Dear Editors and Reviewers:

Thank you for your letter and for the reviewers' comments concerning our manuscript entitled 'Linear passive source surface wave dispersion curve picking based on supervised deep learning and ambient noise tomography for the evolution of the internal structure in landslide area' (egusphere-2026-1329). Those comments are all valuable and very helpful for revising and improving our paper, as well as the important guiding significance to our researches. We have studied comments carefully and have made correction which we hope meet with approval. Revised portion are marked in red in the paper. The main corrections in the paper and the responds to the reviewer's comments are as flowing:

>>>The arrow indicating part are the response of authors to editors and reviewers.

Reviewer 1:

1. The English language requires significant polishing for clarity and conciseness, as many sentences are convoluted and even contain grammatical errors somewhere, i.e., Page 1, Line 1: The current title is too verbose to grab primary contribution.

>>>We highly appreciate the reviewer's valuable comments and suggestions on the language expression. We have carefully revised the whole manuscript, polished sentences and fixed grammatical errors according to your requirements. We sincerely appreciate the reviewer's insightful suggestion on the manuscript title. We have replaced the original title with the new version "Passive-source Surface Wave Dispersion Curve Picking Using Supervised Deep Learning for Ambient Noise Tomography: A Study on Internal Structural Evolution of Landslide Areas". We hope the revised title can better summarize the core research content and highlight the main innovation of this study.

2. Page 2, Lines 27-28: The phrase "attributed to structural evolutions in the very near surface" is vague. Specify the depth range corresponding to "very near surface" based on your resolution limits.

>>> Thank you very much for your reminding. We agree with the comment. Based on the half-wavelength exploration theory of surface waves, the effective detection depth of Rayleigh surface waves is approximately half of their wavelength. In this study, the depth range of the very near surface is defined as >1.0 m according to our detection resolution limits, and we will revise this description clearly in the revised manuscript.

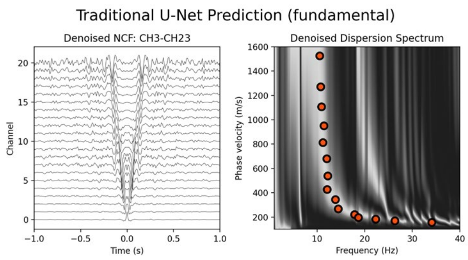
3. Page 6, Lines 111-115: Please add a brief justification for why ESPAC is appropriate for a linear geometry mentioned as this study.

>>> Thank you for the valuable comment. ESPAC is appropriate for a linear geometry mentioned as this study due to the following reasons: 1. The traditional SPAC is extremely prone to being restricted by the terrain as it requires the array to be circularly arranged. However, the ESPAC can calculate the autocorrelation coefficient and then fit the Bessel function to obtain the phase velocity for irregular arrays. It maintains the frequency constant and changes the circular radius to obtain the relationship of the spatial autocorrelation coefficient with distance. Thus, it is more appropriate for the linear or 2D geometry. 2. The ESPAC can simultaneously utilize the spatial autocorrelation

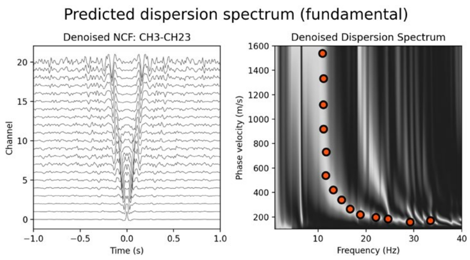
coefficients of different radius stations in the observation array, improving its application convenience. ESPAC also has the advantages of the SPAC method, and it has a higher resolution for dispersion extraction than the F-K method in the low-frequency band with a greater detection depth and more stable dispersion extraction stability. We will supplement concise reasoning to illustrate the suitability of ESPAC for the linear geometry adopted in this research.

4. Page 10, Line 244: Authors state they removed horizontal skip connections. This could be a major architectural deviation from standard U-Net. Provide a clear rationale for this change in the context of dispersion spectra feature propagation.

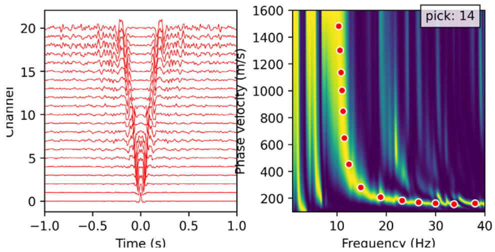
>>> We appreciate the reviewer's insightful comment on the architectural modification of removing horizontal skip connections from the standard U-Net structure. In standard U-Net architectures, initially for medical images, horizontal skip connections are designed to transmit fine-grained shallow spatial details to the decoder, compensating for high-resolution information loss during upsampling and benefiting conventional pixel-wise segmentation tasks. However, for our ambient noise tomography, the dispersion spectrum is often filled with strong, discrete dispersion noise, and the edges of the dispersion energy are too blurry. If the skip connections are not removed, the high-frequency noise features extracted by the encoder's shallow layers will be directly transmitted to the decoder through the skip connections, causing the noise restoration along with the network reconstruction. Removing the skip connections can force the network to perform information compression, which prompts the model to abandon the local noise details at the shallow layers but only learn the macroscopic form of the dispersion energy at the deep layers. Besides, compared with the high precision of medical imaging, the dispersion spectrum adopted generally retains good signal-to-noise ratio and consistency characteristics. Therefore, the excessive parameter introduction and excessive detail retention in the traditional U-Net model may lead to a decrease in accuracy and computational efficiency when executing simple tasks such as the dispersion curve. This is also the reason why LiteU-Net still maintains strong robustness under low signal-to-noise ratios. In order to further demonstrate the rationality of our strategy, Fig. 1 shows a similar precision of the dispersion curve artificial intelligence extraction by the traditional U-Net and LiteU-Net models respectively and the quantitative analysis of the root mean square error (in Fig. 2) from them also demonstrate a perfect achievement.



(a)

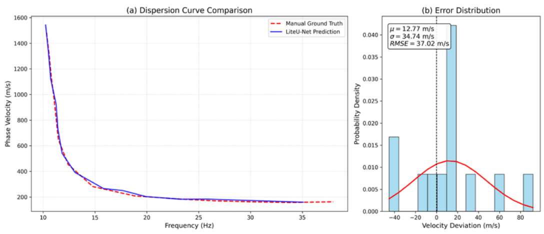


(b)

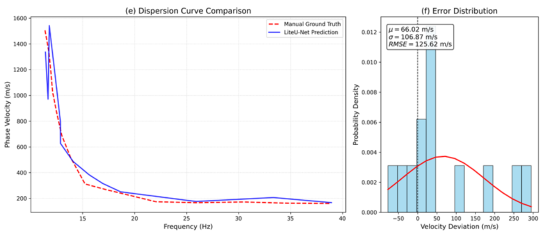


(c)

Fig. 1 A similar precision of the dispersion curve artificial intelligence extraction by the traditional U-Net (a) and LiteU-Net (b) models respectively. The color images are the manual picks (c).



(a)



(b)

Fig. 2 The quantitative analysis of the root mean square error from the traditional U-Net (a) and LiteU-Net (b) models respectively.

5. Page 12, Lines 300-303: The description of labeling three classes (fundamental, higher, noise) with specific gray values (200, 100, 50) is unclear. Are these pixel intensity values? If so, how is the loss calculated on these raw intensities rather than class indices?

>>> We thank the reviewer for pointing out this unclear description. The gray values mentioned in the text (200, 100, 50) are merely for visualization purposes during manual annotation and generation of mask images, with the aim of enabling the human eye to visually distinguish different orders of dispersion curves. Actually, before the data is input into LiteU-Net model for forward propagation and Loss calculation, the code already contains a mapping operation that maps the gray value 50 to the category index 0 (noise), the gray value 200 to the category index 1 (base order), and the gray value 100 to the category index 2 (high order). The loss function is entirely based on the standard category index calculation and conforms to the mathematical norms of deep learning. We will revised and supplemented in the revised manuscript.

6. Page 17, Lines 457-462: The interpretation linking velocity variations to hydro-mechanical triggers would benefit from referencing recent publications, i.e., https://doi.org/10.1016/j.gsf.2023.101773; https://doi.org/10.1016/j.enggeo.2025.108264

>>> Thank you for this valuable comment. We agree that the interpretation linking velocity variations to hydro-mechanical triggers would be strengthened by citing recent publications mentioned. In the revised manuscript, we will add relevant recent publications to better support our hydro-mechanical interpretation.

Zhu, H.H., Ye, X.*, Pei, H.F., Zhang, W., Cheng, G., Li, Z.L., 2024. Probing multi-physical process and deformation mechanism of a large-scale landslide using integrated dual-source monitoring. Geoscience Frontiers, 15(2), 101773. https://doi.org/10.1016/j.gsf.2023.101773

Ruochen Jiang, Limin Zhang, Wenjun Lu, Dalei Peng, Xin He, Shihao Xiao, Yingyue Han, Mingdong Wei. 2025. A thermal-hydro-mechanical coupled analysis model for climate-driven movements of valley glaciers (THM-GA 1.0), Engineering Geology, Volume 355, 108264. https://doi.org/10.1016/j.enggeo.2025.108264.

7. Fig. 15 shows potential sliding surfaces. Could authors add a sentence explaining the specific velocity contour or gradient threshold (e.g., Vs = 350 m/s contour) used to define the interface line shown in the profile?

>>> We appreciate this constructive comment. The specific velocity contour or gradient threshold (e.g., Vs = 350 m/s contour) was determined by the following factors: 1. The landslide body is composed of gravel clay with high permeability in a soft-plastic state, according to the engineering specifications, providing the velocity distribution range between 250 – 350 m/s. 2. The results of the geotechnical mechanics experiments indicate that the velocities of the Quaternary strata and the bedrock are 350 and 800 m/s respectively. 3. The specific velocity contour or gradient threshold (e.g., Vs = 350 m/s contour) used to define the interface line is consistent with the characteristics of the dispersion curves and the landslide structure indicated by the geological borehole profiles. In the revised manuscript, we will explicitly specify the velocity threshold used for defining the interface line in the profile.

8. The comparison between 2024 and 2025 profiles (Figs. 15 and 16) is qualitative. Please provide a difference map or a plot of relative velocity change (dv/v) to support the discussion of spatiotemporal evolution.

>>> We appreciate the valuable comment. As suggested, we have added the difference map (ΔVs) of relative velocity change between the 2024 and 2025 profiles in the revised manuscript. ΔVs of L1 and L3 shows lateral heterogeneity, indicating the structural changes in the landslide mass, likely influenced by shallow groundwater due to rainfall. Besides, the new crack or sliding surface is more distinct characterised with obvious Vs differences. The shear wave velocity of the landslide body shows a general downward trend, especially in L2, and the sliding zone at the bottom of the landslide, affected by groundwater and long-term shear deformation, presenting an expanding trend of low- velocity abnormal area, which indicates that the internal structure of the landslide has undergone irreversible damage and the instability risk has increased. Thus, the temporal dynamic monitoring of the shear wave velocity changes is of great significance for the deformation evolution

process of the landslide. The newly added figures clearly illustrate the spatial and temporal evolution characteristics, which effectively support the relevant discussion.

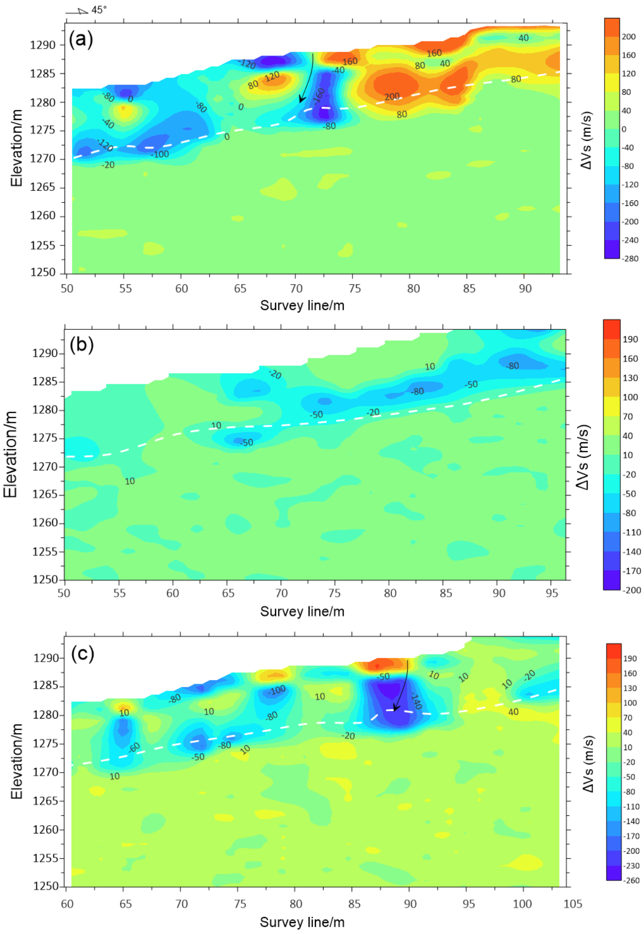


Fig. 3 The difference map (ΔVs) of relative velocity change between the 2024 and 2025 profiles

9. The methodological description of the LiteU-Net is overly verbose and lacks critical technical details regarding network architecture and training and validation splits.

>>> We sincerely thank the reviewer for pointing out the deficiencies in the methodology section of LiteU-Net. As suggested, we have made targeted and substantial revisions as follows: We re-edited the original verbose paragraphs, deleted repetitive and non-essential descriptions, and optimized the layout to enhance conciseness. The elaborate technical details of LiteU-Net shall be carried out as shown in Fig. 5 in our paper. Besides, we have added the missing dataset split description in the corresponding section of the manuscript. In the actual model evaluation, we strictly divided the synthesized dispersion spectrum dataset into training and validation sets in an 80%:20% ratio to monitor potential overfitting of the model. We hope the revised content meets your requirements.

10. The interpretation of velocity variations relies heavily on qualitative visual comparisons with limited quantitative uncertainty analysis. The connection between the deep learning picking accuracy and the final Vs model uncertainty is not sufficiently explored.

>>> Thank you very much. We acknowledge the shortcomings mentioned. We have supplemented the difference map (ΔVs) of relative velocity change between the 2024 and 2025 profiles as comment 8 mentioned. It may provide a basic quantitative uncertainty analysis for velocity

variations in the revision. We are sorry for our insufficient work. We will pay attention to the connection between the deep learning picking accuracy and the final Vs model uncertainty as soon as possible and will carry out more in-depth research on this topic in our future work.