Supplementary material

Application of smartphone-based handheld Raman spectrometer and machine learning for the essential oil evaluation

Leo Lebanov[a,b], Brett Paull[a,b*]

[a] *Australian Centre for Research on Separation Science (ACROSS), School of Natural Sciences, University of Tasmania, Hobart, TAS, Australia*

[b] *ARC Industrial Transformation Research Hub for Processing Advanced Lignocellulosics Products (PALs), School of Natural Sciences, University of Tasmania, Hobart, TAS, Australia*

Brett Paull (Corresponding Author); **E-mail:** brett.paull@utas.edu.au

Full postal address: *School of Natural Sciences, Private Bag 75, University of Tasmania, Hobart, TAS 7001, Australia*

Tel.-Fax: +61 3 6226 6680

ORCID: https://orcid.org/0000-0001-6373-658

Table S1. Performance of RF model based on out-of-bag error (OOBE) for different preprocessing approaches for discrimination of pure EOs and adulterated with BA

	OOBE (%)
Raw	2.78
MSC	2.78
MSC+SG1	2.78

Table S2. Performance of PLS-DA models based on different preprocessing approaches for discrimination of pure EOs and adulterated with BA

		Raw (5 LVs)			MSC (3 LVs)			MSC 1st Ord (6 LVs)		
		Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
BA	Cal	0.750	0.958	0.889	0.750	0.917	0.861	0.917	0.875	0.889
	CV	0.750	0.833	0.806	0.750	0.875	0.833	0.833	0.833	0.833
Pure	Cal	0.958	0.750	0.889	0.917	0.750	0.861	0.875	0.917	0.889
	CV	0.833	0.750	0.806	0.875	0.750	0.833	0.833	0.833	0.833

Table S3. Performance of individual PLS models for BA quantification

		RECV	RMSECV	REP	RMSEP
Cinnamon	Raw	4.64	1.69	2.88	1.04
	MSC	4.13	1.50	2.80	1.01
	MSC+SG1	3.10	1.13	2.41	0.87
Clove bud	Raw	12.8	4.66	9.76	3.51
	MSC	7.89	2.87	6.52	2.34
	MSC+SG1	4.76	1.73	4.09	1.47
Geranium	Raw	16.5	6.00	12.9	4.63
	MSC	10.8	3.92	7.68	2.76
	MSC+SG1	7.22	2.63	6.40	2.30
Rosemary	Raw	7.72	2.81	6.30	2.26
	MSC	2.99	1.09	4.81	1.73
	MSC+SG1	3.29	1.20	2.76	0.99

Table S4. Performance of PLS models for vegetable oil quantification

		RECV	RMSECV	REP	RMSEP
	Raw	4.25	1.55	4.39	1.58
	MSC	6.03	2.19	4.58	1.65
	MSC+SG1	5.20	1.89	4.94	1.77

Table S5. Performance of RF model based on out-of-bag error (OOBE) for different preprocessing approaches for discrimination of pure YY EOs and adulterated with benzyl alcohol, vegetable oil and lower price EO

	OOBE (%)
Raw	0.00
MSC	3.45
MSC+SG1	3.45

Table S6. Performance of PLS-DA models based on different preprocessing approaches for discrimination of pure YY EOs and adulterated with BA, VO and EO

		Raw (7 LVs)			MSC (6 LVs)			MSC 1st Ord	8 LVs	
		Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
YY BA	Cal	1	1	1	1.000	0.833	0.96	1	1	1
	CV	1	1	1	1	0.667	0.92	1	1	1
YY EO	Cal	1	1	1	1	1	1	1	1	1
	CV	0.947	1	0.960	1	1	1	0.947	1	0.96
YY VO	Cal	1	1	1	1	1	1	1	1	1
	CV	1	1	1	0.947	1	0.96	1	1	1
YY Pure	Cal	1	1	1	1	1	1	1	1	1
	CV	1	0.857	0.966	1	0.571	0.88	1	0.714	0.92

Table S7. Performance of PLS models for quantification of BA, VO and cypress EO as adulterants (updated)

		RECV	RMSECV	REP	RMSEP
YY BA	Raw	10.26	3.85	11.36	4.19
	MSC	4.94	1.80	7.59	2.73
	MSC+SG1	9.62	3.50	9.15	2.75
YY VO	Raw	9.34	3.51	3.80	1.40
	MSC	4.31	1.57	3.37	1.21
	MSC+SG1	7.38	2.77	5.43	1.63
YY EO	Raw	8.41	3.16	6.33	2.33
	MSC	2.82	1.17	4.41	1.83
	MSC+SG1	6.78	2.55	5.39	1.99

Fig. S1. Image of the Smart Raman XI™ with liquid vial holder

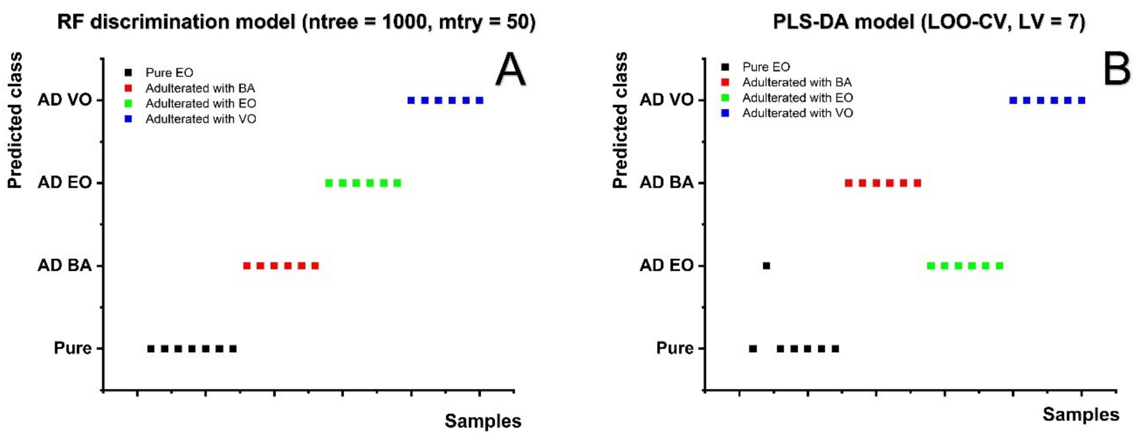

Fig. S2. Prediction performance of (A) Random Forest and (B) PLS-DA models in the cross-validation procedure for the discrimination of pure from adulterated ylang-ylang essential oils and identification of adulterant

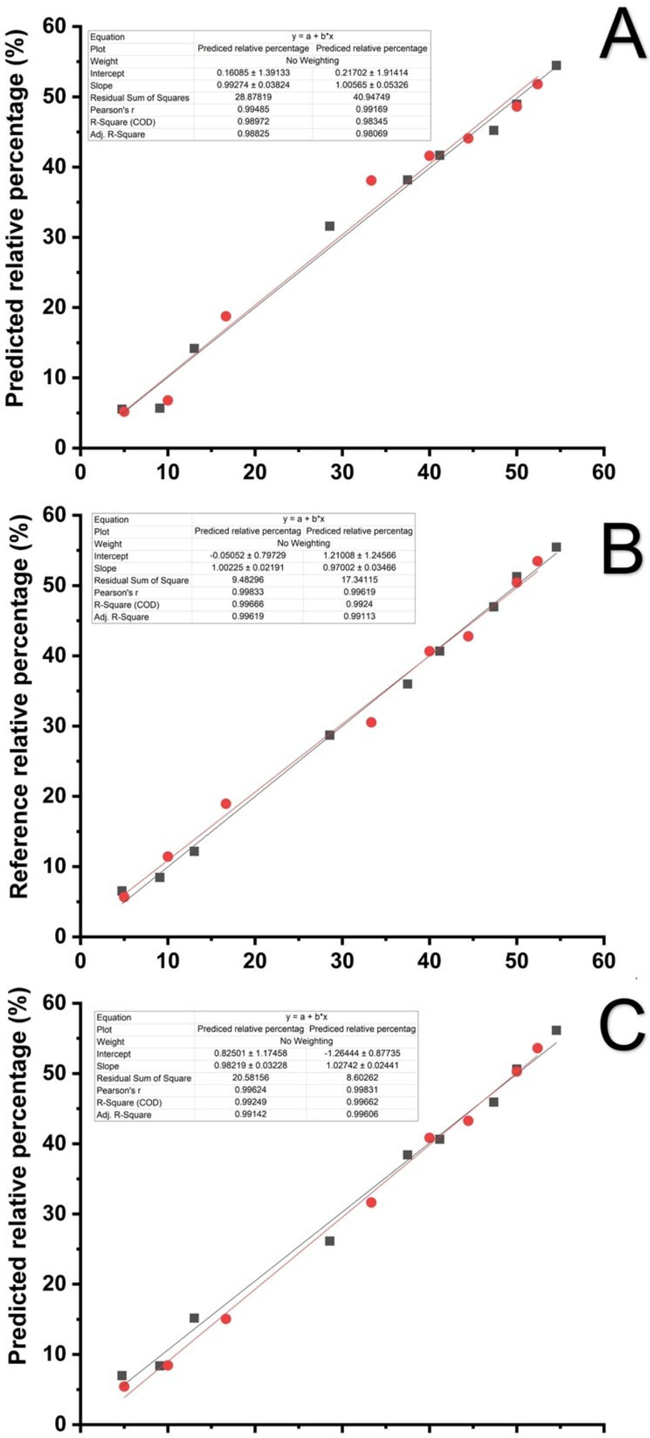

Fig. S3. Predicted relative percentage in the cross-validation (CV) (black square) and in independent test set (double cross validation (DCV)) (red circle) vs reference relative percentage for the quantification of (A) benzyl alcohol (BA), (B) cheaper essential oil (EO) and (C) vegetable oil (VO) in ylang-ylang EOs