of patterns across a landscape and assessing their relative prevalence.

Importantly, this approach does not imply that we are formally modeling the probability of specific landscape events (e.g., landslides, sediment transport, etc.) or their outcomes. Specifically, we are not embedding any a priori assumptions about process in our calculations of entropy, which would be necessary if we wished to formally estimate probabilities of "events" within the model outputs. Instead, we are leveraging the probabilistic perspective to quantify the information content and spatial organization within the landscape. This interpretation is closely tied to the statistical mechanics and entropy-based frameworks that provide a way to measure system diversity or predictability, rather than event-specific probabilities. By adopting this approach, we aim to provide a deeper understanding of spatial patterns without conflating it with event prediction or probability theory in its strictest sense.

Thus, to calculate the entropy for each element in the grid ($E_{i,j}$) for a set of experiments, first it is necessary to calculate the Maximum Modular Difference (which we refer here as Modular R) of all experiments.

The code takes a list of n matrices, each with dimensions matching the grid, and initializes a matrix R with zeros to store the maximum differences. For each element $R_{i,j}$ in the grid, we compute the maximum absolute difference between consecutive matrices. This value serves as a normalization factor that ensures the differences between models are scaled consistently, allowing us to interpret the results within an information-theoretic framework.

To treat all change symmetrically, a wraparound comparison is included: the final matrix is also compared to the first one, effectively making the sequence cyclic. This approach ensures that the full range of spatial variation is captured in the entropy calculation. The final value of k, used for normalization, is defined as the maximum value in matrix R, as shown in Equation 4.

$$R_{i,j} = \max|matrices[t][i,j] - matrices[t+1][i,j]| \quad (4)$$
$$\text{for } t \in \{1, 2, \ldots, n\},$$
$$k = \max(R).$$

If there are only two matrices, the code calculates the probability matrices through Equation 5.

$$P_1 = \frac{|matrices[0] - matrices[1]|}{k\,n} \quad (5)$$
$$P_2 = 1 - P_1$$

For more than two matrices, the code calculates the probability for each pair of consecutive matrices using Equation 6.

$$P_t = \frac{|matrices[t] - matrices[t+1]|}{k\,n} \quad (6)$$
$$\text{for } t \in \{1, 2, \ldots, n\}$$

In the context of calculating Shannon's entropy, having a probability of zero can cause mathematical issues because the logarithm of zero is undefined and results in negative infinity. To avoid these issues, the code replaces any zero probabilities with a very small number, specifically 10^{-10}.

The code then calculates the entropy for each element in the grid using the Shannon's entropy formula with a logarithm base 10. We denote P_t as the matrix of probabilities corresponding to each pair of consecutive matrices. The total entropy matrix is computed as:

$$E = -\sum_{t=1}^{n} P_t * \frac{\log(P_t)}{\log(n)} \quad (7)$$

Applying entropy calculations across an entire landscape can sometimes obscure subtle contrasts, particularly in regions dominated by hillslope processes. To address this issue, it is useful to differentiate between hillslopes and streams, which involves identifying areas dominated by fluvial processes (streams) versus those governed by diffusive processes (hillslopes). This differentiation is important because despite linked, processes acting on hillslopes and those acting within fluvial networks are decidedly different (Montgomery and Dietrich, 1988, 1989) both in terms of behavior, specifically diffusive vs advective processes (Howard, 1997; Perron et al., 2008b, 2009), and scale (Montgomery and Dietrich, 1992; Perron et al., 2008a).

The connection between hillslopes and streams depends on several factors, including catchment wetness, soil properties, and the surface and bedrock topography. Hillslope contributions are often delayed until a storage threshold is exceeded, and spatial variability in throughflow and preferential flow can further obscure these contributions(Spence and Woo, 2003; Buttle et al., 2004; Tromp-van Meerveld et al., 2007; Detty and McGuire, 2010).

The method used here for differentiating streams and hillslopes follow the standard practice of using a threshold drainage area to identify streams (Montgomery and Foufoula-Georgiou, 1993), but other methods present in the literature include applying slope-area analysis to distinguish large contributing areas with lower slopes, typically associated with streams, and small contributing areas with higher slopes characteristic of hillslopes (Montgomery and Foufoula-Georgiou, 1993; Sun et al., 1994; Tucker and Bras, 1998), or employing a curvature-based approach. Threshold slopes can help differentiate hillslopes, which typically exhibit a convex curvature, from streams, which are generally concave (Montgomery, 2001).

To address this, the code includes functions that differentiate between hillslopes and streams, allowing for a more nuanced understanding of spatial variability and connectivity within the landscape. The most straightforward method classifies the landscape by applying user-defined thresholds for drainage area and slope. However, there are also more advanced options, such as a “buffered” version and a “scaled buffered” version. The ”buffered” method incorporates adjacent cells around the target cell to account for local variability and ensure that classifications are not overly sensitive to noise or minor fluctuations in the data. The ”scaled buffered” version extends this concept by adjusting the influence of surrounding cells based on a weighting scheme that accounts for their distance or contributing area. This scaling ensures that closer or more hydrologically significant cells have a greater influence on the classification. Importantly, whether to analyze the entire landscape or to explicitly separate hillslopes from channels is left to the user. Similarly, our approach could easily be adapted to consider an alternative definition of hillslopes vs channels. This flexibility allows the methods to be adapted to different use cases, with separation of hillslopes and channels potentially providing more nuanced insights depending on the analysis.

Regarding the spatial autocorrelation calculation, more specifically the implementation of Moran's I and Geary's C., we used the `libpysal` and `esda` (Rey and Anselin, 2007) libraries to calculate the metrics for a grid over multiple timesteps. Specifically, the `lat2W` function from `libpysal.weights` is used to create a spatial weights matrix based on the dimensions of the input grid, which defines the spatial relationships between grid cells. Moran's I, computed using the `Moran` class from `esda.moran`, measures global spatial autocorrelation by evaluating the similarity of nearby values relative to the overall variance in the dataset. Similarly, Geary's C, computed using the `Geary` class from `esda.geary`, quantifies local spatial autocorrelation by analyzing squared differences between neighboring values.

4 Case Studies

To explore the behavior of these metrics in potential use cases and showcase the utility's effectiveness, we conduct three experiments using Landlab. In all experiments, we selected components that allowed for a more complex representation of landscape processes than a simple stream power model. Specifically, we included both topography and soil depth as metrics to assess how the utility performed with different types of data. To achieve this, we incorporated the following components: "ExponentialWeatherer" (Barnhart et al., 2019) for simulating chemical weathering and soil production, "DepthDependentTaylorDiffuser" (Hobley et al., 2017) for modeling depth-dependent sediment transport, "PriorityFlowRouter" (Barnes, 2017) to determine the flow routing, and "SpaceLargeScaleEroder" (Shobe et al., 2017) for stream power erosion. This combination of components enabled us to simulate interactions between topographic elevation and soil depth over time. Supplementary Table S1 provides a detailed summary of the model parameters used in the modeling experiments. Experiment 1 serves as a simplistic base case, where the model was run for 250 kyr using consistent parameters throughout, maintaining a steady uplift rate. Experiment 2 also spanned 250 kyr and simulated cyclic tectonic activity or base-level changes by alternating the direction of uplift. The uplift rate was set at 1 mm/yr (0.01 m per timestep), and the uplift direction switched every 50 kyr. Specifically, the experiment began with a positive uplift phase (+1 mm/yr) for the first 50 kyr, followed by a negative uplift phase (–1 mm/yr) for the next 50 kyr, and so on, completing five full cycles of alternating uplift and subsidence over the simulation period. While not wholly realistic, this alternation aimed to mimic a simplified and idealized form of the dynamic and episodic nature of tectonic processes and explore the metrics' sensitivity to these types of changes. Fig 1 presents the topographic state and soil depth at each 50 kyr interval for Experiments 1 and 2, which serve as inputs for our subsequent entropy and autocorrelation calculations.

Spatial patterns of variability differ between entropy and Modular R, particularly when comparing Experiments 1 and 2 (Fig. 2). In Experiment 1, the entropy and Modular R maps highlight broadly similar regions of high contrast. In Experiment 2, however, the entropy map reveals a broader zone of elevated variability surrounding the channels, appearing as widespread bright yellow, while the Modular R map shows more localized variability, particularly at the topographic break between the highland and lowland regions.

The code includes functionality to track the behavior of specific pixels, enabling a detailed examination of localized landscape changes over time. This pixel-tracking approach, previously used to measure surface ground motion (Intrieri et al., 2018; Lacroix et al., 2023; Carlà et al., 2019; Handwerger et al., 2019, 2025), is here applied to monitor the pixel with the highest value for each metric independently, specifically, the first pixel in the grid where the maximum value occurs. If multiple pixels share this maximum value, the one closest to the outlet is selected, following Landlab's node ordering from the lower left corner and increasing from left to right across each row, with subsequent rows numbered from bottom to top. This approach allows us to compare areas experiencing the greatest variability across experiments and to assess whether the summary maps capture consistent or divergent landscape dynamics. In Experiment 1 the maximum-value pixels for entropy and modular difference coincide, but in Experiment 2 they differ. For soil depth, while overall spatial patterns are generally consistent across both metrics, the location of the highest-entropy pixel varies again in Experiment 2, underscoring subtle shifts in where variability is concentrated.

By applying a stream mask, we can observe relative differences in the scale of variation between hillslope and fluvial domains, highlighting areas of higher entropy in both. Fig 3 shows that distinguishing between these areas makes it clearer how Experiment 2 exhibits the highest entropy in the intermediate section of the drainage basin, indicating stronger property contrasts driven by cyclic uplift changes. In this experiment, the Buffered Stream Mask Scaled also reveals elevated entropy in the hillslope areas. In contrast, Experiment 1 displays higher entropy primarily in the distal parts of the drainage basin.

We take advantage of the code's pixel tracking capability by applying a mask to highlight the pixels with the highest entropy values within both stream and hillslope areas. In Fig 4 we then plot entropy values for pairs of matrices alongside their corresponding modular differences, enabling a side-by-side assessment of the landscape from both an entropy-based and a modular perspective. The temporal evolution of entropy and modular difference varies between Experiments 1 and 2, showing different responses of stream and hillslope domains to the uplift regimes. In Experiment 1, both the stream and hillslope pixels maintain high entropy values over time, while the modular difference gradually decreases. In contrast, Experiment 2 exhibits a fluctuating pattern in both entropy and modular difference, with alternating phases of increase and decrease, following the uplift patterns. Despite exhibiting similar overall trends, in Experiment 2 the stream pixel shows

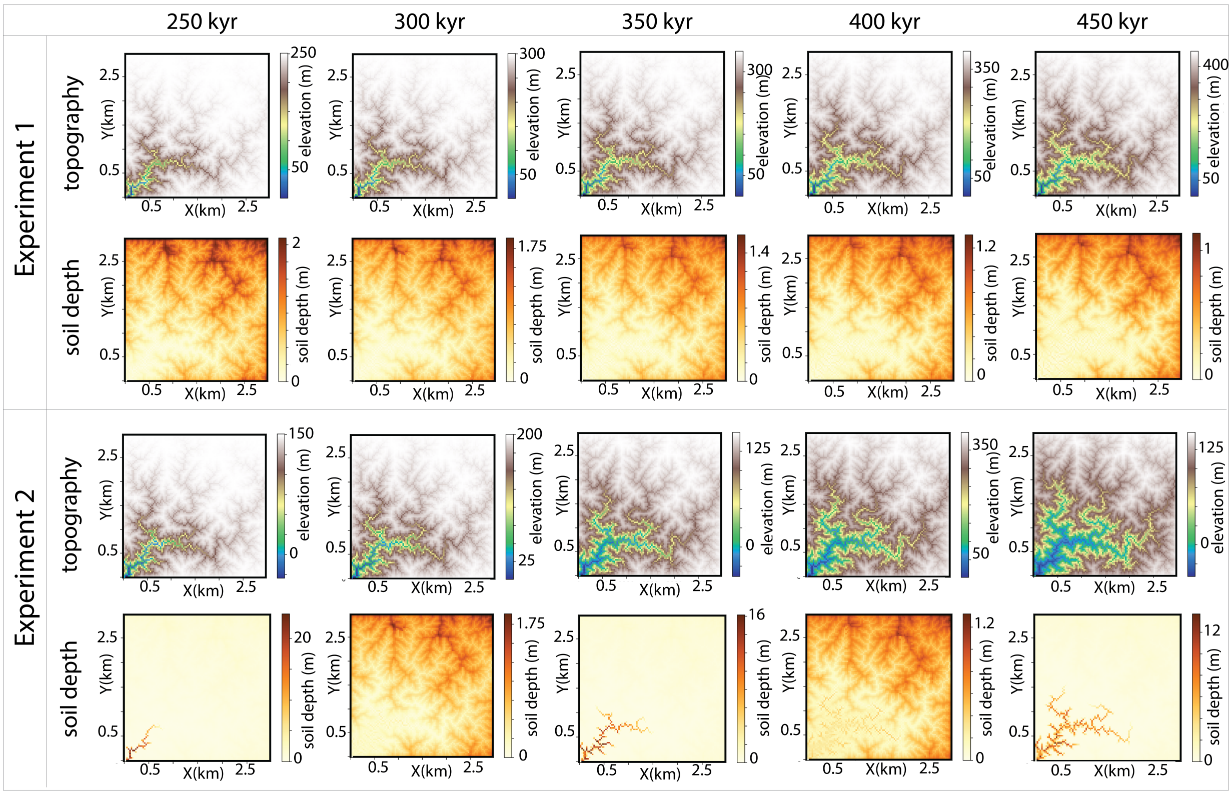

Figure 1 Elevation (top rows) and soil depth (bottom rows) for each output timestep in Experiment 1 (top two rows) and Experiment 2 (bottom two rows). Note that the minimum and maximum of the color scales change between each timestep and model to accentuate the degree of spatial variability of each variable across model runs.

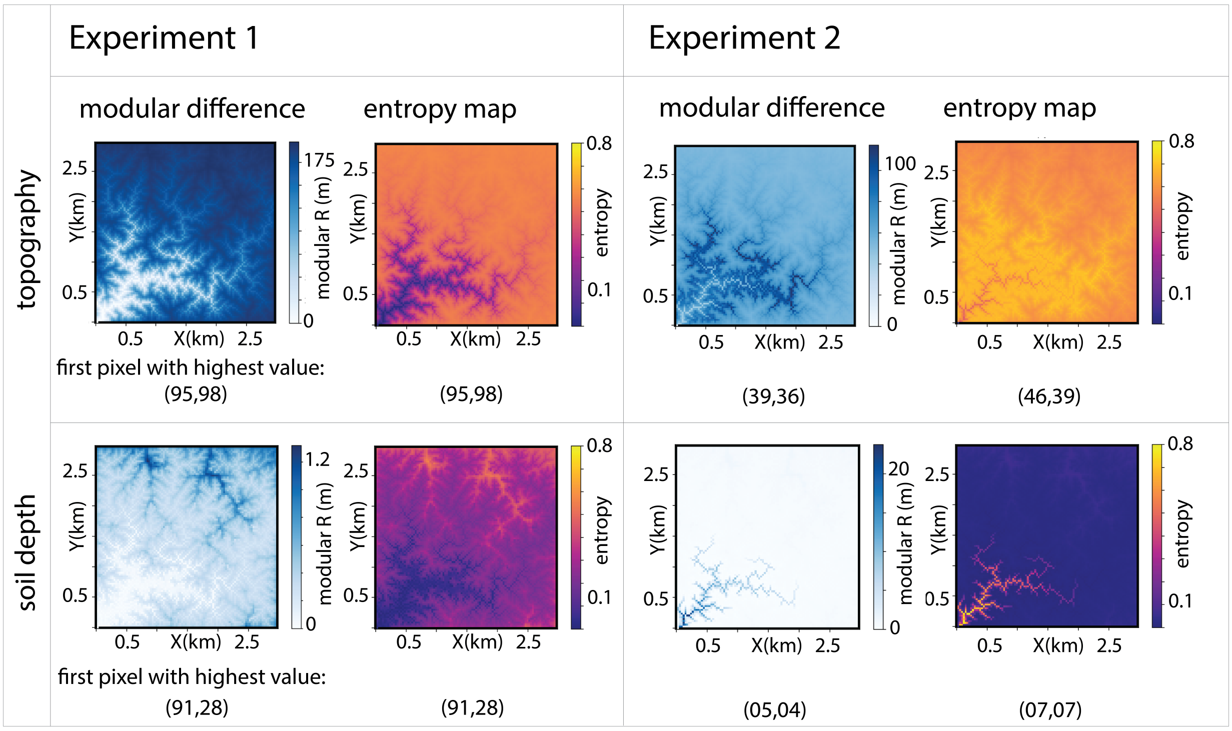

Figure 2 Modular R and entropy map for topography and soil depth analysis of Experiments 1 and 2.

slightly higher entropy for the first and last matrix pairs, while the hillslope pixel shows higher values for the second and third pairs. This pattern aligns with the broader spatial results in Fig. 3, where hillslope areas display higher entropy overall. The maximum entropy values in these pairwise plots can reach 1 because the entropy is computed between individual matrix pairs, allowing some pixels to exhibit the highest possible contrast. In contrast, the aggregated entropy results shown in (Fig. 3) represent the cumulative computation across all matrix pairs, which smooths local variability and reduces the overall maximum values below 1.

Moran's I and Geary's C capture trends in spatial organization between Experiments 1 and 2, reflecting the

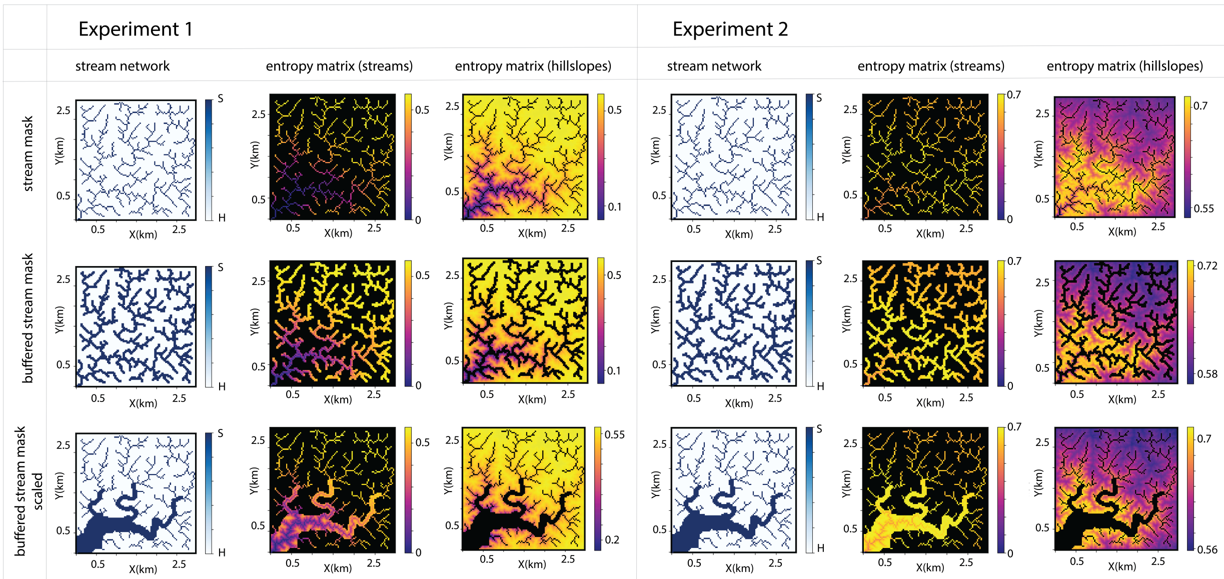

Figure 3 Using stream/hillslope masks to understand entropy variation for topographic analysis of Experiments 1 and 2. The "stream mask" row classifies streams based on a drainage area threshold, where grid nodes with drainage areas larger than 20 m² are considered part of the stream network. The "buffered stream mask" row applies a filter size of 2, extending (buffering) the stream network by two grid cells in all directions. The "buffered stream mask scaled" uses a scaling factor of 0.007, which adjusts the buffer size dynamically based on the drainage area. This means that larger drainage areas result in larger buffers around streams.

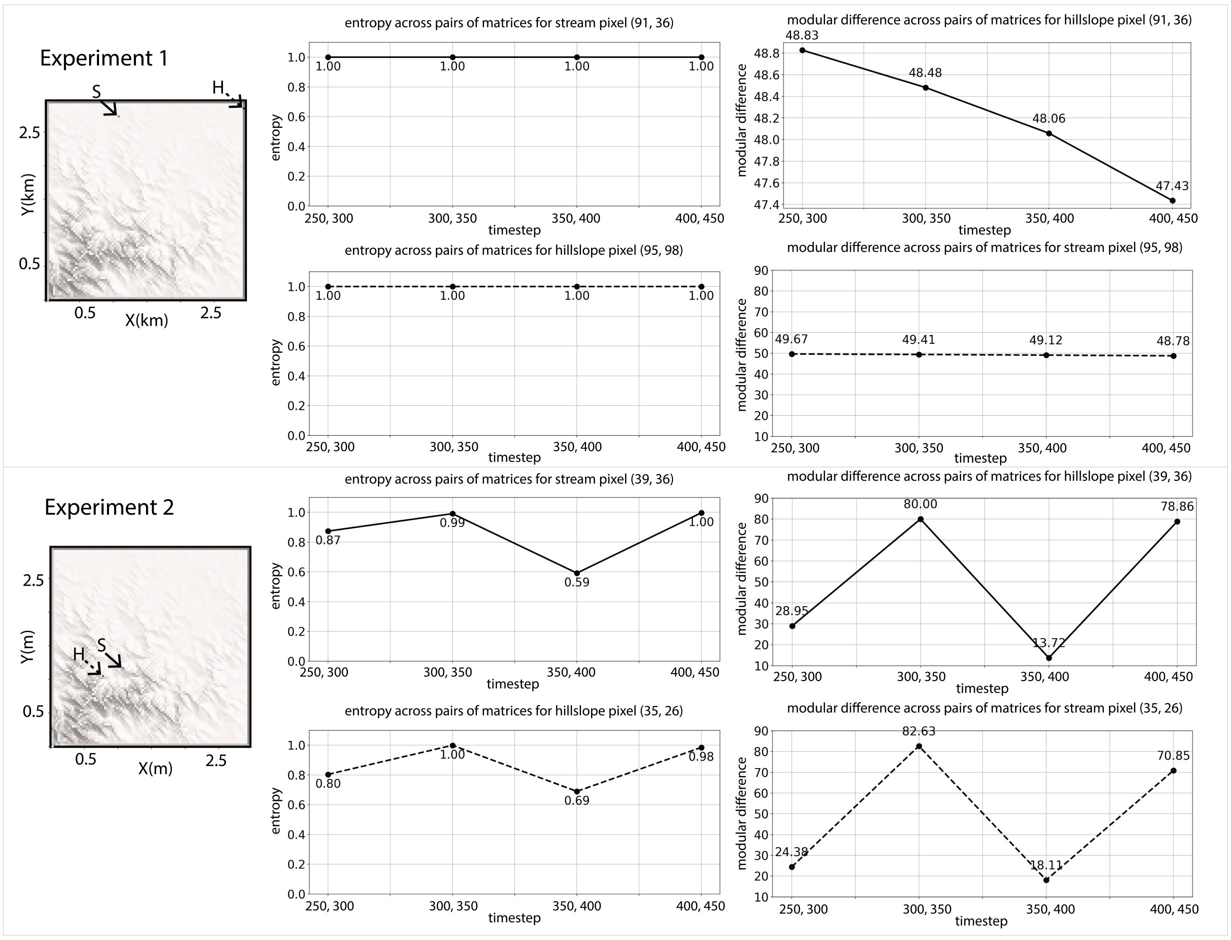

Figure 4 On the left, a hillshade map illustrates the topography, overlaid with the locations of selected pixels in both the drainage and hillslope areas. The dark margins are a visual artifact of the hillshade rendering and boundary conditions and do not represent meaningful data. On the right, we show the temporal evolution of local Shannon's entropy and modular difference for two representative pixels – one from the stream region (S) and one from the hillslope (H), across pairs of experiments.

effects of steady versus alternating uplift on landscape structure. In Experiment 1, where there is steady uplift, the values of Moran's I gradually increase towards 1. This indicates a growing positive spatial autocorrelation, meaning that similar values are increasingly clustered together as the landscape evolves. Concurrently, the values of Geary's C slowly approach 0, signifying a decrease in local variability, with neighboring pixels becoming more similar over time. In contrast, Experiment 2, which features alternating uplift conditions, exhibits a different pattern. The values of Moran's I and Geary's C in this experiment alternate between increasing and decreasing, reflecting the periodic changes in uplift and their impact on spatial relationships within the landscape, but the scale of this variation is relatively small.

To investigate how changes in topographic forcing influence drainage patterns and the behavior of Geary's C, Moran's I, and Shannon's entropy metrics, we now apply a differing suite of spatially varying uplift rate gradients. In detail, Experiment 3 consists of 10 discrete simulations with different spatial and temporal variations in uplift rate. For all the Experiment 3 runs, the uplift rate was defined as a linear gradient that varied across the domain, transitioning from west to east. In some cases, the gradient was reversed mid-simulation, as in Experiments 3a to 3d, creating dynamic scenarios where the regions experiencing the highest uplift shifted spatially. This approach allowed us to observe how drainage reorganization and topographic complexity evolved in response to the changing forcing and in turn were reflected in the metrics. Experiments 3e to 3h maintained a constant uplift gradient throughout the entire simulation, enabling us to study the effects of stable topographic forcing on landscape development. For Experiments 3i and 3j, the uplift gradient changed magnitude during the simulation but without reversing, introducing another variation in how topographic forcing was applied.

Although p-values are not shown for all subsequent figs, all Moran's I and Geary's C values were associated with extremely low p-values (<0.001), confirming that the observed spatial patterns are statistically significant and reflect genuine spatial autocorrelation rather than random variation. This experimental framework provided a comprehensive basis for evaluating how metrics that are sensitive to spatial autocorrelation, such as Moran's I, and dissimilarity, such as Geary's C, respond to evolving drainage patterns and increasing topographic complexity. By combining these metrics with Shannon's entropy, we assessed how varying conditions in uplift gradients influence changes in landscape patterns, offering insights into the dynamic and static responses of spatial metrics to topographic forcing. All experiments evolved from a steady landscape evolved for 200k years with a 0.001 m/yr uplift, and their run time, uplift type and uplift range are described in the Supplementary Table S2.

The resulting metrics shown in (Fig. 6) reveal reorganization of the final drainage patterns, particularly in the distal areas (highlighted in grey) of Experiments 3a through d, which exhibit a more pronounced east–west orientation with higher uplift contrast. In the topographic entropy analysis, increasing uplift range contrast appears to smooth out local variations, leading to lower spatial contrast across the landscape. Examining Moran's I and Geary's C over time reveals that as contrast increases, the maximum value increases as well (Fig. 9). After the first run, there is a subtle inversion in the trend, which is more pronounced with higher uplift contrast. For the soil patterns, we observe an overall reduction in similarity. Entropy exhibits a more contrasting distribution in Experiment 3a but becomes more localized in the subsequent experiments. Moran's I and Geary's C consistently decrease across all experiments, although the trend becomes less smooth from A to D. Furthermore, the rate of change is more pronounced with higher uplift contrasts, while lower contrasts result in a slower rate of decrease.

The uplift gradients were applied in opposite directions in Experiments 3e to 3h, with higher uplift rates in the west during Experiments 3e and 3f, and in the east during Experiments 3g and 3h (Fig. 7). The drainage network exhibits more pronounced reorganization in 3e and 3f, though the values of Moran's I and Geary's C remains lower than 3g and 3h (Fig. 9). Shannon's entropy consistently aligns with regions of more pronounced uplift, being higher in the western portion during 3e and 3f and shifting to the eastern portion in 3g and 3h. This underscores the spatial dependence of landscape complexity on the direction and magnitude of uplift forcing. The soil analysis shows the same autocorrelation patten as seen before, with a steep decrease in Moran's I and Geary's C with higher uplift contrast and if the gradient is higher in the east (3g and 3h). Contrastingly, there were more concentrated higher values entropy values with increasing uplift in the west (3f), while when the contrast is observed in the east, there were more dispersed entropy values.

The uplift gradient contrast was increased during the second run of Experiments 3i and 3j, in which the resulting metrics are shown in Fig. 8. Shannon's entropy aligns with the region experiencing the highest contrast, and drainage reorganization is more pronounced in Experiment 3j, consistent with the trends observed in previous experiments. For both experiments, Moran's I and Geary's C and soil distribution patterns remain very similar to the ones observed in Experiment 3g (for 3i) and Experiment 3e (for 3j).

To evaluate whether the proposed methodology can be applied to other types of gridded data, in Experiment 4 we tested it on three LiDAR-derived digital elevation models (DEMs) covering the Mud Creek landslide site in the California Coast Ranges (Fig. 10). In 2015, the region experienced a historic drought followed by an exceptionally wet year. Although the area had been characterized by long-term, slow-moving slope deformation, these extreme climatic conditions triggered a rapid transition to failure (Handwerger et al., 2019). The DEMs correspond to surveys from 2010 (Data, 2010), 2016 (Survey, 2021b), and 2018 (Survey, 2021a), spanning the period before and after the 2017 failure. The study area was cropped to include the landslide region while excluding problematic DEM pixels, such as those with extreme outlier or zero values. Quality control of the data are essential prior to metric computation, even though the entropy map can help identify anomalous pixels.

As shown in Fig. 10b, the total entropy map lightly delineates the area that later failed in 2018. In Fig. 10c,

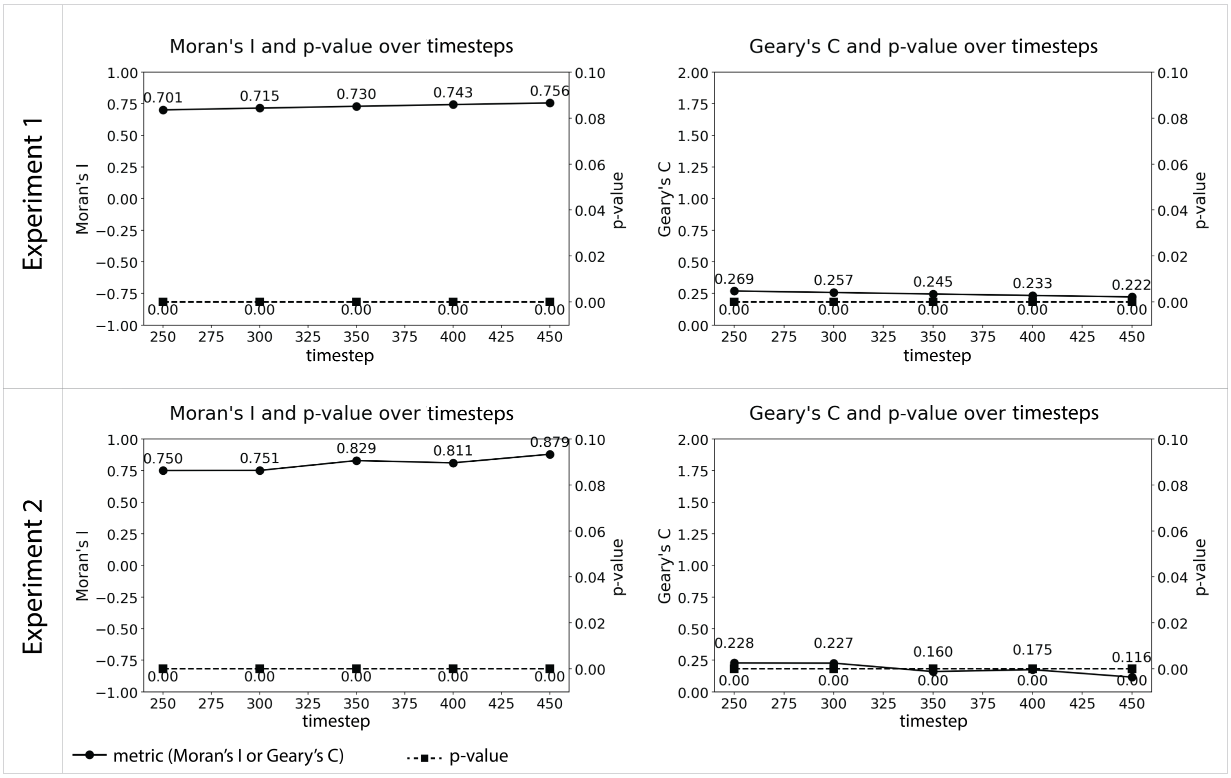

Figure 5 Moran's I and Geary's C values for each timestep for Experiments 1 and 2.

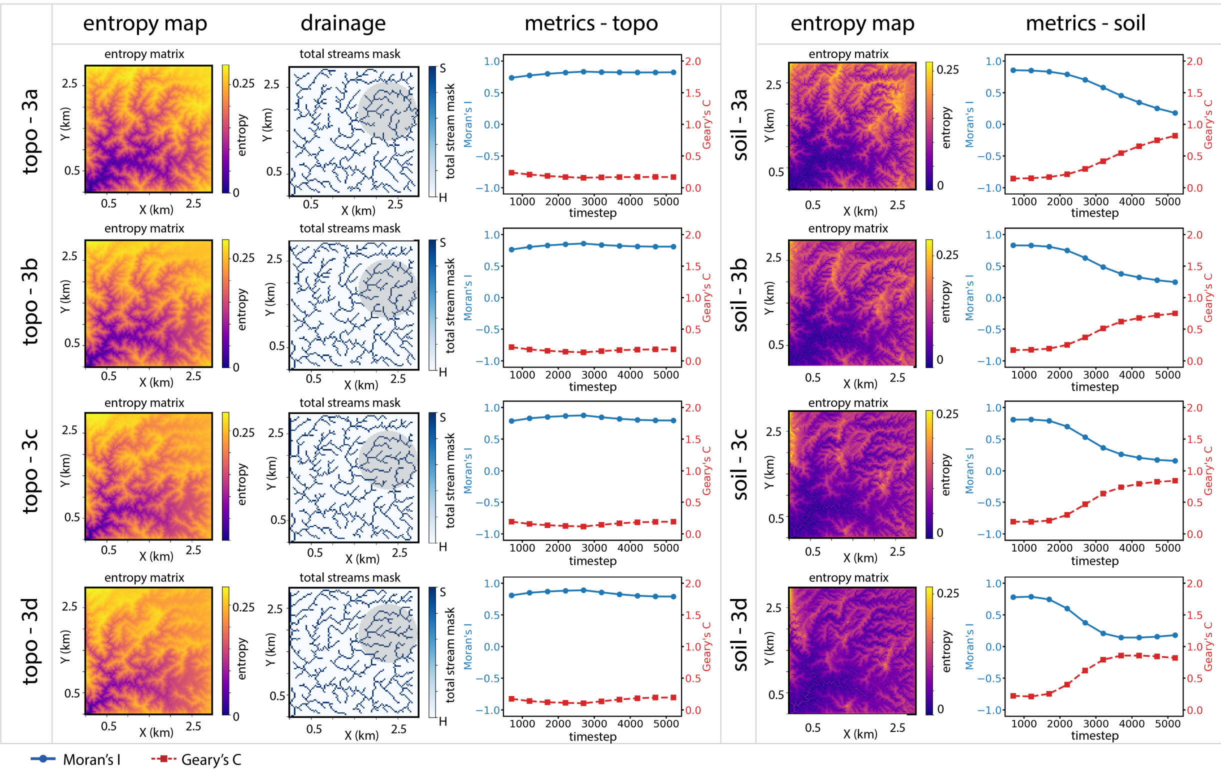

Figure 6 Entropy, drainage, Moran's I, and Geary's C for Experiment 3a–d. The metrics were calculated across all timesteps. In the drainage maps, the section highlighted in grey indicates an area selected for comparison of drainage patterns among the different experiments.

the entropy time series reveals increasing values along NW–SE transverse cracks parallel to the coast — features likely reflecting ongoing earth creep between 2010 and 2016. These same structures exhibit higher total entropy after the failure. Since the scale of analysis can obscure subtle variations — because spatial autocorrelation generally increases at finer resolutions (Turner et al., 1989; Qi and Wu, 1996) — we computed the autocorrelation metrics at multiple resolutions, a process here referred to as downsampling. Downsampling reduces the DEM resolution by aggregating pixels, effectively increasing pixel size, which allowed us to assess how spatial autocorrelation changes across coarser representations of the landscape.

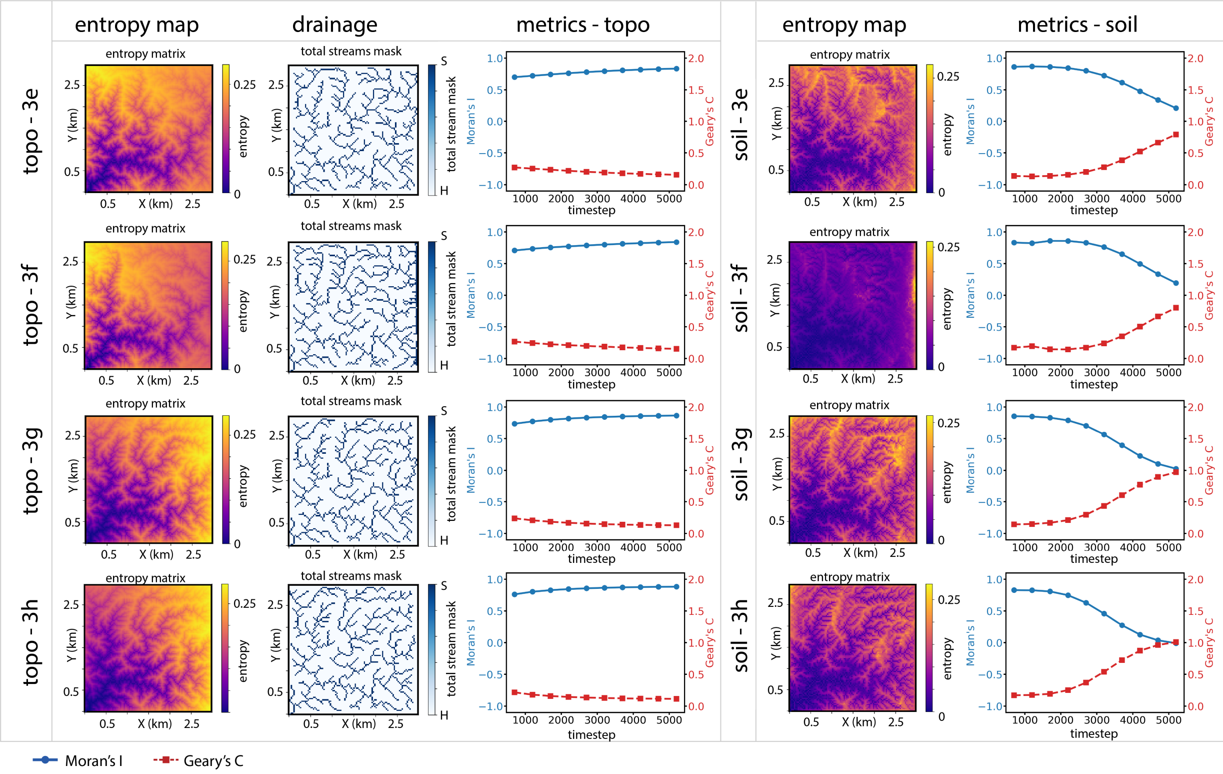

Figure 7 Entropy, drainage, and autocorrelation metrics for Experiments 3e–3h.

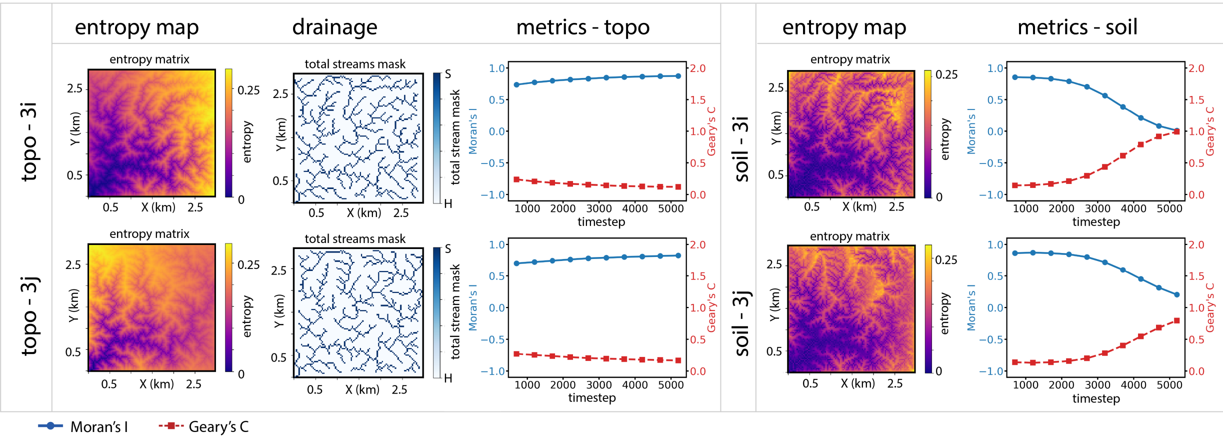

Figure 8 Entropy, drainage, Moran's I, and Geary's C for Experiment 3i–j.

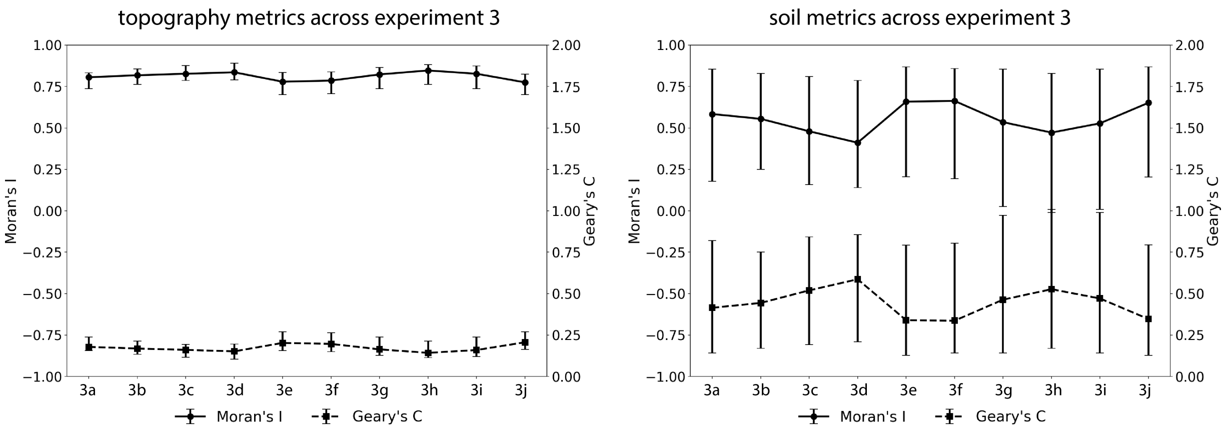

Figure 9 Topography and Soil metrics (Moran's I and Geary's C) for Experiment 3. The bars show the maximum and minimum values across the experiments from 3a to 3j, and the symbol marks the mean value.

Fig. 10d presents Moran's I and Geary's C for the different downsampling factors. At high resolution (1m), autocorrelation values are consistently high, but they decrease as pixel size increases. Notably, at a 30m resolution, the 2018 landscape shows higher spatial autocorrelation than the 2010 and 2016 landscapes.

5 Discussion

5.1 Interpreting entropy results

5.1.1 Comparing different variables and separating domains

An advantage of using entropy to analyze landscape changes is that it allows us to compare different variables within the landscape evolution model outputs that have different scales or magnitudes of variation. For instance, as shown in Fig. 2, entropy allows us to assess variations in both topography and soil depth despite their different magnitude variation. In our experiments, when evaluating just the modular difference, we compare approximately 150 meters of change in topography to only 20 meters in soil depth across both experiments. Traditional difference-based metrics may obscure such relationships because their outcomes are strongly scale-dependent, much like roughness parameters that vary with the scale of measurement (Sanner et al., 2022). This discrepancy in scale could obscure our understanding of landscape dynamics. However, entropy provides a normalized measure that accounts for not just the scale, but the variations within each feature, making it easier to identify patterns and contrasts without being biased by their absolute magnitudes.

Similarly, by separating hillslope domains from streams in the topographic analysis for Experiments 1 and 2, as seen in Fig. 3, we can also assess how the entropy contrasts and nuances in each are expressed. Although delineating the precise transition between hillslopes and streams remains challenging (Bracken and Croke, 2007; Ali and Roy, 2009; Hopp and McDonnell, 2009), a well-informed understanding of the landscape allows users to define drainage area thresholds or employ buffered and scaled-buffered masks to refine this separation, adjusting how streams and hillslopes are classified based on their specific characteristics. The different masks (stream, buffered, scaled) highlight varying spatial domains, yet the consistently high entropy in hillslope areas across all mask options suggests that the exact mask definition is less critical than the broader decision to distinguish hillslopes from streams. This pattern indicates that hillslope processes are a primary driver of sediment heterogeneity in both experiments—particularly in Experiment 2, where stronger contrasts suggest that cyclic uplift amplifies hillslope erosion variability before sediment reaches the stream network.

In Experiment 1, the buffering and scaling around streams may amplify the localized variability of sediment contributions from hillslopes, even if the overall system has less dramatic contrasts than Experiment 2. This could happen because lower uplift or more uniform erosion rates in Experiment 1 mean fluvial transport dominates downstream entropy patterns (hence higher entropy in distal basin areas), but hillslope contributions still have localized variability that the buffered mask highlights.

In Experiment 2, where uplift varies cyclically, the modular difference method can sometimes "mask" variability in topography and soil depth (Fig. 2). This is because it measures net change between two states, and in the case of alternating uplift, opposing changes can cancel out, resulting in values near zero—even in areas undergoing significant change over time. As a result, important spatial patterns may appear subdued. In contrast, entropy captures the degree of disordered or heterogeneous outcomes (Ricotta, 2000; Vranken et al., 2015), independent of the direction of change. This allows it to highlight broader zones of variability that may be less apparent in the modular difference results. Thus, while the modular difference map reveals more spatial detail in localized contrasts, particularly topographic transitions, the entropy map complements this by capturing the full extent of variability across the landscape. This demonstrates the advantage of using both methods side by side: while the modular difference can provide insights into broad patterns of change, entropy offers a more detailed view of variability and providing a more complete picture of landscape dynamics. Additionally, to further enhance interpretation, analyzing the stream network and hillslopes separately can provide more insights into how different landscape processes influence variability.

5.1.2 Pixel tracking

As illustrated in Fig. 4, distinct patterns emerge in Experiments 1 and 2. For Experiment 1, which involved steady uplift, it is expected that the pixel with the highest value consistently remains the highest over time, reaching a maximum due to the uniform uplift. This stable pattern reflects predictable landscape changes. In contrast, Experiment 2, where uplift varied periodically, shows a different trend. Here, the pixel with the highest entropy value did not always reach the maximum value for all pairs of matrices, corresponding to the alternating uplift scenario. This variation caused the landscape to respond differently at different times, leading to a more complex and dynamic pattern of change that entropy tracking captures – highlighting in the landscape where the most change occurs.

By focusing on specific pixels in Fig. 4, we isolate representative locations in the stream and hillslope domains to examine how local entropy and modular difference evolve over time in response to the uplift history applied in each experiment. Rather than mapping spatial variability across the entire landscape (as shown in Figs. 2 and 3), this targeted approach allows us to track how individual points respond to the imposed uplift trends. While the overall temporal trends appear broadly similar—reflecting the shared influence of the external uplift signal—subtle differences emerge across the matrix pairs. For example, in Experiment 2, the stream pixel shows slightly higher entropy in the

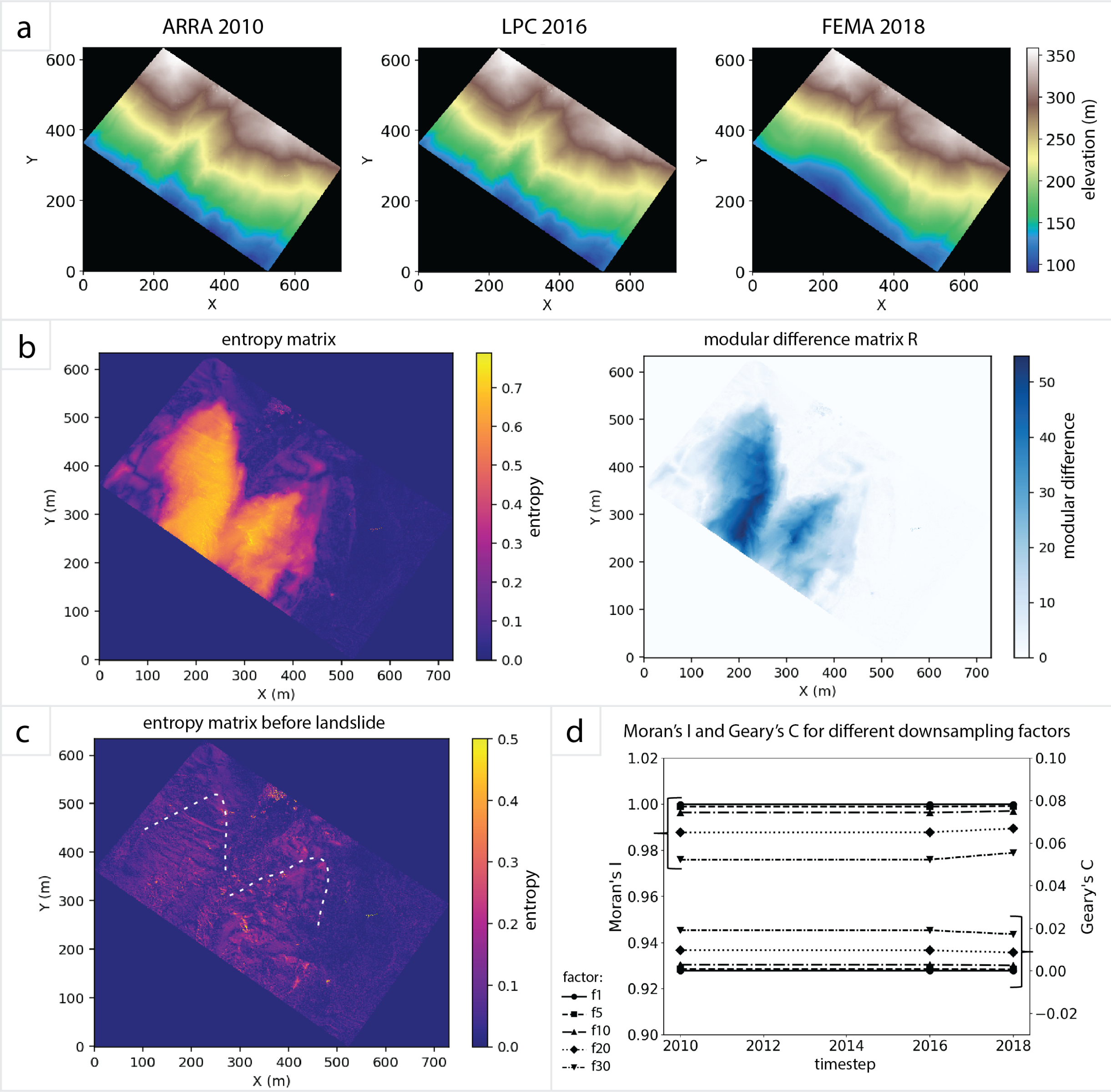

Figure 10 (a) Digital elevation models (DEMs) used for analysis. The LiDAR DEMs were obtained from the USGS d Elevation Program via OpenTopography (datasets: ARRA-CA CentralCoast-Z4 2010, USGS LPC CA WestCoastElNinoUTM10 2016 LAS 2017, and CA FEMA Z4 B2 2018) and are available for download at http://www.opentopography.org. (b) Corresponding Shannon's entropy and modular difference maps illustrating spatial variability through time. (c) Corresponding Shannon's entropy for the DEMs from before the landslide. (d) Moran's I and Geary's C values computed for each timestep, showing spatial autocorrelation across the time series. Different downsampling factors (1, 5, 10, 20, 30m) were applied to the DEMs to evaluate how spatial resolution affects the autocorrelation metrics.

first and last pairs, whereas the hillslope pixel is more variable in the middle pairs. These distinctions align with the spatial patterns observed in Fig. 3 and reinforce how surface processes integrate uplift signals differently across landscape domains.

5.1.3 Experiment with spatially varying uplifts

Evaluating Experiments 3a to d (Fig. 6), the lower contrast in entropy in the eastern part of the drainage basin and higher in the west (Experiment d) suggests that regions experiencing more pronounced uplift contrast exhibit greater variability in elevation and slope. This increasing complexity could reflect the interplay between uplift, erosion, and sediment transport, which creates diverse topographic features. Landscapes with high entropy may be more fragmented or irregular, indicating dynamic processes at work. The inversion of trends after the first run, which becomes more pronounced with higher contrasts, reflects the adaptive response of the landscape to changing uplift patterns. The consistent alignment of Shannon's entropy with regions of pronounced uplift—higher in the west during 3e and 3f and shifting eastward in 3g and 3h—demonstrates the spatial dependence of landscape complexity on the direction and magnitude of tectonic forcing (Burbank and Pinter, 1999; Snyder et al., 2000). Areas experiencing more intense uplift become hotspots of topographic variability, likely due to accelerated erosion and the reorganization of drainage networks. The observations from Experiments 3i and 3j also suggest that the alignment of entropy with the region experiencing the highest contrast indicates that uplift gradients significantly influence landscape complexity. This may reflect accelerated erosion and sediment redistribution in areas with more pronounced forcing, driving localized changes in topography and drainage patterns.

5.1.4 DEM example

The entropy analysis of the Mud Creek landslide site reveals spatial variability and complexity within the landscape, highlighting zones that underwent the greatest change between surveys. Elevated entropy values in the pre-failure DEMs (2010–2016) delineate the section that later collapsed, indicating increased heterogeneity likely associated with incipient deformation and subsurface instability. This suggests that entropy can serve as an early indicator of evolving geomorphic processes, capturing subtle, spatially distributed variations that precede large-scale failures. Identifying such precursory signals has long been a major challenge in landslide prediction and analysis (Saito, 1969; Bell, 2018; Handwerger et al., 2019; Lacroix et al., 2023).

5.2 Interpreting spatial autocorrelation results

The autocorrelation results for Experiments 1 and 2 are presented in Fig. 5. While Moran's I and Geary's C exhibit similar overall patterns, previous studies highlight the importance of using multiple metrics for comparison and verification (Turner et al., 1989; Cullinan and Thomas, 1992; Qi and Wu, 1996). Our experiments provide insight into how these metrics respond to changes in uplift, revealing consistent trends in spatial autocorrelation that reflect underlying landscape dynamics.

In Experiment 1, with steady uplift, Moran's I values gradually increase towards one, indicating that similar landscape features are becoming more clustered over time, leading to greater spatial autocorrelation. Simultaneously, Geary's C values approach zero, suggesting that there is less variability among neighboring pixels, indicating a homogenization of the landscape features. However, in Experiment 2, with periodic uplift variation, the values of Moran's I and Geary's C fluctuate, reflecting alternating patterns of spatial correlation and variability. These fluctuations suggest that the alternating uplift disrupts the uniformity seen in Experiment 1, leading to more complex spatial arrangements and increased landscape diversity. This highlights how different uplift patterns influence landscape dynamics, with steady conditions promoting uniformity and alternating conditions fostering diversity and complexity.

Evaluating Experiments 3a to d (Fig. 6), we observe that in the first half of the experiment, increasing uplift contrast is associated with higher metric values, suggesting a reduction in spatial heterogeneity. However, when the direction of contrast is reversed at higher values, the trend shifts, resulting in decreased spatial autocorrelation. The less smooth trends in Moran's I and Geary's C with higher uplift contrasts in the soil analysis indicate that abrupt or nonlinear changes in spatial patterns are more likely in highly dynamic systems. These metrics reveal how spatial relationships evolve in response to uplift. Higher contrasts disrupt spatial autocorrelation and introduce variability in landscape features, making the patterns more heterogeneous. In higher contrasts, localized uplift intensifies erosion in some areas and deposition in others, leading to more concentrated soil patterns. The faster rate of change in Moran's I and Geary's C under higher contrasts suggests that the landscape rapidly responds to stronger gradients, possibly due to accelerated erosion and drainage reorganization. The slower rate of change under lower contrasts implies a more gradual evolution, where the landscape adjusts more steadily to the forcing conditions. These temporal patterns indicate how uplift gradients influence the rate of landscape adaptation, with stronger contrasts driving more dynamic and immediate responses.

Across all scenarios, Moran's I and Geary's C values decrease with increasing uplift contrast, indicating a general disruption of spatial autocorrelation in soil and sediment patterns. However, this decrease is steeper when the uplift is concentrated in the east (3g and 3h), farther from the outlet. This suggests that when the forcing occurs farther from the outlet, the resulting soil redistribution is more spatially fragmented, likely due to a delayed or less efficient reorganization of the drainage network. In contrast, when uplift is concentrated in the west (3e and 3f), near the outlet, the system may respond more coherently, preserving a slightly higher degree of autocorrelation. Regarding entropy, high uplift in the west (3f) produces more localized variability, leading to lower entropy toward the eastern border. When uplift occurs in the east (3g and 3h), the entropy becomes more widely distributed, reflecting a broader, less concentrated pattern of soil change.

Interestingly, despite the introduction of contrast mid-run, Moran's I and Geary's C in both experiments remain closely aligned with the values observed in their comparable earlier experiments: Experiment 3i with 3g (eastern uplift) and 3j with 3e (western uplift). This suggests that the landscape's spatial autocorrelation structure is relatively resilient once established, and that increasing uplift contrast later in the simulation does not drastically alter soil or sediment distribution patterns.

Soil metrics similarly follow the same trends seen in Experiments 3e and 3g, reinforcing the idea that the timing of the uplift contrast has less influence than its location and magnitude, at least within the timescales and system sensitivity explored in these experiments.

In the Mud Creek case, both Moran's I and Geary's C remained largely stable across the 2010–2018 period, indicating that despite localized instability, the broader spatial structure of the landscape was maintained. The persistence of spatial autocorrelation suggests that landslide initiation influenced the domain without substantially disrupting the overall continuity of elevation values.

However, the scale of analysis influences sensitivity (Turner et al., 1989; Qi and Wu, 1996). In this study, downsampling the DEMs to coarser grids (up to 30 m) revealed decreasing autocorrelation values, and only then showed the distinction between the 2018 post-landslide landscape and the pre-landslide conditions. This highlights the importance of considering scale, so Moran's I and Geary's C can capture spatial organization, whereas entropy is more sensitive to localized variability and small-scale disturbances. Together, these metrics provide complementary insights into how the landscape maintains structural coherence while responding to perturbations.

5.3 Summary and further applications

To improve the interpretability of our results and provide practical guidance for the modeling community, we summarize in Tab 1 the meaning and implications of high, intermediate, and low values for each spatial metric used in this study, together with a summary of the observations derived from the experiments.

Ultimately, while it is beyond the scope of this work to explore all possible use cases, we suggest that these metrics provide a useful framework for quantitative interpretation of landscape evolution models (LEMs) across diverse applications.

One key use case is comparing LEM outputs to identify where models diverge under varying conditions, such as changes in uplift, climate, or other driving variables. In one sense, metrics like Shannon's entropy are most applicable for intra-model comparison, i.e., for a given time-series of model outputs driven by a single history of forcing, it effectively highlights where variability is at a maximum in the landscape. However, this becomes potentially useful for inter-model comparison as well. For example, small changes are made between model scenarios in forcing variables, comparing entropy maps between such model output series may highlight spatial variance that results from changes in the forcing variables. Similarly, by applying pixel tracking and autocorrelation metrics, temporal patterns of divergence can be identified.

Metrics such as Shannon's entropy, Moran's I, and Geary's C also support model validation by revealing patterns that clarify landscape processes. For instance, decreases in entropy values combined with stable autocorrelation metrics may indicate that a landscape is approaching steady state, a challenging concept to capture numerically (Gasparini et al., 2024). These metrics are not limited to data analysis; they also identify regions of interest and expose inconsistencies in model behavior, aiding debugging and refinement. Additionally, they facilitate the comparison of multidimensional data, as demonstrated in our experiments, where entropy helped reconcile differences in the scales of soil depth and topography. While other normalizing methods can also address scale disparities, entropy uniquely quantifies the complexity and variability within the data, providing complementary insights beyond simple normalization.

These metrics also have value in assessing how well LEMs capture patterns observed in natural landscapes. Model outputs often represent idealized or simplified processes over geological timescales, while field data reflect complex, time-integrated surface conditions influenced by numerous local and stochastic factors. This scale mismatch complicates direct comparisons and calls for careful validation strategies that can reconcile these differences. Comparing model outputs with field data, such as DEMs, introduces additional challenges, including discrepancies in spatial and temporal scales, data resolution, and measurement uncertainties. Metrics such as Moran's I and Geary's C, in particular, could help evaluate whether spatial autocorrelation patterns in model outputs match those of the natural landscapes they aim to mimic. Although this comparison lies beyond the scope of the current study, it represents a promising avenue for future applications of this framework. While this study focused on shorter, synthetic landscape experiments, these methods are well-suited for analyzing large datasets by distilling complex spatial patterns into interpretable results. These metrics can inform a range of applications, including model intercomparison, validation, and process-based studies, offering tools to deepen understanding of landscape evolution. Similarly, these metrics can be applied to effectively any matrix-based spatial dataset and as such, may be helpful for interpreting emerging applications such as repeat topographic surveys (McLean and Church, 1999; Wheaton et al., 2010; Schürch et al., 2011; DeLong et al., 2012; James and Robson, 2012; Croke et al., 2013; Anderson and Pitlick, 2014; Smith et al., 2016; Westoby et al., 2018).

6 Conclusion

In this study, we evaluated the efficacy of different analytical metrics for capturing landscape dynamics under varying uplift conditions. By employing entropy calculations, Moran's I, and Geary's C, we were able to capture different aspects of landscape change and variability. Our results show that while modular differences can overlook subtle changes, particularly in scenarios with periodic uplift, entropy was more sensitive, revealing finer details of landscape complexity. This comparison highlights the strengths and limitations of each metric for detecting and characterizing landscape change.

Moran's I and Geary's C metrics further enriched our analysis by offering insights into spatial patterns and relationships within the model outputs. These metrics, calculated for the entire dataset, summarize spatial autocorrelation by indicating how similar or dissimilar landscape features are arranged. Consistent Moran's I and Geary's C values across outputs suggest similar spatial arrangements and model behaviors, while contrasting values reveal notable differences in spatial patterns, underscoring areas where models diverge in their representation of landscape processes.

Our findings also illustrate how landscape sensitivity varies with uplift contrast. Higher contrasts amplify spatial and topographic variability, driving rapid, localized changes in drainage and soil distribution, whereas lower contrasts promote stability and uniformity. The trends observed in Moran's I and Geary's C emphasize their utility in quantifying the interplay between uplift, erosion, and sediment transport.

Beyond landscape models, this methodological framework is applicable to DEMs or any matrix-based dataset, enabling users to quantify differences in information content, assess spatial consistency, and identify divergences among simulation results. It can

Table 1 Interpretative guide of proposed metrics.

Metric	Purpose	Low value	Intermediate value	High value	Metric observations
Shannon's entropy	Measures the uncertainty of information content of a random variable, i.e. the complexity of spatial patterns. Output: 2D spatial map - One entropy value per grid cell.	Low entropy - Indicated a more uniform or stable area with little change across experiments.	Moderate heterogeneity across experiments, indicating some variability in landscape behavior.	High entropy - Indicates a complex area that changed significantly across experiments.	Entropy enables comparison across variables, magnitudes, and spatial domains, revealing finer details of landscape complexity.
Moran's I	Local autocorrelation: quantifies how closely values are clustered together in space using the data's variance. Output: single value per experiment.	-1: Negative spatial autocorrelation, i.e. the variable of interest is perfectly dispersed.	0: No spatial autocorrelation, i.e. the variable of interest is randomly dispersed.	1: Positive spatial autocorrelation, i.e. the variable of interest is perfectly clustered together.	Higher uplift contrast disrupts spatial autocorrelation, increasing landscape variability and causing a faster decline in metric values. Lower uplift contrast results in more gradual changes over time.
Geary's C	Global autocorrelation: quantifies the dissimilarity between neighboring values, but without data standardization. Output: single value per experiment.	$<$ 1: Positive spatial autocorrelation, i.e. similar values (high or low) are spatially clustered.	1: No spatial autocorrelation, i.e. the spatial distribution of the variable is random.	$>$ 1: Negative spatial autocorrelation, i.e. dissimilar values are adjacent to each other.	Uplift near the outlet preserves spatial autocorrelation, whereas uplift farther from the outlet produces a steeper decrease in metric values. When analyzing DEMs, resolution must be considered, as fine scales can obscure landscape change.

also highlight areas of evolving geomorphic processes, potentially serving as an early indicator of landscape change. When interpreting autocorrelation metrics, however, the resolution must be considered, as fine scales can somewhat paradoxically obscure landscape change due to increased local spatial autocorrelation. This comprehensive approach enhances model evaluation, leading to more informed decision-making in landscape modeling and improving our understanding of the factors that shape landscape evolution.

Acknowledgements

We thank Dr. Gregory Tucker, Dr. Matt Rossi, Dr. Yuval Schmilovitz, and Dr. Virginia Iglesias for their valuable discussions, which greatly enriched this work.

Land Recognition

The Mud Creek area in California, from which we obtained Digital Elevation Models to test our methods, is located within the traditional territory of the Salinan people.

Data and Codes Availability

The Zenodo repository (DOI: 10.5281/zenodo.14799227) associated with this manuscript contains the analysis code for each experiment, including figure generation, Jupyter notebooks, and raw Landscape Evolution Model outputs provided as NetCDF files. The raw data include time-resolved gridded fields of topographic elevation and soil depth with full grid metadata.

Funding Statement

This material is based upon work supported by the National Science Foundation under Grant Nos. EAR-2104102 and EAR-1917695.

Conflict of Interest Disclosure

The authors have no competing interests.

Permission to Reproduce Material from Copyrighted Sources

The authors declare that no material from copyrighted sources was reproduced in this manuscript.

Declaration on Artificial Intelligence Use

The authors declare that no artificial intelligence tool was used to assist in or generate any part of the contents in this manuscript.

References

Ali, G. and Roy, A. Revisiting Hydrologic Sampling Strategies for an Accurate Assessment of Hydrologic Connectivity in Humid Temperate Systems. *Geography Compass*, 3(1):350–374, 2009. doi: 10.1111/j.1749-8198.2008.00180.x.

Anderson, S. and Pitlick, J. Using repeat lidar to estimate sediment transport in a steep stream. *Journal of Geophysical Research: Earth Surface*, 119(3):621–643, 2014. doi: 10.1002/2013JF002933.

Anselin, L. A Local Indicator of Multivariate Spatial Association: Extending Geary's c. *Geographical Analysis*, 51(2):133–150, 2019. doi: 10.1111/gean.12164.

Barnes, R. Parallel non-divergent flow accumulation for trillion cell digital elevation models on desktops or clusters. *Environmental Modelling Software*, 92:202–212, 2017. doi: 10.1016/j.envsoft.2017.02.022.

Barnhart, K., Glade, R., Shobe, C., and Tucker, G. Terrainbento 1.0: A Python package for multi-model analysis in long-term drainage basin evolution. *Geoscientific Model Development*, 12 (4):1267–1297, 2019. doi: 10.5194/gmd-12-1267-2019.

Bell, A. Predictability of Landslide Timing From Quasi-Periodic Precursory Earthquakes. *Geophysical Research Letters*, 45(4): 1860–1869, 2018. doi: 10.1002/2017GL076730.

Boots, B. and Getls, A. *Point Pattern Analysis*. Sage Publication, 2020.

Bracken, L. and Croke, J. The concept of hydrological connectivity and its contribution to understanding runoff-dominated geomorphic systems. *Hydrological Processes*, 21(13):1749–1763, 2007. doi: 10.1002/hyp.6313.

Burbank and Pinter. Landscape evolution: The interactions of tectonics and surface processes. *Basin Research*, 11(1):1–6, 1999. doi: 10.1046/j.1365-2117.1999.00089.x.

Buttle, J., Dillon, P., and Eerkes, G. Hydrologic coupling of slopes, riparian zones and streams: An example from the Canadian Shield. *Journal of Hydrology*, 287(1–4):161–177, 2004. doi: 10.1016/j.jhydrol.2003.09.022.

Carlà, T., Intrieri, E., Raspini, F., Bardi, F., Farina, P., Ferretti, A., Colombo, D., Novali, F., and Casagli, N. Perspectives on the prediction of catastrophic slope failures from satellite InSAR. *Scientific Reports*, 9(1):14137, 2019. doi: 10.1038/s41598-019-50792-y.

Chai, T. and Draxler, R. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3):1247–1250, 2014. doi: 10.5194/gmd-7-1247-2014.

Chembolu, V. and Dutta, S. An entropy based morphological variability assessment of a large braided river. *Earth Surface Processes and Landforms*, 43(14):2889–2896, 2018. doi: 10.1002/esp.4441.

Cherkauer, D. Minimization of power expenditure in a riffle-pool alluvial channel. *Water Resources Research*, 9(6):1613–1628, 1973. doi: 10.1029/WR009i006p01613.

Chung, M. and Noguchi, J. Geographic spatial autocorrelation of morphological characters of the Hemerocallis middendorffii complex (Liliaceae. *Annales Botanici Fennici*, 35(3):183–189, 1998.

Cliff, A. and Ord, K. Spatial Autocorrelation: A Review of Existing and New Measures with Applications. *Economic Geography*, 46: 269, 1970. doi: 10.2307/143144.

Croke, J., Todd, P., Thompson, C., Watson, F., Denham, R., and Khanal, G. The use of multi temporal LiDAR to assess basin-scale erosion and deposition following the catastrophic January 2011 Lockyer flood, SE Queensland, Australia. *Geomorphology*, 184: 111–126, 2013. doi: 10.1016/j.geomorph.2012.11.023.

Cullinan, V. and Thomas, J. A comparison of quantitative methods for examining landscape pattern and scale. *Landscape Ecology*, 7(3):211–227, 1992. doi: 10.1007/BF00133311.

Cushman, S. Calculation of Configurational Entropy in Complex Landscapes. *Entropy*, 20(4):298, 2018. doi: 10.3390/e20040298.

Cushman, S. Entropy in Landscape Ecology: A Quantitative Textual Multivariate Review. *Entropy*, 23(11):1425, 2021. doi: 10.3390/e23111425.

Data, N. L. ARRA Lidar: California Coastal Project – Zone 4, 2010, Dataset, https://portal.opentopography.org/noaaDataset?noaaID=4843.

Dejene, G. and Muleta, G. Statistical Analysis of Spatial Distribution of Tuberculosis in Ethiopia, 2020.

DeLong, S., Prentice, C., Hilley, G., and Ebert, Y. Multitemporal ALSM change detection, sediment delivery, and process mapping at an active earthflow. *Earth Surface Processes and Landforms*, 37(3):262–272, 2012. doi: 10.1002/esp.2234.

Detty, J. and McGuire, K. Topographic controls on shallow groundwater dynamics: Implications of hydrologic connectivity between hillslopes and riparian zones in a till mantled catchment. *Hydrological Processes*, 24(16):2222–2236, 2010. doi: 10.1002/hyp.7656.

Fan, C. and Myint, S. A comparison of spatial autocorrelation indices and landscape metrics in measuring urban landscape fragmentation. *Landscape and Urban Planning*, 121:117–128, 2014. doi: 10.1016/j.landurbplan.2013.10.002.

Gao, P. and Li, Z. Computation of the Boltzmann entropy of a landscape: A review and a generalization. *Landscape Ecology*, 34(9): 2183–2196, 2019. doi: 10.1007/s10980-019-00814-x.

Gasparini, N., Forte, A., and Barnhart, K. Short Communication: Numerically simulated time to steady state is not a reliable measure of landscape response time. *Earth Surface Dynamics*, 12(6): 1227–1242, 2024. doi: 10.5194/esurf-12-1227-2024.

Geary, R. The Contiguity Ratio and Statistical Mapping. *The Incorporated Statistician*, 5(3):115, 1954. doi: 10.2307/2986645.

Gupta, H., Kling, H., Yilmaz, K., and Martinez, G. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2):80–91, 2009. doi: 10.1016/j.jhydrol.2009.08.003.

Handwerger, A., Huang, M.-H., Fielding, E., Booth, A., and Bürgmann, R. A shift from drought to extreme rainfall drives a stable landslide to catastrophic failure. *Scientific Reports*, 9(1): 1569, 2019. doi: 10.1038/s41598-018-38300-0.

Handwerger, A., Lacroix, P., Bell, A., Booth, A., Huang, M.-H., Mudd, S., Bürgmann, R., and Fielding, E. Multi-sensor remote sensing captures geometry and slow-to-fast sliding transition of the 2017 Mud Creek landslide. *Scientific Reports*, 15(1):29831, 2025. doi: 10.1038/s41598-025-11399-8.

Hobley, D., Adams, J., Nudurupati, S., Hutton, E., Gasparini, N., Istanbulluoglu, E., and Tucker, G. Creative computing with Landlab: An open-source toolkit for building, coupling, and exploring two-dimensional numerical models of Earth-surface dynamics. *Earth Surface Dynamics*, 5(1):21–46, 2017. doi: 10.5194/esurf-5-21-2017.

Hopp, L. and McDonnell, J. Connectivity at the hillslope scale: Identifying interactions between storm size, bedrock permeability, slope angle and soil depth. *Journal of Hydrology*, 376 (3–4):378–391, 2009. doi: 10.1016/j.jhydrol.2009.07.047.

Howard, A. Badland Morphology and Evolution: Interpretation Using a Simulation Model. *Earth Surface Processes and Landforms*, 22(3):211–227, 1997. doi: 10.1002/(SICI)1096-9837(199703)22:3.

Ijjász-Vásquez, E., Bras, R., and Moglen, G. Sensitivity of a basin evolution model to the nature of runoff production and to initial conditions. *Water Resources Research*, 28(10):2733–2741, 1992. doi: 10.1029/92WR01561.

Intrieri, E., Raspini, F., Fumagalli, A., Lu, P., Del Conte, S., Farina, P., Allievi, J., Ferretti, A., and Casagli, N. The Maoxian landslide as seen from space: Detecting precursors of failure with Sentinel-1 data. *Landslides*, 15(1):123–133, 2018. doi: 10.1007/s10346-017-0915-7.

James, M. and Robson, S. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research: Earth Sur-

face, 117(F3), 2012. doi: 10.1029/2011JF002289.

Joshi, P., Lele, N., and Agarwal, S. Entropy as an indicator of fragmented landscape. *Current Science*, 91(3):276–278, 2006.

Karaca, Y. and Moonis, M. Shannon entropy-based complexity quantification of nonlinear stochastic process. In *Multi-Chaos, Fractal and Multi-Fractional Artificial Intelligence of Different Complex Systems*, page 231–245. Elsevier, 2022. doi: 10.1016/B978-0-323-90032-4.00018-3.

Kissling, W. and Carl, G. Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*, 17(1):59–71, 2008. doi: 10.1111/j.1466-8238.2007.00334.x.

Koenig, W. and Knops, J. Testing for spatial autocorrelation in ecological studies. *Ecography*, 21(4):423–429, 1998. doi: 10.1111/j.1600-0587.1998.tb00407.x.

Kuuluvainen, T., Järvinen, E., Hokkanen, T., Rouvinen, S., and Heikkinen, K. Structural heterogeneity and spatial autocorrelation in a natural mature Pinus sylvestris dominated forest. *Ecography*, 21(2):159–174, 1998. doi: 10.1111/j.1600-0587.1998.tb00670.x.

Lacroix, P., Huanca, J., Albinez, L., and Taipe, E. Precursory Motion and Time-Of-Failure Prediction of the Achoma Landslide, Peru, From High Frequency PlanetScope Satellites. *Geophysical Research Letters*, 50(19):2023 105413, 2023. doi: 10.1029/2023GL105413.

Leopold, L. and Langbein, W. *The concept of Entropy in Landscape Evolution*, volume 500. US Government Printing Office, 1962.

Lin, J. Comparison of Moran's I and Geary's c in Multivariate Spatial Pattern Analysis. *Geographical Analysis*, 55(4):685–702, 2023. doi: 10.1111/gean.12355.

Liu, X. and Zhao, H. Exploring spatial entropy in the analysis of landscape pattern and ecological sustainability in watershed ecosystem. *IOP Conference Series: Earth and Environmental Science*, 1004(1):012022, 2022. doi: 10.1088/1755-1315/1004/1/012022.

MacArthur, R. Fluctuations of Animal Populations and a Measure of Community Stability. *Ecology*, 36(3):533–536, 1955. doi: 10.2307/1929601.

Margalef, R. Information theory in ecology. *Gen Syst*, 3:36–72, 1958.

Mathur, M. Spatial autocorrelation analysis in plant population: An overview. *Journal of Applied and Natural Science*, 7(1):501–513, 2015. doi: 10.31018/jans.v7i1.639.

Mattos, S., Vicente, L., Vicente, A., Júnior, C., and Piqueira, J. Metrics based on information entropy applied to evaluate complexity of landscape patterns. *PLOS ONE*, 17(1):0262680, 2022. doi: 10.1371/journal.pone.0262680.

McLean, D. and Church, M. Sediment transport along lower Fraser River: 2. Estimates based on the long-term gravel budget. *Water Resources Research*, 35(8):2549–2559, 1999. doi: 10.1029/1999WR900102.

Molnár, P. and Ramírez, J. Energy dissipation theories and optimal channel characteristics of river networks. *Water Resources Research*, 34(7):1809–1818, 1998. doi: 10.1029/98WR00983.

Montgomery, D. Slope Distributions, Threshold Hillslopes, and Steady-state Topography. *American Journal of Science*, 301 (4–5):432–454, 2001. doi: 10.2475/ajs.301.4-5.432.

Montgomery, D. and Dietrich, W. Where do channels begin? *Nature*, 336(6196):232–234, 1988. doi: 10.1038/336232a0.

Montgomery, D. and Dietrich, W. Source areas, drainage density, and channel initiation. *Water Resources Research*, 25(8): 1907–1918, 1989. doi: 10.1029/WR025i008p01907.

Montgomery, D. and Dietrich, W. Channel Initiation and the Problem of Landscape Scale. *Science*, 255(5046):826–830, 1992. doi: 10.1126/science.255.5046.826.

Montgomery, D. and Foufoula-Georgiou, E. Channel network source representation using digital elevation models. *Water Resources Research*, 29(12):3925–3934, 1993. doi: 10.1029/93WR02463.

Moran, P. The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society Series B*, 10(2):243–251, 1948. doi: 10.1111/j.2517-6161.1948.tb00012.x.

Nash, J. and Sutcliffe, J. River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3):282–290, 1970. doi: 10.1016/0022-1694(70)90255-6.

Perron, J., Dietrich, W., and Kirchner, J. Controls on the spacing of first-order valleys. *Journal of Geophysical Research: Earth Surface*, 113(F4):2007 000977, 2008a. doi: 10.1029/2007JF000977.

Perron, J., Kirchner, J., and Dietrich, W. Spectral signatures of characteristic spatial scales and nonfractal structure in landscapes. *Journal of Geophysical Research: Earth Surface*, 113(F4):2007 000866, 2008b. doi: 10.1029/2007JF000866.

Perron, J., Kirchner, J., and Dietrich, W. Formation of evenly spaced ridges and valleys. *Nature*, 460(7254):502–505, 2009. doi: 10.1038/nature08174.

Qi, Y. and Wu, J. Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices. *Landscape Ecology*, 11(1):39–49, 1996. doi: 10.1007/BF02087112.

Quijano, J. and Lin, H. Entropy in the Critical Zone: A Comprehensive Review. *Entropy*, 16(6):3482–3536, 2014. doi: 10.3390/e16063482.

Rey, S. and Anselin, L. PySAL: A Python Library of Spatial Analytical Methods. *The Review of Regional Studies*, page 5–27, 2007.

Ricotta, C. From theoretical ecology to statistical physics and back: Self-similar landscape metrics as a synthesis of ecological diversity and geometrical complexity. *Ecological Modelling*, 125(2–3): 245–253, 2000. doi: 10.1016/S0304-3800(99)00185-4.

Rinaldo, A., Rodriguez-Iturbe, I., Rigon, R., Bras, R., Ijjasz-Vasquez, E., and Marani, A. Minimum energy and fractal structures of drainage networks. *Water Resources Research*, 28(9): 2183–2195, 1992. doi: 10.1029/92WR00801.

Rodríguez-Iturbe, I., Rinaldo, A., Rigon, R., Bras, R., Marani, A., and Ijjász-Vásquez, E. Energy dissipation, runoff production, and the three-dimensional structure of river basins. *Water Resources Research*, 28(4):1095–1103, 1992. doi: 10.1029/91WR03034.

Saito, M. Research on forecasting the time of occurrence of slope failure. *Railway Technical Research Institute, Quarterly Reports*, 10, 1969.

Sanner, A., Nöhring, W., Thimons, L., Jacobs, T., and Pastewka, L. Scale-dependent roughness parameters for topography analysis. *Applied Surface Science Advances*, 7:100190, 2022. doi: 10.1016/j.apsadv.2021.100190.

Schürch, P., Densmore, A., Rosser, N., Lim, M., and McArdell, B. Detection of surface change in complex topography using terrestrial laser scanning: Application to the Illgraben debris-flow channel. *Earth Surface Processes and Landforms*, 36(14): 1847–1859, 2011. doi: 10.1002/esp.2206.

Shannon, C. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3):379–423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.

Shobe, C., Tucker, G., and Barnhart, K. The SPACE 1.0 model: A Landlab component for 2-D calculation of sediment transport, bedrock erosion, and landscape evolution. *Geoscientific Model Development*, 10(12):4577–4604, 2017. doi: 10.5194/gmd-10-4577-2017.

Shrestha, S., Bauer, C., Hendricks, B., and Stopka, T. Spatial epidemiology: An empirical framework for syndemics research. *Social Science Medicine*, 295:113352, 2022. doi: 10.1016/j.socscimed.2020.113352.

Smith, M., Carrivick, J., and Quincey, D. Structure from motion photogrammetry in physical geography. *Progress in Physical Geography: Earth and Environment*, 40(2):247–275, 2016. doi: 10.1177/0309133315615805.

Snyder, N., Whipple, K., Tucker, G., and Merritts, D. Landscape response to tectonic forcing: Digital elevation model analysis of stream profiles in the Mendocino triple junction region, northern California. *Geological Society of America Bulletin*, 112(8): 1250–1263, 2000. doi: 10.1130/0016-7606(2000)112.

Sokal, R., Oden, N., and Thomson, B. Local spatial autocorrelation in biological variables. *Biological Journal of the Linnean Society*, 65(1):41–62, 1998. doi: 10.1111/j.1095-8312.1998.tb00350.x.

Spence, C. and Woo, M. Hydrology of subarctic Canadian shield: Soil-filled valleys. *Journal of Hydrology*, 279(1–4):151–166, 2003. doi: 10.1016/S0022-1694(03)00175-6.

Sun, T., Meakin, P., and Jøssang, T. A minimum energy dissipation model for drainage basins that explicitly differentiates between channel networks and hillslopes. *Physica A: Statistical Mechanics and Its Applications*, 210(1–2):24–47, 1994. doi: 10.1016/0378-4371(94)00053-0.

Survey, U. CA FEMA Z4 B2 2018 (Version CA FEMA, 2021a, Dataset, https://portal.opentopography.org/usgsDataset?dsid=CA_FEMA_Z4_B2_2018.

Survey, U. In *USGS LPC CA WestCoastElNinoUTM10 2016 LAS 2017 (Version USGS LPC CA WestCoastElNinoUTM10 2016 LAS 2017*. 2021b. doi: https://portal.opentopography.org/usgsDataset?dsid=USGS_LPC_CA_WestCoastElNinoUTM10_2016_LAS_2017.

Tromp-van Meerveld, H., Peters, N., and McDonnell, J. Effect of bedrock permeability on subsurface stormflow and the water balance of a trenched hillslope at the Panola Mountain Research Watershed, Georgia, USA. *Hydrological Processes*, 21(6): 750–769, 2007. doi: 10.1002/hyp.6265.

Tucker, G. and Bras, R. Hillslope processes, drainage density, and landscape morphology. *Water Resources Research*, 34(10): 2751–2764, 1998. doi: 10.1029/98WR01474.

Tucker, G. and Hancock, G. Modelling landscape evolution. *Earth Surface Processes and Landforms*, 35(1):28–50, 2010. doi: 10.1002/esp.1952.

Turner, M. and Gardner, R. *Landscape Ecology in Theory and Practice: Pattern and Process*. Springer, New York, 2015. doi: 10.1007/978-1-4939-2794-4.

Turner, M., O'Neill, R., Gardner, R., and Milne, B. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology*, 3(3–4):153–162, 1989. doi: 10.1007/BF00131534.

Ulanowicz, R. Information theory in ecology. *Computers Chemistry*, 25(4):393–399, 2001. doi: 10.1016/S0097-8485(01)00073-0.

Upton, G., Fingleton, B., and Upton, G. *Point pattern and quantitative data (Reprinted*. Wiley, 1994.

Vranken, I., Baudry, J., Aubinet, M., Visser, M., and Bogaert, J. A review on the use of entropy in landscape ecology: Heterogeneity, unpredictability, scale dependence and their links with thermodynamics. *Landscape Ecology*, 30(1):51–65, 2015. doi: 10.1007/s10980-014-0105-0.

Walter, S. Assessing spatial patterns in disease rates. *Statistics in Medicine*, 12(19–20):1885–1894, 1993. doi: 10.1002/sim.4780121914.

Westoby, M., Lim, M., Hogg, M., Pound, M., Dunlop, L., and Woodward, J. Cost-effective erosion monitoring of coastal cliffs. *Coastal Engineering*, 138:152–164, 2018. doi: 10.1016/j.coastaleng.2018.04.008.

Wheaton, J., Brasington, J., Darby, S., and Sear, D. Accounting for uncertainty in DEMs from repeat topographic surveys: Improved sediment budgets. *Earth Surface Processes and Landforms*, 35(2):136–156, 2010. doi: 10.1002/esp.1886.

Willmott, C. and Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30:79–82, 2005. doi: 10.3354/cr030079.