Supplemental Materials

Machine learning enabled prediction and process optimization of VFA production from riboflavin-mediated sludge fermentation

Weishuai Li[1], Jingang Huang[1,2,*], Zhuoer Shi[1], Wei Han[1,2], Yuanyuan Lin[3], Jianfang Meng[4], Xiaobing Xu[2], Pingzhi Hou[2]

1. College of Materials and Environmental Engineering, Hangzhou Dianzi University, Hangzhou 310018, China
2. China-Austria Belt and Road Joint Laboratory on Artificial Intelligence and Advanced Manufacturing，Hangzhou Dianzi University, Hangzhou 310018, China
3. Zhejiang Province Environmental Engineering Co. Ltd, Hangzhou 310012, China.
4. M-U-T Maschinen-Umwelttechnik-Transportanlagen GmbH，Stockerau 2000，Austria.

[*]Corresponding author. Address: Room 517, The 6th Building, Hangzhou Dianzi University, Hangzhou 310018, PR China. Tel./fax: +86 13867110375; E-mail: hjg@hdu.edu.cn (J. Huang).

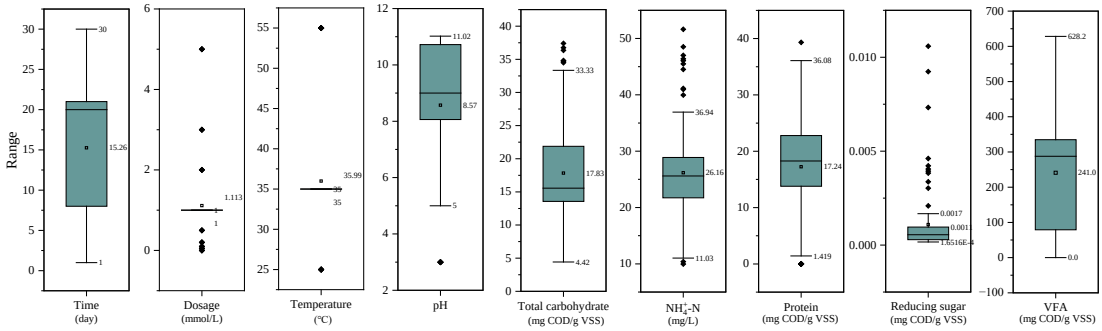

Fig. S1. The statistical analysis of collected datasets for ML modeling. The data range, mean value and the media of each input variables and output target (VFA production) were shown in the boxplot.

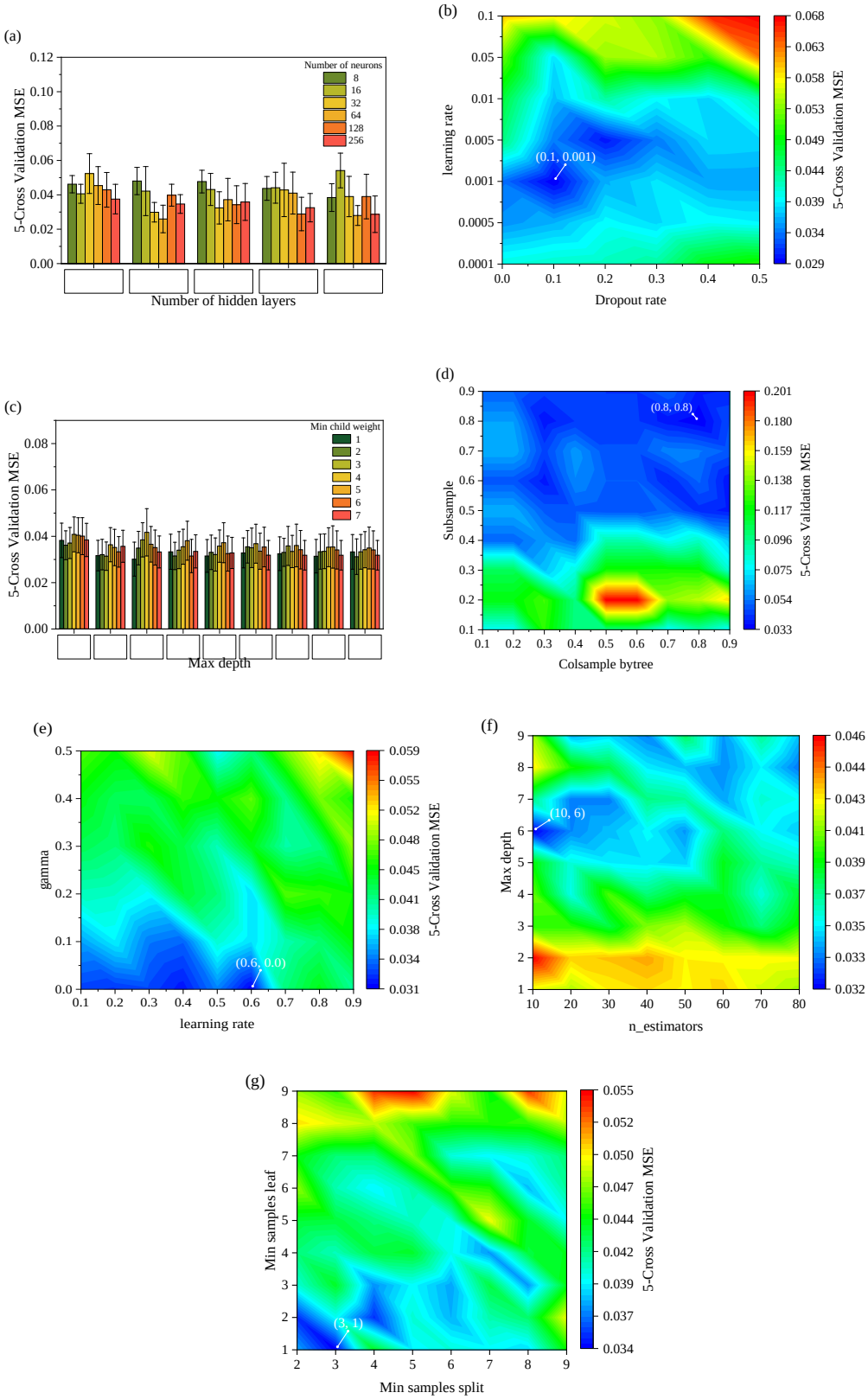

Fig. S2. Tuning process of hyper-parameters for ANN model (a, b), XGBoost model (c, d, e) and RF model (f, g)

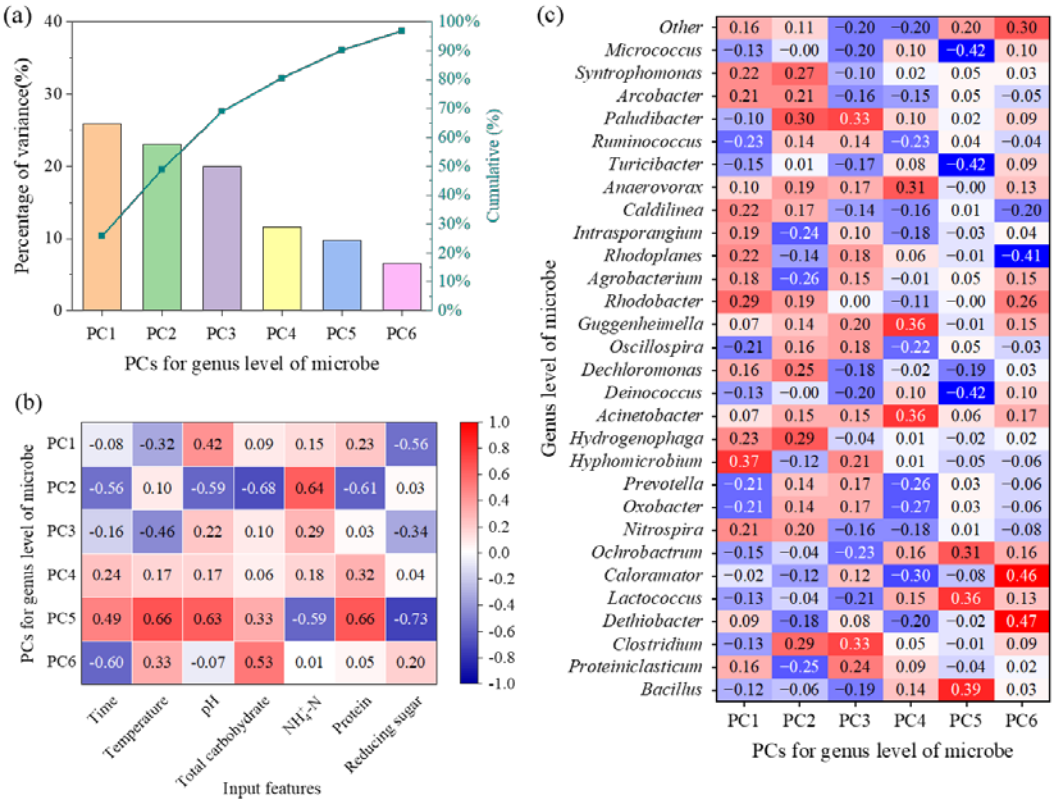

Fig. S3. (a) Compressed principal components (PCs) of high-dimensional taxonomic composition at genus level, (b) the correlation between different PCs and input features including operating conditions and intermediates, and (c) the contributions of various genus on the individual PCs

Table S1. Optimal input features for maximal VFA production based on the well-tuned XGBoost model after 10 runs by GA and PSO.

Models	Population	Item	Time (d)	Dosage (mmol/L)	Temperature (°C)	pH	Total carbohydrate (mg COD/g VSS)	NH_4^+-N (mg /L)	Protein (mg COD/g VSS)	Reducing sugar (mg COD/g VSS)	VFA (mg COD/g VSS)
GA	400	100	4	4.0	31	10.8	19.0	26.6	39.8	0.0011	650.5265
	400	100	5	1.0	38	10.6	22.2	24.8	37.2	0.0009	650.5265
	600	100	8	3.1	43	10.6	21.3	26.3	36.9	0.0010	650.5265
	600	100	6	4.3	33	10.5	20.4	25.5	37.5	0.0010	650.5265
	800	100	5	1.0	30	10.5	14.6	27.3	37.8	0.0009	650.5265
	800	100	8	2.4	30	10.7	14.7	25.1	38.1	0.0009	650.5265
	1000	100	9	3.4	32	10.8	20.7	26.8	35.5	0.0009	650.5265
	1000	100	6	1.3	42	10.5	20.2	29.2	37.0	0.0011	650.5265
	1200	100	9	4.5	35	10.8	18.2	27.8	38.0	0.0010	650.5265
	1200	100	6	0.4	43	10.6	14.4	27.0	39.4	0.0010	650.5265
PSO	400	60	1	4.1	30	10.4	19.6	28.5	36.4	0.0010	650.5265
	400	60	1	0.0	27	10.8	19.1	26.7	36.6	0.0011	650.5265
	600	60	2	4.3	30	10.8	15.4	26.0	39.3	0.0009	650.5265
	600	60	1	0.0	25	10.6	17.7	26.4	40.0	0.0009	650.5265
	800	60	1	5.0	25	10.6	20.0	25.5	40.0	0.0009	650.5265
	800	60	1	1.6	25	10.6	14.4	27.7	38.8	0.0011	650.5265
	1000	60	4	4.5	25	10.9	19.7	26.2	37.7	0.0010	650.5265
	1000	60	1	1.6	26	10.8	21.4	24.6	39.9	0.0009	650.5265
	1200	60	4	4.4	27	10.8	21.6	26.3	37.1	0.0010	650.5265
	1200	60	8	4.8	39	10.8	19.4	30.4	40.0	0.0010	650.5265