Supplementary Material

Towards Sustainable Soybean Waste Management: Bio-Based Solvent Extraction of Isoflavones for Food and Cosmetic Applications

*Corresponding author: eacbat@unicamp.br

TABLES

Table S1. Elution gradient for Quantification of total isoflavones content (TIC).

Mobile phase A	Mobile phase B	Mobile phase C	Elution gradient (min)
20%	0%	80%	0 min
30%	0%	70%	0 – 6.5 min
40%	0%	60%	6.5 – 11.5 min
60%	0%	40%	11.5 – 12.5 min
80%	0%	20%	12.5 – 13.5 min
100%	0%	0%	13.5 – 16 min
100%	0%	0%	16 – 18 min
70%	0%	30%	18 – 19 min
50%	0%	50%	19 – 20 min
20%	80%	0%	20 – 22 min

Mobile phase A = Acetonitrile;
Mobile phase B = Methanol;
Mobile phase C = Water+0.1% acetic acid.

Table S2. Evaluation of total phenolic compounds yield (TPC yield) and total isoflavone compounds yield (TIC yield) in okara extracts. Extracts obtained with water or ethanol mixtures of bio-based solvents (70 wt% water or ethanol) and extractions with methanol and ethanol (30 wt% water). Extraction conditions: S/L ratio 10 mg/mL, 60 min, 500 rpm and 40 ºC.

	TIC yield yield (µg TI/g DW)					
Solvents	**Daidzin**	**Daidzein**	**Genistin**	**Genistein**	**Total**	**TPC yield**
1,2,4-Butanotriol+Water	87.42	657.34	174.93	324.79	1244.48±76.14	9.03±0.5
Methanol 70%	123.70	462.89	167.88	377.37	1131.84±4.29	1.58±0.12
Ethanol 70%	101.19	446.12	165.67	363.44	1076.42±30.97	2±0.16
Glycerol+Water	133.77	428.30	172.66	322.10	1056.83±84.97	11±0.53
Citric acid+Water	120.29	425.08	177.70	327.95	1051.02±96.24	7.68±0.67
Butyleneglycol+Water	104.26	428.86	154.91	339.91	1027.94±37.91	8.31±0.11
Malic acid+Water	114.55	423.05	158.56	323.11	1019.27±161.16	7.24±0.51
Propyleneglycol+Water	108.83	406.90	166.03	318.62	1000.38±24.05	9.56±0.18
Formic acid+Water	77.40	421.92	154.25	340.25	993.83±7.63	5.24±0.45
Lactic acid+Ethanol	126.78	406.57	116.08	340.10	989.53±2.86	5.72±0.18
Tartaric acid+Water	152.77	387.19	158.87	289.65	988.50±41.44	8.51±0.84
Malonic acid+Water	84.79	404.78	144.99	321.35	955.91±19.07	6.8±0.62
Acetic acid+Water	90.56	396.59	150.39	315.94	953.48±84.07	8.18±0.53
GVL+Water	0.00	417.41	162.93	355.50	935.85±20.50	2.57±0.27
Lactic acid+70%Water	90.95	380.06	157.96	299.50	928.47±4.68	9.24±0.87
Malic acid+Ethanol	62.79	320.07	256.78	282.74	922.38±21.14	5.58±0.40
Formic acid+Ethanol	79.22	460.77	0.00	379.23	919.22±69.68	4.49±0.40
Citric acid+Ethanol	66.62	374.55	115.53	329.36	886.07±68.36	7.71±0.19
Malonic acid+Ethanol	111.47	363.07	100.72	303.63	878.90±0.60	4.89±0.58
Acetic acid+Ethanol	71.16	359.07	99.76	318.17	848.16±10.96	3.39±0.04
1,2,4-Butanotriol+Ethanol	68.37	337.34	125.91	292.66	824.28±26.78	3.11±0.21
Glycerol+Ethanol	83.60	306.45	108.95	274.87	773.86±6.08	3.84±0.16
Ethyl Lactate+Ethanol	0.00	303.99	113.36	276.58	693.92±2.67	5.5±0.36
Propyleneglycol+Ethanol	67.02	275.34	83.70	255.14	681.20±22.89	3.04±0.04
Butyleneglycol+Ethanol	75.46	265.66	83.93	242.96	668.01±36.68	2.82±0.29
GVL+Ethanol	0.00	260.86	81.92	241.78	584.57±14.93	4.03±0.33
Ethyl Lactate+Water	0.00	0.00	0.00	0.00	0.00	8.57±0.5
Benzoic acid+Ethanol	0.00	0.00	0.00	0.00	0.00	3.67±0.0

Table S3. Experimental and model-predicted extraction values obtained for TPC yield and TPC and their respective relative and absolute deviations.

				TPC yield (mg GAE/g DW)				TPC concentration (mg GAE/100 mL)			
Test	**Temperature (°C)**	**S/L ratio (m/mL)**	**%Water**	**Observed**	**Predicted**	**Absolute deviation**	**Relative deviation (%)**	**Observed**	**Predicted**	**Absolute deviation**	**Relative deviation (%)**
1	-1	-1	-1	8.71	9.5	0.8	9.16	17.52	17.42	0.11	0.61
2	-1	-1	1	9.18	9.44	0.26	2.86	18.47	18.6	0.13	0.7
3	-1	1	-1	6.18	6.14	0.04	0.6	30.82	32.72	1.9	6.16
4	-1	1	1	6.17	6.08	0.09	1.5	30.77	30.68	0.09	0.3
5	1	-1	-1	10.08	10.22	0.14	1.34	20.29	19.93	0.36	1.76
6	1	-1	1	9.16	8.73	0.43	4.7	18.43	16.09	2.35	12.73
7	1	1	-1	7.18	6.85	0.33	4.55	35.82	35.23	0.59	1.64
8	1	1	1	5.72	5.37	0.35	6.16	28.51	28.16	0.35	1.23
9	-1.68	0	0	6.99	6.88	0.11	1.52	24.46	24.38	0.08	0.32
10	1.68	0	0	6.76	6.88	0.12	1.81	23.66	24.38	0.72	3.05
11	0	-1.68	0	12.93	12.47	0.46	3.56	12.92	14.78	1.87	14.46
12	0	1.68	0	6.34	6.82	0.48	7.5	38.05	37.8	0.24	0.64
13	0	0	-1.68	7.81	7.46	0.34	4.4	27.33	26.6	0.72	2.64
14	0	0	1.68	5.8	6.16	0.36	6.2	20.29	21.65	1.36	6.72
15	0	0	0	7.74	7.57	0.16	2.12	27.08	26.29	0.79	2.92
16	0	0	0	7.42	7.57	0.15	2.09	25.97	26.29	0.33	1.26
17	0	0	0	7.86	7.57	0.28	3.59	27.5	26.29	1.2	4.37
18	0	0	0	7.65	7.57	0.08	1.05	26.792	26.29	0.5	1.86
19	0	0	0	7.21	7.57	0.37	5.11	25.22	26.29	1.07	4.26

TPC yield (mg GAE/g DW) = Extraction yield of total phenolic compounds;
TPC concentration (mg GAE/100 mL extract) = Total phenolic compounds concentration.

Table S4. Experimental and model-predicted extraction values obtained for TIC yield and TIC and their respective relative and absolute deviations.

				TIC yield (μg TI/g DW)				TIC concentration (ug TI/mL of extract)			
Test	Temperature(°C)	S/L ratio (m/mL)	%Water	Observed	Predicted	Absolute deviation	Relative deviation (%)	Observed	Predicted	Absolute deviation	Relative deviation (%)
1	-1	-1	-1	1109.05	1020.9	88.15	7.95	23.29	20.86	2.43	10.43
2	-1	-1	1	864.32	898.05	33.73	3.9	18.15	18.45	0.29	1.62
3	-1	1	-1	886.46	875.23	11.24	1.27	44.03	43.36	0.67	1.53
4	-1	1	1	722.02	752.37	30.35	4.2	35.86	37.27	1.41	3.94
5	1	-1	-1	983.37	982.52	0.85	0.09	20.65	19.66	0.99	4.79
6	1	-1	1	858.98	859.67	0.69	0.08	17.75	17.25	0.51	2.85
7	1	1	-1	857.68	836.85	20.83	2.43	42.6	42.16	0.44	1.04
8	1	1	1	712.57	713.99	1.42	0.2	35.39	36.07	0.68	1.93
9	-1.68	0	0	857.75	848.85	8.9	1.04	30.02	29.76	0.26	0.86
10	1.68	0	0	802.56	784.31	18.26	2.27	28.09	27.75	0.34	1.22
11	0	-1.68	0	1005	1030.83	25.84	2.57	10.05	11.38	1.33	13.26
12	0	1.68	0	792.28	785.84	6.44	0.81	47.54	46.13	1.41	2.96
13	0	0	-1.68	906.63	972.01	65.37	7.21	31.73	34.11	2.37	7.48
14	0	0	1.68	811.36	765.39	45.97	5.67	28.4	26.96	1.44	5.06
15	0	0	0	803.71	816.58	12.86	1.6	28.13	28.76	0.63	2.22
16	0	0	0	846.75	816.58	30.17	3.56	29.64	28.76	0.88	2.97
17	0	0	0	804.05	816.58	12.53	1.56	28.14	28.76	0.61	2.18
18	0	0	0	802.14	816.58	14.44	1.8	28.07	28.76	0.68	2.42
19	0	0	0	27.41	28.76	1.35	4.93	27.41	28.76	1.35	4.93

TIC yield (ug TI/g DW) = Extraction yield of total isoflavones.
TIC concentration(ug TI/mL extract) = Total isoflavone compounds concentration.

Table S5. Analysis of variance for the TPC yield model.

Variation source	SS	DF	MS	Fcalc.	Ftab.
Regression	52.22	6	8.70	48.99	3.00
Residual	2.13	12	0.18		
Lack of fit	1.86	8	0.23	3.41	6.04
Pure error	0.27	4	0.07		
Total	54.35	18			
R^2 (%)	96.08%				

R^2 = % explained variation; SS = Sum of Squares; DF = Degree of Freedom;
MS = Mean Square; $F_{calc.}$ = calculated F-value; $F_{tab.}$ = tabulated F-value.

Table S6. Analysis of variance for the TPC model.

Variation source	SS	DF	MS	Fcalc.	Ftab.
Regression	699.95	6	116.66	70.82	3.00
Residual	19.77	12	1.65		
Lack of fit	16.43	8	2.05	2.46	6.04
Pure error	3.34	4	0.83		
Total	719.72	18			
R^2 (%)	97.25%				

R^2 = % explained variation; SS = Sum of Squares; DF = Degrees of Freedom;
MS = Mean Square; $F_{calc.}$ = calculated F-value; $F_{tab.}$ = tabulated F-value.

Table S7. Analysis of variance for the TIC yield model.

Variation source	SS	DF	MS	Fcalc.	Ftab.
Regression	146585.26	5	29317.05	18.62	3.03
Residual	20467.95	13	1574.46		
Lack of fit	18274.18	9	2030.46	3.70	6.00
Pure error	2193.77	4	548.44		
Total	167053.20	18			
R^2 (%)	87.75%				

R^2 = % explained variation; SS = Sum of Squares; DF = Degree of Freedom;
MS = Mean Square; $F_{calc.}$ = calculated F-value; $F_{tab.}$ = tabulated F-value.

Table S8. Analysis of variance for the TIC model.

Variation source	SS	DF	MS	Fcalc.	Ftab.
Regression	1536.27	5	307.25	154.82	3.03
Residual	25.80	13	1.98		
Lack of fit	23.11	9	2.57	3.82	6.00
Pure error	2.69	4	0.67		
Total	1562.07	18			
R^2 (%)	98.35%				

R^2 = % explained variation; SS = Sum of Squares; DF = Degree of Freedom;
MS = Mean Square; $F_{calc.}$ = calculated F-value; $F_{tab.}$ = tabulated F-value.

Table S9. Reparametrized models obtained for optimizing the concentration of total phenolic compounds and total isoflavones (models 1 and 3, respectively) and optimizing the extraction yields (models 2 and 4, respectively).

Model 1	TPC yield (mg GAE/g DW) = $7{,}57 - 1{,}68*R - 0{,}39*W - 0{,}24*T^2 + 0{,}73*R^2 - 0{,}27*W^2 - 0{,}36*T*W$	**R^2=96.08%**
Model 2	TPC (mg GAE/100 mL extract) = $26.29 + 6.84*R - 1.47*W - 0.68*T^2 - 0.77*W^2 - 1.26*T*W - 0.8*R*W$	**R^2=97.25%**
Model 3	TIC yield (µg TI/g DW) = $816.58 - 19.19*T - 72.84*R - 61.43*W + 32.43*R^2 + 18.43*W^2$	**R^2=87.75%**
Model 4	TIC (µg TI/mL extract) = $28.76 - 0.6*T + 10.33*R - 2.12*W + 0.63*W^2 - 0.92*R*W$	**R^2=98.35%**

TPC yield (mg GAE/g DW) = Extraction yield of total phenolic compounds;
TPC (mg GAE/100 mL extract) = Total phenolic compounds;
TIC yield (µg TI/g DW) = Extraction yield of total isoflavones;
TIC (µg TI/mL extract) = Total isoflavones;
T (Temperature), R (Solid-liquid ratio) and W (Water%) are linear terms of the model, T^2, R^2, W^2 are the quadratic terms of the model and T*W and R*W are the terms of interactions.

Table S10. Optimal conditions for each response, observed and predicted values, relative deviations and characterization of extracts obtained under optimal conditions.

	Optimal condition			Values			*Other results			
Response	**Temperature (°C)**	**S/L ratio**	**Water%**	**Observed**	**Predicted**	**Relative deviation (%)**	**TPC**	**Y1**	**TI**	**Y2**
TPC concentration (mg GAE/100 mL)	66.0	60	40	43.2	41.98	2.82	---	7.16	58.31	966.41
TPC yield (mg GAE/g DW)	61.9	10	40	14.64	12.7	13.25	14.88	---	11.64	1164.26
TIC concentration (µg TI/mL)	30	60	40	59.18	55.05	6.98	30.46	4.94	---	980.88
TIC yield (µg TI/g DW)	30	10	40	1112.5	1217.95	9.48	11.02	11.02	11.12	---

*Other results – relating to other analyses for the characterization of extracts obtained under optimal conditions.

Table S11. Estimated Inhibitory concentrations (IC_{50}, IC_{25} and IC_{10}) to cell lines HCT116 and A375 after 24h of exposure to extracts and the corresponding extract's solvent. IC's were calculated after interpolation of a nonlinear regression with four-parameter dose-response curve. Values missing indicate the values were out of the curve range.

		HCT116			A375		
	Solvent (%)	**IC_{50}**	**IC_{25}**	**IC_{10}**	**IC_{50}**	**IC_{25}**	**IC_{10}**
Solvent Control	**Glycerol**	6.7	3.75	2.16	9.27	5.15	2.55
	Ethanol	18.815	5.9	3.38	-	63.3	7.21
	Methanol	-	8.3	3.49	55.21	16.81	7.99
30 ºC	**60% Glycerol**	6.72	0.61	0.07	12.52	2.58	-
	60% Ethanol	7.5	3.8	2.64	14.52	3.77	2.68
	60% Methanol	-	6.27	3.22	-	-	-
66 ºC	**60% Glycerol**	5.21	0.85	0.19	16.07	5.15	0.507
	60% Ethanol	7.53	3.21	1.95	16.16	3.85	2.70
	60% Methanol	21.75	6.52	3.22	-	-	-

Table S12. Concentration of the isoflavones Daidzin, Daidzein, Genistin e Genistein (µg/mL) present in the extracts at the computed IC_{50} for HCT116 and A375 cell lines.

		HCT116				A375			
Solvent	**Temperature (°C)**	**Daidzin**	**Daidzein**	**Genistin**	**Genistein**	**Daidzin**	**Daidzein**	**Genistin**	**Genistein**
60 % Glycerol	30	0.029	0.154	0.101	0.112	0.054	0.288	0.189	0.210
	66	0.021	0.121	0.072	0.089	0.065	0.374	0.222	0.276
60 % Ethanol	30	0.030	0.187	0.074	0.155	0.055	0.351	0.139	0.291
	66	0.034	0.182	0.070	0.151	0.072	0.389	0.150	0.325
60 % Methanol	30	-	-	-	-	-	-	-	-
	66	0.127	0.544	0.204	0.442	-	-	-	-

Table S13. Concentration of total phenolic content (TPC) (mg GAE/100 mL) and total isoflavones (TI) (µg/mL) present in the extracts at the computed IC_{50} for HCT116 and A375.

Solvent	Temperature (°C)	HCT116		A375	
		TPC	TI	TPC	TI
60% Glycerol	30	0.00205	0.397	0.00382	0.741
	66	0.00225	0.304	0.00694	0.937
60% Ethanol	30	0.00063	0.446	0.00118	0.836
	66	0.00074	0.437	0.00160	0.937
60% Methanol	30	-	-	-	-
	66	0.00209	1.316	-	-

Table S14. Values of the two-way ANOVA test to evaluate differences between fitted curves of extracts regarding the solvent (F (DFn, DFd), F – F distribution, DFn – degrees of freedom of numerator, DFd – degrees of freedom of the denominator; significant differences in bold, $p < 0.05$).

	HCT116			**A375**		
Solvent	**Methanol**	**Ethanol**	**Glycerol**	**Methanol**	**Ethanol**	**Glycerol**
Interaction	F (14, 48) = 0.4451 p = 0.9502	F (14, 48) = 1.125 p = 0.3622	F (14, 48) = 1.752 p = 0.0757	F (14, 48) = 1.559 p = 0,1267	F (14, 48) = 16,28 p < 0.0001	F (14, 48) = 34.95 p < 0,0001
Concentration	F (7, 48) = 23.23 p < 0.0001	F (7, 48) = 60.93 P < 0.0001	F (7, 48) = 35.59 p < 0.0001	F (7, 48) = 69,07 p < 0,0001	F (7, 48) = 133.1 p < 0.0001	F (7, 48) = 596.5 p < 0,0001
Solvent	F (2, 48) = 0.1964 p = 0.8224	F (2, 48) = 2,498 p = 0.0929	F (2, 48) = 19.54 p < 0.0001	F (2, 48) = 68,82 p < 0,0001	F (2, 48) = 3.207 p = 0.0493	F (2, 48) = 327.2 p < 0.0001

Table S15. Values of the two-way ANOVA test to evaluate differences between fitted curves of extracts regarding the extraction temperature (F (DFn, DFd), F – F distribution, DFn – degrees of freedom of numerator, DFd – degrees of freedom of the denominator; significant differences in bold, $p < 0.05$).

	HCT116		**A375**	
Temperature (ºC)	**30**	**66**	**30**	**66**
Interaction	F (14. 48) = 2.681 p = 0.0056	F (14. 48) = 1.711 p = 0.0846	F (14. 48) = 41.13 p < 0.0001	F (14. 48) = 21.53 p < 0.0001
Concentration	F (7. 48) = 31.11 p < 0.0001	F (7. 48) = 37.83 p < 0.0001	F (7. 48) = 183.3 p < 0.0001	F (7. 48) = 86.98 p < 0.0001
Solvent	F (2. 48) = 51.78 p < 0.0001	F (2. 48) = 40.50 p < 0.0001	F (2. 48) = 479.7 p < 0.0001	F (2. 48) = 158.0 p < 0.0001

FIGURES

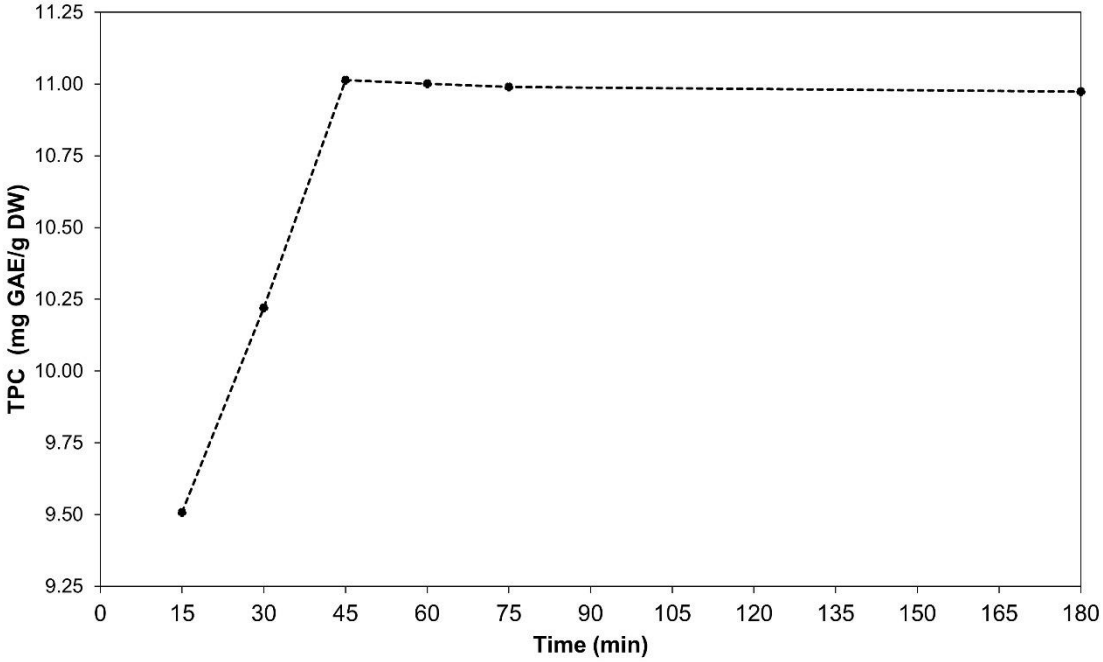

Figure S1. Evaluation of the effect of extraction time of TPC yield from okara using glycerol+70% water.

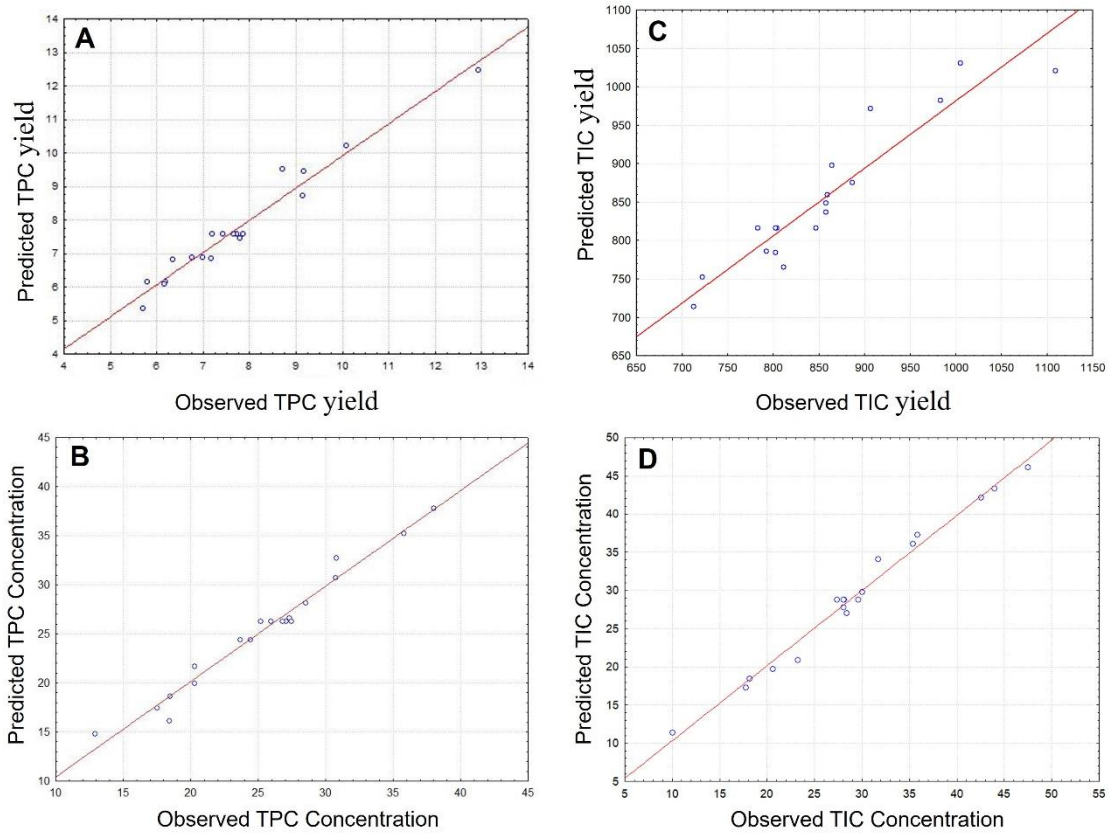

Figure S2. Observed and predicted values for the responses of (A) extraction yield of total phenolic compounds (TPC yield) (B) total phenolic compounds in the extracts (TPC concentration) (C) extraction yield of total isoflavone compounds (TIC yield) and (D) total isoflavone compounds in the extracts (TIC concentration).

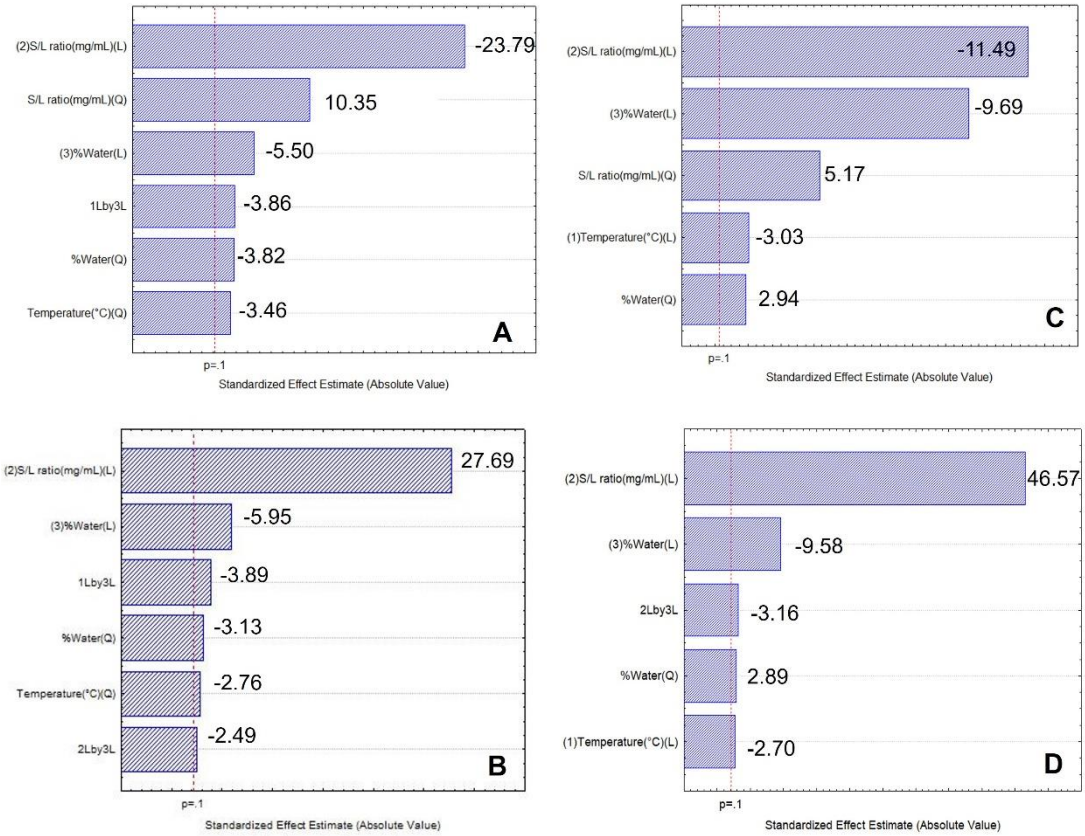

Figure S3. Pareto diagrams for the significant terms of the models for the results of (A) extraction yield of total phenolic compounds (TPC yield) (B) total phenolic compounds in the extracts (TPC concentration) (C) extraction yield of total isoflavone compounds (TIC yield) and (D) total isoflavone compounds in the extracts (TIC concentration).