Reply to referee 1

Dear Prof. Miller

Thank you for carefully reading our manuscript and providing valuable comments. Your suggestions will make our manuscript more complete.

We reply to each of your comments below (in blue text)

Comment: This manuscript presents a novel approach for quantifying soil complexity using digital soil mapping products. It demonstrates the utility for mapping a Δ complexity metric for soil properties across depths. I particularly appreciate the integration of digital soil mapping with soil-landscape analysis, as this approach has the potential to reveal connections between spatial patterns in soil properties and the processes that shape them. The methodology is innovative and has the potential to advance quantitative assessments of soil heterogeneity.

Authors: We sincerely appreciate the reviewer's thorough, insightful and constructive feedback on our manuscript, as well as your positive recognition of the novelty and value of our research work. We are especially encouraged by your positive remarks regarding the novelty of our approach and the potential of integrating digital soil mapping with soil-landscape analysis. We agree with your thoughtful evaluations and critical suggestions, and we have carefully studied every comment point by point to revise and strengthen our manuscript comprehensively.

Despite these strengths, I found the manuscript difficult to follow conceptually. Much of the confusion stems from ambiguity in the use of key terms and confounding factors that cloud the interpretation of connections between the proposed metric and soil evolutionary concepts. In my view, the manuscript is methodologically innovative but theoretically underdeveloped. The work would benefit from a clearer conceptual framework that distinguishes soil heterogeneity, complexity, development, convergence, divergence, and sensitivity to change.

We completely acknowledge your helpful comments that the current manuscript lacks a rigorous, well-organized conceptual framework, which leads to obscure logical reading and ambiguous definition of core terminologies. Accordingly, we will provide clear definitions and elaborations for all core terminology adopted throughout this manuscript:

1. Complexity: We calculate complexity as the spatial standard deviation of a soil property x at a certain depth d within the catchment.
2. Δ Complexity: The Δ complexity is calculated as the difference in complexity value between two layers, divided by the corresponding depth interval between the two layers.
3. Divergence: Negative values of the Δ complexity indicate a divergent change in soil properties, meaning that the lower layer is more heterogeneous than the upper layer and that its complexity value is larger than that of the upper layer.
4. Convergence: Positive values of the Δ complexity indicate the convergent change in soil properties, meaning that the lower layer is more homogeneous than the upper layer and that its complexity value is smaller than the upper layer. Convergent and divergent, the Δ complexity they describe is merely a superficial manifestation, and it only serves as a metric for determine sensitive areas.
5. Soil heterogeneity: Soil heterogeneity in this manuscript encompasses two dimensions. The first dimension is vertical heterogeneity, which refers to the positive and negative variations of Δ

complexity along the depth profile. The more frequent the shifts between positive and negative values, the greater the heterogeneity. The other dimension is horizontal heterogeneity. Horizontal heterogeneity is, in essence, vertical heterogeneity over a larger depth interval (specifically involving the shallow layer: the difference in complexity between 0-5 cm and 20-25 cm, and the deep layer: the difference in complexity between 25-30 cm and 35-40 cm). We visualize disparities in △complexity between shallow and deep soil layers via horizontal spatial maps, which enables the analysis of soil redistribution processes.

6. Sensitive area: The sensitive area is determined by horizontal heterogeneity and underlying surface conditions. Specifically, △complexity is calculated for both shallow (subtract 20-25 cm from 0-5 cm) and deep (subtract 35-40 cm from 25-30 cm) layers across all catchments in the study area. Catchments exhibiting different transformation directions with larger discrepancies between deep and shallow layers are then evaluated in combination with underlying surface conditions—including parent material, topography, and vegetation—to ultimately determine the three sensitive areas.

7. Highly sensitive area: Highly sensitive area is further identified from the previously determined sensitive areas based on vertical heterogeneity. Specifically, △complexity is calculated layer by layer from 0-40 cm at 5 cm intervals. More frequent shifts between positive and negative △complexity values from topsoil to subsoil indicate stronger vertical heterogeneity. Catchment within sensitive areas that exhibited strong vertical heterogeneity were designated as highly sensitive area.

8. Soil landscape evolution: Soil landscape evolution refers to the a certain period of time, dynamic transformation of the geomorphic surface (the landscape) and its 0-40 cm soil cover spatially distributed (the pedosphere) under the interplay of soil forming factors. Due to the lack of chronological data in this study, the manuscript tries to avoid mentioning "evolution" as much as possible, but will cautiously mention time related matters in the future, such as our cautious belief that the deeper the soil, the older the age. Some of the "evolution" mentioned in the manuscript is used to describe Van der Meij's method, while others refer to "soil landscape evolution". The term "evolution" in the discussion section may be inappropriate, and we will make the necessary changes.

If I understand the authors correctly, the manuscript implies that increasing vertical complexity reflects greater evolutionary sensitivity. However, I do not think this interpretation is fully supported by the string of literature that has developed this framework. I suspect some aspects of these concepts have been slowly lost as they have been applied in different ways across a sequence of papers The present manuscript would benefit from stronger engagement with the work of Johnson and Watson-Stegner (1987) and Johnson et al. (1987), as well as revisiting the approaches in Phillips (1999, 2000, 2017).

We appreciate the reviewer's careful reading with our manuscript and your call for stronger connections to the foundational literature. We have carefully revisited the work of Johnson and Watson-Stegner (1987), Johnson et al. (1987), and Phillips (1999, 2000, 2017) to clarify our conceptual framework and address the concern that we may be conflating heterogeneity with pedogenic advancement. The detailed reply is presented as follows.

The meaning expressed in the manuscript is that the stronger the vertical heterogeneity of a catchment, that is, the more frequent shifts between positive and negative △complexity values along the depth profile, the more sensitive the catchment will be. This sensitivity does not refer to advanced pedogenesis, but rather to the soil in this region being more sensitive to environmental changes.

The core insights drawn from literature are as follows:

Johnson and Watson-Stegner (1987) proposed a comprehensive model of soil genesis based on the observation that soils evolve along two coacting pathways: progressive (P) and regressive (R). Progressive pedogenesis promotes organized, differentiated profiles through developmental upbuilding and deepening. Regressive pedogenesis promotes simplified, rejuvenated profiles through retardant upbuilding and removals. A critical insight of their model is that both pathways operate concurrently in all soils, and that either may predominate at different times. The soil evolution model captures the multi gene characteristics of soil and integrates these factors. Our method also identified multiple factors that affect soil.

Johnson et al. (1987) refined and extended the concept of pedoturbation (soil mixing) within the evolutionary framework established by Johnson and Watson-Stegner (1987). This paper said that "Mixing processes may function either proisotropically or proanisotropically. This can be true not only for different soils on a landscape, but also for the same soil at different times during its evolution." The paper notes that the general categories of pedoturbation not only reflect the functioning by which mixing agents and vectors operate, but also embody the adjustments made by the soil system in response to such mixing processes. This is consistent with the viewpoint in our manuscript that "sensitive areas are areas that are more sensitive to environmental changes (pedoturbation)", that is, the soil system will respond and adjust to external changes. The passage also stresses that whether a given soil has undergone proisotropic (regressive) or proanisotropic (progressive) pedoturbations can only be ascertained from its observable and quantifiable morphological features and properties, supplemented as needed by laboratory-derived characterization data. This aligns well with our manuscript's approach of analyzing through measured field data.

Phillips (1999) identifies at least eleven distinct definitions of self-organization relevant to landscapes. "Despite these differences, most self-organization concepts can be broadly aggregated into two categories: those concerned with the evolution of order and regularity in the aggregate or ensemble properties of the landscape, and those concerned with the differentiation of landscapes into more diverse spatial units."

Phillips (2000) applies the conceptual framework developed in Phillips (1999) to soil profiles, using Kolmogorov entropy (K-entropy) as a measure of self-organization. Because entropy is an indicator of "organization" rather than "developmental degree". Accordingly, our Δ complexity also serves as an indicator for quantifying self-organization. Pedogenic development may either increase or decrease entropy and Δ complexity. Meanwhile, the direction of entropy change is not equal to the direction of evolution. It breaks the simple linear thinking of "entropy increase=progressive, entropy decrease=regressive". Thus, an entropy increase can be regressive (e.g., erosion, bioturbation), but an entropy decrease can also be progressive (e.g., horizon development concentrating thickness). Phillips (2000) essentially reminds us that entropy—specifically, Kolmogorov entropy or profile entropy—is a statistical measure of order (or information content) in the vertical arrangement of horizons, rather than a value judgment about the direction of evolution. A complex soil may exhibit either high or low entropy; an increase in entropy could reflect regressive pedogenesis (e.g., homogenization or haploidization), or it could be a progressive pedogenesis (e.g., the formation of new horizons through horizon differentiation). Therefore, the high or low entropy cannot be simply equated with the "progressive" and "regressive" of soil evolution.

Phillips (2017) defines soil complexity in multiple dimensions. The primary focus is on "the complexity of soil types within soil landscapes". More broadly, the paper states that "soil complexity is inextricably intertwined with soil variability, irregularity, and predictability" and is "fundamental to pedodiversity". Complexity encompasses both "complexity of individual pedons or weathering profiles" and "complexity of the soil cover (soil maps or landscapes)". The paper explicitly acknowledges that complexity has "many important aspects and implications, including pedodiversity, spatial variability, and pedometrics, as well as evolutionary pathways".
After reviewing the aforementioned literature, we recognize that the concepts of convergence and divergence adopted in our manuscript differ substantially from those defined in Phillips' framework. Our use of these two terms merely describes statistical patterns of soil properties across distinct soil layers, rather than denoting evolution direction.

The above represents our understanding after revisiting the relevant papers, and we will incorporate their definitions in our Introduction.

The emphasis on vertical heterogeneity strongly echoes Johnson and Watson-Stegner's evolutionary model of pedogenesis. They argued that soils simultaneously experience progressive differentiation, regressive simplification, proanisotropic (horizon-forming), and proisotropic (mixing and homogenizing) processes. The present manuscript does a good job of treating both positive and negative changes in complexity as meaningful. However, I think the present manuscript sometimes conflates heterogeneity with more advanced pedogenesis. Profile complexity can be increased or decreased through pedogenic processes, erosion, deposition, faunal activity, tree throw, and anthropogenic disturbance. For example, the authors mention the interception of wind-blown sediments that occurs to different degrees within the study area. A slow rate of this process would result in a progressive pathway of developmental upbuilding as pedogenesis assimilates the influx of sediment. However, if the rate of the same process overwhelms pedogenesis, then the soil shifts to a regressive pathway of retardant upbuilding. While many of these mechanisms are acknowledged, the interpretation often treats Δ complexity as direct evidence of soil evolution rather than simply evidence of changing spatial organization.
Thank you for reading carefully. Sorry for the confusion caused by not defining important terms in the manuscript. More frequent shifts between positive and negative Δ complexity values from topsoil to subsoil indicate stronger vertical heterogeneity. High heterogeneity indicates that the area exhibits a stronger response to environmental changes, rather than reflecting advanced pedogenic processes. The overextended interpretation in the Discussion section may have led to misunderstandings, and we will restructure the framework and revise the content of Discussion section accordingly. Indeed, you are correct. Δ complexity is merely an index for quantifying spatial self-organization. In other words, Δ complexity only reflects one "appearance", and there

are different possibilities for the reasons that lead to this appearance. To further determine, more soil investigations and measurements of other soil properties are required. The content mentioned in the manuscript is based on literature review.

Phillips' framework distinguishes complexity from convergence and divergence. For Phillips, divergence refers to similar initial conditions producing increasingly different outcomes through time, whereas convergence refers to different initial conditions producing increasingly similar outcomes. The present study measures contemporary spatial structure and infers evolutionary tendencies from that structure. Without temporal constraints (observations made over time or via soil dating for space-time substitution) to support independent reconstruction of developmental pathways, the manuscript does not directly demonstrate convergent or divergent evolution. The authors hint at this in the section on potential refinements for future work, but I think without those controls, the interpretation of trajectories from Δ complexity becomes overextended. According to Phillips, complexity cannot be equated to divergence or convergence in the absence of evidence of trajectories.

Yes, the present study measures contemporary spatial structure. But "convergent" and "divergent" in the manuscript do not mean temporal evolutionary tendencies, as we have indicated in the definition of convergence and divergence on page 1. In this article, "convergent" and "divergent" have little correlation with time, only refer to how complexity changes from upper to lower layers. We will amend any improper applications of the terms "convergent" and "divergent" that appear in the manuscript.

Indeed, the absence of chronological data in this manuscript renders our inferences regarding why soil properties exhibit such a state somewhat overreaching. We will accordingly curtail the speculative content in our discussion. The term "evolutionary tendencies" is more like a prediction for the future, but this manuscript focuses more on analyzing contemporary spatial data of the soil.

The manuscript appears to assume that highly heterogeneous landscapes are more sensitive. I'm not sure I am thinking about sensitivity the same way the authors are, but I don't think heterogeneity always indicates sensitivity. Highly heterogeneous systems may represent stable long-term equilibria, while relatively homogeneous systems may be close to critical thresholds and highly responsive to external forcing. Consequently, the conceptual leap from high Δ complexity to high sensitivity doesn't seem fully justified.

Thank you for reading carefully. The high heterogeneity in this manuscript refers to the frequent positive and negative changes in the value of Δ complexity along the depth. While high heterogeneity is not synonymous with high sensitivity, in the context of our study region it nevertheless reflects more intense and more frequent soil responses to environmental forcing. Δ Complexity quantifies the structural complexity difference of soil profiles under varying environmental drivers (water erosion deposition, wind erosion deposition, moisture fluctuation, bioturbation, human activities etc.). High vertical heterogeneity means Δ complexity alternating positive and negative values at short depth intervals. Mechanistically, this vertical fluctuation arises when environmental forcing factors do not act uniformly through the soil profile. While high heterogeneity is not synonymous with high single-factor sensitivity, in the context of our study region it nevertheless reflects more intense and more frequent soil responses to environmental forcing, as alternating depth-specific processes. Such as erosion, deposition, bioturbation repeatedly push Δ complexity to swing between positive and negative values across narrow soil intervals, driving constant reconfiguration of soil structural complexity along the profile.

Much of the justification for the proposed approach and the assumptions underlying its interpretation appear only in the discussion. I recommend moving more of this rationale into the introduction and methods, particularly regarding why Δ complexity should be interpreted as an indicator of sensitivity and whether greater heterogeneity necessarily implies greater sensitivity to future change. I also think it would help to establish how the proposed method will be validated, such as through a falsifiable hypothesis or a predefined test that allows for more objective evaluation.
Thank you for this valuable reminder. We apologize for failing to clarify the definitions and interrelationships of key terms in the Introduction. We will relocate some contents from the Discussion section to the Introduction and Method sections. Additionally, we will add detailed elaborations in the Introduction to explain why Δ complexity can serve as an indicator of sensitivity, as well as whether greater heterogeneity corresponds to higher sensitivity. However, in our study, we use only the current Δ complexity values to indicate the spatial heterogeneity of soil properties, in order to assess whether soils in a given region respond more strongly to environmental changes. We do not attempt to infer future evolutionary trends.

Overall, the methodology is innovative and potentially valuable for soil-landscape analysis. The Δ complexity metric appears capable of revealing meaningful patterns in the spatial organization of soil properties. However, the manuscript would benefit from a stronger theoretical foundation and a clearer distinction among complexity, development, convergence-divergence, and sensitivity, with additional theoretical development to address multiple developmental pathways. In interpreting the Δ complexity metric, there needs to be a better explanation, preferably with data, that constrains the potential for the current state of complexity to have been produced by different pathways.
We appreciate your favorable evaluation of our methods. We will further strengthen the theoretical framework of this manuscript and clarify the definitions and interrelations of all relevant terminology in the Introduction. Nevertheless, we sincerely apologize that we are unable to provide additional supporting data to validate the interpretation of Δ complexity.

General Comments

Comment: The paper relies on several concepts, including heterogeneity, complexity, convergence, divergence, soil development, and sensitivity, but not all are defined in the introduction. Adding consideration of progressive and regressive pathways would also help flesh out the complex interactions among concepts in this study. It would be beneficial to clearly distinguish these terms.
Authors: Thank you for your reminder. We will define the main terms in the Introduction section of the manuscript. The meanings of each term in the manuscript we have indicated in pages 1-2.

In addition to better defining terms and concepts, the introduction should do more to establish the research gap the study addresses and provide a clearer foundation for the three research objectives.
Authors: Thank you for reading carefully. We will clarify the shortcomings of existing research in the introduction to provide a clearer foundation for the three research objectives.

Be explicit throughout the manuscript when complexity refers to lateral variation versus vertical differentiation.
Authors: We appreciate the reviewer's careful reading. The previous manuscript did not clearly explain horizontal variation and vertical differentiation. We will specify in the manuscript the scenarios corresponding to horizontal variation and vertical differentiation, respectively.

Additional wording is sometimes needed to ensure that terms such as "shift" are interpreted as the right kind of locational difference rather than a temporal change.
Authors: We appreciate your reminder. We apologize for not clearly defining the term "shift" in our manuscript. In the revision, we will add a clear explanation. Specifically, "shift" in our study refers to a change in the sign of △complexity, such as, a transition from positive to negative.

Please define the depths that were considered to be shallow versus deep.
Authors: Thank you for your reminder. The manuscript involves the spatial heterogeneity of soil △complexity between shallow and deep layers. The soil △complexity was calculated for shallow layer 0-25 cm (a, 0 – 5 cm and 20 – 25 cm, two layers) and deep layer 25-40 cm (b, 25 – 30 cm and 35 – 40 cm, two layers).

For the catchment analysis, the order of operations wasn't obvious to me in the methods. Was the lateral complexity calculated for the catchments first and then compared vertically, or was the vertical complexity calculated on a cell-by-cell basis and then summarized for the catchment?
Authors: We appreciate your comments. We apologize for the unclear description of the operational sequence in our methodology, and we will revise this part accordingly. △complexity is calculated on a per catchment basis. Horizontal heterogeneity and vertical heterogeneity are essentially the difference in complexity between the upper and lower layers of the same catchment divided by the depth interval. However, the calculation interval for horizontal heterogeneity is large (Specifically, △complexity is calculated for both shallow (subtract 20-25 cm from 0-5 cm) and deep (subtract 35-40 cm from 25-30 cm) layers across all catchments in the study area), and the △complexity of all shallow and deep catchment areas in the study area is presented in the form of a horizontal map, which facilitates the analysis of soil redistribution. Vertical heterogeneity is calculated every five centimeters for △complexity, and only three typical catchment trajectories are shown.

Specific Comments

Comment: L39 – Would be useful to define the type of scale being referred to here. Are the differences in these studies the extent or the unit of analysis?
Authors: Thank you for your reminding. Temme's research is indeed at the landscape scale. We

will define the types of scales in the introduction. Whether a study involving soil profiles can be considered to operate at the landscape scale depends on whether there are landscape-level linkages among the profiles—for example, through a soil catena.

Comment: L40 – Is a chronosequence study not a study of the landscape?
Authors: We appreciate the reviewer's careful reading. This is both a chronosequence study and a landscape scale study. Whether a study involving soil profiles can be considered to operate at the landscape scale depends on whether there are landscape-level linkages among the profiles, for example, through a soil catena.

Comment: L48 – Is this implying space-time substitution or something else?
Authors: We appreciate the reviewer's careful reading. We will add corresponding descriptions here. This sentence will be changed to "With appropriate adjustments, the evolutionary pathway method can be used to identify soil spatial heterogeneity at multiple depths, which to some extent characterizes a few trajectories about time, such as our cautious belief or assumption that the deeper the soil, the older the age. And then determine both horizontal and vertical soil heterogeneity."

Comment: L51 – Please elaborate on how understanding soil heterogeneity is helpful in regions with significant soil variation.
Authors: We appreciate for your comment. It will change to "Where significant soil variation drives both soil landscape and ecosystem processes, a nuanced understanding of soil heterogeneity is particularly valuable—for instance, scientific evaluation of soil quality and development trends, guiding site-specific land management."

Comment: L67 – The study area figure is helpful, but this text could be more descriptive to help us readers better understand the context. For example, pointing out that the QD are limited to some upland flats in the southeast portion of the study area.
Authors: Thank you for your reminder. A description of the parent material distribution will be included in the study area section.

Comment: L69 – Does this create a potential for developmental or even retardant upbuilding of the soil with influx of new sediment?
Authors: Thank you for your valuable comments. Indeed, there is aeolian accumulation of sand here, which supplies fresh parent materials for pedogenesis. We will include this in the manuscript.

Comment: L86 – Please add a note about why not all samples could be taken. This could be especially important if some locations had bedrock at depths less than this study's soil core depth of 40 cm.
Authors: We appreciate the reviewer's careful reading. Indeed, as you said, there are some sample points with bedrock depths less than 40 cm, so not all depths were sampled for all locations.

Comment: L95 – Suggest replacing "laser particle sizer" with "laser diffraction" to be more explicit about the method.

Authors: We will change “laser particle sizer” to “laser diffraction”.

Comment: L96 – Please describe the pretreatment method for removing the organic matter.

Authors: The procedure for removing soil organic matter is described as follows: Weigh approximately 0.5 g of soil sample and transfer it into a beaker. Moisten the sample with distilled water, then add 30% hydrogen peroxide. Heat the mixture on a hot plate. Supplement additional 30% hydrogen peroxide when gas bubbling subsides. The complete removal of organic matter is confirmed when no further bubbles are generated upon peroxide addition. After oxidation is complete, add distilled water to the soil suspension and let it stand for sedimentation. Siphon off the supernatant, and repeat the washing process 4 to 5 times until the supernatant reaches neutral pH.

Comment: L98 – Was this three measurements of the same subsample, or three subsamples pulled separately from the whole sample?

Authors: It is three measurements of a sample.

Comment: L99-100 – Do the authors have any concern about compatibility with pipette method results? Many studies have shifted the clay-silt break from 0.002 to 0.006 mm (American system) to match the results of those methods better.

Authors: Thank you very much for your careful reading of this manuscript. Although the compatibility between laser diffraction and pipette methods was not addressed in the present study, however, our research group is well aware of the measurement discrepancies between these two approaches and has conducted corresponding comparative studies (Zhu et al., 2018). With 1,317 samples requiring soil particle size distribution and magnetic susceptibility measurements, the pipette method would entail considerable time and labor. Therefore, after comprehensive consideration, we opted for the laser diffraction method, which has gained increasingly widespread adoption in soil research in recent years. And we did not deliberate extensively on the classification criteria for soil particle size distribution.

Comment: L122 – The elevation data probably wasn't derived from the DEM, because the DEM is elevation data. In any case, please cite the source of the DEM.

Authors: The source of DEM is from NASA.

Comment: L114-131 – Results of the validation for the digital soil mapping are presented in the results, but the methods of that validation should be provided in the methods section. Please provide the conditions of the validation data set, including how the validation points were selected and how many there were.

Authors: Thank you for your careful reading. We will add the validation method in the Methods section. For the mapping results of soil particle size distribution and magnetic susceptibility, this study selected the coefficient of determination (R^2) and root mean square error (RMSE) as evaluation metrics. The calculation formulas are as follows:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n-1}}$$

Where $\hat{y}_i$ represents the predicted values of the QRF model on the validation set, y_i represents the observed soil property values, $\bar{y}$ represents the mean of the observed values, and n represents the number of sampling points.

Thank you for your reminder. We will add the selection method for the validation set to the Methods section. In this study, soil samples from each layer were divided into training and validation sets at a ratio of 4:1. The mapping accuracy was evaluated by comparing the predicted values with the measured values for the validation set. The validation samples were selected by taking into account parent material, topography, vegetation, and other relevant factors.

Comment: L134 – In connecting the digital soil maps with analysis of variability, is there any danger of the digital soil mapping product having more or less variability due to modeling constraints? For example, could there be a tendency towards producing more homogeneous maps at greater depths due to fewer samples and weaker relationships with covariates measured on the surface?

Authors: Thank you for your careful reading and thoughtful feedback. We agree with your perspective. While environmental covariates such as land use and NDVI are more closely associated with the surface soil layer, other covariates—including slope, aspect, and topographic positions are more tightly linked to soil at the 40 cm depth. Notably, parent material exerts a far more profound influence on subsoil at 40 cm. This is clearly reflected in the covariate importance rankings for predicting particle size distribution and magnetic susceptibility at the 35 – 40 cm interval, where parent material ranked second and third, respectively.

Comment: L142-143 – I'm not sure "divergent" and "convergent" are appropriate terms for the vertical heterogeneity. For lateral heterogeneity, we could assume some homogeneity with the original parent material, with spatially variable soil-forming factors driving soil properties to diverge. Thinking about changes in soil-forming factors, I would guess that stronger pedogenesis drives greater heterogeneity at the surface. At the same time, the deposition of new sediment dampens pedogenesis and makes the shallower depths more homogeneous. The authors should explain their rationale for using these interpretive terms, rather than simply extending the use of terms from previous papers to describe different things.

Authors: Thank you for your comment. Negative values of the △complexity indicate a divergent change in soil properties (the lower layer is more heterogeneous than the upper layer, the lower layer complexity value is larger than the upper layer). Positive values of the △complexity indicate the convergent change in soil properties (the lower layer is more homogeneous than the upper layer, the lower layer complexity value is smaller than the upper layer). The convergent and divergent in this manuscript tend to express how the indicator of △complexity, which describes the degree of spatial self-organization of soil properties, changes. We will amend any improper applications of the terms "convergent" and "divergent" that appear in the manuscript.

Comment: L152 – How were catchments determined? What were their resulting sizes?

Authors: Thank you for your comment. The catchment boundaries in this manuscript were delineated by processing DEM data via the Spatial Analyst Tools ->Hydrology within ArcGIS. Most catchment areas are between 1-4.9 km^2.

Comment: L156 – What are these underlying surface conditions?

Authors: The underlying surface conditions include vegetation, terrain, and parent material.

Comment: L159 – Should "changes" be "change"?

Authors: Thank you very much for your reminder. The word "changes" will change to "change" in the text.

Comment: L159-160 – I don't understand this sentence.

Authors: Thank you for your comment. Sorry for the confusion caused by the unclear expression of this sentence. We will reorganize the language. Let me rephrase: the statement "large vertical heterogeneity" means that Δ complexity exhibits frequent positive and negative fluctuations or shifts, which indicates that the soil properties of different layers in the same catchment vary greatly.

Comment: L160 – Please remove all extra periods, here and throughout the manuscript.

Authors: We will further check the language and punctuation in the manuscript.

Comment: L170-175 – Is there a potential connection between the uncertainty index and variability?

Authors: We appreciate the reviewer's careful reading. The calculation formula for Uncertainty Index (UI) in the manuscript is:

$$UI = \frac{qp_{0.9} - qp_{0.1}}{qp_{0.5}}$$

Here, $qp_{0.1}$, $qp_{0.5}$, $qp_{0.9}$ denote the predicted values at the 0.1, 0.5, and 0.9 quantiles, respectively. A larger UI indicates a higher level of uncertainty.

The numerator $qp_{0.9} - qp_{0.1}$, represents that under the current characteristic conditions, 80% of the true dependent variable will fall within this interval, which is the absolute uncertainty amplitude. The larger the value, the wider the fluctuation range of the variable itself. Denominator $qp_{0.5}$ represents the central level of the conditional distribution. Dividing the numerator by the denominator normalizes the result using the median, thereby eliminating the effects of dimensionality and numerical magnitude. The essence of UI is the proportion of 80% of the data's fluctuation amplitude to the median predicted value at the current prediction level. The smaller the UI, the lower the uncertainty.

Comment: L185 – No need to change terminology so often. In several instances in this paper, I have been uncertain whether different terms were describing the same thing or whether there was an important difference. In this case, I am more confident that there is no difference between environmental covariates and environmental predictors, so why use different terms?

Authors: Thank you for your careful reading. I apologize for using different terms for the same concept, which made it difficult to follow. I will use the same term “environmental covariates” in the manuscript.

Comment: L190 – Please define abbreviations that appear in Figure 3.

Authors: Thanks for your feedback. Abbreviations in Figure 3 are provided in Table 1. For ease of reading, we will restate all abbreviations in the caption below Figure 3. Chang to “The abbreviations are defined as follows: PM stands for parent material, Ele represents elevation, Asp denotes slope aspect, Slope refers to slope gradient, Plancur stands for plan curvature, Profcur represents profile curvature, WI denotes wetness index, TP refers to topographic positions, LU is short for land use, and NDVI stands for normalized difference vegetation index.”

Comment: L209 – How does this relate to our understanding of soil geomorphology in floodplains? As a starting point, let's consider that the parent material in floodplains tends to have many layers. Also, evidence for changes in erosion rates in catchments is often found in the morphology of floodplain soils.

Authors: We appreciate the reviewer’s careful reading. The discussion on the divergent transform of shallow and deep layers in the western and central areas due to the accumulation of new parent materials will be added to section 4.3 (around L350). That is, rivers transported new parent material (soil particles of similar size) during different periods, leading to divergent transform in the shallow and deep layers.

Comment: L211 – I think it would be useful to define what is meant by “transform” I think this is the first time the term is used in the results section of this manuscript.

Authors: We appreciate for your comment. In our study, “transform”is used to denote the change in soil properties from an upper layer to a lower layer—for example, from 0-5 cm to 5-10 cm. Specifically, complexity is defined as the standard deviation of soil properties within each layer for a catchment, and Δ complexity is calculated as the difference between the upper layer’s complexity and the lower layer’s complexity, divided by the vertical interval in centimeters. Thus, “transform” captures the shift in complexity from one layer to the next. This study mainly focuses on the sign (positive or negative) of Δ complexity to identify how the complexity derived from soil properties changes vertically from upper to lower layers within the same catchment. For example, when Δ complexity is positive values indicate the convergent change in soil properties (the lower layer is more homogeneous than the upper layer), while negative values indicate a divergent change in soil properties (the lower layer is more heterogeneous than the upper layer).

Comment: L229 – How the analysis of different environmental conditions was done should be described in the methods.

Authors: We will add the method for analyzing different environmental conditions to the Methods section. The method is to apply Python code in ArcGIS to generate masks under different environmental conditions, extract the values of Δ complexity according to the masks, and calculate the minimum, maximum, mean, 25% percentile, median, and 75% percentile of each layer partition. Finally, the obtained data will be used to create a box plot using MATLAB.

Comment: L247-248 – I think a stronger argument for how analysis of current vertical heterogeneity indicates a timeline of different directions of soil processes. This isn’t the same as a chronosequence.

Authors: Since this study lacks chronological dating, we avoided making excessive references to time. However, we will cautiously speculate on temporal evolution within a limited scope. We tentatively consider that deeper soil layers are older than shallower ones. As you noted, this is a timeline of different directions of soil processes. This isn’ the same as a chronosequence.

Comment: L272 – Is it the visualization that is missing? Don't we want quantitative measures?

Authors: The visualization referred to in the manuscript emphasizes spatial visualization, specifically presented in the form of maps.

Comment: L288-289 – Were spatial connections among the digital soil map’s cells actually revealed? Does variability substitute for connectivity? Wouldn't some sort of topology be needed?

Authors: We appreciate for your thorough reading. Sorry for the ambiguity caused by unclear expression. The connectivity you mentioned indeed refers to links between cells. However, what we describe is not a cell-to-cell approach, but rather a process of calculating the standard deviation of soil properties using catchments as the analytical units, which is called complexity. This takes into account the degree of dispersion in soil property data distribution with the same catchment in the same layer. The difference between two layers divided by the depth interval is Δcomplexity, and the positive or negative value of Δcomplexity indicates the direction of change in soil property data across different layers within the same catchment. This differs from previous digital soil mapping that focused solely on grid cell values.

Comment: L294 – Because pedogenic processes for vertical complexity do not have a 1:1 relationship with redistribution processes, I'm not convinced by this statement. Mainly, I can think of both depositional and erosional processes that result in a more homogeneous soil profile (at least between depths of 0 to 40 cm). I encourage the authors to either acknowledge their assumption or clearly argue why the vertical complexity metric does not suffer from confounding factors.

Authors: We will change the sentence to “This approach effectively bridges the gap between detailed 3D soil models and accessible 2D visualization, enabling the identification of zones with profiles of different self-organizing characteristics that would otherwise be obscured in traditional layer-by-layer mapping approaches.”

Comment: L296 – I think the presumption of trajectory based on one point in time needs some additional support. I'm concerned that the assumption is being made that greater complexity equates to a trajectory of greater development. If that is not the case, I think the text could be improved to clarify the basis for the interpretations.

Authors: Thank you for your reminder. The trajectory in the manuscript is not based on one point, but on the catchment. We will change the original text to “Soil complexity reveals distinct spatial self-organization characteristics of soil properties within the 0–40 cm layer across catchments under varying environmental conditions.”

Comment: L342 – Perhaps specify the listed soil redistribution processes as primary in this study area.
Authors: Thank you for your comment. We will make soil redistribution the main process in the study area based on your suggestions.

Comment: L369 – Okay, yes, but usually there is some measure of age to determine where the location is on the timeline relative to other locations studied.
Authors: Thank you for your valuable comment. We agree wholeheartedly with your perspective. While we are unable to incorporate temporal analysis at this stage due to the absence of age data, we recognize its importance and will advance chronosequence research in our future endeavors.

Comment: L388-390 – These are well-established processes that influence the vertical heterogeneity of soil. However, I'm not sure this study conclusively connects the results of the complexity metric with them. These things coincided, but the case for their connection was largely made in the discussion section. Perhaps different wording could soften these conclusions to align with the data presented in this paper.
Authors: We appreciate for your constructive suggestion. What you said makes perfect sense. We will revise the wording to soften these conclusions.

Comment: L396-397 – This sentence illustrates why I am questioning the definition of sensitive areas. I keep thinking a sensitive area is one that responds more to environmental forces. If a soil is highly susceptible to erosion and loses its surface 40 cm, couldn't the result be a vertically homogeneous profile?
Authors: Thank you for your careful reading. In our manuscript, “sensitive areas” refer to catchments where the sign of Δ complexity differs between shallow and deep layers—indicating opposite transformation directions. Our results are based on the current soil properties within the 0-40 cm profile; some erosional or depositional processes may have occurred in the past, which we are unable to determine. We will revise the manuscript to clarify this logic for readers.

Technical Corrections

Comment: L35 – Should this sentence be simplified? The current text seems somewhat circular with “information of soil heterogeneity for revealing soil variation”, unless the authors want to differentiate heterogeneity from variation, or maybe better, variability.
Authors: Thank you for your comment. We will rewrite this sentence. It will change to “provide distinct information of soil properties for revealing soil variation and development”.

Comment: L43 – This citation appears to be missing from the reference list. I realize this is a companion paper to the 2022 paper, but I think it should still be included in the reference list.
Authors: Both Van der Meij’s 2020 and 2022 papers are included in the references.

Comment: L64 – The number of decimal places for the area number seems excessive.
Authors: We will only keep one decimal place and change it to "165.6 km^2".

Comment: L117 – Fix typo. "b" to "be"
Authors: We will change "b" to "be" .

Comment: L125 – Clean-up authors' note to themselves to add a needed reference.
Authors: Thank you for your reminder. We will check the manuscript.

Comment: L170 – Typo. Change "cay" to "clay"
Authors: We will change "cay" to "clay" .

Comment: L210 – Should "catchment" be "catchments" ?
Authors: We will change "catchment" to " catchments" .

Comment: L315 – Change "human" to "humans"
Authors: We will change "human" to "humans"

Comment: L388 – This sentence is awkwardly written. Please reword.
Authors: Thank you for your reminder. It will change to "The vertical heterogeneity of soil Δ complexity for clay is fundamentally shaped by pedogenesis, including parent material, topography, and vegetation, and is further modified by tree throw and human activities. However, the relative importance of these drivers varies with topography type. In lowland catchments, vertical heterogeneity is primarily driven by sediment accumulation and human activities. In mountainous catchments, it is mainly caused by tree throw and topographic effects. In mixed-relief catchments, there are both lowland and mountains, topography and vegetation play dominant roles; notably, these catchments are concentrated in the southeast of the study area, where water erosion is more severe, suggesting that erosional processes may also contribute to their vertical differentiation."

References mentioned in this review that are not already in the manuscript's reference list

Zhu Y, Zhang Z D, Liu C, et al. Comparison of laser diffraction method and pipette method on soil particle size distribution determination： A case study of variously degraded kastanozem [J]. Research of Soil and Water Conservation, 2018, 25(3): 62-67, 204. DOI: 10.13869/j.cnki.rswc.2018.03.011. (in Chinese with English abstract)

Johnson, D.L., and Watson-Stegner, D.: Evolution model of pedogenesis, Soil Science, 143(5), 349–366, https://doi.org/10.1097/00010694-198705000-00005, 1987.

Johnson, D.L., Watson-Stegner, D., Johnson, D.N., and Schaetzl, R.J.: Proisotropic and proanisotropic processes of pedoturbation, Soil Science, 143(4), 278–292, https://doi.org/10.1097/00010694-198704000-00005, 1987.

Phillips, J.D.: Divergence, convergence, and self-organization in landscapes, Annals of the Association of American Geographers, 89(3), 466–488, https://doi.org/10.1111/0004-5608.00158, 1999.

Phillips, J.D.: Signatures of divergence and self-organization in soils and weathering profiles, The Journal of Geology, 108(1), 91–102, https://doi.org/10.1086/314386, 2000.

Phillips, J. D.: Soil Complexity and Pedogenesis, Soil Science, 182, 117–127, https://doi.org/10.1097/SS.0000000000000204, 2017.