Supplementary Materials for

ChinaTidalVSC: a 10 m wall-to-wall dataset of vegetation structural complexity in China's tidal wetlands derived from LiDAR-based relative entropy and AlphaEarth Foundation data.

Guanpu Qi[1]†, Jiawei Wei[2]†, Zaizhou Xin[3], Longcheng Yao[1], Huilan Yuan[1], Ming Nie[1], Bo Li[1,4], Xiong Chen[2*], Jun Ma[1*]

[1] State Key Laboratory of Wetland Conservation and Restoration, National Observations and Research Station for Wetland Ecosystems of the Yangtze Estuary, Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, and Institute of Eco-Chongming, School of Life Sciences, Fudan University, # 2005 Songhu Road 10, Shanghai, 200438, China.

[2] School of Information Science and Engineering, Fudan University, # 2005 Songhu Road 10, Shanghai, 200438, China.

[3] Key Laboratory of Computational Physical Sciences (Ministry of Education), Institute of Computational Physical Sciences, State Key Laboratory of Surface Physics, and Department of Physics, Fudan University, Shanghai 200433, China.

[4] Ministry of Education Key Laboratory for Transboundary Ecosecurity of Southwest China; Yunnan Key Laboratory of Plant Reproductive Adaptation, Evolutionary Ecology and Centre for Invasion Biology, Institute of Biodiversity, School of Ecology and Environmental Science, Yunnan University, Kunming, 650504, China.

†These authors contributed to the work equally.

*Corresponding author

E-mail:

Jun Ma: ma_jun@fudan.edu.cn

Xiong Chen: **chenxiong@fudan.edu.cn**

Tel. Jun Ma: +86-21-31246690.

Tel. Xiong Chen: +86-21-65643231.

This file includes:

- Figs. S1 to S9
- Tables S1 to S4
- Text S1-S3

Figures

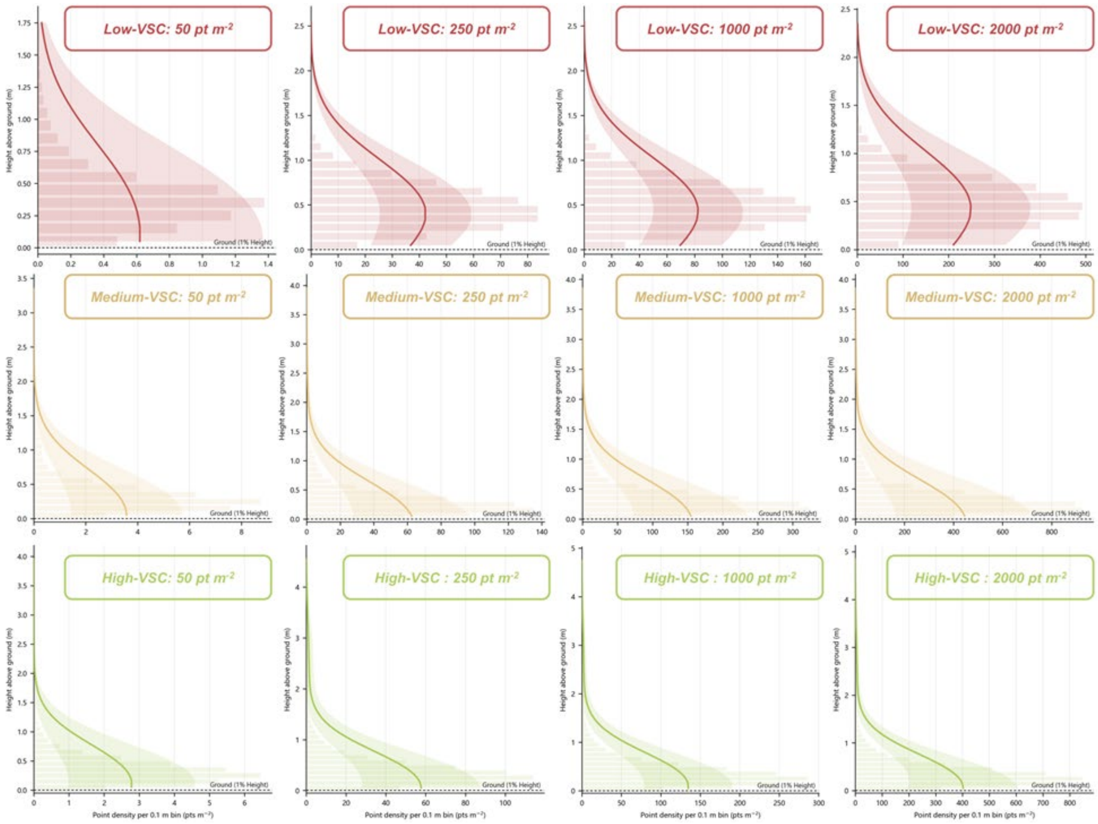

Fig. S1 | Vertical point-density profiles of fixed-plot samples across VSC groups and target LiDAR point-cloud densities. Vertical point-density profiles are shown for 25×25 m fixed-plot samples grouped by VSC and target PCD. Rows represent low-, medium-, and high-VSC groups, and columns represent target PCDs of 50, 250, 1000, and 2000 pt m^{-2}. In each panel, horizontal bars show the mean point density within 0.1 m vertical bins, the solid curve represents the Gaussian-smoothed mean profile, and the shaded envelope indicates variability among samples. Heights were normalized relative to local ground, defined as the 1st percentile of point height and indicated by the black dashed line. The x-axis represents point density per 0.1 m height bin, and the y-axis represents height above ground.

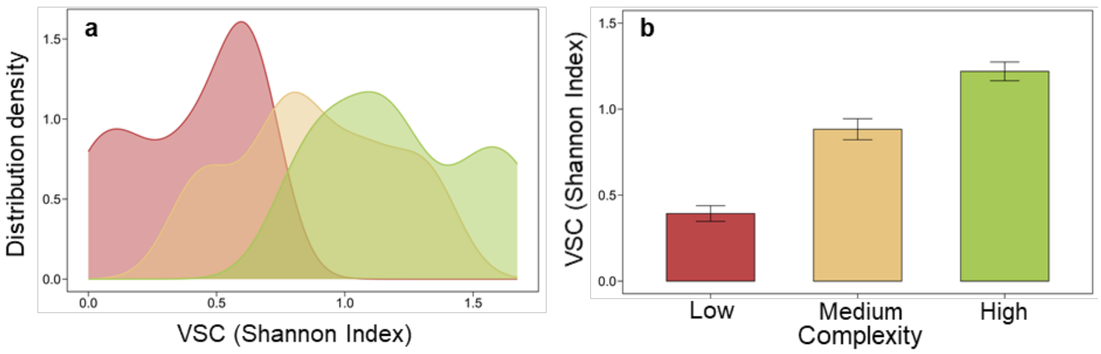

Fig. S2 | Biomass-weighted Shannon index across treatment-derived VSC groups. **a**, Distribution of the biomass-weighted Shannon index, fitted using kernel density estimation. **b**, Mean biomass-weighted Shannon index for the low-, medium-, and high-VSC groups. Error bars indicate standard deviations.

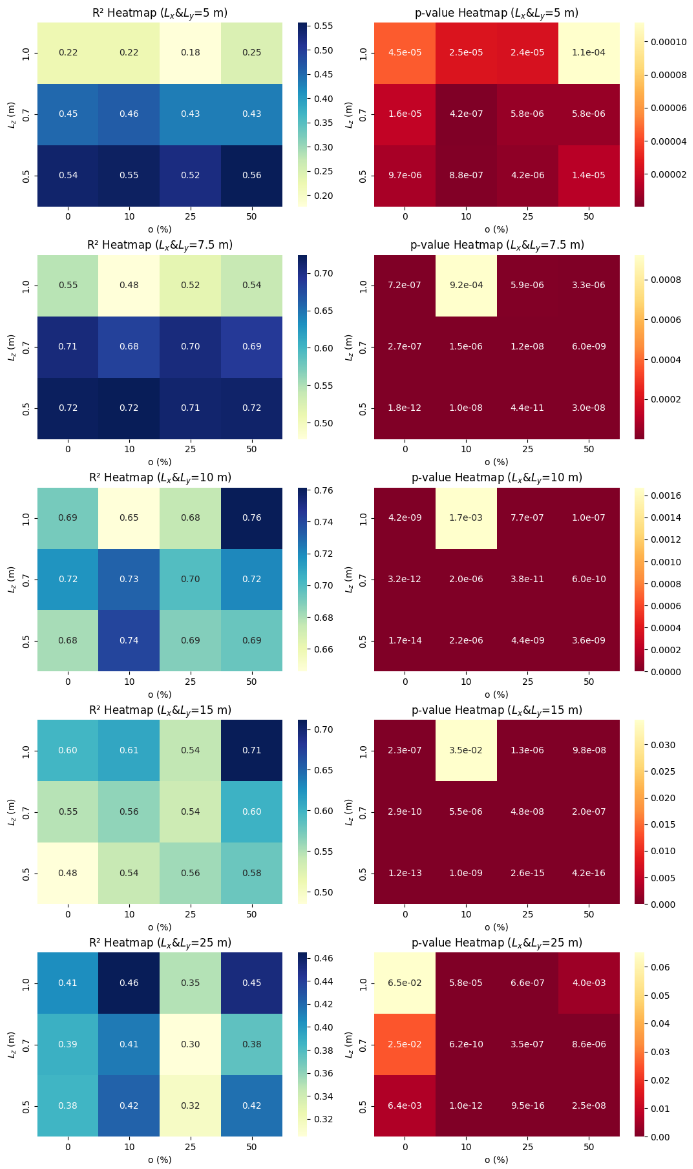

Fig. S3 | Heatmap of R^2 and p-values for different hyperparameter combinations. We conducted a grid search over the hyperparameters of the proposed method to find the optimal combination for the application. In the tests, we performed regression analysis with the Shannon index on a 60 m × 50 m plot and ANOVA analysis on a 25 m × 25 m plot. In these experiments, the R^2 value and p-value for each parameter combination were selected as evaluation metrics. Specifically, we explored window side lengths (L_x and L_y) of 5 m, 7.5 m, 10 m, 15 m, and 25 m; window heights (L_z) of 0.5 m, 0.7 m, and 1 m; and overlap ratios of 0, 0.1, 0.25, and 0.5 (o). **Figure S3** shows heatmaps of the two metrics under different parameter combinations.

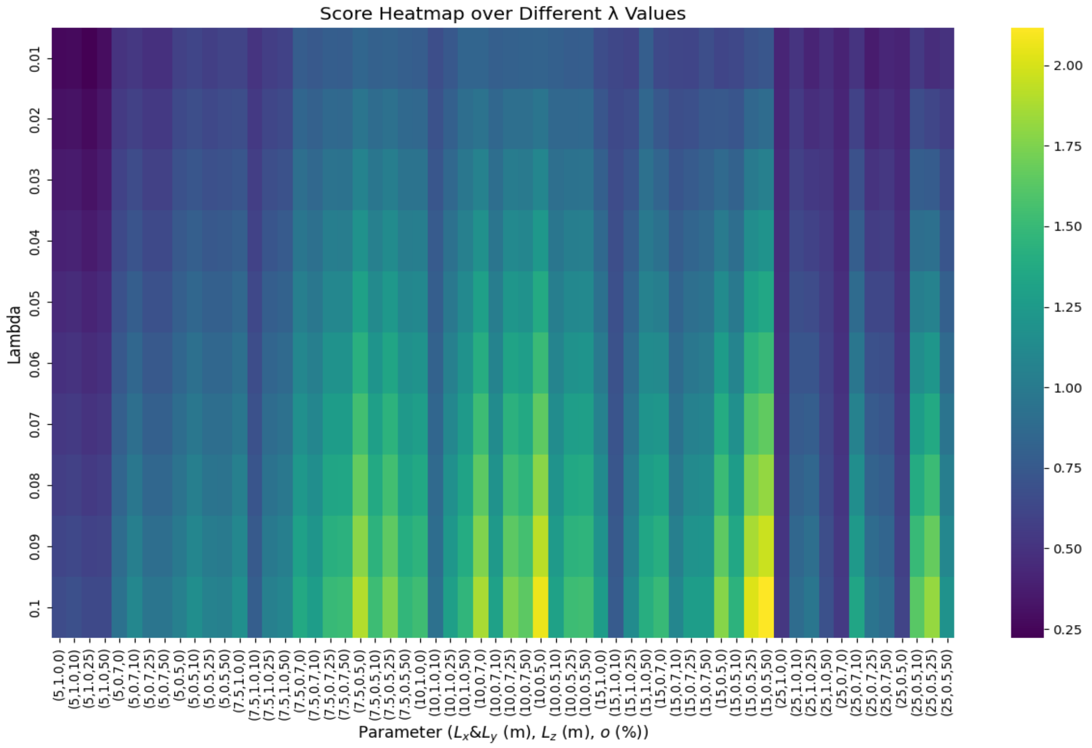

Fig. S4 | Heatmap of scores for different hyperparameter combinations under varying trade-off factors.

We calculated the final score for each parameter combination according to the following formula:

$$\text{Score} = R^2 - \lambda \log_{10}^{\text{p-value}} \tag{S1}$$

where λ is a weighting factor used to balance the two evaluation metrics, we varied λ from 0.01 to 0.1 and computed the scores for each case, as shown in **Figure S4**. Among all parameter combinations, those with a window height of L_z=0.5 m achieved higher scores, which may be related to the average height of low-lying vegetation. In particular, combinations of small windows without overlap (10, 0.5, 0) and large windows with overlap (15, 0.5, 25) and (15, 0.5, 50) achieved the best scores across all tested settings. Considering the additional terrain normalization required during processing and the pre-trimming of water body regions in the nationwide field dataset, we selected the optimal combination (10, 0.5, 0) under λ = 0.02 to ensure computational efficiency.

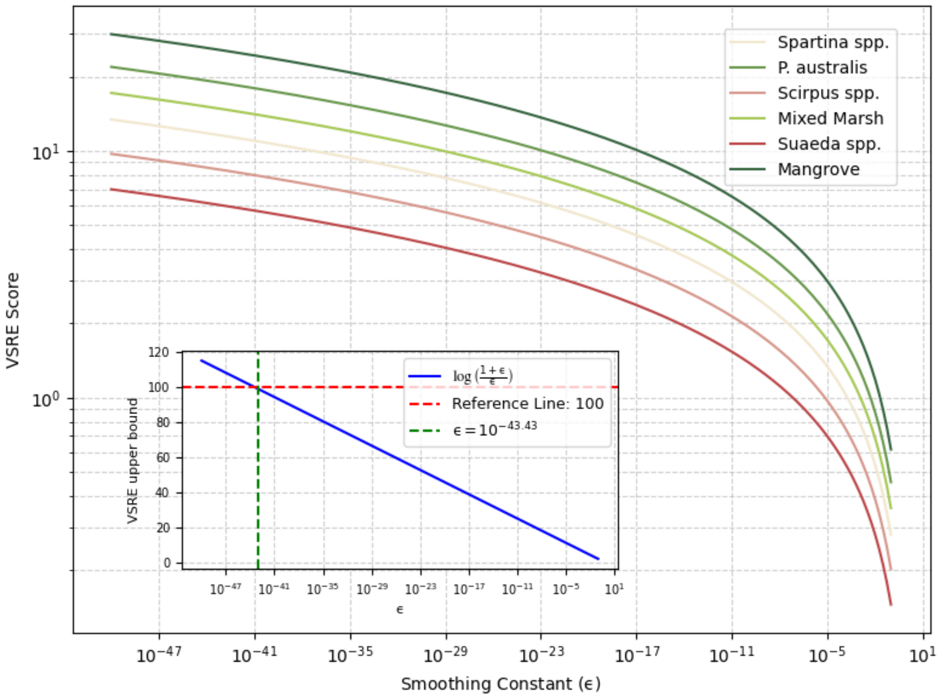

Fig. S5 | The variation trend of average VSRE scores for different vegetation communities in field data concerning the smoothing constant and the theoretical upper bound of VSRE as a function of the smoothing constant.

To ensure numerical stability in the computation of RE between voxel-based 3D histograms, we adopt a modified formulation:

$$RE(\Phi||\Psi) = \sum_{i,j,k} \Phi_{ijk} log \frac{\Phi_{ijk} + \epsilon}{\Psi_{ijk} + \epsilon} \tag{S2}$$

Where, Φ and Ψ are the normalized point cloud distributions in the two windows. ϵ is a small positive constant introduced to prevent divergence caused by zero entries in either distribution. This adjustment stabilizes the computation and implicitly constrains the upper bound of the VSRE. Specifically, when considering the extremal case where a single bin in Φ concentrates the entire probability mass (i.e., $\Phi_a = 1$ for one index and zero elsewhere), and Ψ is uniform or zero in that bin, the RE divergence simplifies to:

$$RE_{max} = log\left(\frac{1+\epsilon}{\epsilon}\right) \tag{S3}$$

is expression shows that the maximum divergence is directly determined by the choice of ϵ. The smaller the value of ϵ, the higher the potential upper bound for the RE divergence. In our experiments with field data, we set:

$$RE_{max} = log\left(\frac{1+10^{-43.43}}{10^{-43.43}}\right) \approx log(10^{43.43}) = 100 \quad \text{(S4)}$$

Figure S5 presents the distributions of the VSRE scores across various plant species from field data in China, under different smoothing constants ϵ. The figure also shows the theoretical upper bound of VSRE as a function of ϵ. As illustrated, the choice of ϵ exerts a significant influence on the numerical stability and the sensitivity of the VSRE metric. When ϵ is set too large (e.g., $\epsilon > 10^{-2}$), the added smoothing overwhelms the original differences between distributions, compressing the RE values into a narrow range. Consequently, the VSRE scores fail to distinguish structural complexity across vegetation types, resulting in a substantial loss of discriminative power. Conversely, when ϵ is excessively small (e.g., $\epsilon < 10^{-50}$), although the sensitivity to structural differences increases, the computation becomes highly susceptible to numerical instability due to finite-precision floating-point arithmetic. Minor numerical errors may be amplified in such cases, resulting in unreliable VSRE estimates.

By setting $\epsilon = 10^{-43.43}$, the theoretical upper bound of the VSRE score is approximately 100, placing the indicator within a practical, interpretable [0, 100] range. This choice ensures that the VSRE metric maintains sufficient sensitivity to capture structural variations in the point cloud while preserving numerical stability across different datasets. Moreover, normalizing VSRE within [0, 100] facilitates intuitive interpretation and straightforward comparison across studies and ecosystems.

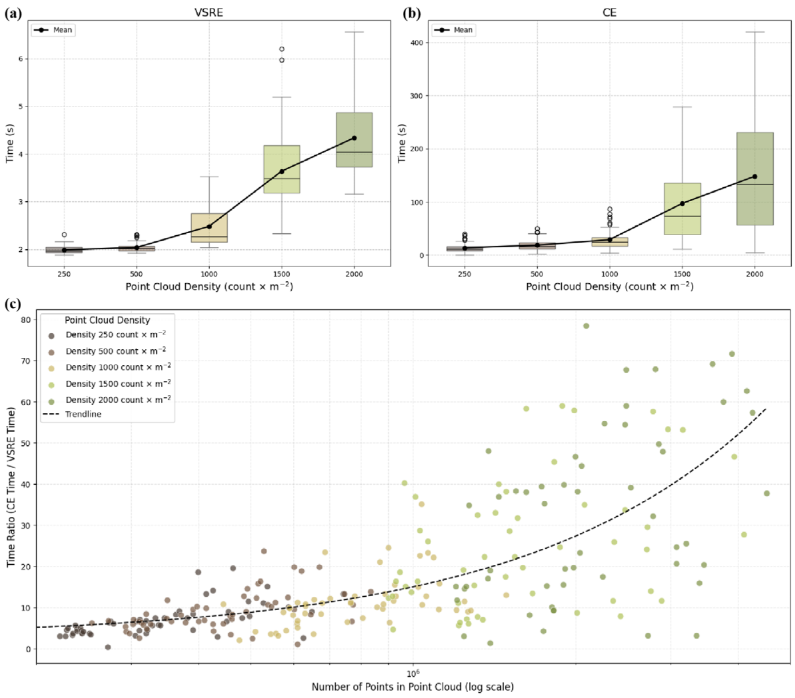

Fig. S6 | Comparison of Computational Efficiency Between VSRE and CE Methods. a-b, Variation trends of computation time for VSRE and CE concerning point cloud density. **c**, The time ratio (CE Time/VSRE Time) as a function of point cloud size. All statistics were derived from 25m×25m samples in fixed plots.

In **Figure S6**, we compare the computational efficiency between the VSRE and CE methods using 25m×25m plot data. Panels (a) and (b) display box plots of processing time for both methods across varying point cloud densities. As density increases, the computational time rises for both approaches due to the inherent processing demands of larger point clouds. Nevertheless, our VSRE method consistently requires less than one-tenth the time of CE. Panel (c) presents the time ratio (CE/VSRE) as a function of point count, revealing that VSRE's efficiency advantage becomes more pronounced with increasing point numbers, reaching a 70-fold improvement at approximately 5×10^6 points. This significant gain stems from VSRE's bin-wise voxel distribution statistics, which directly characterize local point cloud distributions rather than CE's computationally intensive point-wise resampling and projection operations. The bin-wise approach maintains robustness against variations in point count while substantially reducing computational overhead during distribution construction. It can be expected that as the point cloud grows larger, this efficiency gap will continue to widen.

Tables

Table S1 | Coordinates and VSRE of the forest LiDAR samples

ID	Longitude	Latitude	Elevation Min	Elevation Max	VSRE
1	119.5685625	28.1306505	183.24	202.88	51.65
2	119.5684155	28.13112325	187.47	207.67	42.50
3	119.5715543	28.12946875	186.32	209.44	41.50
4	119.5709303	28.12837475	210.87	230.49	57.84
5	119.5717565	28.128487	188.57	210.84	44.95
6	119.572456	28.13168525	216.79	244.13	35.64
7	119.5608293	28.12190925	193.30	225.95	35.48
8	119.5612788	28.12004175	187.43	216.56	36.28
9	119.5847853	28.22370175	664.73	690.83	22.01
10	119.5853543	28.22377425	686.36	703.42	43.18
11	119.5738523	28.1344965	239.03	263.80	63.46
12	119.582558	28.22422925	664.00	693.83	33.45
13	119.5827285	28.22397425	642.43	669.91	39.11
14	119.5851405	28.22349825	664.55	690.21	27.85
15	119.5854633	28.22349775	666.43	692.79	21.71
16	119.5796698	28.08116525	210.64	246.07	27.44
17	119.57957	28.08140575	213.75	244.79	36.95
18	119.5799563	28.08170375	242.39	271.01	32.08
19	119.5189663	28.0811495	230.55	270.53	14.22
20	119.5828998	28.0871795	255.99	276.02	47.33
21	119.5826563	28.08743325	235.27	262.78	61.26
22	119.5823155	28.08752925	216.27	245.22	50.07
23	138.7030558	25.45691525	205.01	234.99	47.97
24	119.5822858	28.08710225	232.16	256.44	43.09
25	119.5814625	28.08639575	212.10	235.55	42.61
26	119.5811858	28.086243	209.90	240.85	36.33
27	119.5807738	28.08612875	197.18	231.28	34.12
28	119.5807183	28.08586275	207.02	232.85	43.42
29	119.580713	28.08558425	208.65	232.27	31.59
30	119.5807443	28.08524175	208.14	230.94	28.19

Table S2 | VSRE of different vegetation communities in different provinces

Province	*Spartina* spp.	*P. australis*	*Scirpus* spp.	Mixed Marsh	*Suaeda* spp.	Mangrove
Liaoning	12.80	14.05	/	/	4.38	/
Shandong	10.75	15.43	/	/	5.42	/
Jiangsu	7.15	7.73	9.49	/	12.12	/
Shanghai	/	21.84	6.04	14.49	/	/
Zhejiang	13.79	11.53	/	13.94	/	/
Fujian	12.87	20.12	9.04	15.86	6.28	55.39
Guangdong	23.31	/	/	22.66	/	28.52
Guangxi	15.96	17.65	/	/	/	20.92

Table S3 | Tidal marsh area in different provinces (km^2).

Province	*Mudflat*	*Spartina* spp.	*P. australis*	*Scirpus* spp.	*Suaeda* spp.	*Mixed marsh*	*Mangrove*	Vegetated	Total
Fujian	1048.7	68.6	13.6	93.0	/	/	13.1	175.2	1223.9
Guangdong	712.5	105.8	/	/	/	1.0	67.5	106.8	819.3
Guangxi	573.3	127.7	/	/	/	/	61.6	127.7	701
Hebei	453.7	1.2	1.2	/	1.0	/	/	3.4	457.1
Jiangsu	3196.0	141.6	77.1	46.2	39.1	0.7	/	304.7	3500.7
Liaoning	1539.1	/	66.4	/	75.6	20.8	/	162.8	1701.9
Shandong	1752.8	97.3	65.5	/	63.1	13.0	/	238.9	1991.7
Shanghai	422.9	151.9	98.0	151.5	/	14.9	/	416.3	839.2
Tianjin	92.1	1.6	1.4	/	0.7	/	/	3.7	95.8
Zhejiang	894.6	112.2	98.5	52.7	/	/	0.0	263.4	1158

Table S4 | Model Comparison Based on Cross-Validation MAE

Category	Method	CV MAE Mean
Statistical	Ridge	6.6151
Machine Learning	Random Forest	4.6750
Machine Learning	Gradient Boosting	5.0676
Machine Learning	SVR	7.8506
Machine Learning	XGBoost	4.5278
Machine Learning	KNN	4.2419
Machine Learning	Decision Tree	5.8195
Deep Learning	MLP	3.8087
Deep Learning	Residual	3.8815
Deep Learning	Auto Encoder	3.9343
Deep Learning	Transformer	4.2711

MAE (mean absolute error) is the average magnitude of prediction errors and serves as an indicator of overall model accuracy. All metrics in this table were calculated as averages from tenfold cross-validation on the dataset. To select the most appropriate model for the task, a set of machine learning and deep learning algorithms was compared using their default parameter settings.

Text

1 Terrain simulation

To focus on structural diversity, the input point cloud data must first undergo terrain undulation removal. A Gaussian mixture model in the horizontal plane efficiently fits rugged terrain due to its multi-modal nature. Given a set of LiDAR point cloud data $\mathcal{P} = \{(\mathrm{x_I}, \mathrm{y_I}, \mathrm{z_i})\}_{\mathrm{i=1}}^{\mathrm{N}}$, where x_i, y_i, and z_i denote the 3D coordinates of the i^{th} point, the fitted surface $\hat{z} = f(x, y)$ can be expressed as(Wolfe, 1970):

$$\hat{z}(x, y) = \sum_{k=1}^{K} \omega_k \cdot G_k(x, y) \tag{S5}$$

Where $G(x, y) \sim \mathcal{N}(\vec{\mu}, \Sigma)$ is a bivariate Gaussian distribution. For the k^{th} distribution:

$$G_k(x, y) = A_k exp\left(-\frac{1}{2}\begin{bmatrix} x - \mu_k^{(x)} \\ y - \mu_k^{(y)} \end{bmatrix}^T \Sigma_k^{-1} \begin{bmatrix} x - \mu_k^{(x)} \\ y - \mu_k^{(y)} \end{bmatrix}\right) \tag{S6}$$

adopt learnable A_k as the amplitude, $\left(x - \mu_k^{(x)}, y - \mu_k^{(y)}\right)$ as the distribution center. Σ_k is the covariance matrix defining the distribution's orientation and spread:

$$\Sigma_k = \begin{bmatrix} {\sigma_k^{(x)}}^2 & \rho_k \sigma_k^{(x)} \sigma_k^{(y)} \\ \rho_k \sigma_k^{(x)} \sigma_k^{(y)} & {\sigma_k^{(y)}}^2 \end{bmatrix} \tag{S7}$$

Where ρ_k is the correlation coefficient and $\sigma_k^{(\cdot)}$ denotes the standard deviation in a given direction. To effectively fit real terrain, independent distributions with weights ω_k are linearly combined to form peaks ($\omega_k > 1$) or valleys ($\omega_k < 1$). The point cloud $\mathcal{P}$ is divided into grids (P × Q) along the x and y directions, extracting the minimum z-coordinate points in each grid as the ground morphology data S:

$$\mathrm{S} = \left\{\arg \min_{(x,y,z)\in\mathcal{P}} \mathrm{z} \mid (\mathrm{x}, \mathrm{y}) \in (\mathrm{x_p}, \mathrm{x_{p+1}}] \times (\mathrm{y_p}, \mathrm{y_{p+1}}]\right\}_{\mathrm{p=1,q=1}}^{\mathrm{P,Q}} \tag{S8}$$

Residuals are defined as:

$$r_i = z_i - \hat{z}(x_i, y_i) \forall z_i \in S \tag{S9}$$

To further enhance the robustness of outlier fitting, we adopt the Tukey loss function(Beaton and Tukey, 1974; Belagiannis et al., 2015), which suppresses the influence of mild outliers without compromising overall fitting performance. This improves the robustness of terrain normalization:

$$\mathcal{L}(r_i) = \begin{cases} \frac{\delta^2}{6}\left(1-\left(1-\frac{{r_i}^2}{\delta^2}\right)^3\right), |r_i| \leq \delta \\ \frac{\delta^2}{6}, |r_i| > \delta \end{cases} \tag{S10}$$

To prevent degeneration of the covariance matrix and improve model stability, a regularization term is added to the final loss(Grave et al., 2011). Consequently, the optimization objective for the fitting stage is summarized as:

$$\arg\min_{\Theta} \sum_{\mathrm{i}=1}^{\mathrm{P}\times\mathrm{Q}} \mathcal{L}(r_i) + \lambda \sum_{\mathrm{k}=1}^{\mathrm{K}} \mathrm{Tr}(\Sigma_k^{-1}) \tag{S11}$$

Where the model parameters are defined as:

$$\Theta = \{(\omega_k, A_k, \vec{\mu}_k, \Sigma_k)\}_{\mathrm{k}=1}^{\mathrm{K}} \tag{S12}$$

2 Description of conventional LiDAR-derived vegetation structural metrics used for comparison with VSRE

To evaluate whether VSRE captures vegetation structural complexity beyond existing LiDAR-derived metrics, we compared it with five representative structural metrics that have been widely used in forest and vegetation-structure studies. These metrics were selected because they describe different and complementary dimensions of vegetation structure, including vertical development, horizontal occupancy, vertical heterogeneity, canopy-surface roughness, and integrated three-dimensional occupancy.

- Mean vegetation height was used as a simple descriptor of vertical vegetation development. It was calculated as the mean height of all terrain-normalized LiDAR returns within each sample. This metric provides a direct estimate of the average vertical position of vegetation elements and has been widely used in studies of canopy height, biomass, productivity, and habitat structure.
- Canopy cover was used to represent horizontal vegetation occupancy. In this study, canopy cover was calculated as the proportion of LiDAR returns above a height threshold of 0.2 m. This metric reflects the degree to which vegetation elements occupy the horizontal ground area and is commonly used in vegetation-cover mapping, canopy-closure assessment, and habitat-structure studies (Atkins et al., 2018).
- Foliage height diversity was used to quantify vertical structural heterogeneity. It is based on the distribution of LiDAR returns among vertical height layers and is calculated using a Shannon diversity formulation. In this study, point heights were grouped into 0.5 m vertical layers before calculating the proportional occupancy of each layer. Foliage height diversity has been widely used to characterize vertical habitat complexity and canopy stratification (Macarthur and Macarthur, 1961).
- Canopy rugosity was used to describe canopy-surface complexity. It was calculated from the canopy height model generated from terrain-normalized LiDAR point clouds and reflects the degree of spatial variation in canopy height. Higher canopy rugosity generally indicates a more irregular canopy surface and has been used to

characterize forest stand development, canopy roughness, and habitat heterogeneity (Parker and Russ, 2004).

- Canopy entropy was used as an integrated entropy-based descriptor of three-dimensional vegetation occupancy. This metric quantifies the uncertainty or evenness of point-cloud occupancy within canopy space and has recently been used to describe forest canopy structural complexity and its links to ecosystem functions. Compared with height- or surface-based metrics, canopy entropy incorporates more information from the three-dimensional point-cloud distribution (Liu et al., 2022).

3 Detailed comparison of VSRE with conventional LiDAR-derived VSC metrics

The main text summarizes the performance of VSRE across hierarchical VSC groups, mixed PCDs, and a continuous VSC gradient (**Fig. 4**). Here, we provide the complete comparisons with five conventional LiDAR-derived structural metrics: mean vegetation height, canopy cover, foliage height diversity (FHD), canopy rugosity, and canopy entropy. All metrics were calculated from the same terrain-normalized LiDAR samples and evaluated using the same statistical framework described in the Methods.

3.1 Performance across VSC groups at individual PCD levels

Across the five PCD levels, the conventional metrics showed significant group differences in many cases but did not consistently recover the expected low-to-high VSC gradient (**Fig. S7**). Mean vegetation height and FHD repeatedly assigned the medium-VSC group the lowest values and failed to consistently distinguish the low- and high-VSC groups. Canopy cover significantly separated the three groups but produced an ordering inconsistent with the treatment-derived VSC gradient. Canopy rugosity distinguished the high-VSC group but did not separate the low- and medium-VSC groups. Canopy entropy also underestimated the medium-VSC group, and at 2000 pt m^{-2} it no longer significantly separated the low- and high-VSC groups. In contrast, VSRE preserved the expected low < medium < high ordering and significant pairwise separation across all five PCD levels ($P < 0.001$).

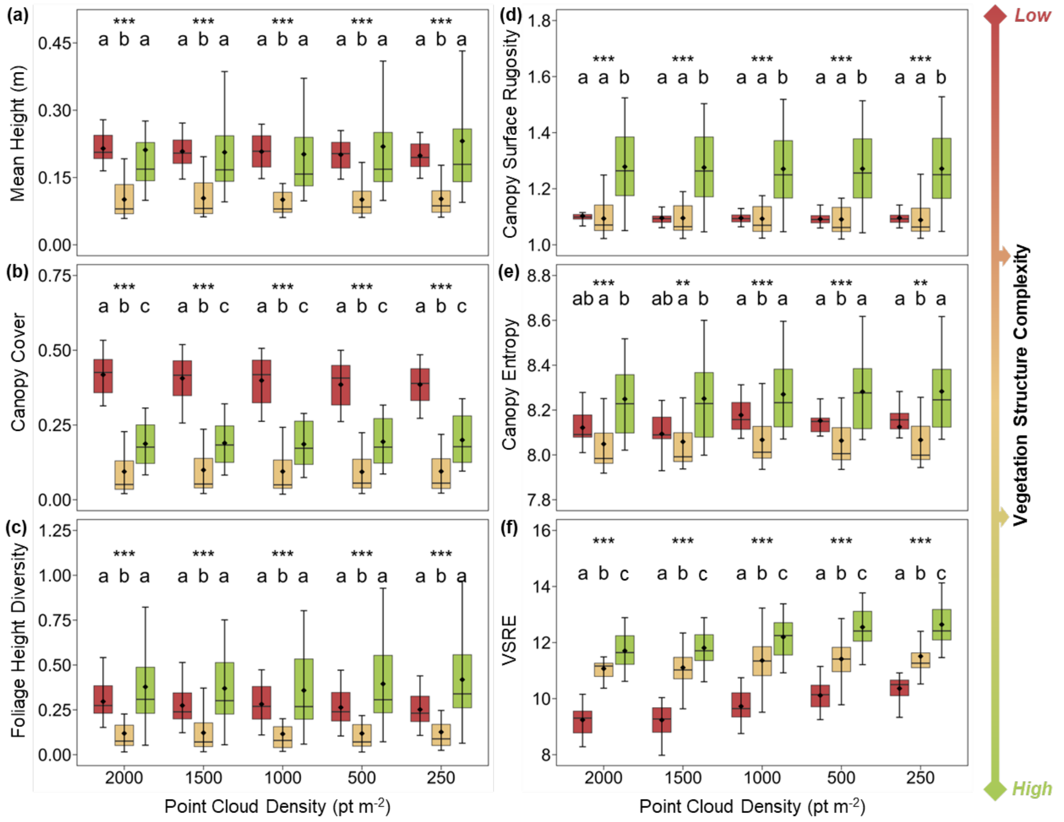

Fig. S7 | Comparison of VSRE and conventional VSC metrics in distinguishing hierarchical coastal VSC across point cloud densities. **a**, mean vegetation height; **b**, canopy cover; **c**, FHD (foliage height diversity); **d**, canopy rugosity; **e**, canopy entropy; and **f**, VSRE. Statistical differences among the low-, medium-, and high-VSC groups were tested separately at each of the five PCD levels using Kruskal–Wallis tests followed by Dunn's post hoc tests with Benjamini–Hochberg correction. Boxplots show the median, interquartile range, and non-outlier range; outliers are not displayed, and black diamonds denote group means. Different letters indicate significant pairwise differences among VSC groups within the same PCD. Asterisks indicate the significance of the corresponding Kruskal–Wallis test. Low-, medium-, and high-VSC groups are shown in red, yellow, and green, respectively.

3.2 Performance under mixed point cloud densities

To evaluate metric performance when PCD was not controlled across samples, data from all five PCD levels were pooled before comparison among the three VSC groups. All six metrics showed significant overall differences among groups ($P < 0.001$; **Fig. S8**), although their responses differed substantially. Mean vegetation height, canopy cover, FHD, and canopy entropy again assigned the medium-VSC group the lowest values and therefore did not reproduce the expected VSC ordering. Canopy rugosity showed higher values in the high-VSC group but did not significantly distinguish the low- and medium-VSC groups. VSRE retained the expected low < medium < high ordering, with significant differences among all three groups ($P < 0.001$). These results indicate that VSRE is less affected by between-sample variation in PCD than the conventional metrics evaluated here.

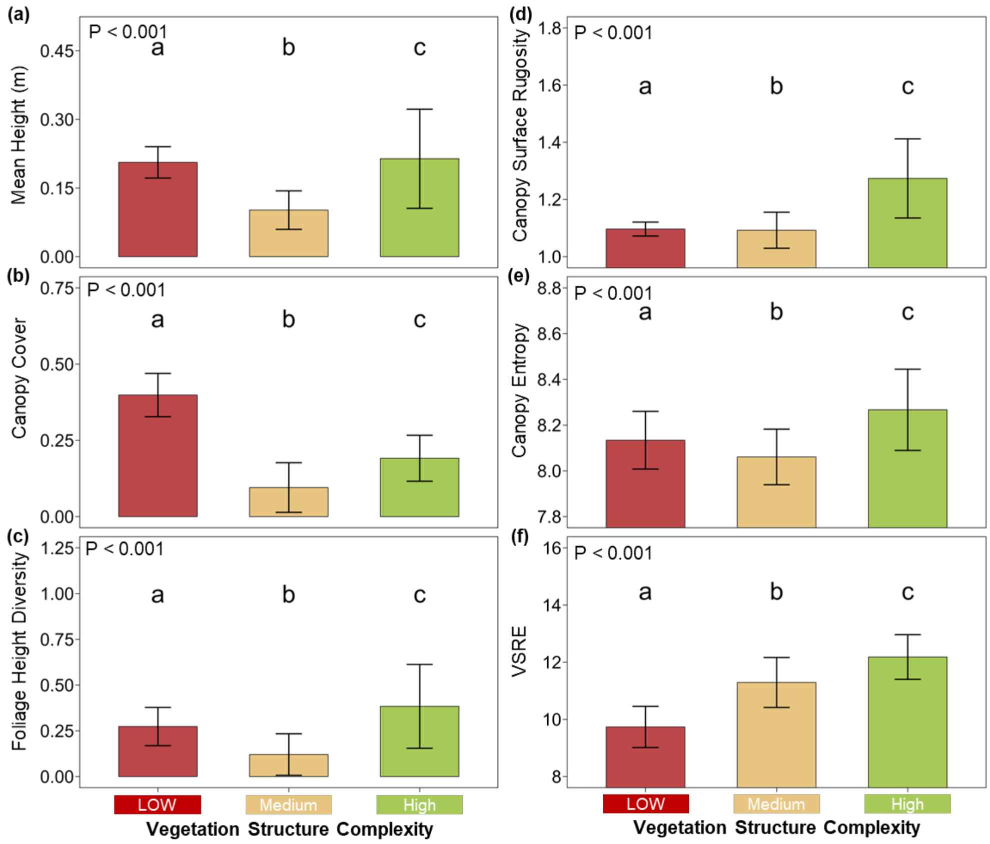

Fig. S8 | Comparison of VSRE and conventional VSC metrics under mixed point cloud densities. **a**, mean vegetation height; **b**, canopy cover; **c**, FHD (foliage height diversity); **d**, canopy rugosity; **e**, canopy entropy; and **f**, VSRE. Samples from all five PCD levels were pooled before statistical analysis. Bars indicate group means and error bars denote standard deviations. Overall differences among VSC groups were evaluated using the Kruskal–Wallis test, followed by Dunn's post hoc tests with Benjamini–Hochberg correction. Different letters indicate significant pairwise differences among groups. Low-, medium-, and high-VSC groups are shown in red, yellow, and green, respectively.

3.3 Performance along a continuous VSC gradient

To complement the categorical comparisons, we used the biomass-weighted Shannon index as a continuous proxy for VSC and fitted separate linear regressions at each PCD (**Fig. S9**). Mean vegetation height, canopy cover, and FHD did not show significant relationships with the continuous VSC proxy ($P > 0.05$). Canopy entropy showed density-dependent performance: its fit reached $R^2 = 0.41$ at 1500 pt m^{-2} but decreased to $R^2 = 0.23$ at 250 pt m^{-2}. Canopy rugosity showed significant positive relationships across all PCD levels ($P < 0.01$), with R^2 values ranging from 0.44 to 0.47. VSRE showed stronger and consistently significant relationships across all five PCD levels, with R^2 values ranging from 0.57 to 0.71 (all $P < 0.001$). The continuous-gradient analysis therefore supports the categorical evaluations and further indicates that VSRE retains sensitivity to structural variation across changes in PCD.

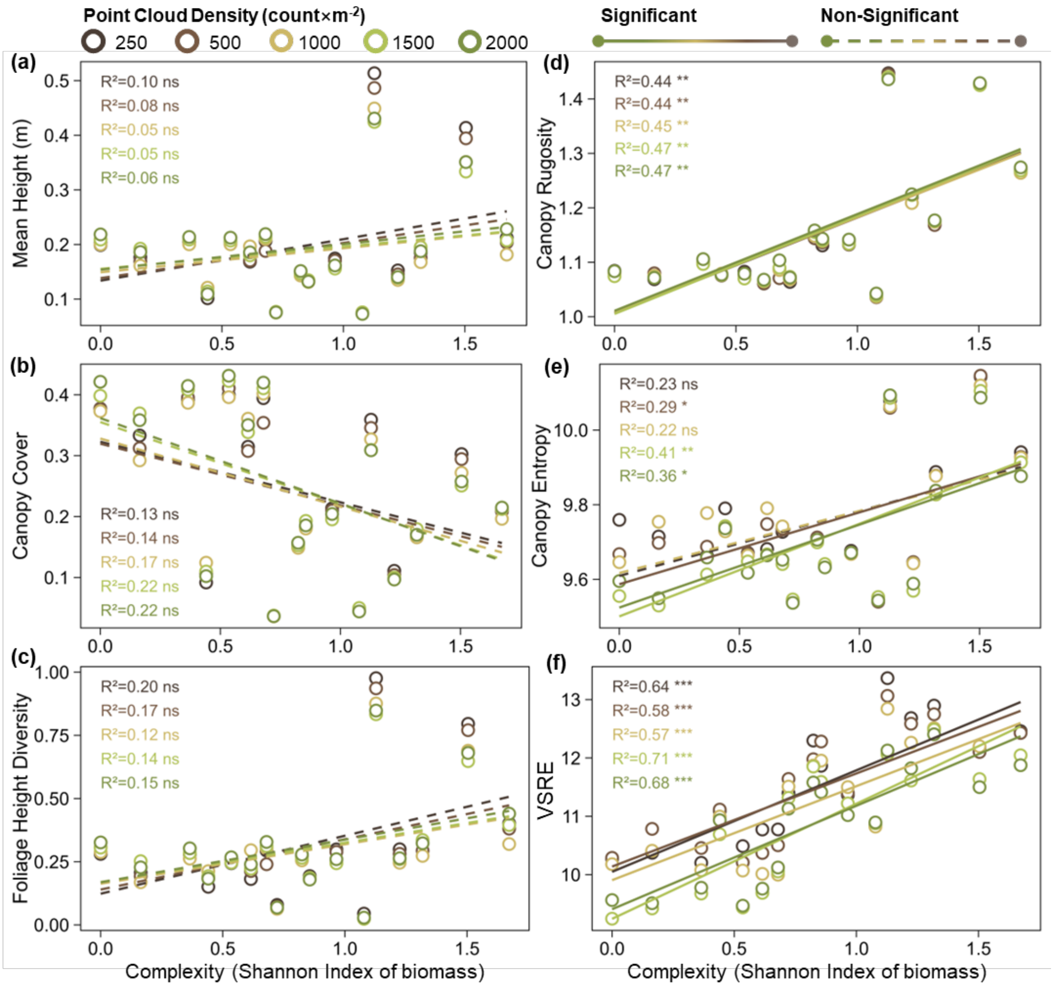

Fig. S9 | Relationships between the continuous VSC proxy and LiDAR-derived vegetation structural metrics across point cloud densities. **a**, mean vegetation height; **b**, canopy cover; **c**, FHD (foliage height diversity); **d**, canopy rugosity; **e**, canopy entropy; and **f**, VSRE. The biomass-weighted Shannon index was used as a continuous proxy for VSC. Each circle represents one plot-level UAV-LiDAR sample from a secondary plot at a given PCD. Separate linear regressions were fitted for each PCD level, and the corresponding R^2 values and significance levels are shown in each panel. Colors denote PCD levels ranging from 250 to 2000 pt m^{-2}.

References

Atkins, J. W., Bohrer, G., Fahey, R. T., Hardiman, B. S., Morin, T. H., Stovall, A. E., Zimmerman, N., and Gough, C. M.: Quantifying vegetation and canopy structural complexity from terrestrial Li DAR data using the forestr r package, Methods Ecol. Evol., 9, 2057-2066, 2018.

Beaton, A. E. and Tukey, J. W.: The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data, Technometrics, 16, 147-185, 1974.

Belagiannis, V., Rupprecht, C., Carneiro, G., and Navab, N.: Robust optimization for deep regression, Proceedings of the IEEE international conference on computer vision, 2830-2838,

Grave, E., Obozinski, G. R., and Bach, F.: Trace lasso: a trace norm regularization for correlated designs, Advances in neural information processing systems, 24, 2011.

Liu, X., Ma, Q., Wu, X., Hu, T., Liu, Z., Liu, L., Guo, Q., and Su, Y.: A novel entropy-based method to quantify forest canopy structural complexity from multiplatform lidar point clouds, Remote Sens. Environ., 282, 113280, 2022.

MacArthur, R. H. and MacArthur, J. W.: On bird species diversity, Ecology, 42, 594-598, 1961.

Parker, G. G. and Russ, M. E.: The canopy surface and stand development: assessing forest canopy structure and complexity with near-surface altimetry, For. Ecol. Manage., 189, 307-315, 2004.

Wolfe, J. H.: Pattern clustering by multivariate mixture analysis, Multivar. Behav. Res., 5, 329-350, 1970.