Review Report for

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

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

*Corresponding author: vgrom1@lsu.edu
Paper DOI: 10.59236/geomorphica.v2i1.51

Table of Contents:

Round 1 Review

Comments by reviewer 1 ("anonymous")

Completed: 2025-04-11 06:35 PM
Recommendation: Major revisions
Reviewer Comments
For author and editor
This article proposes methods to evaluate and compare outputs from landscape evolution models (LEMs). This question is critical as no formal evaluation parameters currently exist, leaving the scientific community without established standards to assess and compare different landscape evolution scenarios and models. The methodology employs three quantitative metrics: Shannon's Entropy, Moran's I, and Geary's C, applied to outputs from a Landlab LEM across multiple scenarios.

The paper is generally well-written and presents compelling results that contribute meaningfully to the field. However, two key aspects require attention to strengthen the work.

Regarding methodology, I have significant concerns about the implementation of Shannon's Entropy in Section 3. The current approach using Equations 5 and 6 generates what is termed a "probability matrix," but this appears mathematically problematic. As the sum of matrix elements can exceed unity (one), it cannot properly function as a probability distribution in the Shannon's Entropy framework. This fundamental issue either needs correction through revised methodology or much more substantial theoretical justification explaining why this approach remains valid despite violating probability axioms.

The Discussion section would benefit from substantial reorganization to better serve readers. While the results are interesting, their practical interpretation remains unclear. Specifically, the significance of obtaining high versus low values for each metric (Entropy, Moran's I, and Geary's C) needs explicit explanation - what do these values actually tell us about landscape evolution models? For a modeling community eager to apply these standardized methods, this interpretive guidance is crucial yet currently lacking (a table could help).

I strongly recommend restructuring the Discussion into clear subsections, each dedicated to:

- Interpreting entropy results and their implications

- Explaining spatial autocorrelation findings (Moran's I and Geary's C)

This separation would provide much-needed clarity while maintaining rigorous analysis.

These changes would transform the paper from presenting interesting results to providing genuinely useful, actionable standards for the modeling community. The current version makes important progress, but realizing its full potential requires addressing these methodological and presentation issues.

I recommend major revisions to address key methodological and presentation issues.

ADDITIONAL COMMENTS

Line 24: I recommend adding a paragraph describing the challenges regarding the comparison between landscape evolution model outputs and field data such as DEMs. This should address issues of scale compatibility, temporal resolution, and validation approaches.

Line 213: The "max value" mentioned here - is this referring to the total number of matrices? Please clarify this in the text.

Line 217: What does "n" represent here? Is this the logarithm base? Please explicitly define this parameter in the text.

Line 233: There is ambiguity in this equation. Please specify whether this should be interpreted as: "log(Pk)" (logarithm of Pk), or "log(P)*k" (logarithm of P multiplied by k)

Additionally, please clearly indicate the base of the logarithm (is it base n, base e, or another base?).

Note that the right side of the equation appears to yield a scalar value rather than a matrix, which contradicts the matrix notation on the left side. This equation requires a more detailed explanation.

Comments by reviewer 2 (“anonymous”)

Completed: 2025-04-21 02:39 PM
Recommendation: Major revisions
Reviewer Comments
For author and editor
Thank you for the opportunity to review this paper. The authors have proposed an interesting suite of techniques for analyzing landscape evolution model outputs using an information theory lens. I appreciate that the authors are applying these techniques that are common in other fields to landscape evolution modeling, and I think the idea has a lot of merit. Despite the neat idea, I am started and ended the paper unsure of the knowledge gap addressed, or question answered that cannot be addressed with other topographic analyses. I am sure there is more to learn from the information theory lens, but it is not yet well communicated in this paper. I suggest a major revision of the introduction and discussion to reframe the paper around the way that the results provide an advance in our ability to interpret the results of landscape evolution models. This reframing should be specific to the results and analyses presented in the paper, rather than saying "this code could be used for a purpose..." focus on demonstrating exactly how the code was effectively used for one of those purposes. The discussion has a lot of hypothetical uses of the code, without effectively walking through how a researcher would use the code.

The paper has a high number of typos and messy writing (for example, repeated variables in the equations with different meanings, this is very hard to follow for the reader and requires a major revision).

Many of the figures do not yet meet a standard for publication.

I have detailed some general and specific comments below.

The authors emphasize throughout the paper that they can separate the hillslopes from the stream to analyze how different parts of the landscapes vary. However, I’m not sure what we learned from the exercise. The authors use non-specific language to say, for example (Line 455-456) “we can more effectively assess the entropy contrasts and the nuances in each are expressed” or Line 472-473 (“analyzing the stream network and hillslopes separately can provide more insight into how different landscape processes influence variability”.
Please consider providing specific examples of what you did learn by separating the regimes, rather than just stating one could learn something through these exercises.
Separating the hillslopes and streams was done using standard methods (threshold drainage area), as the paper reads now, distinguishing between these parts of the landscape is not an emergent outcome of the statistical analyses. Thus, this is not a particularly novel method of analysis. Could these analyses be used by a geomorphologist to assess which parts of an empirical landscape are operating under hillslope or fluvial processes?
The authors state throughout that one of the benefits of the method is to reconcile differences in scales of soil depth and topography (e.g. the absolute magnitude of elevation and soil depth are

different making comparison hard). While the entropy statistic does indeed help with this, so would other normalizing techniques. I find this one of the least convincing arguments for the use of these indices.

The authors also state that the entropy method is more sensitive to the modular difference method. But, looking at the maps, in Fig. 2 for example, I'm seeing a lot of patterns in the modular difference maps that I'm not seeing in the Entropy maps. For example, in Experiment 2 the entropy map is pretty much all yellow except for the bottom left corner while the Modular Difference method is showing more variability, highlighting this topographic break between the lowland and the highland that you can see in Fig. 1. So, I'm not quite following using the spatial data at least, how "entropy offers a more detailed view of variability" (line 471).

Figure 4: it is a bit concerning that the outline of these grids (which I assume are showing entropy) are all one color (e.g. dark gray around the top and right). This looks like something that is the result of a boundary condition. Can the authors describe why we observe this pattern – is it a model artefact?

Line-by-line comments:

Line 33: Grammar issue - Especially relevant is the need for mechanisms to compare multidimensional models...

Line 70: I am surprised that work by Ignacio Rodriguez-Iturbe and/or Andrea Rinaldo is not mentioned here.

Line 81: typo near end of line ("where while...")

Line 138 and Equation 2: isn't x_bar the average of x(i,j); not just x(i) if in this equation the spatial units are indexed by (i,j) as opposed to Eq. 1.

Line 205: missing word: "first IT is necessary"

Line 211: Delete "Equation 1: maximum modular difference" (in this manuscript it is Equation 4).

Line 211-214: Are there two different uses of k in this equation? It is used to index the consecutive matrices (k {0,1,....}). Then it is calculated as the maximum value of R. I suggest redefining one of these k's.

Line 212: Also, in Eq. 4 the second term may have an extra m: matrices(k+1)*m[i,j].

Line 221 (Eq. 6): should the left side of this equation by P(xi,j) as is noted when this section is introduced in line 205?

Line 220-221: Here i and j are defined as j = i+1. But, everywhere else in the paper (i,j) is used to spatially index elements within the grids. It seems in the rest of the paper you are using either x or k to index which matrix you are using.

Line 231: Is Xi,j in this equation the same as matrices[k][i,j] in equation 4?

Line 231: Why is Xi,j capitalized on one side of the equation and not the other?

Line 227: you aimed to include or you did include?

Table 1: Please provide information on the boundary conditions for the model.

Figure 2 and Line 217: What is meant by "first highest value pixel"? Just the highest pixel value in the summary statistic? Perhaps the "highest value pixel at the first time step"?

Figure 2: The difference in scales on the Entropy maps make it hard to compare across calculations.

Line 328: "thus enhancing our understanding of landscape dynamics". This belongs in the discussion.

Line 336: a drainage area threshold of 3 m^2 is extremely small!

Figure 4: What are the values shown in the images? Elevation? Entropy? The color/scale bar is nearly illegible and doesn't seem to match the gray background of the image. I cannot interpret what these grids are showing from the figure or caption.

Lines 347: this belongs in the discussion. I'm not sure how Figure 4 is demonstrating any spatial variability in the landscape as is claimed in this paragraph (rather Figure 2 and 3 do so).

Figure 4: It is interesting that the hillslopes and streams look the same in terms of the temporal trends (middle and right-hand columns) in Figure 4. Would you expect more of a difference? If not, why iså it important that we are looking at these two patterns separately?

Figure 5: Please define what the dashed lines are supposed to mean on the Figure or in the caption.

Recommendation by Associate Editor 2 (Roberto Fernández)

Completed: 2025-05-15 02:43 AM
Recommendation: Major revisions

Authors' responses to review comments

August 26, 2025

Summary:

Reviewer 1 felt the paper lacked a clear articulation of the specific knowledge gap or question addressed that couldn't be answered by other topographic analyses. They recommended a major revision of the introduction and discussion to better frame the paper around concrete advances demonstrated by the results, focusing less on hypothetical uses of the code and more on how it was effectively applied.

To address the need of a more concrete analysis of the experiments, the manuscript now provides specific examples showing how entropy patterns differ between hillslope and stream domains, clarifying insights gained from separating these landscape parts (e.g., elevated entropy in hillslope areas during cyclic uplift).

We also added explanation acknowledging that while other normalization methods exist, entropy provides a complementary metric capturing data complexity and heterogeneity beyond simple scale adjustments. We clarified that entropy captures cumulative variability (degree of disorder) while modular difference detects directional change, which may be masked under cyclic processes.

Reviewer 2 raised significant concerns about the mathematical validity of the Shannon entropy implementation, specifically questioning the construction of the "probability matrices" used in the calculation. They also found the Discussion section lacking clear interpretive guidance on what the metric values reveal about landscape evolution processes, suggesting a more structured and detailed explanation is needed. Additional clarifications were requested regarding terminology, equation notation, and the challenges of comparing model outputs with field data.

To address those issues we reviewed the methodology, adding detailed explanation clarifying that the "probability matrices" used in entropy calculations are normalized by dividing the absolute difference between consecutive matrices by the global maximum difference (k) times the number of matrices (n). This scaling ensures all values are in [0,1] and the sum does not exceed 1, making the entropy calculation mathematically valid within a probability-like framework.

We also reorganized the Discussion into three clear subsections: one on interpreting entropy results, one on spatial autocorrelation metrics (Moran's I and Geary's C), and a concluding subsection on metric summaries and broader applications. We added a new summary table (Table 3) explaining what high, intermediate, and low values mean for each metric, the nature of the output they produce, and their implications for interpreting landscape evolution models.

We thank the reviewers and believe that these changes and restructuring provided clearer, actionable interpretive guidance for the modeling community.

Anonymous Associate Editor
Recommendation: Major revisions

Reviewer 1:

Thank you for the opportunity to review this paper. The authors have proposed an interesting suite of techniques for analyzing landscape evolution model outputs using an information theory lens. I appreciate that the authors are applying these techniques that are common in other fields to landscape evolution modeling, and I think the idea has a lot of merit. Despite the neat idea, I am started and ended the paper unsure of the knowledge gap addressed, or question answered that cannot be addressed with other topographic analyses. I am sure there is more to learn from the information theory lens, but it is not yet well communicated in this paper. I suggest a major revision of the introduction and discussion to reframe the paper around the way that the results provide an advance in our ability to interpret the results of landscape evolution models. This reframing should be specific to the results and analyses presented in the paper, rather than saying "this code could be used for a purpose..." focus on demonstrating exactly how the code was effectively used for one of those purposes. The discussion has a lot of hypothetical uses of the code, without effectively walking through how a researcher would use the code.

The paper has a high number of typos and messy writing (for example, repeated variables in the equations with different meanings, this is very hard to follow for the reader and requires a major revision).

Many of the figures do not yet meet a standard for publication.

I have detailed some general and specific comments below.

1. The authors emphasize throughout the paper that they can separate the hillslopes from the stream to analyze how different parts of the landscapes vary. However, I'm not sure what we learned from the exercise. The authors use non-specific language to say, for example (Line 455-456) "we can more effectively assess the entropy contrasts and the nuances in each are expressed" or Line 472-473 ("analyzing the stream network and hillslopes separately can provide more insight into how different landscape processes influence variability".
    1. Please consider providing specific examples of what you *did* learn by separating the regimes, rather than just stating one could learn something through these exercises.
    2. Separating the hillslopes and streams was done using standard methods (threshold drainage area), as the paper reads now, distinguishing between these parts of the landscape is not an emergent outcome of the statistical analyses. Thus, this is not a particularly novel method of analysis. Could

these analyses be used by a geomorphologist to assess which parts of an empirical landscape are operating under hillslope or fluvial processes?

Response: Regarding the insights gained from separating hillslope and stream domains:

We agree that the original wording was general. To be more specific, our analysis revealed that entropy patterns differ not only spatially but also between domains. For example, in Experiment 2, separating the domains showed that the intermediate basin section has slightly elevated entropy within the hillslope areas, indicating that cyclic uplift influences hillslope erosion variability before sediment enters the stream network. In contrast, Experiment 1 exhibited elevated entropy mainly in the distal basin and fluvial domain, with higher values for the buffered zone in the hillslopes. We added that both in the results (lines 332-340) and discussion session (lines 477-489).

On the novelty and applicability of the separation method, we emphasized in the implementation section that separation by drainage area threshold is a standard approach (lines 246-267) and that there is potential utility in examining them separately, shown by the results.

2. The authors state throughout that one of the benefits of the method is to reconcile differences in scales of soil depth and topography (e.g. the absolute magnitude of elevation and soil depth are different making comparison hard). While the entropy statistic does indeed help with this, so would other normalizing techniques. I find this one of the least convincing arguments for the use of these indices.

Response: We understand that the ability of the entropy statistic to reconcile differences in scale between soil depth and topography may not be entirely unique compared to other normalizing techniques. However, we believe that the entropy-based approach offers a complementary and intuitive metric that captures the complexity and variability of the data in a way that simple normalization does not fully address. While other methods can adjust for scale differences, entropy specifically quantifies the degree of disorder or heterogeneity within the landscape variables, providing an additional layer of insight. We recognize this interpretation may not be universally convincing, but we consider it a valid and useful tool within the broader analytical framework for characterizing landscape variability and we added some explanation of this point in lines 618-625.

3. The authors also state that the entropy method is more sensitive to the modular difference method. But, looking at the maps, in Fig. 2 for example, I'm seeing a lot of patterns in the modular difference maps that I'm not seeing in the Entropy maps. For example, in Experiment 2 the entropy map is pretty much all yellow except for the bottom left corner while the Modular Difference method is showing more variability, highlighting this topographic break between the lowland and the highland that you can see in Fig. 1. So, I'm not quite following using the spatial data at least, how "entropy offers a more detailed view of variability" (line 471).

Response: Entropy captures cumulative variability, whereas modular difference is sensitive to unidirectional change, meaning changes over time that consistently increase or decrease rather than oscillate. Modular difference effectively detects these consistent changes but may underrepresent variability when changes are cyclic, as positive and negative fluctuations can cancel out.

The modular difference maps, especially in Experiment 2, display clear localized patterns, such as the topographic break between lowland and highland regions. We have clarified in the manuscript the distinction between the sensitivities of the two metrics (lines 314–318).

In the revised interpretation, we note that modular difference highlights sharp spatial transitions due to unidirectional change, while entropy captures overall complexity and heterogeneity regardless of direction. In the case of alternating uplift (Experiment 2), cyclic changes can produce low net modular difference even in areas with substantial fluctuations. In contrast, entropy remains sensitive to these variations because it quantifies disorder rather than net change. This explains why the entropy map appears more uniform (broadly high values) around channel zones, reflecting widespread variability that is not directional (lines 490–505).

4. Figure 4: it is a bit concerning that the outline of these grids (which I assume are showing entropy) are all one color (e.g. dark gray around the top and right). This looks like something that is the result of a boundary condition. Can the authors describe why we observe this pattern – is it a model artefact?

Response: To clarify, the left panel in Figure 4 is a hillshade map representing topography, overlaid with red markers indicating the locations of two selected pixels—one in the drainage network and one in the hillslope domain. It does not display entropy or modular difference values. The uniform dark margins are a visual artifact caused by boundary conditions in the model domain and the rendering of the hillshade, but they do not convey any meaningful data or affect the interpretation.
The analytical results come from the plots on the right, which show the temporal evolution of local Shannon entropy and modular difference for the selected pixels across the experiment pairs. We have updated the figure caption to clearly explain the purpose of each panel and to clarify that the dark borders in the hillshade map are not relevant to the analysis (lines 365-369).

Line-by-line comments:

Line 33: Grammar issue - Especially relevant <u>is</u> the need <u>for</u> mechanisms <u>to compare</u> multidimensional models…

Response: Suggested change made.

Line 70: I am surprised that work by Ignacio Rodriguez-Iturbe and/or Andrea Rinaldo is not mentioned here.

Response: We agree that the work by Rodriguez-Iturbe and Rinaldo provides important theoretical context for understanding landscape and river network evolution. While our study focuses on Shannon entropy, their development of Optimal Channel Networks offers a complementary framework, as both approaches are grounded in principles of minimizing local and global energy expenditure. We have now incorporated a citation to their work.

Line 81: typo near end of line ("where while…")

Response: Changed the sentence for clarification.

Line 138 and Equation 2: isn't x_bar the average of x(i,j); not just x(i) if in this equation the spatial units are indexed by (i,j) as opposed to Eq. 1.

Response: Yes, the term refers to the mean of i, j.

Line 205: missing word: "first IT is necessary"

Response: Suggested change made.

Line 211: Delete "Equation 1: maximum modular difference" (in this manuscript it is Equation 4).

Response: Suggested change made.

Line 211-214: Are there two different uses of *k* in this equation? It is used to index the consecutive matrices (*k {0,1,….}).* Then it is calculated as the maximum value of *R.* I suggest redefining one of these *k's.*

Response: Used "t" for indexing.

Line 212: Also, in Eq. 4 the second term may have an extra *m: matrices(k+1)*m[i,j].*

Response: You are right, extra "m" removed.

Line 221 (Eq. 6): should the left side of this equation by *P(xi,j)* as is noted when this section is introduced in line 205?

Response: We replaced P(xi,j) with Pt to emphasize that the calculation is being performed in a matrix-wise, elementwise fashion across each Pt rather than referencing a single scalar value.

Line 220-221: Here *i* and *j* are defined as *j = i+1.* But, everywhere else in the paper *(i,j)* is used to spatially index elements within the grids. It seems in the rest of the paper you are using either *x* or *k* to index which matrix you are using.

Response: We understand that may have been confusing, so we changed the matrix index to "t."

Line 231: Is *Xi,j* in this equation the same as *matrices[k][i,j]* in equation 4?

Response: No, they are not the same. In Equation 4, matrices[t][i,j] refers to the value of a specific element at position (i,j) within the k-th matrix, whereas xi,j is a more general notation used to refer to individual elements in the grid. Equation 4 uses these values to compute the maximum absolute difference between consecutive matrices. I have removed Equation 7, as Equation 8 already provides a clearer and more complete explanation of the entropy calculation.

Line 231: Why is *Xi,j* capitalized on one side of the equation and not the other?

Response: I have removed Equation 7, as Equation 8 already provides a clearer and more complete explanation of the entropy calculation.

Line 227: you aimed to include or you did include?

Response: If that comment was for line 276, we changed the phrasing.

Table 1: Please provide information on the boundary conditions for the model.

Response: Now included in Table 1.

Figure 2 and Line 217: What is meant by "first highest value pixel"? Just the highest pixel value in the summary statistic? Perhaps the "highest value pixel at the first time step"?

Response: We have revised the text to clarify that "first highest value pixel" refers to the first occurrence of the maximum entropy value in the grid. If multiple pixels share this maximum value, the one closest to the outlet—due to Landlab's node ordering starting from the lower left—is selected. We now include this explanation earlier in the text, before Figure 2 (lines 319–322), to improve clarity.

Figure 2: The difference in scales on the Entropy maps make it hard to compare across calculations.

Response: We changed the entropy maps to the same

Line 328: "thus enhancing our understanding of landscape dynamics". This belongs in the discussion.

Response: We removed the sentence from paragraph.

Line 336: a drainage area threshold of 3 m^2 is extremely small!

Response: The experiments were rerun with a larger domain ($9km^2$).

Figure 4: What are the values shown in the images? Elevation? Entropy? The color/scale bar is nearly illegible and doesn't seem to match the gray background of the image. I cannot interpret what these grids are showing from the figure or caption.

Response: We rewrote the caption to explain in a clearer way what is being shown. On the left, we are showing a hillshade map to illustrate the topography, overlaid in red with the locations of selected pixels in both the drainage and hillslope areas. On the right, we show the temporal evolution of local Shannon entropy and modular difference for two representative pixels—one from the stream region and one from the hillslope—across pairs of experiments. We also removed the color bar, since it is not meaningful.

Lines 347: this belongs in the discussion. I'm not sure how Figure 4 is demonstrating any spatial variability in the landscape as is claimed in this paragraph (rather Figure 2 and 3 do so).

Response: We agree that Figure 4 does not depict spatial variability across the full landscape as Figures 2 and 3 do. Instead, Figure 4 is intended to illustrate how local metrics—specifically entropy and modular difference—evolve over time at selected stream and hillslope pixels. We have revised the paragraph to more accurately reflect the purpose of this figure, emphasizing that it shows the temporal response of specific points to the uplift history rather than spatial patterns. The revised text has been moved to the discussion section (lines 516 – 522).

Figure 4: It is interesting that the hillslopes and streams look the same in terms of the temporal trends (middle and right-hand columns) in Figure 4. Would you expect more of a difference? If not, why is it important that we are looking at these two patterns separately?

Response: This resemblance is expected to some degree, as both domains are responding to the same externally imposed uplift signal and operate within a coupled landscape system. However, subtle differences—such as the relative timing and magnitude of entropy changes—still emerge across the selected matrix pairs and align with broader spatial patterns seen in Figure 3, where hillslope areas exhibit higher entropy in Experiment 2.

Analyzing these two domains separately remains important because it allows us to validate the spatial insights from earlier figures using temporal data from representative locations. While the trends may appear broadly similar, the differences in specific periods help us interpret how variability propagates differently through hillslope versus fluvial components. This approach also strengthens our interpretation of process dynamics by showing how localized signals relate to broader spatial patterns.

It is also worth emphasizing that from the perspective of potential users of these metrics as we have implemented them, whether the user wants to analyze the whole landscape or analyze it with the explicit separation of hillslopes from channels (regardless of which exact separation method is chosen) is up the user. In this sense, we wish to provide flexibility and depending on the underlying use case, separating hillslopes and channels may be useful, we added that point to the Implementation section (lines 264-267).

We added more specific sentences in the results section (lines 356-360) and added to the discussion (lines 477-489).

Figure 5: Please define what the dashed lines are supposed to mean on the Figure or in the caption.

Response: Changed the figure and took out the dashed line in zero, since it indeed did not have a purpose. The other dashed lines are referring to the p-value, with its label being shown in the right axis.

Recommendation: Major revisions

Reviewer 2:

This article proposes methods to evaluate and compare outputs from landscape evolution models (LEMs). This question is critical as no formal evaluation parameters currently exist, leaving the scientific community without established standards to assess and compare different landscape evolution scenarios and models. The methodology employs three quantitative metrics: Shannon's Entropy, Moran's I, and Geary's C, applied to outputs from a Landlab LEM across multiple scenarios.

The paper is generally well-written and presents compelling results that contribute meaningfully to the field. However, two key aspects require attention to strengthen the work.

Regarding methodology, I have significant concerns about the implementation of Shannon's Entropy in Section 3. The current approach using Equations 5 and 6 generates what is termed a "probability matrix," but this appears mathematically problematic. As the sum of matrix elements can exceed unity (one), it cannot properly function as a probability distribution in the Shannon's Entropy framework. This fundamental issue either needs correction through revised methodology or much more substantial theoretical justification explaining why this approach remains valid despite violating probability axioms.

Response: We appreciate the reviewer's concern regarding the use of Shannon's Entropy and the construction of what we refer to as "probability matrices." We would like to clarify that our methodology ensures that all values used in the entropy calculation lie within the [0,1] interval and behave as normalized weights.

Specifically, we define k as the maximum absolute difference between any two consecutive matrices at any grid point (as shown in Equation 4). This value represents the maximum possible change observed in the system. By dividing the absolute difference between matrices by k*n (where n is the number of matrices), we scale all values such that:

$$P_t = \frac{|matrices\,[t] - matrices\,[t+1]|}{k * n}$$

This ensures that:

- All Pt values are in the range [0, 1],
- The sum over all Pt for each element in the grid will not exceed 1 (since each component is divided by the global maximum change k, and aggregated over n elements
- Thus, the entropy calculation is consistent with a valid probability-like framework.

We also added clarifying text explaining the reason why we use "k" to Section 3 of the manuscript (lines 205-213) to better justify our approach.

The Discussion section would benefit from substantial reorganization to better serve readers. While the results are interesting, their practical interpretation remains unclear. Specifically, the significance of obtaining high versus low values for each metric (Entropy, Moran's I, and Geary's C) needs explicit explanation - what do these values actually tell us about landscape evolution models? For a modeling community eager to apply these standardized methods, this interpretive guidance is crucial yet currently lacking (a table could help).

I strongly recommend restructuring the Discussion into clear subsections, each dedicated to:

- Interpreting entropy results and their implications

- Explaining spatial autocorrelation findings (Moran's I and Geary's C)

This separation would provide much-needed clarity while maintaining rigorous analysis.

Response: We appreciate the reviewers comment and agree that the Discussion would benefit from clearer structure and more explicit interpretive guidance regarding the spatial metrics. To address this, we have reorganized the Discussion into three subsections: one focused on entropy (which we also subdivided further to enhance clarity), one on spatial autocorrelation (Moran's I and Geary's C), and a final section dedicated to summarizing the metrics and discussing broader applications. In this concluding section, we have added a new summary table (now Table 3) that explains the meaning and implications of high, intermediate, and low values for Shannon's Entropy, Moran's I, and Geary's C. The table also clarifies the type of output produced by each metric, helping to contextualize their use and interpretation in landscape evolution modeling.
These changes would transform the paper from presenting interesting results to providing genuinely useful, actionable standards for the modeling community. The current version

makes important progress, but realizing its full potential requires addressing these methodological and presentation issues.

I recommend major revisions to address key methodological and presentation issues.

ADDITIONAL COMMENTS

Line 24: I recommend adding a paragraph describing the challenges regarding the comparison between landscape evolution model outputs and field data such as DEMs. This should address issues of scale compatibility, temporal resolution, and validation approaches.

Response: We have added a paragraph (Lines 631-635) discussing the challenges of comparing landscape evolution model outputs with field data such as DEMs, including issues of scale compatibility, temporal resolution, and validation strategies.

Line 213: The "max value" mentioned here - is this referring to the total number of matrices? Please clarify this in the text.

Response: As explained in lines 206–208, the "max value" refers to the maximum value in the matrix R, which is constructed by computing the maximum absolute difference between each element across consecutive pairs of matrices. It is not related to the number of matrices.

Line 217: What does "n" represent here? Is this the logarithm base? Please explicitly define this parameter in the text.

Response: "n" refers to the total number of matrices in the list, not a logarithm base. This is first introduced in line 204, where we state: "The code takes a list of n matrices."

Line 233: There is ambiguity in this equation. Please specify whether this should be interpreted as: "log(Pk)" (logarithm of Pk), or "log(P)*k" (logarithm of P multiplied by k)

Response: I have removed Equation 7, as Equation 8 already provides a clearer and more complete explanation of the entropy calculation.

Additionally, please clearly indicate the base of the logarithm (is it base n, base e, or another base?).

Response: The logarithm is base 10. I have now explicitly indicated this in line 230 to clarify the expression and avoid ambiguity.

Note that the right side of the equation appears to yield a scalar value rather than a matrix, which contradicts the matrix notation on the left side. This equation requires a more detailed explanation.

Response: I have revised the equation so that the right-hand side now clearly represents a matrix calculation, consistent with the matrix notation on the left-hand side.

Recommendation: Major revisions

Round 2 Review

Comments by reviewer 1

Completed: 2025-09-11 06:06 PM
Recommendation: Minor revisions
Reviewer Comments
For author and editor
General Comments

Thank you for the revised version of the article. The quality of the text and figures has improved significantly. The methods section is now much clearer, and the discussion section is better organised.

After a careful reading, my impression is that the proposed methods can indeed serve as a complementary approach for analysing Landscape Evolution Model (LEM) outputs, such as modular differences in topography (or other parameters) through time.

That said, there remain substantial gaps in understanding what the values of entropy or autocorrelation truly represent for LEM outputs. While the scenarios presented illustrate examples of what these values might indicate in the model results, the interpretations are not yet conclusive from a broader perspective. Future work should expand on this analysis to make the proposed methods more accessible and interpretable for the wider community. I suggest including a statement to this effect in the discussion section.

Specific Comments

Line 203: This is the first appearance of the parameter Eij. I recommend describing its meaning here.

Line 315: Although intuitive, I recommend explicitly using the term Modular R in the methods section to avoid misinterpretations. Please also indicate that you are referring to Fig. 2 in the text.

Line 357: Why is the maximum entropy pixel value 1 here, while in Fig. 3 the maximum values are below 1? Please clarify.

Line 412: It is difficult to identify drainage reorganisation in Fig. 6. Please provide a more specific description in the text or add supplementary images.

Line 414: Does “range” here refer to uplift range in the scenario? Please specify.

Figure 10: This table is well done, but much of the information repeats what is already presented in the Methods section. I recommend moving it there. What would be more suitable for the Discussion is

an interpretation of what the low/moderate/high values mean for LEM results, or at least what was observed in your simulations.

Line 634: I understand this is not the focus of the paper, but such an analysis would be very interesting and could strengthen the contribution of this work.

Line 667: I recommend rephrasing this sentence. As currently written, it implies that you tested your method on datasets beyond the LEMs, which was not the case. Unless you have a reference to support this, I suggest a softer formulation, such as stating that the method has the potential to be applied beyond LEMs but still requires testing.

Comments by reviewer 2

Completed: 2025-09-23 06:50 PM
Recommendation: Major revisions
Reviewer Comments
For author and editor
The figures and equations are improved from the first draft which improves the readability of the manuscript significantly. The introduction and discussion are also improved. I have three categories of suggestions, which are detailed in line-by-line comments below.

Category 1: The organization of the Discussion has improved but I strongly suggest the Discussion be revisited and strengthened prior to publication. There are few links made in the Discussion between the results presented and existing state-of-the art modelling or surface process literature (indeed, no references in the discussion until the last section). Section 5.3 begins to do this, but these insights should be integrated within sections 5.1-5.2.

Category 2: The results and figure captions remain challenging to follow. The results read not as a narrative but as a bulleted list, each paragraph starting with "Figure X shows…" or "in Experiment 1 we did this…". The manuscript would be easier to follow instead each paragraph in the results started with a topic sentence introducing the concept to explore and then the Figures and results from the experiments were cited as supporting evidence.

Category 3: The p-values for Moran's I and Geary's C are shown in Figure 5, seem to indicate that there is no significant spatial autocorrelation, and then are not shown in any subsequent figure displaying these analyses or described in the figure caption, results, or discussion.

Line comments:

Line 314-318 – cite the figure you are referring to here

Line 319-322: "the first pixel in the grid where the maximum value occurs" – this is still confusing. The explanation in the response statement makes sense but the corresponding full explanation is not included in the revision. Clarify in the manuscript how nodes are ordered in Landlab and how that determines which node you choose to track. Indeed, the node ordering must be communicated for the reader to understand the pixel indices cited in Figure 2 (more on that below).

Figure 2 – I do not follow why the pixels chosen for Experiment 2 are different between the Modular Difference map and the Entropy Map. From the explanation, it seems like the pixel should be chosen based on the Entropy value (and not the greatest value in the Modular Difference analysis)? Is there a typo on the figure – or is something missing in the explanation for how these pixels were chosen? For example Figure 2, Experiment 2, Topography has the pixel (39,36) under the Modular difference map and (46,39) under the Entropy map.

Line 324-325: Here you state that the location of the highest entropy pixel varies for soil depth between experiments. But that appears to also be true for topography given the pixels listed below Exp 1 and Exp 2 for topography: Experiment 1, Topography Entropy is (95,98) and Experiment 2 Entropy is (46,39).

Line 335-336: This is the first place the model domain is referred to as a 'basin'. The nomenclature seems strange because you are not modelling sedimentation or stratigraphy in a depositional basin. I suppose what is meant her is 'drainage basin' because the model domain has only one outlet. I recommend rephrasing to avoid confusion.

Figure 3 caption: Awkward wording or typo ("The 'Stream Mask' row classifies streams based on a drainage threshold of, where grid nodes…). Remove the word 'of'.

Lines 339: "consistently slightly higher values in hillslope areas across all masks" -- higher than what? The fluvial part of the domain or as compared to Exp 2?

Figure 5: What are we supposed to get out of the p-value here? Why is there a negative p-value for Moran's I? With the current display of data we cannot really tell whether the p-values indicate significant autocorrelation. It looks like they do not ($p>0.05$) for both Moran's I and Geary's C. One option is to put an asterisk or a different symbol next to each of the points with a p-value <.05. Regardless of modifications to the figure the p-value is not currently discussed anywhere in the caption, results or discussion.

Figure 6 and Figure 7: No p-values are presented or shown in this figure for Moran's I or Geary's C.

The discussion is improved, but I still found the insight to be rather shallow. There are no papers cited until the final section of the discussion which still makes it challenging to link the results here to existing knowledge in surface process modelling. One example is below, but other paragraphs could be revisited in this manner:

Lines 520-526: "For example, in Experiment 2, the stream pixel shows slightly higher entropy in the first and last pairs, whereas the hillslope pixel is more variable in the middle pairs. These distinctions align with the spatial patterns observed in Figure 3 and reinforce how surface processes integrate uplift signals differently across landscape domains."

What knowledge about surface processes has been reinforced here? I think the author is assuming the reader will make a link to surface process literature, but that should be made explicit. Consider connecting this result to the tectonic geomorphology literature on the timescale of hillslope and fluvial responses to external pertubations.

Figure 10: Typo in Row "Shannon's Entropy" and Column "High Value" (entropu -- entropy).

Figure 10: Typo in Row "Geary's C" and Column" High Value" (adjecent -- adjacent).

Lines 658-661: “contrasting values reveal significant differences in spatial patterns...” Significance is not established statistically in this paper.

Recommendation by Associate Editor 2 (Roberto Fernández)

Completed: 2025-10-05 05:54 PM
Recommendation: Major revisions
Reviewer Comments
For author and editor
The authors have addressed concerns raised by the reviewers during the first review and both reviewers now agree that the manuscript has improved. The first reviewer still has some minor comments but reviewer 2 encourages major revisions to further improve the results and discussion sections. Category 1 must be addressed as linking the results to the wider literature should be done throughout the discussion. The second category mentioned by R2 needs to be addressed carefully as I believe that results sections should stand on their own without any interpretation (which belongs in the discussion). When addressing R2s request for a topic sentence, make sure it is objective in nature and not interpretive of the results in any way.
The manuscript is looking good and I encourage the authors to do the extra work recommended by the reviewers to go from a good paper to a great one.
I believe that the issues raise by R2 are in line with a minor revision but their number warrants a 'major revisions' recommendation to give the authors the necessary time to work on them.

Authors' responses to review comments

November 20, 2025

Summary:

Both reviewers acknowledge clearer text, better-organized figures, and an improved discussion. Reviewer 1 notes that the methods are now well-presented and the proposed metrics (Shannon's Entropy, Moran's I, Geary's C) are useful for analyzing Landscape Evolution Model (LEM) outputs. They highlight that interpretations of these metrics are already clearly stated in the conclusion and that future work using these metrics is discussed. Minor clarifications were suggested, including explanations of parameters, maximum entropy values, pixel selection in Figure 2, and additional guidance on drainage reorganization and Figure 10 interpretation. These were addressed by clarifying text, figure captions, and explanations in the methods.

Reviewer 2 emphasizes the need to link results to broader surface process and modeling literature (Category 1), restructure the results to begin with objective topic sentences rather than figure-driven lists (Category 2), and clarify statistical significance of Moran's I and Geary's C (Category 3). These issues were addressed by: integrating relevant literature throughout the discussion, restructuring the results section with clear topic sentences, and adding clarifications that p-values were extremely low, confirming statistically significant spatial autocorrelation. Additional line edits and corrections to figures and captions were implemented as requested.

The Associate Editor recognizes the overall improvements and agrees that addressing Reviewer 2's suggestions will strengthen the manuscript, particularly in the results and discussion sections. The manuscript now provides clearer interpretations of the metrics, incorporates an example of potential repeat DEM applications, and improves linkage to existing literature.

Anonymous Associate Editor:

The authors have addressed concerns raised by the reviewers during the first review and both reviewers now agree that the manuscript has improved. The first reviewer still has some minor comments but reviewer 2 encourages major revisions to further improve the results and discussion sections. Category 1 must be addressed as linking the results to the wider literature should be done throughout the discussion. The second category mentioned by R2 needs to be addressed carefully as I believe that results sections should stand on their own without any interpretation (which belongs in the discussion). When addressing R2s request for a topic sentence, make sure it is objective in nature and not interpretive of the results in any way.

The manuscript is looking good and I encourage the authors to do the extra work recommended by the reviewers to go from a good paper to a great one.

I believe that the issues raise by R2 are in line with a minor revision but their number warrants a 'major revisions' recommendation to give the authors the necessary time to work on them.

Recommendation: Major revisions

Reviewer 1:

General Comments

Thank you for the revised version of the article. The quality of the text and figures has improved significantly. The methods section is now much clearer, and the discussion section is better organised. After a careful reading, my impression is that the proposed methods can indeed serve as a complementary approach for analysing Landscape Evolution Model (LEM) outputs, such as modular differences in topography (or other parameters) through time.
That said, there remain substantial gaps in understanding what the values of entropy or autocorrelation truly represent for LEM outputs. While the scenarios presented illustrate examples of what these values might indicate in the model results, the interpretations are not yet conclusive from a broader perspective. Future work should expand on this analysis to make the proposed methods more accessible and interpretable for the wider community. I suggest including a statement to this effect in the discussion section.
Response: We thank the reviewer for this comment. In response, we believe that the additional experiment using DEMs acquired before and after the 2017 Mud Creek landslide in California broadens the applicability of the metrics and demonstrates their relevance beyond synthetic landscape evolution models. We also expanded the discussion to include clearer guidelines for interpreting the metrics, emphasizing how they capture distinct yet complementary aspects of landscape organization and change.
Specific Comments
Line 203: This is the first appearance of the parameter *Eij*. I recommend describing its meaning here.
Response: Suggested change made.
Line 315: Although intuitive, I recommend explicitly using the term *Modular R* in the methods section to avoid misinterpretations. Please also indicate that you are referring to Fig. 2 in the text.
Response: Suggested change made, and term reiterated in implementation (methods) section.
Line 357: Why is the maximum entropy pixel value 1 here, while in Fig. 3 the maximum values are below 1? Please clarify.
Response: In this case, the entropy is calculated between pairs of matrices, allowing the maximum entropy pixel to reach the upper limit of 1. In contrast, Fig. 3 shows the final aggregated entropy matrices, which represent the computation of entropy across all matrix pairs. This aggregation process smooths local contrasts, resulting in maximum values below 1. We have added a clarification to the text to make this distinction explicit.
Line 412: It is difficult to identify drainage reorganisation in Fig. 6. Please provide a more specific description in the text or add supplementary images.
Response: We have added a shaded orange area in Figure 6 to clearly identify the region we refer to in the text. This highlights the distal areas where drainage reorganization is pronounced, making it easier to visually assess the changes in orientation across Experiments 3A–3D.
Line 414: Does "range" here refer to uplift range in the scenario? Please specify.
Response: Yes, suggested change made.
Figure 10: This table is well done, but much of the information repeats what is already presented in the Methods section. I recommend moving it there. What would be more suitable for the Discussion is an interpretation of what the low/moderate/high values mean for LEM results, or at least what was observed in your simulations.
Response: We thank the reviewer for this suggestion. The intent of Figure 10 was to provide a clear summary of the meaning and implications of low, intermediate, and high values for Shannon's Entropy, Moran's I, and Geary's C, serving as a concise reference for interpretation within the landscape evolution modeling context. While some information overlaps with the Methods section, we

added a new column summarizing the observed patterns from our experiments, providing interpretation of how these metric values correspond to actual landscape dynamics. This addition emphasizes the practical significance of the metrics and supports reader understanding of the results.

Line 634: I understand this is not the focus of the paper, but such an analysis would be very interesting and could strengthen the contribution of this work.

Response: To illustrate the potential broader applicability, we have included an example (Experiment 4) using real-world DEMs (Mud Creek landslide site) and briefly discuss the observations in the results and discussion.

Line 667: I recommend rephrasing this sentence. As currently written, it implies that you tested your method on datasets beyond the LEMs, which was not the case. Unless you have a reference to support this, I suggest a softer formulation, such as stating that the method *has the potential* to be applied beyond LEMs but still requires testing.

Response: Since we now demonstrate its application using real-world DEMs (Mud Creek landslide site), we kept the paragraph as it was. This example illustrates the method's broader applicability, though further testing would help fully assess its generalizability.

Recommendation: Minor revisions

Reviewer 2:

The figures and equations are improved from the first draft which improves the readability of the manuscript significantly. The introduction and discussion are also improved. I have three categories of suggestions, which are detailed in line-by-line comments below.

Category 1: The organization of the Discussion has improved but I strongly suggest the Discussion be revisited and strengthened prior to publication. There are few links made in the Discussion between the results presented and existing state-of-the art modelling or surface process literature (indeed, no references in the discussion until the last section). Section 5.3 begins to do this, but these insights should be integrated within sections 5.1-5.2.

Response: We appreciate this suggestion and have revised the discussion to better situate our findings within the broader context of landscape evolution research.

Category 2: The results and figure captions remain challenging to follow. The results read not as a narrative but as a bulleted list, each paragraph starting with "Figure X shows…" or "in Experiment 1 we did this…". The manuscript would be easier to follow instead each paragraph in the results started with a topic sentence introducing the concept to explore and then the Figures and results from the experiments were cited as supporting evidence.

Response: We restructured the results section so that each paragraph now begins with a topic sentence that introduces the main concept being addressed, followed by descriptions of the corresponding figures and experimental outcomes as supporting evidence. We believe this should improve the clarity of the text and thank the reviewer for the suggestion.

Category 3: The p-values for Moran's I and Geary's C are shown in Figure 5, seem to indicate that there is no significant spatial autocorrelation, and then are not shown in any subsequent figure displaying these analyses or described in the figure caption, results, or discussion.

Response: In Figure 5, p-values are included to indicate the statistical significance of the Moran's I and Geary's C metrics. In subsequent figures, we focus on the metric values themselves to highlight spatial patterns and trends over time, as these are the primary variables of interest. All p-values in our analyses were extremely low (effectively 0), indicating that the observed spatial autocorrelation patterns are highly significant. We have added a clarification in the results section to note that p-values were consistently very low, confirming that the patterns observed are statistically significant.
Line comments:
Line 314-318 – cite the figure you are referring to here
Response: Suggested change made.
Line 319-322: "the first pixel in the grid where the maximum value occurs" – this is still confusing. The explanation in the response statement makes sense but the corresponding full explanation is not included in the revision. Clarify in the manuscript how nodes are ordered in Landlab and how that determines which node you choose to track. Indeed, the node ordering must be communicated for the reader to understand the pixel indices cited in Figure 2 (more on that below).
Response: We have added the explanation of Landlab's node ordering to the manuscript.
Figure 2 – I do not follow why the pixels chosen for Experiment 2 are different between the Modular Difference map and the Entropy Map. From the explanation, it seems like the pixel should be chosen based on the Entropy value (and not the greatest value in the Modular Difference analysis)? Is there a typo on the figure – or is something missing in the explanation for how these pixels were chosen? For example Figure 2, Experiment 2, Topography has the pixel (39,36) under the Modular difference map and (46,39) under the Entropy map.
Response: The purpose of Figure 2 is to illustrate the differing behaviors of the "maximum value pixels" when comparing entropy and modular difference. For each map, the pixel with the highest value is tracked independently, rather than selecting pixels based solely on the entropy map. In Experiment 1, the maximum value pixels happen to coincide, but in Experiment 2 they differ, as noted in the figure. We have clarified this in both the figure caption and the main text to ensure that this distinction is clear.
Line 324-325: Here you state that the location of the highest entropy pixel varies for soil depth between experiments. But that appears to also be true for topography given the pixels listed below Exp 1 and Exp 2 for topography: Experiment 1, Topography Entropy is (95,98) and Experiment 2 Entropy is (46,39).
Response: We now included that observation in the text.
Line 335-336: This is the first place the model domain is referred to as a 'basin'. The nomenclature seems strange because you are not modelling sedimentation or stratigraphy in a depositional basin. I suppose what is meant her is 'drainage basin' because the model domain has only one outlet. I recommend rephrasing to avoid confusion.
Response: Suggested change made.
Figure 3 caption: Awkward wording or typo ("The 'Stream Mask' row classifies streams based on a drainage threshold of, where grid nodes…). Remove the word 'of'.
Response: Suggested change made.
Lines 339: "consistently slightly higher values in hillslope areas across all masks" -- higher than what? The fluvial part of the domain or as compared to Exp 2?
Response: When the domain was smaller (in a previous version of the manuscript), the hillslope entropy values were higher than the stream values, which is not observed anymore. We removed that sentence.

Figure 5: What are we supposed to get out of the p-value here? Why is there a negative p-value for Moran's I? With the current display of data we cannot really tell whether the p-values indicate significant autocorrelation. It looks like they do not ($p>0.05$) for both Moran's I and Geary's C. One option is to put an asterisk or a different symbol next to each of the points with a p-value <.05. Regardless of modifications to the figure the p-value is not currently discussed anywhere in the caption, results or discussion.

Response: Comment addressed in "Category 3" observation.

Figure 6 and Figure 7: No p-values are presented or shown in this figure for Moran's I or Geary's C.

Response: Comment addressed in "Category 3" observation.

The discussion is improved, but I still found the insight to be rather shallow. There are no papers cited until the final section of the discussion which still makes it challenging to link the results here to existing knowledge in surface process modelling. One example is below, but other paragraphs could be revisited in this manner:

Lines 520-526: *"For example, in Experiment 2, the stream pixel shows slightly higher entropy in the first and last pairs, whereas the hillslope pixel is more variable in the middle pairs. These distinctions align with the spatial patterns observed in Figure 3 and reinforce how surface processes integrate uplift signals differently across landscape domains."*

What knowledge about surface processes has been reinforced here? I think the author is assuming the reader will make a link to surface process literature, but that should be made explicit. Consider connecting this result to the tectonic geomorphology literature on the timescale of hillslope and fluvial responses to external pertubations.

Response: We included links to the literature throughout the discussion, as suggested.

Figure 10: Typo in Row "Shannon's Entropy" and Column "High Value" (entropu -- entropy).

Response: Suggested change made.

Figure 10: Typo in Row "Geary's C" and Column" High Value" (adjecent -- adjacent).

Response: Suggested change made.

Lines 658-661: "contrasting values reveal significant differences in spatial patterns…" Significance is not established statistically in this paper.

Response: We revised the text to clarify that the term "significant" was used in a descriptive, not statistical, sense. We replaced the term by "notable."

Recommendation: Major revisions

Round 3 Review

Recommendation by Associate Editor 2 (Roberto Fernández)

Completed: 2025-11-25 12:54 AM
Recommendation: Accept
Reviewer Comments
For author and editor
The authors have addressed all comments and concerns expressed in the previous round. The manuscript may be accepted but the authors need to address a few minor comments:

Certain figures rely on color to convey meaning and this must be avoided. The graphs where the two vertical axes represent two different variables rely on color to distinguish between variables. Some have a solid line and dashed line to differentiate but no legend so a reader with color vision deficiency would not be able to understand the figures.

Certain sentences need a through read to make sure they are correct. Maybe after the track changes, some words were left behind or removed and a couple of sentences need to be corrected. For example, line 417. Missing word before sensitive?

Variables in the text (Line 91) do not have a sub index but the equation does. Check for these issues throughout.

Figure 11 is a Table even if not labeled that way. Either the authors need to move Table 1 or Table 2 to the appendix to keep this one or change it to a proper figure. Given its importance, I would try to create a better visual. Perhaps the authors can explore options with Napkin.ai and use it as a source of inspiration to summarize the contents of Figure 11 (Table 3) and present them in a unique way that is engaging for readers.