

Remote sensing using unmanned aerial vehicles and ensemble learning for rice aboveground biomass prediction: insights into selecting base and meta learners

Jikai LIU^{1,2}, Weiqiang WANG¹, Xueqing ZHU¹, Jun LI¹, Bo LI (✉)¹, Xinwei LI (✉)^{1,2}

1 College of Resource and Environment, Anhui Science and Technology University, Fengyang, 233100, China.

2 Anhui Province Crop Intelligent Planting and Processing Technology Engineering Research Center, Fengyang, 233100, China.

Received August 21, 2025;

Accepted January 21, 2025.

Correspondence: lixw@ahstu.edu.cn, libo@ahstu.edu.cn

© The Author(s) 2026. Published by Higher Education Press. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0>)

SUPPLEMENTARY MATERIALS

Table S1 Specific parameters of unmanned aerial vehicles platform

Parameter	Value	Parameter	Value
Weight	1487 g	Equivalent focal length	40 mm
Diagonal wheelbase	35 cm	Resolution	1600 × 1300 pixels
Endurance time	43 min	Photo format	TIFF
RTK GNSS accuracy	Vertical: 1.5 cm Horizontal: 1.0 cm	Sensor	1/2.9 CMOS
Field of view	62.7°	Pixels	2.08 megapixels

Table S2 The types of base learners

Type	Base learner
Linear-kernel models	PLSR, SMR, LAASO, RR
Tree-kernel models	RFR, XGBoost, GBM
Other models	SVR, GPR

Note: SMR, stepwise multiple regression; PLSR, partial least squares regression; RR, ridge regression; LASSO, least absolute shrinkage and selection operator; RFR, random forest regression; XGBoost, extreme gradient boosting; GBM, gradient boosting machine; SVR, support vector regression; GPR, Gaussian process regression.

Table S3 Hyperparameter configuration for machine learning base learners

Base learner	Hyperparameter	Value
RFR	mtry	Floor(sqrt(ncol(M)))
	ntree	500
	nodesize	5
SVR	sigma	Automatic calculation using sigest(X) function
	C	1
	kernel	Radial basis function (RBF)
XGBoost	nrounds	100
	max_depth	6
	eta	0.3
	gamma	0
	colsample_bytree	1
	min_child_weight	1
	subsample	1
LASSO	alpha	1
	lambda	Automatic selection via 3-fold internal cross-validation.
RR	alpha	0
	lambda	Automatic selection via 3-fold internal cross-validation.
PLSR	ncmp	min(ncol(x), 15)
	n.trees	100
GBM	interaction.depth	1
	shrinkage	0.1
	n.minobsinnode	10
GPR	sigma	Automatic calculation using sigest(X) function
	kernel	Radial basis function (RBF)
SMR		None

Note: SMR, stepwise multiple regression; PLSR, partial least squares regression; RR, ridge regression; LASSO, least absolute shrinkage and selection operator; RFR, random forest regression; XGBoost, extreme gradient boosting; GBM, gradient boosting machine; SVR, support vector regression; GPR, Gaussian process regression; X, feature matrix of the experimental dataset with dimensions $N \times M$ (N is number of samples, M is number of features).

Table S4 Sensitive features derived from recursive feature elimination method

Feature	Name
VIs	GDNVI, mNDblue, RVI, NDVI, CIrededge, NDRE, MTCI
TFs	PC1Cor, PC2Mean, PC2Cor, PC1SM, PC1Ent, PC2SM, PC1Con, PC2Ent

Note: GNDVI, green normalized difference vegetation index ; mNDblue, modified normalized difference blue index; RVI, ratio vegetation index; NDVI, normalized difference vegetation index; CIrededge, red edge chlorophyll index; NDRE, normalized difference red edge index; MTCI, MERIS terrestrial chlorophyll index; PC1, first principal component; PC2, two principal component; Cor, correlation; Mean, mean; SM, second moment; Ent, entropy; Con, contrast.

Table S5 The performance of the nine machine learning base learners based on the VIs, TFs and VIs+TFs in predicting aboveground biomass

Base learner	Dataset	Vegetation indice (VIs)		Texture feature (TFs)		VIs+TFs	
		R^2	RMSE (t·ha ⁻¹)	R^2	RMSE (t·ha ⁻¹)	R^2	RMSE (t·ha ⁻¹)
RFR	Test	0.672	2.18	0.713	2.05	0.763	1.85
SVR		0.700	2.08	0.712	2.07	0.760	1.87
SMR		0.525	2.62	0.485	2.76	0.734	1.96
PLSR		0.223	3.36	0.229	3.37	0.516	2.65
GPR		0.719	2.02	0.736	1.97	0.769	1.83
XGBoost		0.655	2.26	0.705	2.09	0.727	1.99
GBM		0.614	2.37	0.639	2.31	0.699	2.09
LASSO		0.487	2.73	0.426	2.92	0.700	2.08
RR		0.532	2.60	0.486	2.76	0.750	1.90
RFR	Training	0.704	2.08	0.676	2.16	0.770	1.84
SVR		0.711	2.07	0.703	2.07	0.757	1.90
SMR		0.507	2.70	0.505	2.68	0.753	1.91
PLSR		0.216	3.41	0.225	3.35	0.506	2.70
GPR		0.730	1.99	0.726	1.9	0.767	1.86
XGBoost		0.692	2.13	0.668	2.19	0.752	1.91
GBM		0.635	2.32	0.616	2.36	0.703	2.09
LASSO		0.476	2.79	0.435	2.86	0.692	2.14
RR		0.508	2.70	0.510	2.66	0.759	1.89

Note: SMR, stepwise multiple regression; PLSR, partial least squares regression; RR, ridge regression; LASSO, least absolute shrinkage and selection operator; RFR, random forest regression; XGBoost, extreme gradient boosting; GBM, gradient boosting machine; SVR, support vector regression; GPR, Gaussian process regression.