

**Table S1.** List of pre-processing methods for hyperspectral data of basil plants.

1 <sup>st</sup> pre-processing	2 <sup>nd</sup> pre-processing	3 <sup>rd</sup> pre-processing	Abbreviation
Raw reflectance	None	None	Raw reflectance
		SNV	SNV
		MSC	MSC
	Savitzky–Golay filter	None	SG filter
		SNV	SG filter + SNV
		MSC	SG filter + MSC
	1st derivative	None	1st Der
		SNV	1st Der + SNV
		MSC	1st Der + MSC
	2nd derivative	None	2nd Der
		SNV	2nd Der + SNV
		MSC	2nd Der + MSC
Normalization	None	None	Norm
		SNV	Norm + SNV
		MSC	Norm + MSC
	Savitzky–Golay filter	None	Norm + SG filter
		SNV	Norm + SG filter + SNV
		MSC	Norm + SG filter + MSC
	1st derivative	None	Norm + 1st Der
		SNV	Norm + 1st Der + SNV
		MSC	Norm + 1st Der + MSC
	2nd derivative	None	Norm + 2nd Der
		SNV	Norm + 2nd Der + SNV
		MSC	Norm + 2nd Der + MSC
Logarithmic transformation	None	None	Log (1/R)
		SNV	Log (1/R) + SNV
		MSC	Log (1/R) + MSC
	Savitzky–Golay filter	None	Log (1/R) + SG filter
		SNV	Log (1/R) + SG filter + SNV
		MSC	Log (1/R) + SG filter + MSC
	1st derivative	None	Log (1/R) + 1st Der
		SNV	Log (1/R) + 1st Der + SNV
		MSC	Log (1/R) + 1st Der + MSC
	2nd derivative	None	Log (1/R) + 2nd Der
		SNV	Log (1/R) + 2nd Der + SNV
		MSC	Log (1/R) + 2nd Der + MSC

**Table S2.** Hyperparameter tuning results of final model according to ensemble algorithms and feature selection methods

Prediction model	Hyperparameter and selected value for each model					
	Tested values	Feature selection method				
		Full band	RF	AdaBoost	XGBoost	LightGBM
RF	n_estimators: 10, 50, 100, 300	300	50	<b>100</b>	100	100
AdaBoost	n_estimators: 10, 50, 100, 500	100	10	50	500	<b>500</b>
	learning_rate: 0.0001, 0.001, 0.01, 0.1, 1.0	1	0.01	1	0.1	<b>0.1</b>
XGBoost	n_estimators: 50, 100, 200, 300	300	300	<b>100</b>	50	50
	learning_rate: 0.01, 0.03, 0.05, 0.07, 0.1	0.01	0.01	<b>0.1</b>	0.07	0.1
	max_depth: 3, 5, 7	3	3	<b>3</b>	3	3
	subsample: 0.7, 0.8, 0.9, 1.0	0.8	0.7	<b>1</b>	0.7	0.7
LightGBM	n_estimators: 50, 100, 200, 300	100	100	<b>100</b>	100	300
	learning_rate: 0.005, 0.01, 0.03, 0.05	0.05	0.05	<b>0.03</b>	0.03	0.03
	max_depth: 3, 5, 7	5	5	<b>3</b>	3	5
	subsample: 0.7, 0.8, 0.9, 1.0	0.7	0.7	<b>0.7</b>	0.7	0.7
	colsample_bytree: 0.8, 0.9, 1.0	0.9	1	<b>0.8</b>	0.8	1

n\_estimators, the number of trees or maximum number of estimators at which boosting is terminated;  
learning\_rate (eta), weight applied to each regressor at each boosting iteration; max\_depth, maximum  
depth of a tree; subsample, subsample ratio of the training instances.