Supplementary Information

Enhancing *Verbena officinalis* L. antioxidant yield through natural deep eutectic solvents and MOF synergetic application

Haixiang Li [a,b,c], Kunze Du [b,c,#], Chong Liu [d], Yumin Yang [a,b,c], Cailin Ye [a,b,c], Xiaoxia Li [a,b,c,*], Yanxu Chang [b,c,*]

[a] School of Chinese Materia Medica, Tianjin University of Traditional Chinese Medicine, Tianjin, 301617, China

[b] State Key Laboratory of Chinese Medicine Modernization, Tianjin University of Traditional Chinese Medicine, Tianjin, 301617, China

[c] Tianjin Key Laboratory of Therapeutic Substance of Traditional Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin, 301617, China

[d] Department of Chemical & Materials Engineering, University of Auckland, 0926, New Zealand

*** Correspondence:**

Xiaoxia Li, School of Chinese Materia Medica, Tianjin University of Traditional Chinese Medicine, Tianjin 300193, China
E-mail: lixx@tjutcm.edu.cn (X. Li)

Yanxu Chang, State Key Laboratory of Modern Chinese Medicine, Tianjin University of Traditional Chinese Medicine
E-mail: Tcmcyx@tjutcm.cn (Y. Chang)

# The authors contributed equally to the first author.

2. Materials and methods

2.2. Preparation of NADES and MOF

2.2.1. Synthesis of MOF-199

A total of 2.200 g of copper nitrate hydrate was dissolved in 15.0 mL of ultra-pure water, and 1.100 g of trimeric acid was dissolved in 15.0 mL of ethanol. The two solutions were then mixed and stirred under magnetic agitation for 30 minutes. The mixture was transferred to an autoclave and heated in a drying oven at 110 °C for 20 hours. After cooling, the product was filtered, collected, and washed three times each with ethanol and distilled water. Finally, the product was activated in a drying oven at 170 °C for 48 hours, yielding the blue powder product.

2.2.2. Synthesis of MIL-101(Cr)

A total of 2.800 g of chromic nitrate nonahydrate and 1.100 g of p-phthalic acid were dissolved in 49.0 mL of ultra-pure water and 0.2 mL of hydrofluoric acid using ultrasound for 10 minutes. The resulting solution was then transferred to an autoclave and heated in a drying oven at 220 °C for 12 hours. After cooling, the product was filtered, collected, and sequentially washed with ethanol, DMF, and tetramethylammonium. Finally, the product was dried in a drying oven at 60 °C for 12 hours, yielding the green powder product.

2.2.3. Synthesis of MIL-101(Fe)

A total of 3.300 g of *p*-phthalic acid and 5.400 g of ferric chloride hexahydrate were dissolved in 100.0 mL of DMF solution. The mixture was then transferred to an autoclave and heated in a drying oven at 110 °C for 20 hours. After cooling, the product was filtered and washed twice each with ethanol and DMF. Finally, the product was dried in a drying oven at 100 °C for 8 hours, yielding the brown powder product.

2.2.4. Synthesis of ZIF-67

A total of 2.000 g of 1,2-dimethylimidazole was dissolved in 20.0 mL of methanol, and 1.700 g of cobaltous nitrate hexahydrate was dissolved in 20.0 mL of ethanol. The two solutions were then mixed and stirred magnetically for 48 hours. Finally, the mixture was centrifuged and dried, yielding the purple powder product.

2.11. Characterization

Nile red was used as a probe to determine the polarity of DESs. A volume of 1.5 mL of DES was mixed with 4.5 mL of Nile red solution at a concentration of 10 μg/mL and then vortexed in the dark for 5 min. A UV–vis spectrophotometer (Shanghai, China) was used to detect the maximum absorption wavelength within the range of 350–800 nm. The polarity parameter, E_{NR} (kJ/mol), was calculated using the following formula:

$$E_{NR} = hcN_A/\lambda_{max} = 28591/\lambda_{max} \quad (1)$$

In this equation, h, c, N_A and λ_{max} represent the Planck constant, the speed of light in a vacuum, the Avogadro constant, and the maximum absorption wavelength, respectively.

2.13. Theoretical computations

All structures were geometrically optimized, and frequency analyses were conducted prior to calculations to ensure that the structures represent true energy minimum, without any imaginary frequencies. For structural optimization, the Stuttgart-Dresden-Bonn (SDD) basis set was employed to describe the extranuclear electron distribution of the Cu atom, while the 6-31G (d) basis set was used for other elements, including C, H, O, and N. To accurately account for weak interaction forces, DFT-D3BJ empirical dispersion corrections and Basis Set Superposition Error (BSSE) were applied throughout the calculations. Given the large size of MOF-199, a cluster of MOF-199 was selected as the model for this study to reduce computational cost and improve computational efficiency.

During the computational process, the interaction Gibbs free energy of DES formation was calculated using the following formula:

$$\Delta G_{Int} = G_{DES} - nG_{HBA} - nG_{HBD} \quad (2)$$

In this equation, G_{DES}, G_{HBA} and G_{HBD} represented the Gibbs free energy of DES, HBA and HBD, respectively. ΔG_{Int} value indicated the interaction Gibbs free energy of the selected DES, with n denoting the number of HBA and HBD molecules.

Table S1. 32 kinds of NADES and its molar ratio (corresponding picture from left to right).

Abbreviation	HBA	HBD	Molar ratio			
Bet-Lac	betaine	lactic acid	1:1	1:2	1:3	1:4
ChCl-Lac	choline chloride	lactic acid	1:1	1:2	1:3	1:4
Bet-Gl	betaine	glycerin	1:1	1:2	1:3	1:4
ChCl-Gl	choline chloride	glycerin	1:1	1:2	1:3	1:4
Bet-PG	betaine	propylene glycol	1:1	1:2	1:3	1:4
ChCl-PG	choline chloride	propylene glycol	1:1	1:2	1:3	1:4
ChCl-Vit C	choline chloride	ascorbic acid	3:1	2:1	1:1	1:2
ChCl-H_2MA	choline chloride	malic acid	1:1	1:2	1:3	1:4

Table S2. Box-Benhnken design and experimental values of total contents of analytes.

Run	A (mg)	B (min)	C (%)	D (mL)	E (min)	Total contents of analytes (mg/g)
1	10	2	1	2.5	7.5	20.3284
2	10	2	2	2	7.5	19.5654
3	10	2	0	2.5	10	14.7912
4	0	2	1	2	7.5	20.3173
5	20	3	1	2.5	7.5	18.5548
6	0	2	0	2.5	7.5	14.0253
7	10	2	1	2.5	7.5	19.8459
8	20	2	1	2.5	5	19.7431
9	0	2	2	2.5	7.5	20.6688
10	10	2	1	2.5	7.5	19.7373
11	10	2	1	2.5	7.5	19.9549
12	10	2	1	3	5	20.0365
13	0	2	1	2.5	5	20.474
14	10	2	0	2	7.5	13.3254
15	10	2	1	3	10	21.3799
16	10	3	1	2.5	10	19.8052
17	10	1	1	2	7.5	19.1908
18	10	2	2	3	7.5	19.7143
19	20	2	2	2.5	7.5	18.4608
20	10	1	2	2.5	7.5	20.0679
21	10	2	1	2	5	20.7116
22	20	2	0	2.5	7.5	13.1499
23	10	1	1	3	7.5	21.1428
24	20	2	1	2	7.5	19.5898
25	20	2	1	2.5	10	19.8563
26	10	1	0	2.5	7.5	13.4674
27	0	3	1	2.5	7.5	21.1949
28	20	1	1	2.5	7.5	19.5256
29	10	2	0	2.5	5	14.4679
30	10	2	0	3	7.5	14.1296
31	10	3	2	2.5	7.5	20.0926
32	10	1	1	2.5	10	21.4306
33	10	3	0	2.5	7.5	13.4851
34	10	1	1	2.5	5	19.1027
35	0	1	1	2.5	7.5	20.3681
36	10	2	1	2.5	7.5	20.1231
37	10	3	1	3	7.5	20.1449

38	20	2	1	3	7.5	18.3262
39	10	2	2	2.5	10	20.9271
40	10	2	1	2	10	20.88
41	10	3	1	2.5	5	21.1883
42	0	2	1	2.5	10	21.4351
43	0	2	1	3	7.5	21.7021
44	10	2	1	2.5	7.5	20.0504
45	10	2	2	2.5	5	19.9202
46	10	3	1	2	7.5	20.136

A: adsorbent dosage; B: grinding time; C: DES dosage;

D: extractant volume; E: extraction time.

Table S3. Details of six machine learning models.

ML model	ML principle	Hyperparameters	Best Hyperparameters
AdaBoost	The basic principle of AdaBoost is to iteratively train a series of weak classifiers, adjust sample weights based on previous classification results, and combine these weak classifiers into a strong classifier through weighted voting to improve overall performance.	'n_estimators': (50, 200) 'learning_rate': (0.01, 0.1) 'max_depth': (3, 10) 'min_samples_split': (5, 20) 'min_samples_leaf': (3, 15)	'n_estimators': 100 'learning_rate': 0.01 'max_depth': 6 'min_samples_split': 5 'min_samples_leaf': 3
CatBoost	The basic principle of CatBoost is to handle categorical features efficiently by applying an ordered boosting algorithm, which avoids data leakage, and to build a robust ensemble of decision trees by iteratively minimizing loss while preventing overfitting through techniques like gradient-based optimization and overfitting detectors.	'iterations': (100, 200) 'depth': (1, 16) 'learning_rate': (0.05, 0.2) 'subsample': (0.5, 0.9) 'colsample_bylevel': (0.5, 0.9) 'l2_leaf_reg': (1.0, 10.0)	'iterations': 189 'depth': 1 'learning_rate': 0.0900 'subsample': 0.8895 'colsample_bylevel': 0.8577 'l2_leaf_reg': 4.2707
GBDT	The basic principle of GBDT is to build an ensemble of decision trees sequentially, where each tree corrects the residual errors of the previous trees by minimizing a loss function through gradient descent, resulting in a strong predictive model.	'n_estimators': (100, 200) 'max_depth': (1, 20) 'learning_rate': (0.01, 0.1) 'subsample': (0.5, 0.9) 'min_samples_split': (2, 10) 'min_samples_leaf': (2, 10) 'max_features': (0.5, 0.9)	'n_estimators': 200 'max_depth': 1 'learning_rate': 0.05 'subsample': 0.9 'min_samples_split': 9 'min_samples_leaf': 3 'max_features': 0.9

HistGB	The basic principle of HistGradientBoosting is to speed up traditional gradient boosting by discretizing continuous features into bins (histograms), reducing memory usage and computation time, while iteratively building decision trees to minimize the loss function.	'learning_rate': (0.01, 0.1) 'max_iter': (50, 150) 'max_depth': (3, 10) 'min_samples_leaf': (5, 20) 'l2_regularization': (0.1, 1.0)	'learning_rate': 0.01 'max_iter': 150 'max_depth': 3 'min_samples_leaf': 5 'l2_regularization': 0.1
LightGBM	The basic principle of LightGBM is to improve the efficiency of gradient boosting by using histogram-based feature discretization, a leaf-wise tree growth strategy with depth constraints, and optimizations like Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB) to reduce computation time and memory usage while maintaining high predictive performance.	'n_estimators': (100, 200) 'max_depth': (1, 20) 'learning_rate': (0.05, 0.2) 'subsample': (0.5, 0.9) 'colsample_bytree': (0.5, 0.9) 'min_child_samples': (1, 10) 'reg_alpha': (0.0, 0.5) 'reg_lambda': (0.0, 0.5)	'n_estimators': 200 'max_depth': 20 'learning_rate': 0.05 'subsample': 0.9 'colsample_bytree': 0.9 'min_child_samples': 1 'reg_alpha': 0.0 'reg_lambda': 0.5
XGBoost	XGBoost enhances gradient boosting with regularization, tree pruning, and optimized parallel computation, achieving high efficiency and accuracy.	'n_estimators': (100, 200) 'max_depth': (1, 20) 'learning_rate': (0.05, 0.2) 'subsample': (0.5, 0.9) 'colsample_bytree': (0.5, 0.9) 'min_child_weight': (1, 10) 'gamma': (0.0, 0.5)	'n_estimators': 196 'max_depth': 1 'learning_rate': 0.2 'subsample': 0.7521 'colsample_bytree': 0.7973 'min_child_weight': 3 'gamma': 0.0

Table S4. Physicochemical properties of NADES.

NADES	Viscoity (mPa.S)	E_{NR} (kJ/mol)	PH
Betaine: propylene glycol=1:4	98	53.74	7.75
Choline chloride: propylene glycol=1:3	109	52.65	5.42
Betaine: glycerol=1:2	3542	48.46	7.81
Choline chloride: glycerol=1:2	465	51.06	5.10
Choline chloride: ascorbic acid=2:1	54715	53.84	2.14
Choline chloride: malic acid=1:1	24798	55.95	0.26
Betaine: lactic acid=1:4	99	55.84	2.82
Choline chloride: lactic acid=1:4	46	53.95	0.74

Table S5. Assignment of characteristic FT-IR bands for the key functional groups.

Functional group	Vibration mode	Typical wavenumber range (cm^{-1})	Observed wavenumber (cm^{-1})
O-H stretch (lactic acid)	Stretching	3200-3600	3380
O-H stretch (after DES formation)	Stretching (H-bonded)	3200-3600	3360
C=O stretch (lactic acid)	Stretching	1700-1750	1702
C=O stretch (betaine)	Stretching	1600-1650	1625
C=O stretch (after DES formation)	Stretching (H-bonded)	1700-1750	1720
C=N stretch (betaine)	Stretching	1450-1550	1480

Table S6. ANOVA analysis of BBD-RSM regression model.

Source	Sum of Squares	df	Mean Square	F-Value	*p*-value	significant
Model	287	20	14.35	155.1	< 0.0001	**
A-A	10.53	1	10.53	113.8	< 0.0001	**
B-B	0.0058	1	0.0058	0.0632	0.8035	
C-C	147.47	1	147.47	1593.93	< 0.0001	**
D-D	0.5112	1	0.5112	5.53	0.0269	
E-E	1.48	1	1.48	15.96	0.0005	**
AB	0.8078	1	0.8078	8.73	0.0067	**
AC	0.444	1	0.444	4.8	0.038	
AD	1.75	1	1.75	18.95	0.0002	**
AE	0.1797	1	0.1797	1.94	0.1757	
BC	0	1	0	0.0001	0.9909	
BD	0.9439	1	0.9439	10.2	0.0038	**
BE	3.44	1	3.44	37.21	< 0.0001	**
CD	0.1074	1	0.1074	1.16	0.2917	
CE	0.1168	1	0.1168	1.26	0.2718	
DE	0.3452	1	0.3452	3.73	0.0648	**
A^2	0.1903	1	0.1903	2.06	0.164	
B^2	0.0043	1	0.0043	0.0464	0.8312	
C^2	91.73	1	91.73	991.48	< 0.0001	**
D^2	0.0772	1	0.0772	0.8345	0.3697	
E^2	2.96	1	2.96	31.97	< 0.0001	**
Residual	2.31	25	0.0925			
Lack of Fit	2.09	20	0.1046	2.38	0.1711	Not significant
Pure Error	0.2201	5	0.044			
Cor Total	289.31	45				
R^2	0.9920					
Adjusted R^2	0.9856					
Predicted R^2	0.9700					
CV %	1.60					

Factor A: adsorbent dosage; B: grinding time; C: DES dosage; D: extractant volume; E: extraction time; ** indicate the level of significance at $P < 0.01$.

Table S7. Regressive equations, linearity, limits of detection (LOD) and limit of quantitation (LOQ) of five target analytes (n = 6).

Compound	Regressive equation	R^2	Linear range (μg/mL)	LOD (μg/mL)	LOQ (μg/mL)
Hastatoside	y=16.709x-117.43	0.9952	5-100	0.235	0.783
Cornin	y=11.347x-51.98	0.9953	7.5-150	0.363	1.210
Acteoside	y=7.6143x-37.345	0.9945	5-100	0.600	2.000
Luteolin	y=37.622x-9.2121	0.9982	2.5-100	0.088	0.292
Apigenin	y=21.072x-1.1208	0.9958	2.5-100	0.151	0.503

Table S8. Precision, stability and repeatability of five target analytes (n = 6).

Compound	Concentration (μg/mL)	Intra-day precision RSD (%)	Inter-day precision RSD (%)	Stability RSD (%)	Repeatability RSD (%)
Hastatoside	5	1.00	2.69	1.45	2.51
	50	0.74	2.60	0.92	
	100	2.18	1.84	1.18	
Cornin	7.5	2.79	1.51	2.84	2.02
	75	1.14	2.99	1.73	
	150	1.51	1.62	2.49	
Acteoside	5	1.24	2.73	2.59	2.75
	50	1.07	2.71	2.57	
	100	2.46	2.46	1.46	
Luteolin	2.5	2.02	1.88	2.93	2.09
	10	0.73	2.69	1.08	
	100	0.76	2.40	1.15	
Apigenin	2.5	0.97	2.83	1.2	2.48
	10	0.70	2.98	1.13	
	100	0.40	2.96	1.12	

Table S9. Recovery of five target analytes (n = 6).

Compound	Sample (μg/mL)	Spiked (μg/mL)	Recovery (%)	RSD (%)
Hastatoside	168.7	170	108.54	2.19
Cornin	142.3	150	98.80	2.33
Acteoside	166.35	100	101.82	2.24
Luteolin	3.22	5	105.73	1.14
Apigenin	4.42	5	105.01	0.90

Table S10. Results of UPLC-Q-TOF-MS compositional analysis.

Serial number	Compound name	Ion mode	Retention time	Molecular formula	m/z	Quality error ppm	Secondary fragment ion	category
1	Hastatoside	+	6.47	$C_{17}H_{24}O_{11}$	427.1207	7.13	405.1309, 243.0871, 225.0753, 207.0641, 193.0506	b
2	Cornin	+	7.03	$C_{17}H_{24}O_{10}$	389.143	0.58	357.178, 195.0658, 177.0545	b
3	Acteoside	-	9.34	$C_{29}H_{36}O_{15}$	623.2004	0.23	461.1654, 315.1093, 179.0346, 161.0224, 153.0523, 135.0434, 113.0227	c
4	Luteolin	+	9.02	$C_{15}H_{10}O_6$	287.054	8.09	133.0293, 151.0038	a
5	Apigenin	-	13.26	$C_{15}H_{10}O_5$	269.0459	8.34	197.0610, 224.1741, 225.0558, 269.0456	a
6	2'-Acetylverbascoside	-	1.05	$C_{31}H_{38}O_{16}$	665.2145	5.69	503.1590, 179.0550	c
7	Cistanoside F	-	5.48	$C_{21}H_{28}O_{13}$	487.147	-2.63	179.0346, 161.0233, 135.0433	c
8	Cistanoside C	-	10.01	$C_{30}H_{38}O_{15}$	637.2137	0.15	175.0378, 193.0464, 461.1588	c
9	Rehmannitin	-	11.31	$C_{31}H_{40}O_{15}$	651.2278	2.52	193.0505, 475.169, 651.2278	b
10	Quercetin	-	7.46	$C_{15}H_{10}O_7$	301.0372	-6.04	151.0066, 175.9871, 283.9995	a

a. Flavonoids; b. Iridoids; c. Phenylethanolic glycosides.

Table S11. Parameters of ESP.

	Positive area ($Å^2$)	Negative area ($Å^2$)	Vmax (kcal/mol)	Vmin (kcal/mol)	MPI (kcal/mol)
Betaine	103.89	52.62	43.84	-68.39	31.97
Lactic acid	73.4	49.78	48.55	-41.42	16.32
Hastatoside	246.47	140.65	59.89	-59.9	15.07
Cornin	215.54	162.13	57.58	-48.86	12.55
DES	196.77	253.04	47.06	-43.85	16.12
MOF	386.52	289.41	33.5	-21.48	10.57

Table S12. Comparison of various methods.

No.	Extracted compounds	Sample Amount (g)	Type of solvent	Solvent volume (mL)	Extracted method	Extracted time (min)	Analytical method	Estimated energy consumption	Ref.
1	Acteoside, Isoverbascoside	0.3	Methanol	15 (total)	Ultrasonic	150	HPLC	~2.5-3.0 kWh/kg (high)	[88]
2	Hastatoside, Cornin, Acteoside	0.1	Ethanol：water =1:1	1.0	Ultrasonic	10	UHPSFC	~2.5 kWh/kg (moderate)	[89]
3	Hastatoside, Cornin, Acteoside	100	Water	Excess water	Boiled	55	HPLC	~5.0 kWh/kg (high)	[90]
4	Acteoside, Geniposide	10	Water	200	Macerated	20	HPLC	~0.2 kWh/kg (very low)	[91]
5	Hastatoside, Cornin, Acteoside, Luteolin, Apigenin	0.02	Water with 1.52 % DES (Bet:Lac = 1:4)	2.9	UA-MSPD	9.6	HPLC	~0.8 kWh/kg (low)	This work

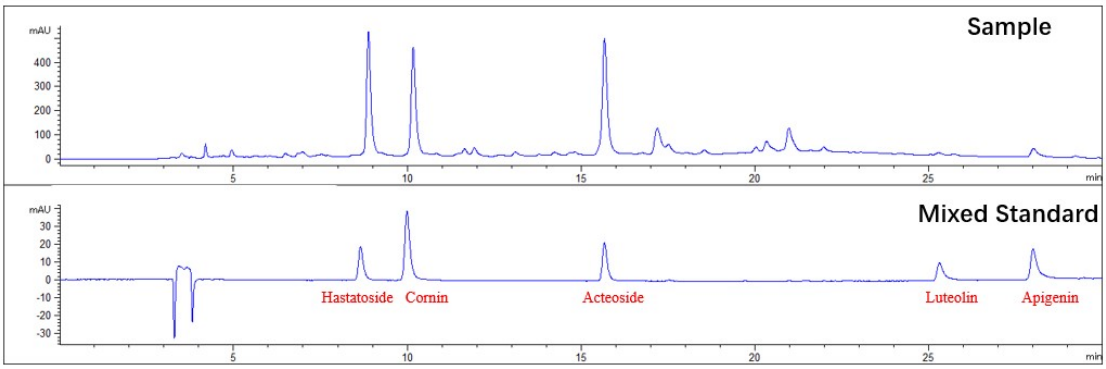

Fig. S1. Chromatogram of five target objects.

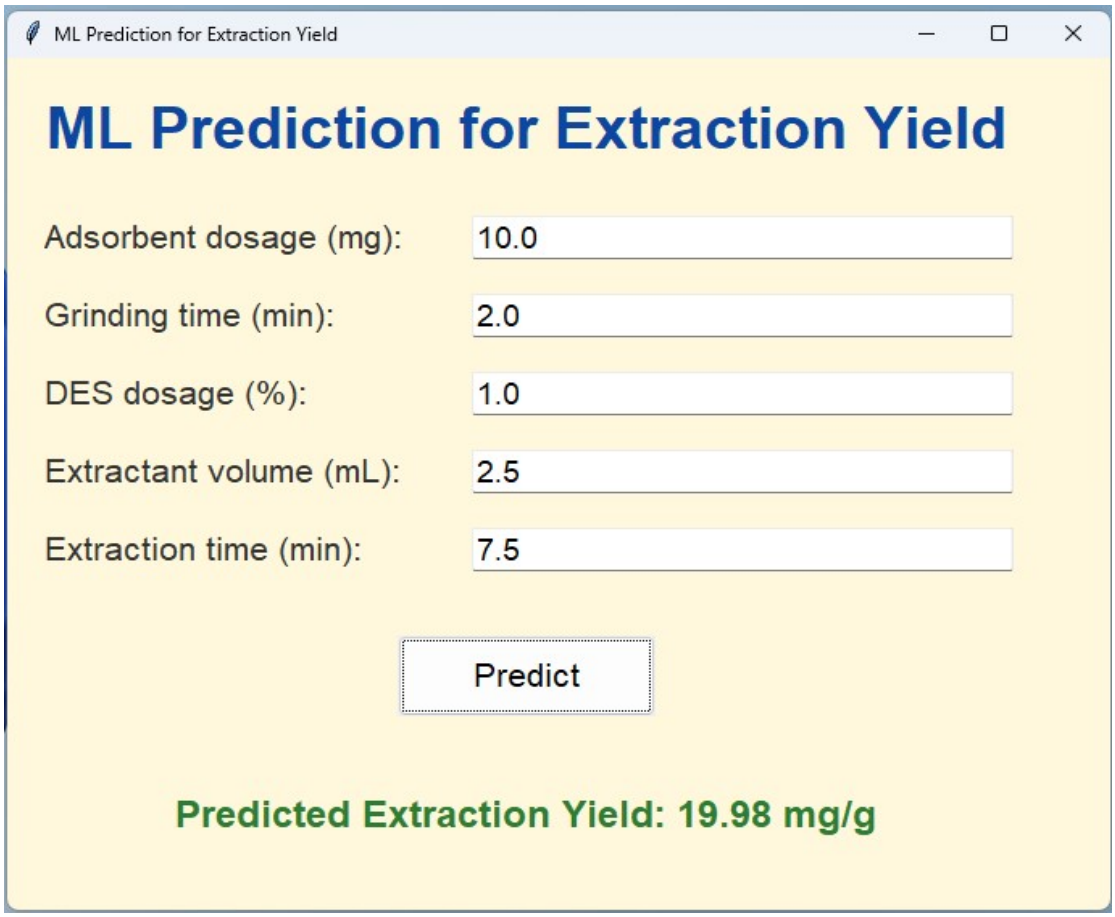

Fig. S2. The graphical user interface for predicting yield was based on the CatBoost model.

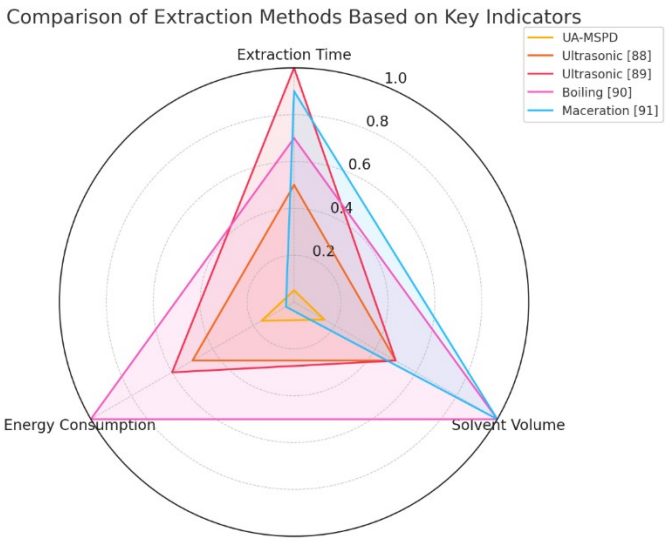

Fig. S3. Radar chart comparison of different extraction methods based on extraction time, solvent volume, and energy consumption.