

# A Systematized Review on the Applications of Hyperspectral Imaging in the Quality of Potatoes.

Carlos Miguel Peraza-Alemán<sup>1</sup>, Ainara López-Maestresalas<sup>1</sup>, Carmen Jarén<sup>1</sup>, Niuton Rubio-Padilla<sup>1</sup>, Silvia Arazuri<sup>1</sup>.

<sup>1</sup>Department of Engineering, Institute on Innovation and Sustainable Development in Food Chain (IS-FOOD), Universidad Pública de Navarra (UPNA), Pamplona, Spain.

Silvia Arazuri (silvia.arazuri@unavarra.es)

## Supplementary Material

**Scale S1** Quality scale designed for studies using machine or deep learning models in the field of food technology and agronomy\*

Author and year:

Name of article:

“low risk”: >75 % scores in relation to the total scores

“moderate risk”: 45 - 74%

“high risk”: <44%

### Sample related:

1. Do the samples capture the full variability of the population analysed?
  - a. Yes
  - b. No
  - c. Unclear (u)
  - d. Not applicable (n/a)
2. Was the sample size adequate? ( $n \geq 100$ )\*\*
  - a. Yes
  - b. No
  - c. Unclear (u)
  - d. Not applicable (n/a)

### Concerning data collection methodology:

3. Were the materials and methods correctly described to allow for reproducibility?
  - a. Yes
  - b. No
  - c. Unclear (u)
  - d. Not applicable (n/a)
4. Were the most important factors that can affect the results of the analysis controlled?
  - a. Yes
  - b. No
  - c. Unclear (u)
  - d. Not applicable (n/a)

### Concerning data analysis

5. Were appropriate statistical techniques used?
  - a. Yes
  - b. No
  - c. Unclear (u)
  - d. Not applicable (n/a)
6. Were all the relevant performance metrics reported?
  - a. Yes
  - b. No
  - c. Unclear (u)
  - d. Not applicable (n/a)

7. Was external validation of the models carried out?
  - a. Yes
  - b. No
  - c. Unclear (u)
  - d. Not applicable (n/a)

**Concerning the results**

8. Were the results adequately discussed?
  - a. Yes
  - b. No
  - c. Unclear (u)
  - d. Not applicable (n/a)
9. Are the conclusions relevant and do they advance knowledge in the field of study?
  - a. Yes
  - b. No
  - c. Unclear (u)
  - d. Not applicable (n/a)

\*The description of how to correctly apply the quality scale is presented in the article.

\*\*We set a value of 100 samples as adequate due to our own experience and based on the fact that obtaining good models requires a considerable number of samples.

**Table S2** Detailed quality assessment of the studies included ( $n=52$ )

Study	Question 1	Question 2	Question 3	Question 4	Question 5	Question 6	Question 7	Question 8	Question 9	Total scores	%	Risk of bias
Al-Mallahi et al. (2008)	n/a	1	1	1	1	1	0	1	1	7	88	“low risk”
Nguyen Do Trong et al. (2011)	1	0	1	1	1	1	1	1	1	8	89	“low risk”
Rady et al. (2014)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Rady et al. (2015)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Su and Sun (2016a)	1	0	1	1	1	1	1	1	1	8	89	“low risk”
Su and Sun (2016b)	1	0	1	1	1	1	1	1	1	8	89	“low risk”
López-Maestresalas et al. (2016)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Kjær et al. (2016)	1	0	1	1	1	1	0	1	1	7	78	“low risk”
Kjær et al. (2017)	1	1	1	1	1	1	0	1	1	8	89	“low risk”
Ye et al. (2018)	1	1	1	1	1	0	1	1	1	8	89	“low risk”
Gao et al. (2018)	1	1	1	1	1	1	1	1	1	6	67	“low risk”
Amjad et al. (2018)	1	0	1	1	1	1	0	1	1	7	78	“low risk”
Moscetti et al. (2018)	1	0	1	1	1	1	1	1	1	8	89	“low risk”
Ji et al. (2019a)	1	1	1	1	1	0	1	1	1	8	89	“low risk”
Ji et al. (2019b)	1	1	1	1	1	0	1	1	1	8	89	“low risk”
Polder et al. (2019)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Rady et al. (2019)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Lu et al. (2019)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Song et al. (2020)	1	0	n/a	0	1	0	1	0	1	4	50	“moderate risk”
Bai et al. (2020)	1	1	1	1	1	1	1	1	1	9	100	“low risk”

Study	Question 1	Question 2	Question 3	Question 4	Question 5	Question 6	Question 7	Question 8	Question 9	Total scores	%	Risk of bias
Rady et al. (2020)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Li et al. (2020)	1	n/a	1	1	1	1	1	1	1	8	100	“low risk”
Xiao et al. (2020)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Garhwal et al. (2020)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Van De Vijver et al. (2020)	1	1	1	1	1	1	u(0)	1	1	8	89	“low risk”
Žibrat et al. (2021)	1	0	1	1	1	1	0	1	1	7	78	“low risk”
Rady et al. (2021)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Wang et al. (2021a)	1	0	1	1	1	1	1	1	1	8	89	“low risk”
Wang et al. (2021b)	1	0	1	1	1	1	1	1	1	8	89	“low risk”
Liu et al. (2021)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Zhao et al. (2021)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Abdelbaki et al. (2021)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Logan et al. (2021)	1	u (0)	1	1	1	0	0	1	1	6	67	“moderate risk”
Duarte-Carvajalino et al. (2021)	1	1	1	1	1	1	0	u (0)	1	7	78	“low risk”
Li et al. (2021a)	1	1	1	1	1	1	1	0	1	8	89	“low risk”
Li et al. (2021b)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Cui et al. (2022)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Lapajne et al. (2022)	1	1	1	1	1	1	0	1	1	8	89	“low risk”
Appeltans et al. (2022)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Shi et al. (2022)	1	n/a	1	1	1	1	1	1	1	8	100	“low risk”
Wang and Wang (2022)	1	0	1	1	1	1	1	1	1	8	89	“low risk”

Study	Question 1	Question 2	Question 3	Question 4	Question 5	Question 6	Question 7	Question 8	Question 9	Total scores	%	Risk of bias
Liu et al. (2022a)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Liu et al. (2022b)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Liu et al. (2022c)	1	1	1	1	1	0	1	1	1	8	89	“low risk”
López-Maestresalas et al. (2022)	1	0	1	1	1	1	1	1	1	8	89	“low risk”
Wang et al. (2022)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Zhou et al. (2022)	1	1	1	1	1	1	0	1	1	8	89	“low risk”
Zou et al. (2022)	1	1	1	1	1	1	0	1	1	8	89	“low risk”
Qi et al. (2023)	1	1	1	1	1	1	0	1	1	8	89	“low risk”
Tunny et al. (2023)	1	n/a	1	1	1	1	1	1	1	7	88	“low risk”
Gao et al. (2023)	1	1	1	1	1	1	1	1	1	9	100	“low risk”
Muruganantham et al. (2023)	1	0	0	1	1	1	0	1	0	5	55	“moderate risk”

**Table S3** Overview of applications of hyperspectral imaging to measure different components in potatoes through regression models ( $n=28$ )

Studies	Number of varieties	Number of potatoes / leaves	Range	Compound to predict	Measure	Analysis	R <sup>2</sup> <sub>cal</sub>	RMSE <sub>cal</sub>	R <sup>2</sup> <sub>cv</sub>	RMSE <sub>cv</sub>	R <sup>2</sup> <sub>val</sub>	RMSE <sub>val</sub>	RPD <sub>val</sub>	Type of sample	
Rady et al. (2014)	1	200	400-1000 nm	glucose	%	PLSR	0.78	-	-	0.1557	0.74	0.1643	1.49	sliced	
				sucrose	%		0.78	-	-	0.0636	0.62	0.058	1.23		
				specific gravity	g/cm <sup>3</sup>		0.45	-	-	0.0107	0.26	0.0097	0.99		
				primordial leaf count	%		0.96	-	-	0.3256	0.94	0.2492	2.92		
				soluble solids	Brix		0.46	-	-	0.3755	0.36	0.3234	1.07		
				glucose	%	PLSR	0.75	-	-	0.3669	0.52	0.3259	1.19		whole
				sucrose	%		0.18	-	-	0.0817	0.14	0.0702	1.02		
				specific gravity	g/cm <sup>3</sup>		0.3	-	-	0.0107	0.2	0.0095	1.01		
				primordial leaf count	-		0.49	-	-	13.124	0.47	11.7014	1.14		
				soluble solids	Brix		0.34	-	-	0.4629	0.24	0.4602	1.03		
Rady et al. (2015)	1	540 180	400-1000 nm	sucrose	%	PLSR	0.68	0.0248	-	0.0269	0.58	0.0268	1.23	sliced	
				glucose	%		0.98	0.0219	-	0.0267	0.97	0.0251	3.7		
Su and Sun (2016a)	-	48 -	900-1700 nm	volatility of tuber compositions	-	FMCIA-TBPANN	0.97	0.009	-	-	0.951	0.018	-	sliced	
				tubercooking degree	-		0.986	0.166	-	-	0.992	0.145	-		
Kjaer et al. (2016)	10	60	380-925 nm	density	g/L	PLSR	-	-	0.71	0.006	-	-	-	whole	
				dry matter	%		-	-	0.66	1.54	-	-	-		
				starch	%		-	-	0.69	1.6	-	-	-		
				sucrose	mg/g		-	-	0.42	0.37	-	-	-		
				asparagine	mg/g		-	-	0.7	0.88	-	-	-		
				aspartate	mg/g		-	-	0.46	0.219	-	-	-		
				glutamate	mg/g		-	-	0.42	0.166	-	-	-		
				tyrosine	mg/g		-	-	0.47	0.195	-	-	-		
				valine	mg/g		-	-	0.6	0.207	-	-	-		
				cell area	μm <sup>2</sup> ×0.001		-	-	0.24	2522	-	-	-		
				cell perimeter	μm×0.01		-	-	0.21	69	-	-	-		
				conductivity	mS/cm		-	-	0.45	2.14	-	-	-		
				400-1050 nm	glucose		mg/g	PLSR	-	-	0.33	2.22	-		-
			fructose		mg/g	-	-		0.39	1.6	-	-	-		
			glutamine		mg/g	-	-		0.58	1.11	-	-	-		
			tryptophan		mg/g	-	-		0.26	0.07	-	-	-		

Studies	Number of varieties	Number of potatoes / leaves	Range	Compound to predict	Measure	Analysis	R <sup>2</sup> <sub>cal</sub>	RMSE <sub>cal</sub>	R <sup>2</sup> <sub>cv</sub>	RMSE <sub>cv</sub>	R <sup>2</sup> <sub>val</sub>	RMSE <sub>val</sub>	RPD <sub>val</sub>	Type of sample
Kjær et al. (2017)	4	318	433-968 nm	chlorophyll	mg/kg	PLSR	-	-	0.916	0.0377	-	-	-	peeled tuber
				glycoalkaloids (TGA)	mg/kg		-	-	0.227	261.46	-	-	-	
				dry matter	%		-	-	0.726	1.9	-	-	-	
Amjad et al. (2018)	1	-	500-1000 nm	moisture	%	PLS	0.99	0.10	0.98	0.14	-	-	-	sliced
				CIELAB a*	-		0.93	0.49	0.93	0.51	-	-	-	
				CIELAB b*	-		0.92	2.31	0.94	1.72	-	-	-	
Moscetti et al. (2018)	1	-	500-1010 nm	moisture	%	PLSR	0.986	0.13	0.985	0.13	0.99	0.11	-	sliced
				hue angle (h)	-		0.967	0.78	0.964	0.82	0.961	0.88	-	
				L*/b*	-		0.952	0.08	0.948	0.08	0.957	0.07	-	
Lu et al. (2019)	1	300	500-1000 nm	solanine	mg/100 g	SVR	-	-	-	-	0.9143	0.0296	-	whole
Rady et al. (2019)	1	200	400-1000 nm	primordial leaf count	-	PLSR	0.49	-	-	0.3055	0.47	11.7014	1.14	whole
	1						0.78	-	-	0.4183	0.43	7.8047	1.1	
Li et al. (2020)	2	-	454-950 nm	chlorophyll (bud stage)	µg cm <sup>-2</sup>	PLS	0.86	2.21	-	-	0.85	2.05	-	field
Xiao et al. (2020)	1	6	477-947 nm	L*	-	LSSVM	0.932	0.865	0.834	1.353	0.851	1.305	2.345	sliced
				a*	-		0.964	0.248	0.95	0.29	0.957	0.271	4.731	
				b*	-		0.962	0.383	0.909	0.597	0.924	0.546	3.56	
				BI (Browning index)	-		0.958	0.685	0.924	0.922	0.94	0.823	4.047	
				L* / b*	-		0.954	0.068	0.922	0.088	0.948	0.072	4.093	
				water content	%		0.803	0.009	0.653	0.012	0.794	0.01	2.018	
Li et al. (2021b)	-	-	400-1000 nm	<i>E. coli</i>	log CFU g <sup>-1</sup>	BP-NN	0.968	0.079	0.98	0.096	0.976	0.065	-	sliced
Wang et al. (2021a)	2	96	382.23-1004.78 nm	starch	g/kg	PLSR	0.9276	1.76	-	-	0.9467	1.63	2.95	sliced
Zhao et al. (2021)	1	44	382.2-1026.7 nm	Fv/Fm	-	PLS	0.9118	0.0101	-	-	0.8218	0.0174	2.3247	leaves
Liu et al. (2021)	4	-	400-2350 nm	petiole NO <sub>3</sub> -N	10 <sup>4</sup> mg/kg	PLSR	-	-	-	-	0.71	6622.56	-	field
				whole leaf N	mg/g		-	-	-	-	0.82	4.77	-	
				whole N	mg/g		-	-	-	-	0.73	6.05	-	
Li et al. (2021a)	1	-	454-950 nm	chlorophyll (tuber formation)	µg cm <sup>-2</sup>	RF	0.63	3.94	-	-	0.69	2.7	-	field

Studies	Number of varieties	Number of potatoes /leaves	Range	Compound to predict	Measure	Analysis	R2cal	RMSEcal	R2cv	RMSEcv	R2val	RMSEval	RPDval	Type of sample
Rady et al. (2021)	2	400	400-1000 nm	glucose sucrose	% %	PLSR	- -	- -	0.918 0.612	0.0204 0.0243	- -	- -	- -	whole
Wang et al. (2021b)	2	96	382-1004 nm	starch (umbilicus)	g/kg	SVR	0.9415	15.9	-	-	0.9346	17.4	2.69	whole
Abdelbaki et al. (2021)	1	-	300-2500 nm	LAI (leaf area index) fCover (fractional vegetation cover) CCC (canopy chlorophyll content)	m <sup>2</sup> /m <sup>2</sup> - g/m <sup>2</sup>	RF <sub>exp</sub>	- - -	- - -	- - -	- - -	0.77 0.82 0.81	- - -	- - -	field
Cui et al. (2022)	1	145 146	1080-2200	water content firmness	% Newton (N)	PLSR	0.9133 0.931	0.0028 0.0216	- -	- -	0.9313 0.9317	0.0025 0.0218	2.7453 2.7531	whole
Wang and Wang (2022)	2	96	382-1004 nm	starch	g/100 g	PLSR	0.8911	2.17	-	-	0.8832	2.29	2.05	intact
Liu et al. (2022b)	2	-	450-950 nm	aboveground biomass	kg/hm <sup>2</sup>	GPR	0.76	199.68	-	-	0.82	130.31	-	field
Liu et al. (2022c)	2	144 (plants)	450-950 nm	aboveground biomass	kg/hm <sup>2</sup>	PLSR	0.72	160.77	-	-	0.74	125.48	-	field
Wang et al. (2022)	1	185	450-998 and 917-1717 nm	dry matter	%	PLSR	0.892	0.788	-	-	0.875	0.801	-	whole
Zhou et al. (2022)	2	160	400-1000 nm	leaf total N concentration petiole nitrate-N concentration	% %	PLSR	- -	- -	0.78 0.87	0.41 0.13	- -	- -	- -	field
Zou et al. (2022)	1	105	387-1035 nm	water content	%	XGBoost	0.908	0.0610	0.8448	0.0544	-	-	-	sliced
Liu et al. (2022a)	2	-	450-950	aboveground biomass	kg/hm <sup>2</sup>	PLSR	0.74	210.55	-	-	0.78	131.91	-	field
Muruganantham et al. (2023)	-	47	400-1000nm	Moisture content	%	PLSR			0.66	4.31				whole



**Table S4** Overview of applications of hyperspectral imaging to measure some components on potatoes through classification models ( $n=26$ )

Studies	Number of varieties	Number of potatoes	Range	Compounds to classify	Analysis	Recall <sub>cat</sub>	Recall <sub>cv</sub>	Recall <sub>val</sub>	Specificity <sub>cat</sub>	Specificity <sub>cv</sub>	Specificity <sub>val</sub>	Accuracy <sub>cat</sub> (%)	Accuracy <sub>cv</sub> (%)	Accuracy <sub>val</sub> (%)	Precision <sub>cat</sub>	Type of sample
Al-Mallahi et al. (2008)	3	230	360-1010 nm	potatoes and clods (wet conditions)	LDA	-	-	-	-	-	-	99.8	-	-	-	whole
				potatoes and clods (dry conditions)		-	-	-	-	-	-	97.4	-	-		
Nguyen Do Trong et al. (2011)	-	33	400-1000 nm	raw pixels	PLSDA	-	-	-	-	-	-	-	-	95.82	-	peeled potatoes
				cooked pixels		-	-	-	-	-	-	-	97.91	-		
Rady et al. (2015)	1	540	400-1000 nm	glucose	Knn	-	-	-	-	-	-	83	-	86	-	sliced
	1	180		sucrose	PLSDA	-	-	-	-	-	-	83	-	70	-	
López-Maestresalás et al. (2016)	3	188	1000-2500 nm	healthy potatoes	PLSDA	1	-	1	0.976	-	0.971	100	-	100	-	whole
				bruised potatoes		0.976	-	0.971	1	-	1	97.63	-	97.12	-	
Su and Sun (2016b)	4	48	944-1678 nm	moisture content (77-82 %)	PLSDA	1	1	-	1	1	-	-	-	-	-	sliced
				moisture content (64-73 %)		1	1	-	1	1	-	-	-	-		
				moisture content (41-57 %)		1	1	-	1	1	-	-	-	-		
				organic potatoes		1	-	1	1	-	1	-	-	-		
Ye et al. (2018)	1	220	450-1000 nm	bruised potatoes (SVM1)	SVM	-	-	-	-	-	-	100	-	95.45	-	whole
				level II bruised potatoes (SVM2)		-	-	-	-	-	-	100	-	95.45		
				level II & III bruises (SVM3)		-	-	-	-	-	-	100	-	90.91		
Gao et al. (2018)	1	150	600-750 nm	pre-sprouting eyes and by-sprouting eyes	SPA-SFA-FDA	-	-	-	-	-	-	96.5	-	97.6	-	whole
Ji et al. (2019b)	-	144	450-1000 nm	healthy and bruised potatoes	AdaBoost-FLD	-	-	-	-	-	-	-	-	99.82	-	whole
Ji et al. (2019a)	-	600	400-1000 nm	intact ones	SVM	-	-	-	-	-	-	91.11	-	-	-	whole
				green husk		-	-	-	-	-	-	90	-	-		
				germination		-	-	-	-	-	-	90.89	-	-		
				dry rot		-	-	-	-	-	-	93.33	-	-		
				wormhole		-	-	-	-	-	-	88.89	-	-		
damage	-	-	-	-	-	-	90	-	-							
Polder et al. (2019)	2	-	400-1000 nm	healthy and infected potatoes with virus Y	FCN	0.92	-	-	0.99	-	-	98.9	-	-	0.92	field



Studies	Number of varieties	Number of potatoes	Range	Compounds to classify	Analysis	Recall <sub>cal</sub>	Recall <sub>ev</sub>	Recall <sub>val</sub>	Specificity <sub>cal</sub>	Specificity <sub>ev</sub>	Specificity <sub>val</sub>	Accuracy <sub>cal</sub> (%)	Accuracy <sub>ev</sub> (%)	Accuracy <sub>val</sub> (%)	Precision <sub>cal</sub>	Type of sample
Shi et al. (2022)	2	-	450-950 nm	healthy potato potato late blight soil background	CropdocNet	0.988 0.976 0.978 0.986	- - - -	- - - -	0.992 0.988 0.996 1	- - - -	- - - -	99.1 98.5 99.1 99.7	- - - -	- - - -	0.976 0.965 0.989 1	leaves
López-Maestresalas et al. (2022)	10	80	900-1700 nm	cooking frying as crisps.	PLSDA	- -	0.889 0.968	0.818 1	- -	0.968 0.889	1 0.818	- -	91.38	90	- -	sliced
Qi et al. (2023)	4	-	-	potato late blight date 1 (July 9) potato late blight date 2 (July 13) potato late blight date 3 (July 15) potato late blight date 4 (July 18)	PLB-2D-3D-A	0.888 0.679 0.700 0.922	- - - -	- - - -	- - - -	- - - -	- - - -	79	- - - -	- - - -	0.796 0.720 0.764 0.878	leaves
Gao et al. (2023)	-	126	400-1000 nm	Anthrax blight early blight	1D-CNN	- - -	1 0.9992 0.9923	0.9997 0.9942 0.971	- - -	- - -	- - -	- - -	1 0.9969 0.9979	0.9987 0.9895 0.9842	- - -	leaves
Tunny et al. (2023)	-	-	894-2504 nm	potatoes and foreign materials	PLSDA	-	-	-	-	-	-	-	100	-	-	fresh-cut potatoes