

TREBALL FINAL DE GRAU

★ **GUILLEM MIRALLES GADEA**

*Estimació Mitjançant Anàlisi No Destructiu Dels Nivells Nutricionals De
Fulles De Cítrics Aplicant Tècniques De Machine i Deep Learning*



ÍNDEX

01
INTRODUCCIÓ

03
MATERIALS I MÈTODES

05
CONCLUSIONS I FUTURA PROJECCIÓ

02
OBJECTIUS

04
RESULTATS I DISCUSSIÓ

ELS CÍTRICS: UNA INDÚSTRIA EN EXPANSIÓ I IMPACTE ECONÒMIC A LA COMUNITAT VALENCIANA



➤ USOS

Utilitzats **no sols per a l'alimentació**. Altres usos com mèdics, biocombustibles, farmàcia, etc..

➤ PRODUCCIÓ MUNDIAL 2020

Es van produir **143.755.600 tones** de cítrics a tot el món, i es nota una tendència positiva de creixement.

➤ APLICACIÓ DE LES TIC

S'han millorat les pràctiques agrícoles en els **últims seixanta anys**, augmentant en un **70%** el rendiment.

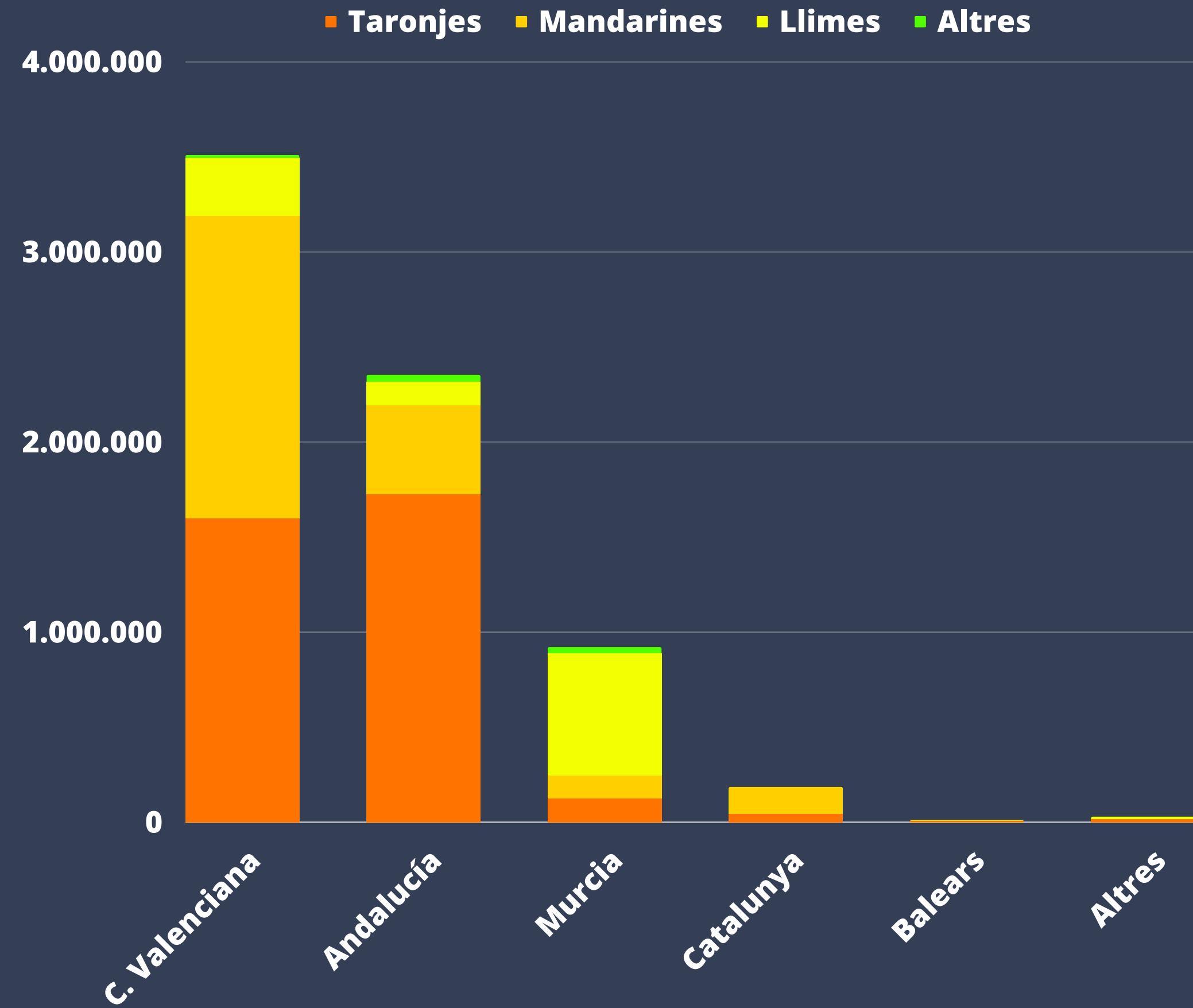
INTRODUCCIÓ

ELS CÍTRICS: UNA INDÚSTRIA EN EXPANSIÓ I IMPACTE ECONÒMIC A LA COMUNITAT VALENCIANA

A la Comunitat Valenciana la Campanya 2020/2021 es tanca amb un balanç de producció de cítrics de: **3.508.051 tones**

*Balanç de la campanya de cítrics
2020/2021 a la Comunitat Valenciana*

PRODUCCIÓ CÍTRICA NACIONAL DE LA CAMPANYA 2020/2021



INTRODUCCIÓ

LA IMPORTÀNCIA DELS NUTRIENTS EN EL CREIXEMENT SALUDABLE DELS ARBRES



SOLUCIÓ

SUPLEMENTS PER A LA CORRECTA NUTRICIÓ, ANOMENATS FERTILITZANTS

PROBLEMA

- ALLIBERAMENT DE CASOS D'EFFECTE HIVERNACLE
- EUTROFITZACIÓ



NUTRIENTS

► **MACRONUTRIENTS PRIMARIS**

Necessaris en grans quantitats i són essencials per al creixement saludable.

- **Nitrogen (N)**
- **Fòsfor (P)**
- **Potassi (K)**

► **MACRONUTRIENTS SECUNDARIS**

Menors en comparació amb els primaris. Tot i això, també són necessaris.

- **Calci (Ca)**
- **Magnesi (Mg)**
- **Sofre (S)**

► **MICRONUTRIENTS**

Fan falta en petites quantitats, tot i això, també són importants.

- **Sodi (Na)**
- **Coure (Cu)**
- **Ferro(Fe)**
- **Bor (B)**
- **Mangenès (Mn)**
- **Molibdè (Mo)**
- **Zinc (Zn)**

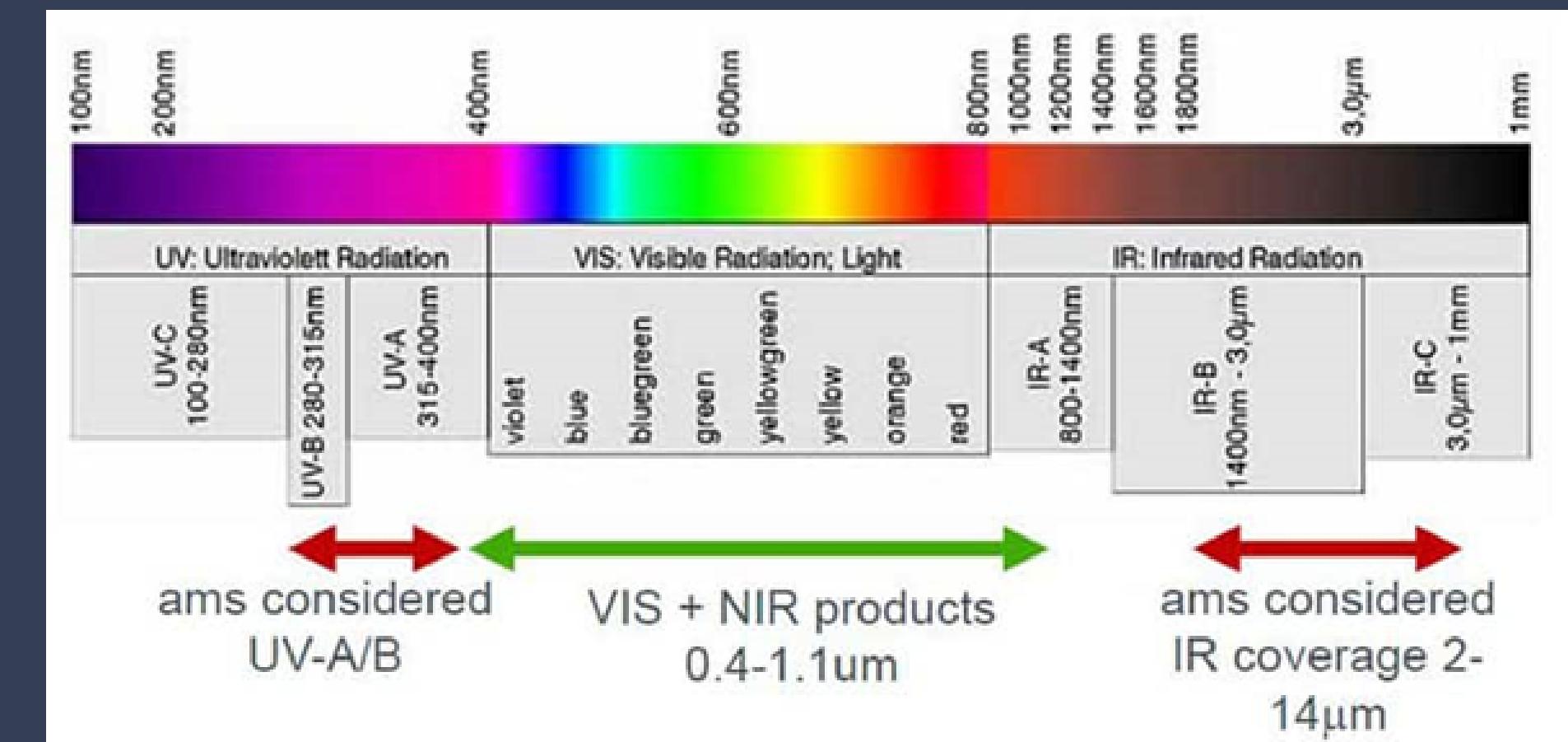
PER A CONÈIXER ELS NUTRIENTS A LES FULLES: LA IONÒMICA FOLIAR



Laboratori d'ionòmica del CEBAS-CSIC

- **ALT COST ECONÒMIC**
- **ALT COST TEMPORAL**
- **TÉCNICA DESTRUCTIVA**
- **REQUEREIX DOMINI**

ESTIMACIÓ MITJANÇANT L'ANÀLISI ESPECTRAL



Espectre electromagnètic VIS-NIR. Font: Blog Digi-Key Spectral Sensors

- Les dades espectrals tenen **àmplies aplicacions** en diverses àrees, incloent **el sector agrícola**.
- L'espectroscòpia VIS-NIR és una **tècnica fàcil, ràpida, no destructiva i econòmica**, utilitzada per avaluacions de qualitat sense preparació prèvia de la mostra.

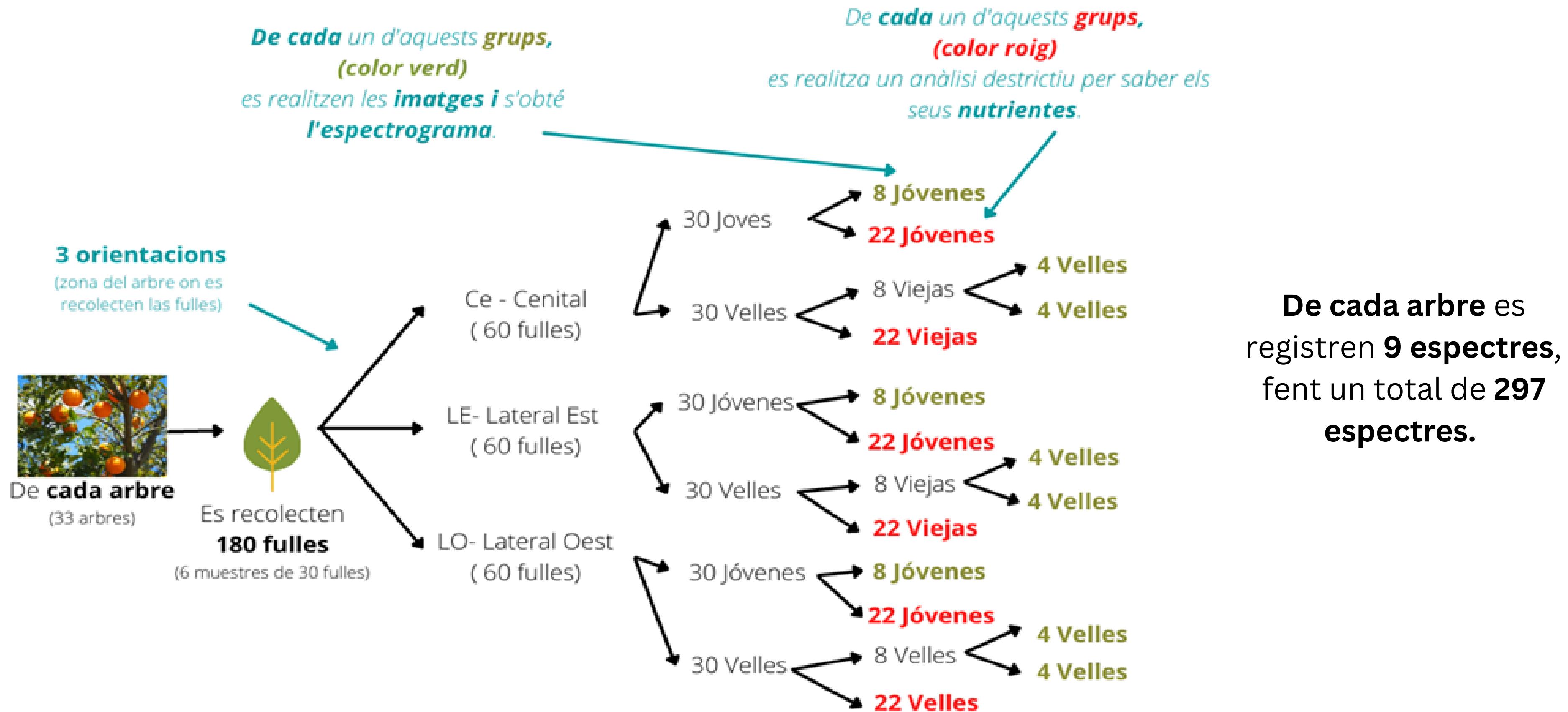
OBJECTIUS

➤ **CREACIÓ DE MODELS DE REGRESSIÓ D'APRENENTATGE AUTOMÀTIC PER ESTIMAR ELS NIVELLS DE NUTRIENTS D'UN ARBRE CÍTRIC A PARTIR DE L'ESPECTRE DE LES SEVES FULLES.**

EN QUE POT SOCIALMENT BENEFICIAR?

- Monitorització de l'estat dels cultius
- Detecció de malalties o estressos
- Optimització dels programes de reg i fertilització
- Ajuda al medi ambient i al sector agrari.

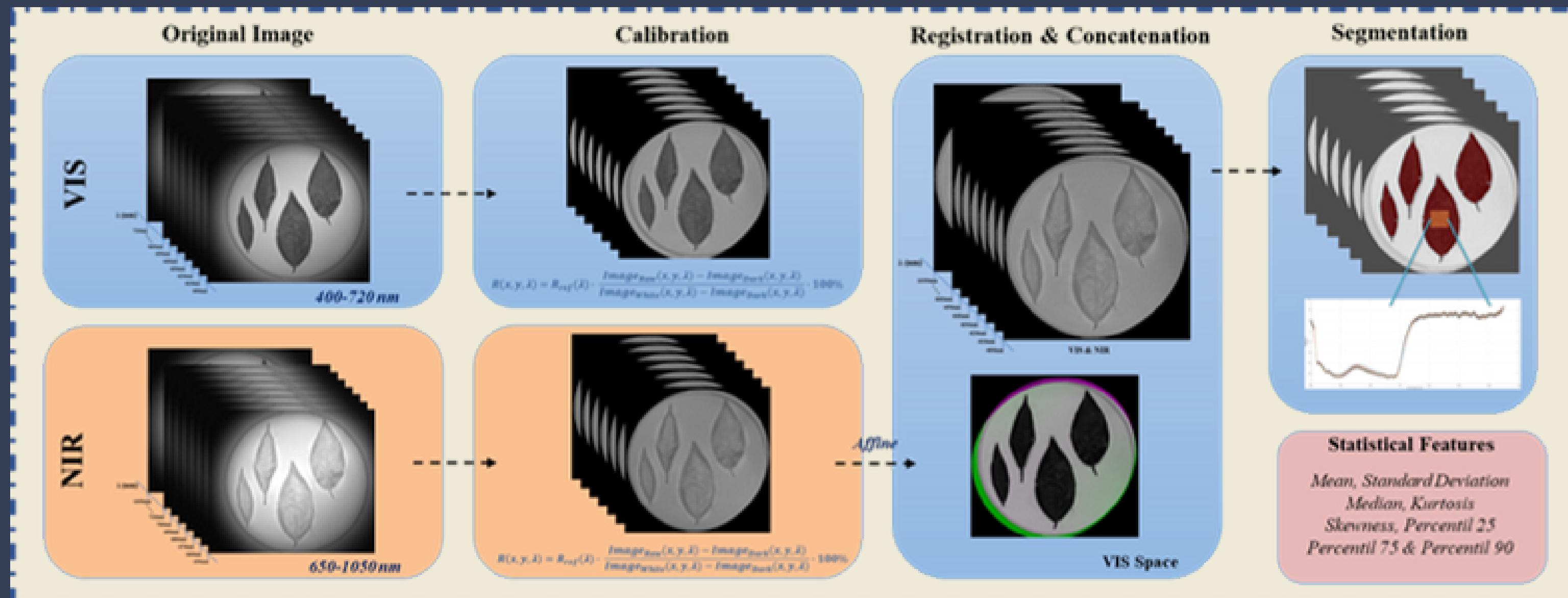
PAS 1 DE L'ADQUISICIÓ DEL CONJUNT DE DADES: DESENVOLUPAMENT EXPERIMENTAL



PAS 2 DE L'ADQUISICIÓ DEL CONJUNT DE DADES: ADQUISICIÓ DE LES IMATGES HIPERESPECTRALS

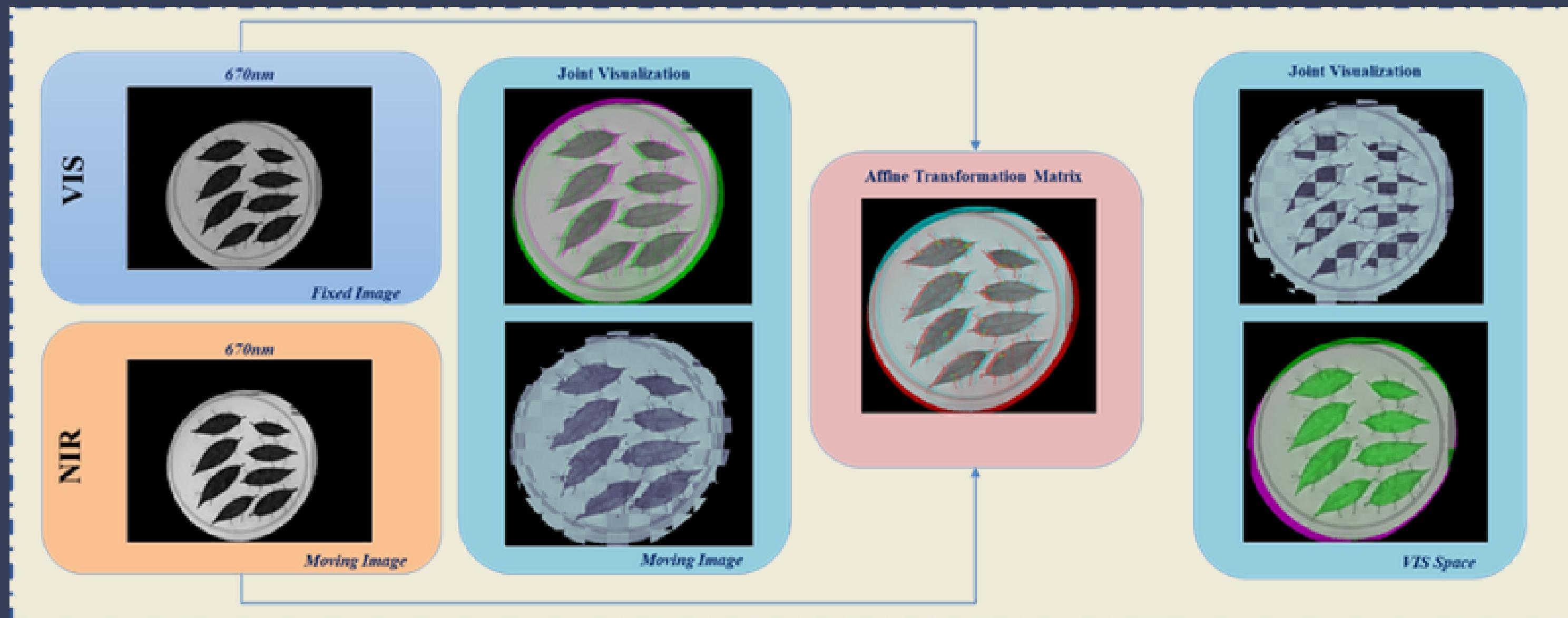


PAS 3 DE L'ADQUISICIÓ DEL CONJUNT DE DADES: ADQUISICIÓ DE L'ESPECTRE DES DE LES IMATGES HIPERESPECTRALS



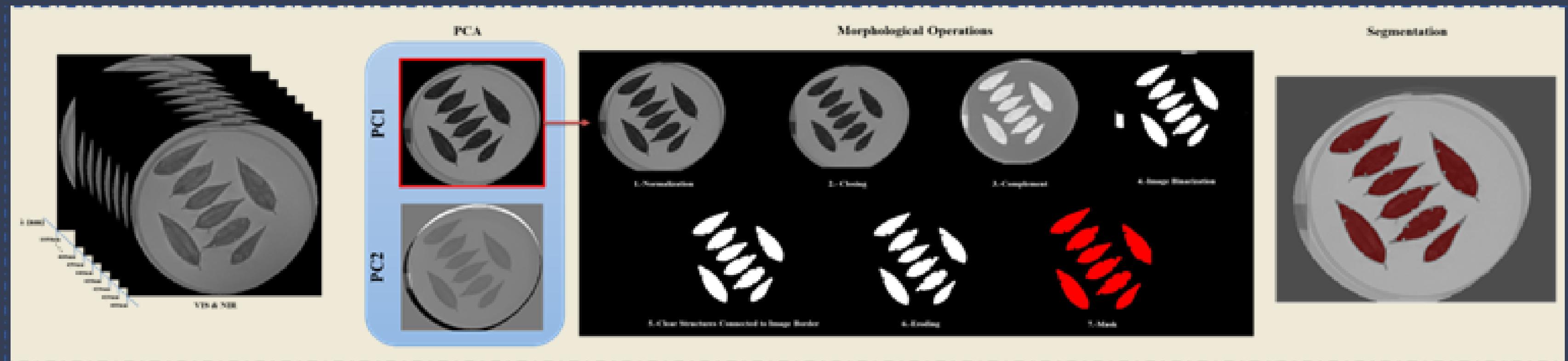
Preprocesament a les imatges hiperespectrals adquirides

PAS 3 DE L'ADQUISICIÓ DEL CONJUNT DE DADES: ADQUISICIÓ DE L'ESPECTRE DES DE LES IMATGES HIPERESPECTRALS



Correcció del desplaçament i concatenació

PAS 3 DE L'ADQUISICIÓ DEL CONJUNT DE DADES: ADQUISICIÓ DE L'ESPECTRE DES DE LES IMATGES HIPERRESPECTRALS



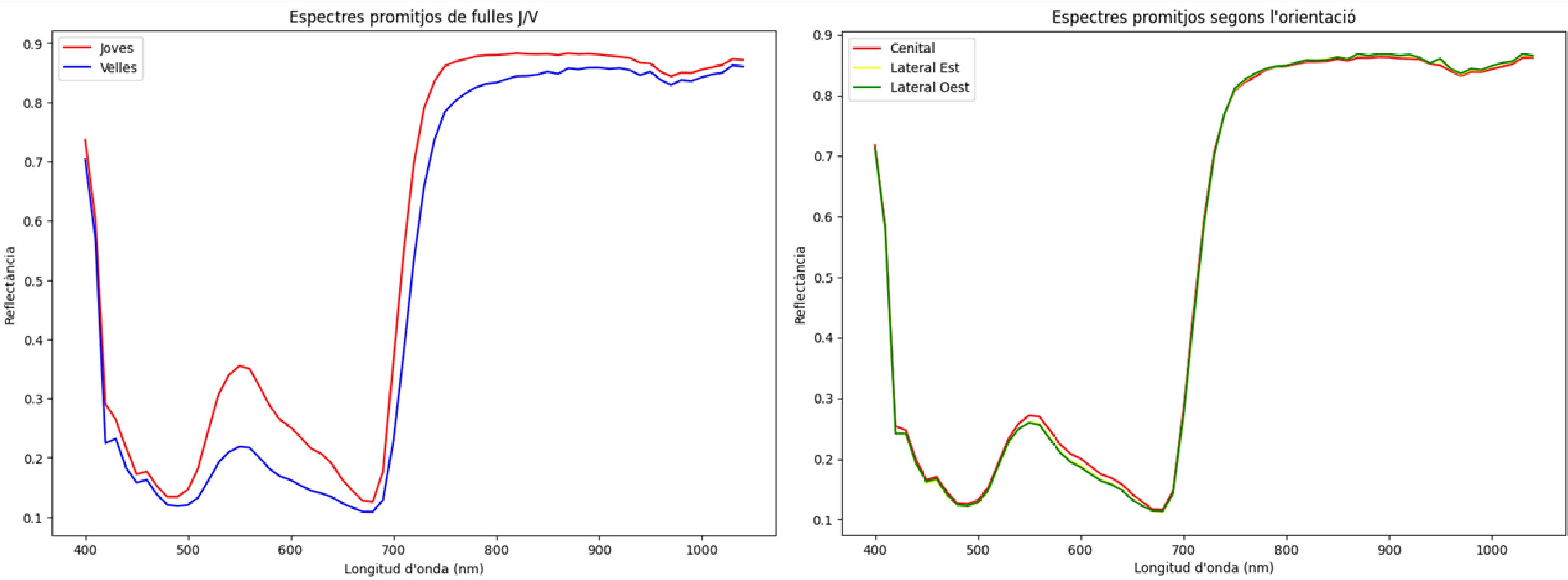
PCA i operacions morfològiques de la imatge

ANÀLISI EXPLORATORI DE LES DADES: NUTRIENTS

<i>Nutrients</i>	<i>Tipus Full (J/V)</i>	<i>Orientació (Ce, LE, LO)</i>
N	37.94 %	6.38 %
P	35.99 %	4.58 %
K	25.49 %	8.02 %
Ca	39.57 %	1.91 %
Mg	34.88 %	7.57 %
Na	17.12 %	4.83 %
S	28.27 %	9.04 %
Fe	23.99 %	15.81 %
Mn	32.88 %	3.55 %
Zn	22.97 %	10.21 %
Cu	35.73 %	4.00 %
B	26.53 %	5.54 %
Mo	12.41 %	5.88 %

Diferència estandarditzada entre el tipus de full i les orientacions segons el nutrient

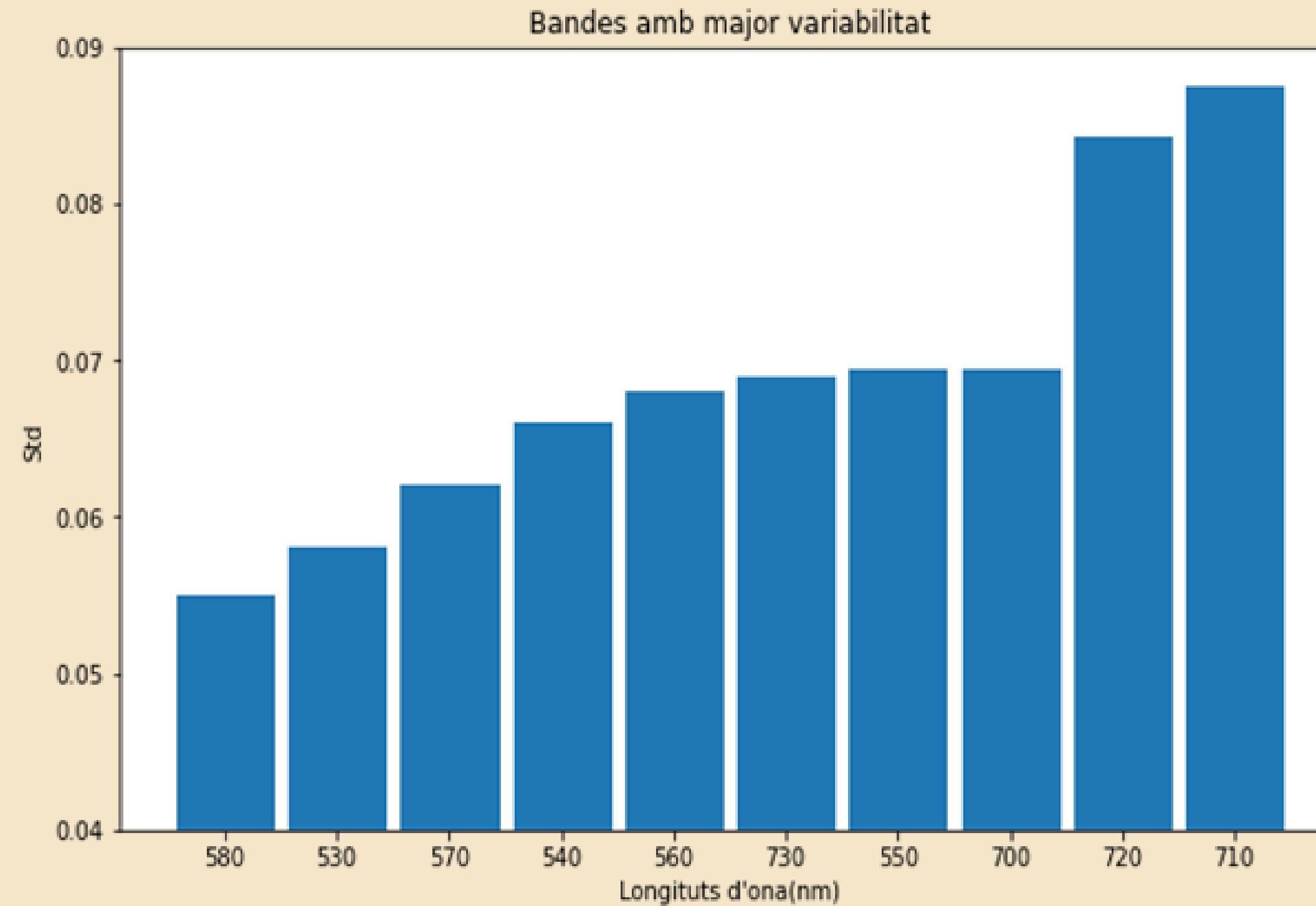
ANÀLISI EXPLORATORI DE LES DADES: EPECTRES



Espectre mitjà de les fulles joves i velles registrades

Espectre mitjà segons l'orientació en la que es va registrar la mostra

ANÀLISI EXPLORATORI DE LES DADES: EPECTRES



Bandes espectrals que registren major variabilitat

1

ELIMINAR DADES FALTANTS

4

OVERSAMPLE

- Per abordar el desequilibri de classes.
- Hi ha un major nombre d'espectres de fulles velles.
- Pot millorar la generalització dels models.

2

DIVISIÓ CALIBRACIÓ I PROVA

- 75% per a entrenar els models.
- 25% per a veure el rendiment dels models.

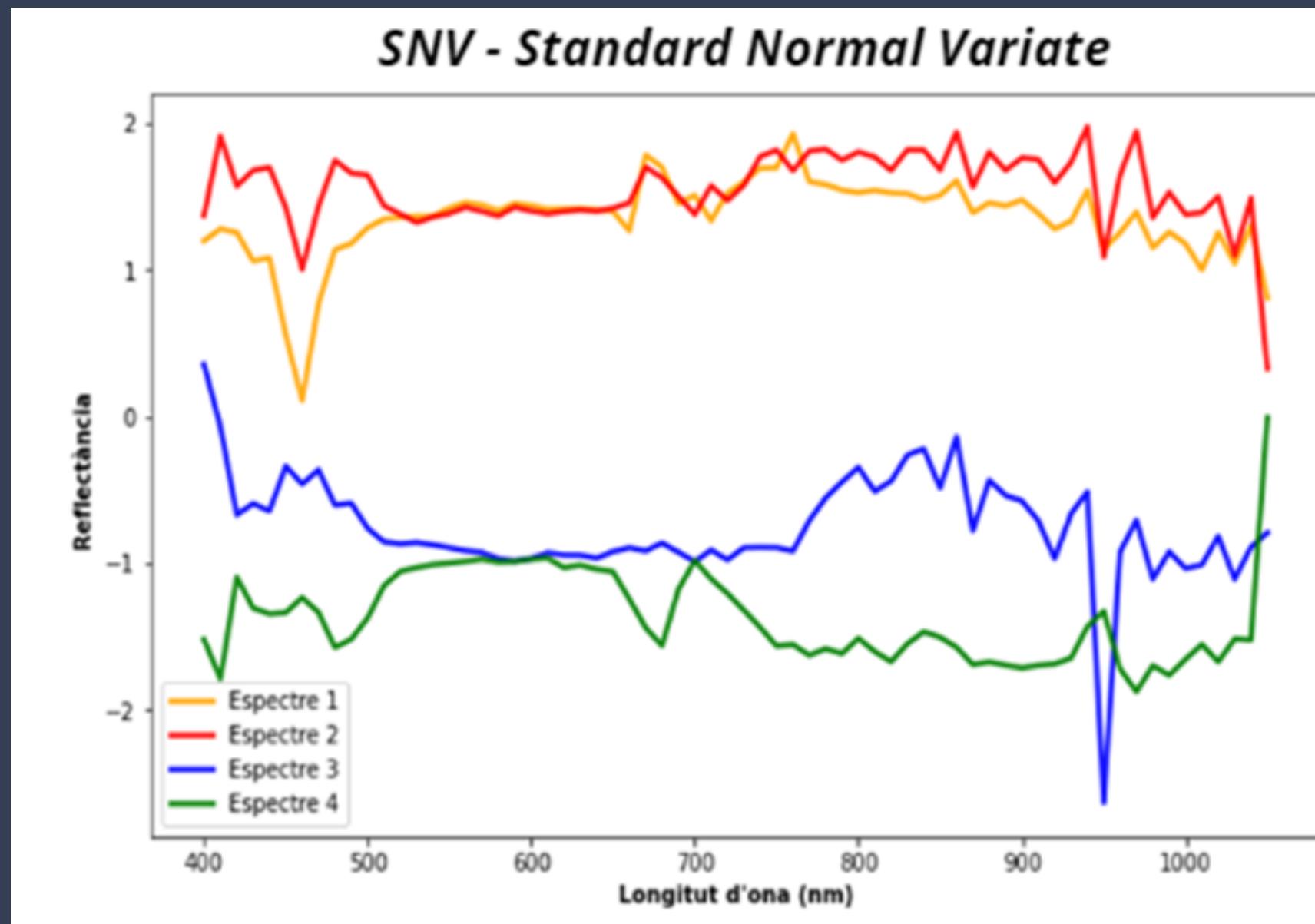
Problema: El **sobreajust** en models d'aprenentatge automàtic és un problema que es produeix quan el **model** és **massa complex per a les dades d'entrenament**, afectant negativament la capacitat de generalització del model en noves dades.

3

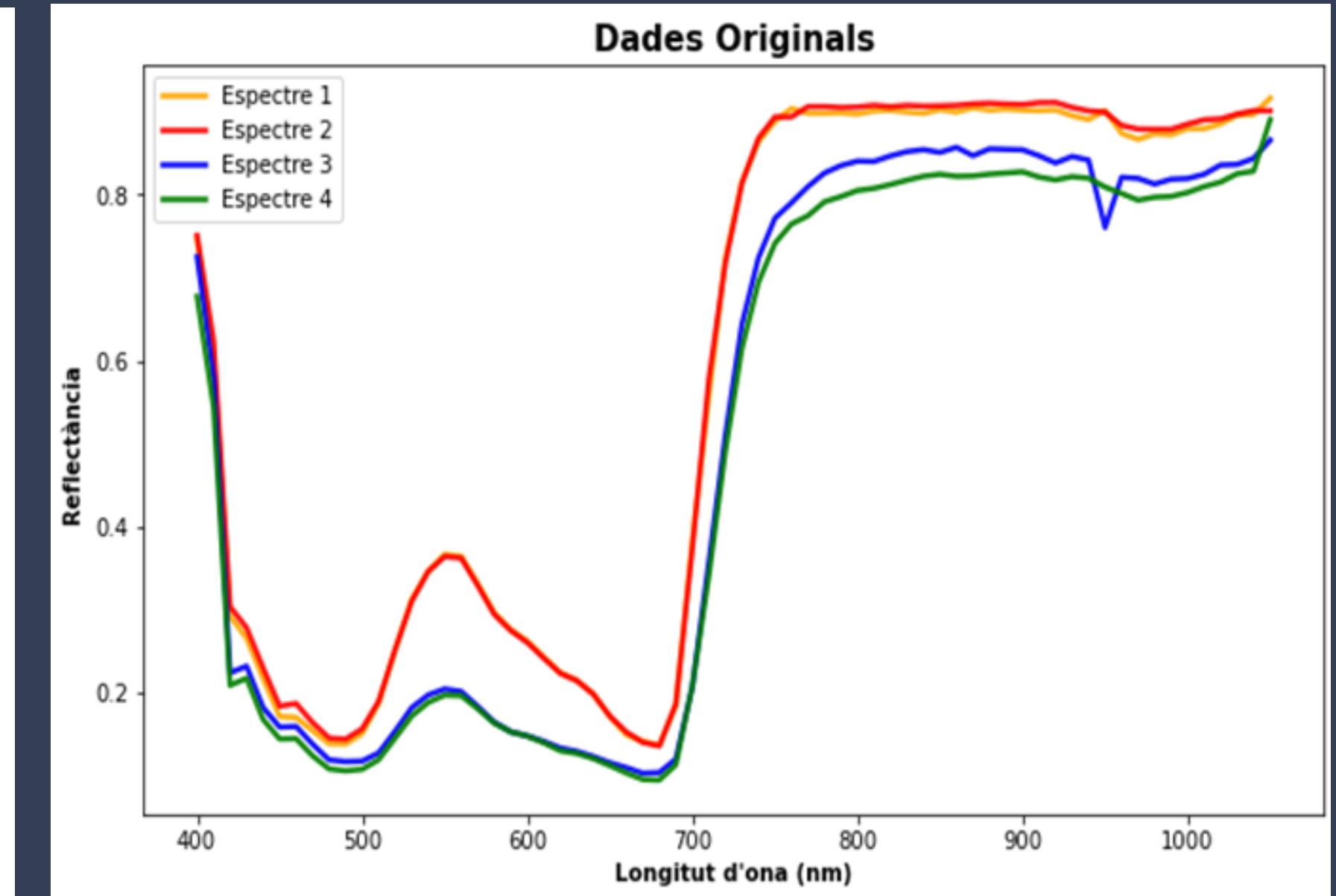
VALIDACIÓ CREUADA

La **validació creuada**, utilitzant la tècnica K-FOLD amb 3 particions en aquest projecte, **és utilitzada per evitar el sobreajust i obtenir resultats més òptims**, permetent l'avaluació mitjançant la mitjana de les mètriques en el conjunt de validació.

PREPROCESSAMENT ESPECTRAL: STANDARD NORMAL VARIATE (SNV)

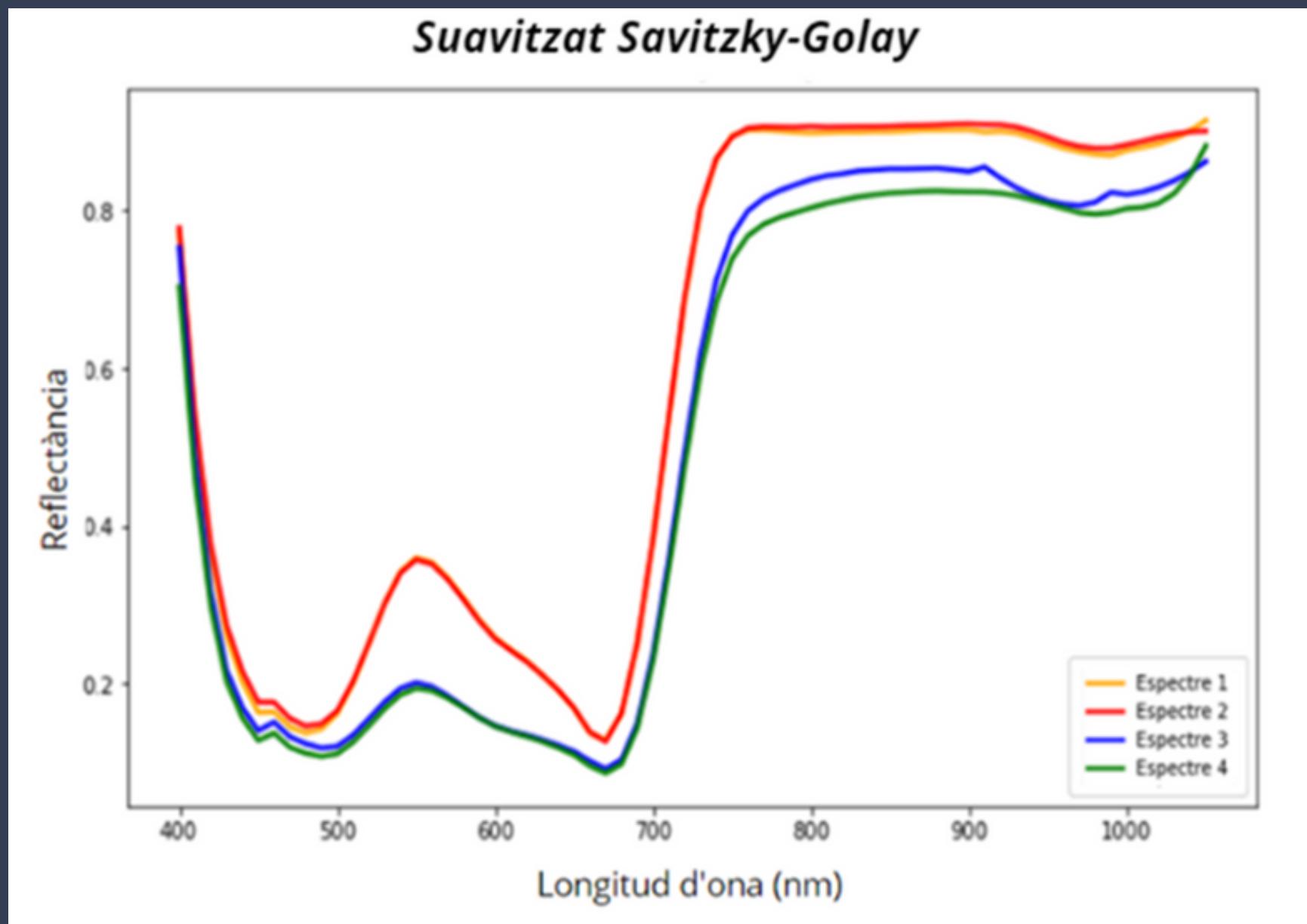


Representació gràfica de l'estandardització de 4 mostres espectrals

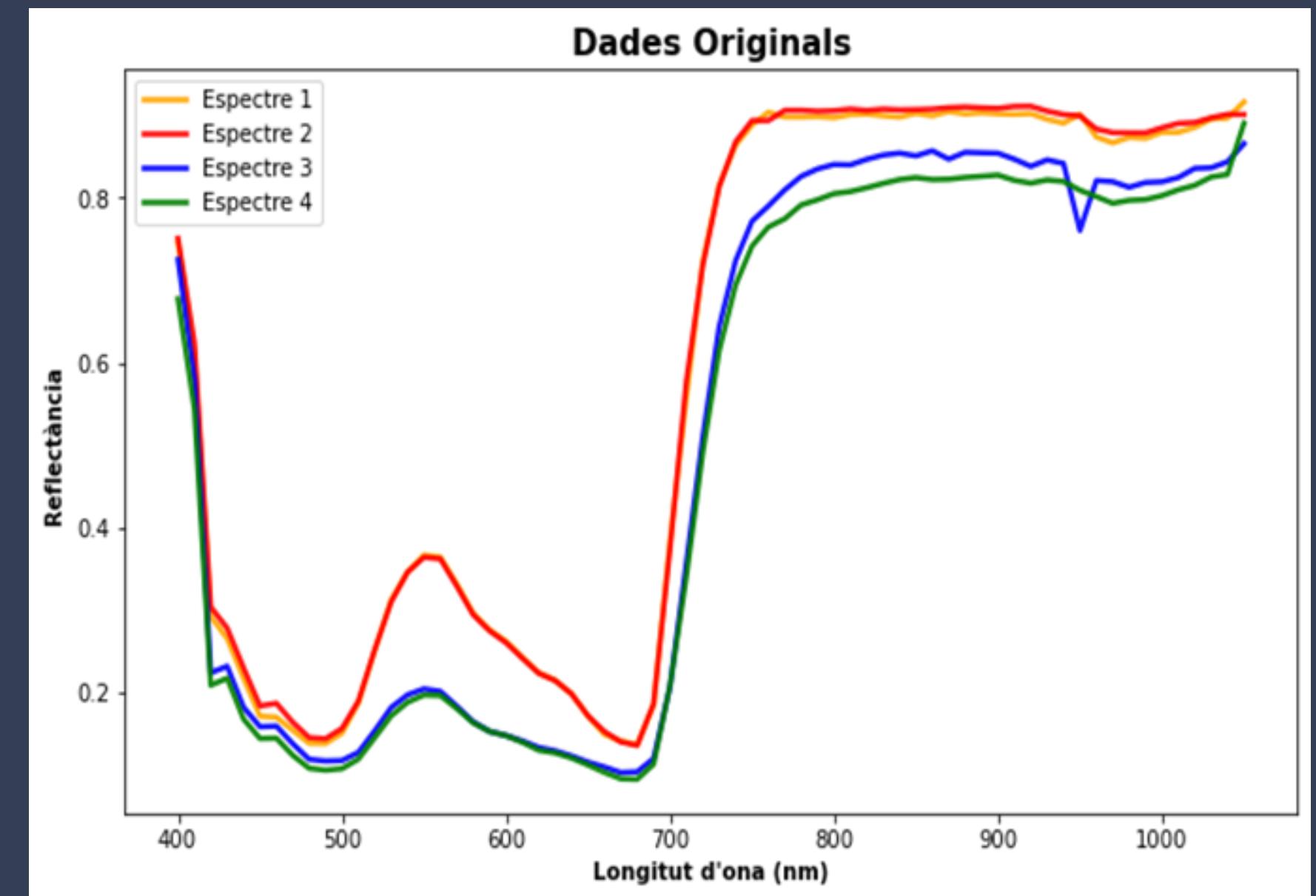


Representació gràfica de 4 mostres espectrals originals

PREPROCESSAMENT ESPECTRAL: SAVITZKY-GOLAY



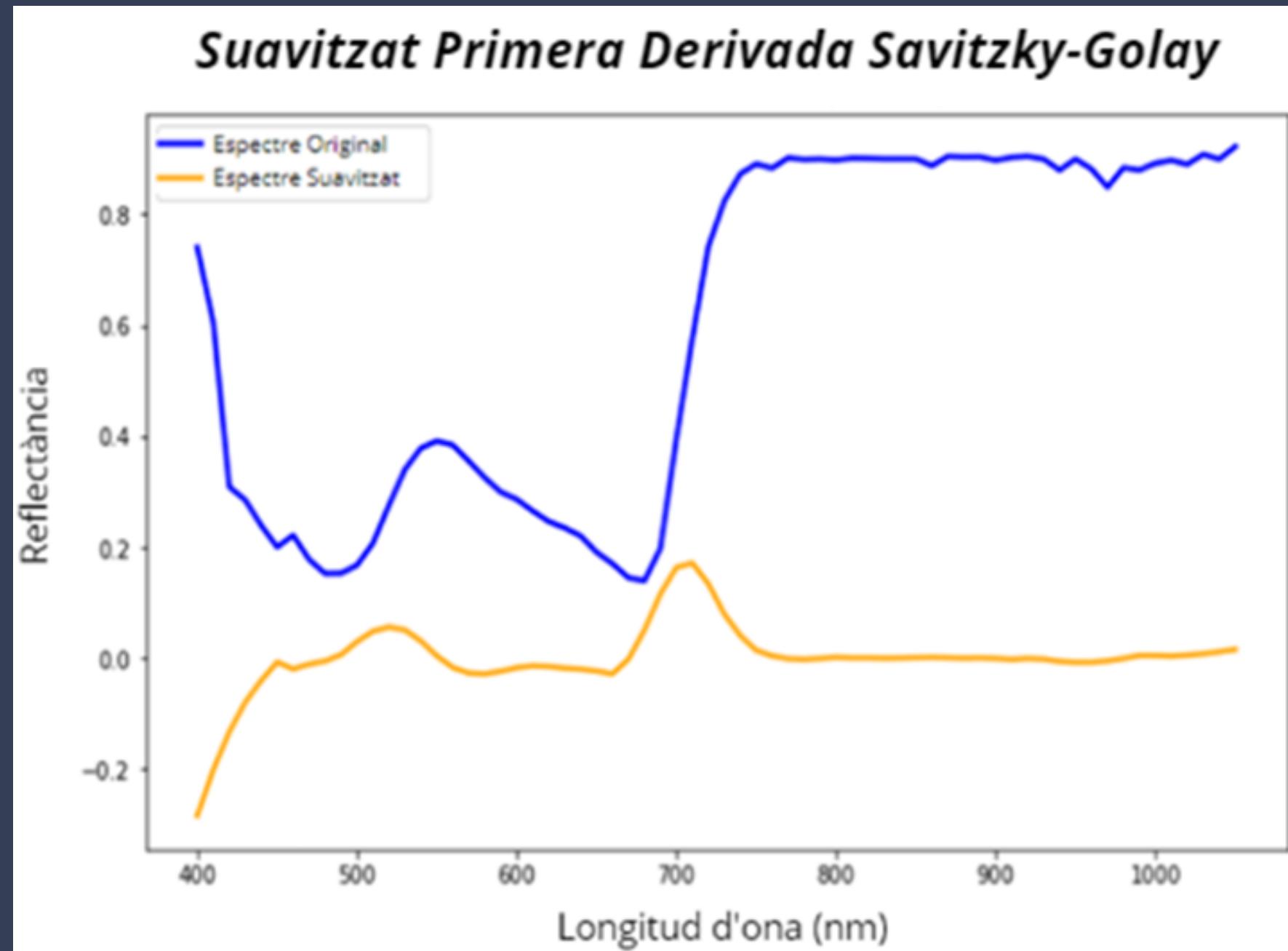
Representació de 4 mostres espectrals transformades amb Savitsky-Golay.



Representació gràfica de 4 mostres espectrals originals

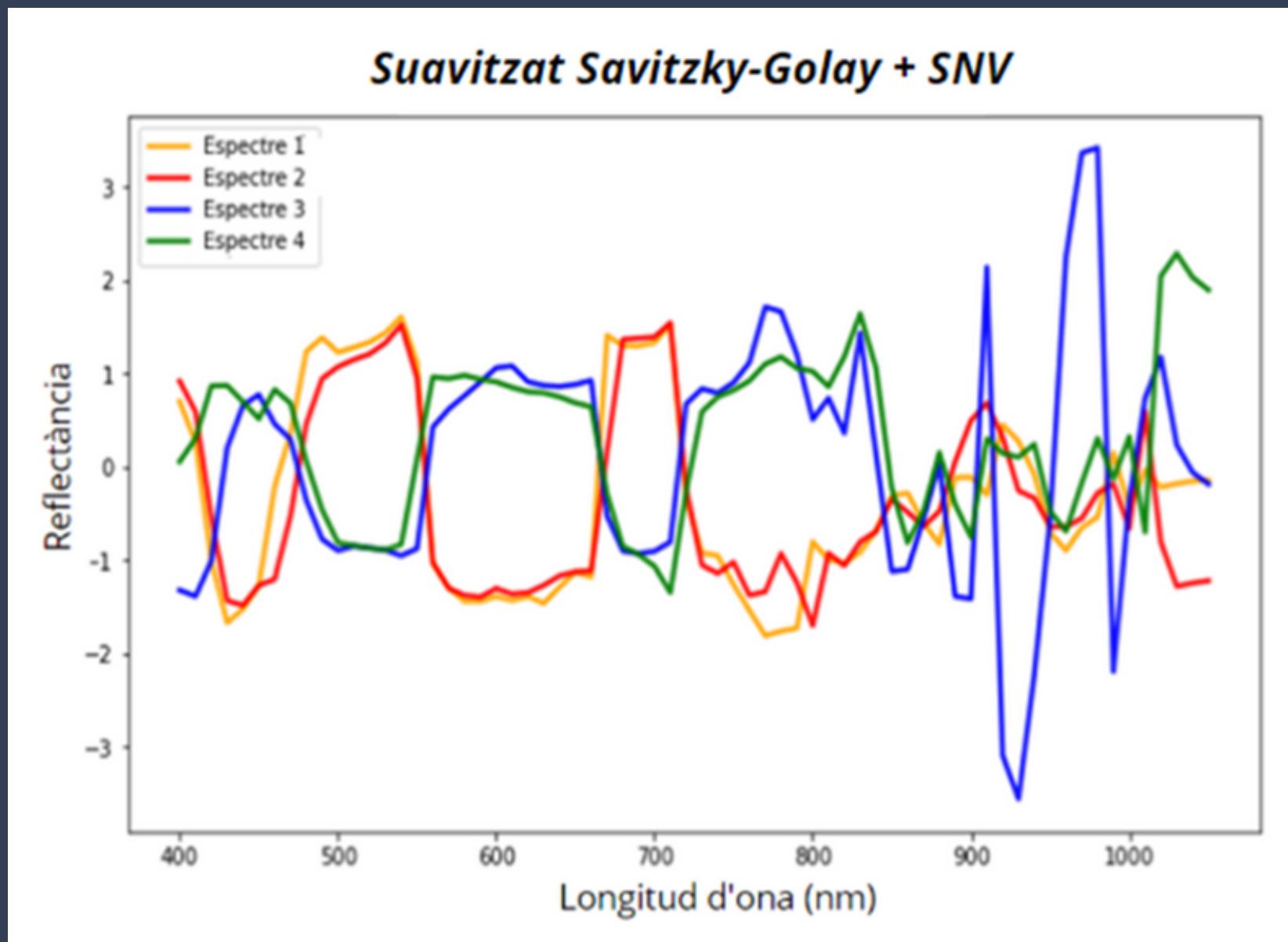
PREPROCESSAMENT ESPECTRAL: SAVITZKY GOLAY

PRIMERA DERIVADA

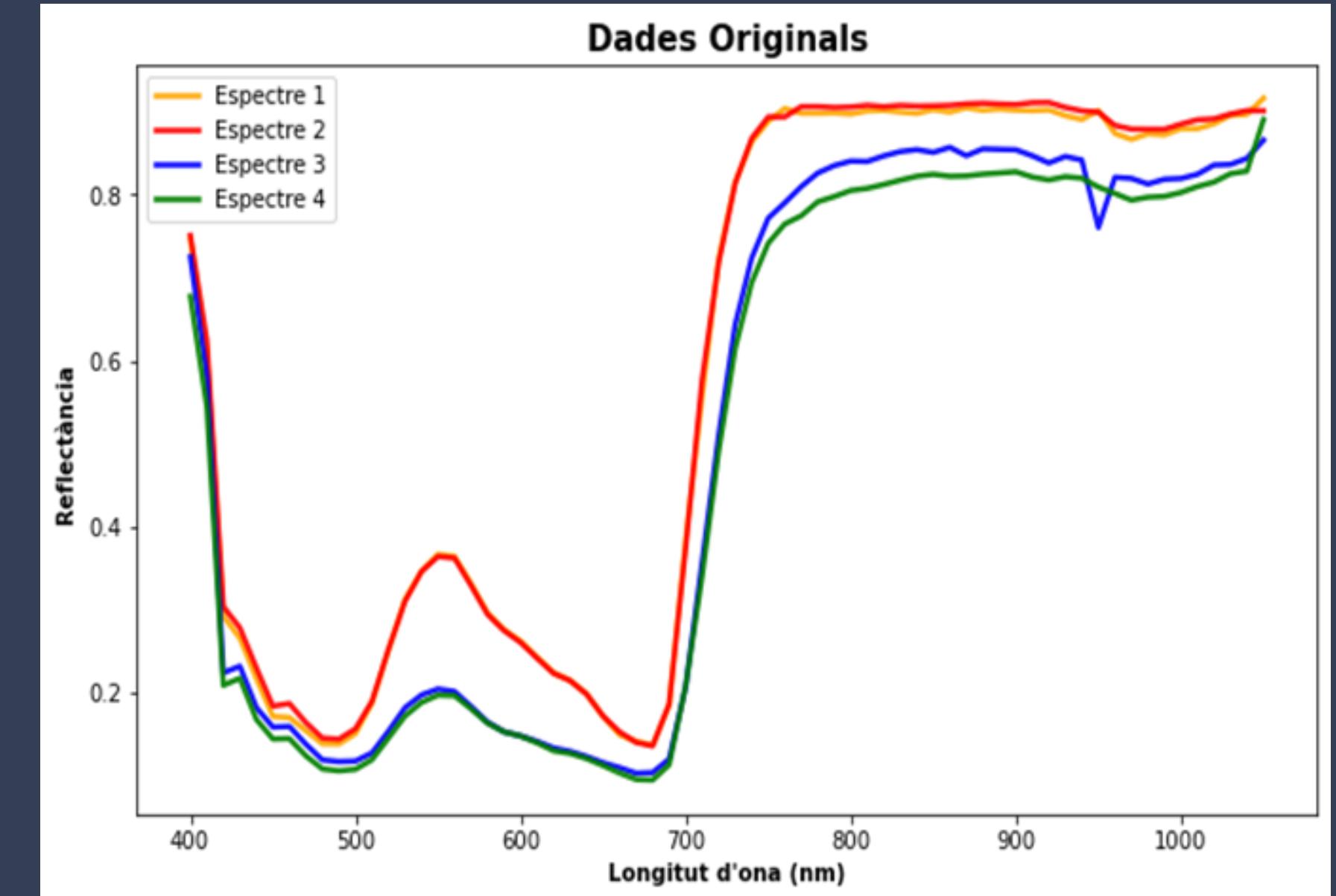


Representació gràfica de una mostra espectral transformada amb la primera derivada de Savitzky-Golay i la mateixa mostra espectral original

PREPROCESSAMENT ESPECTRAL: SAVITZKY GOLAY I SNV

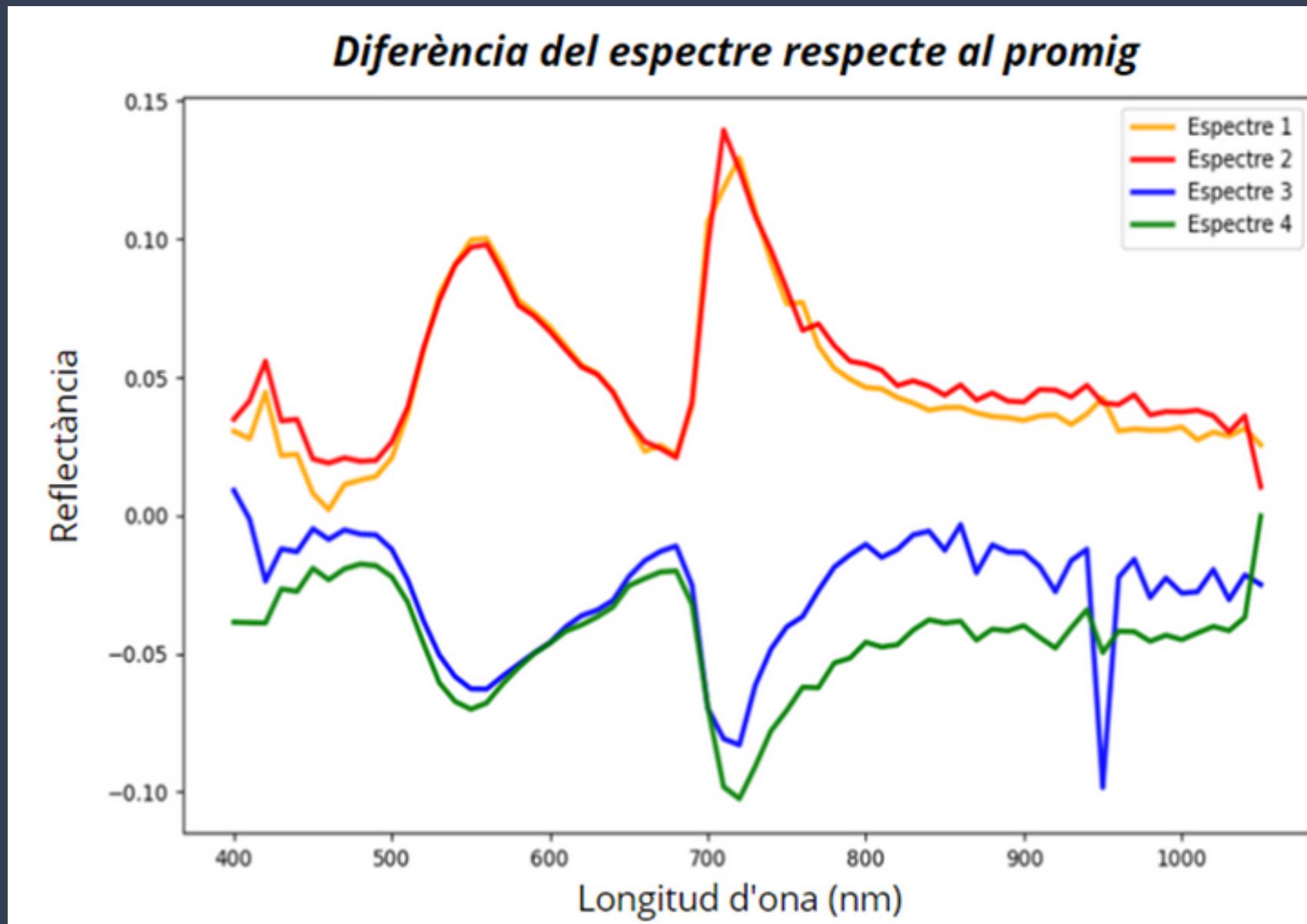


Representació gràfica de 4 mostres espectrals transformades amb Savitsky-Golay i la estandardització estàndard

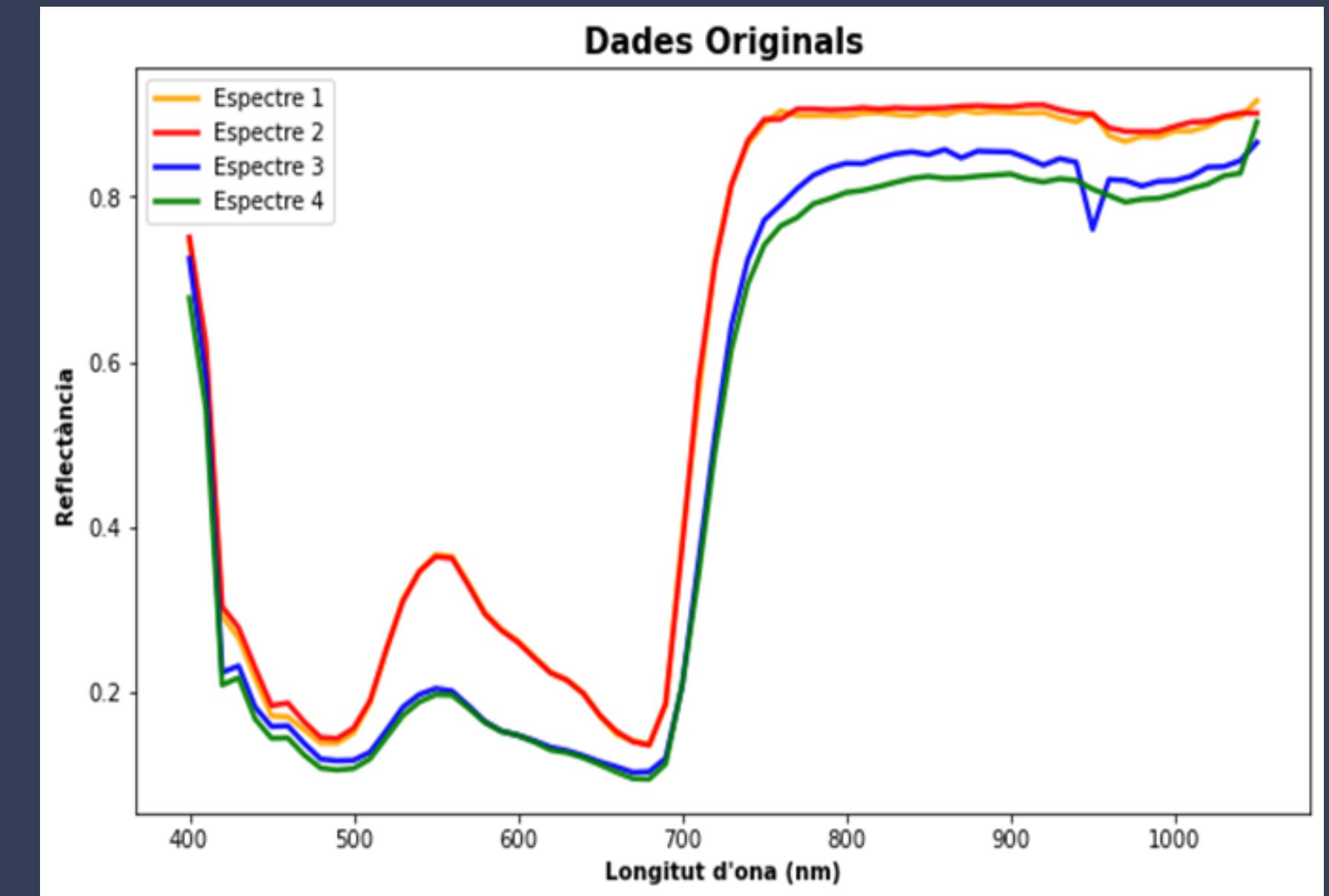


Representació gràfica de 4 mostres espectrals originals

PREPROCESSAMENT ESPECTRAL: TERME MITJÀ DEL ESPECTRE

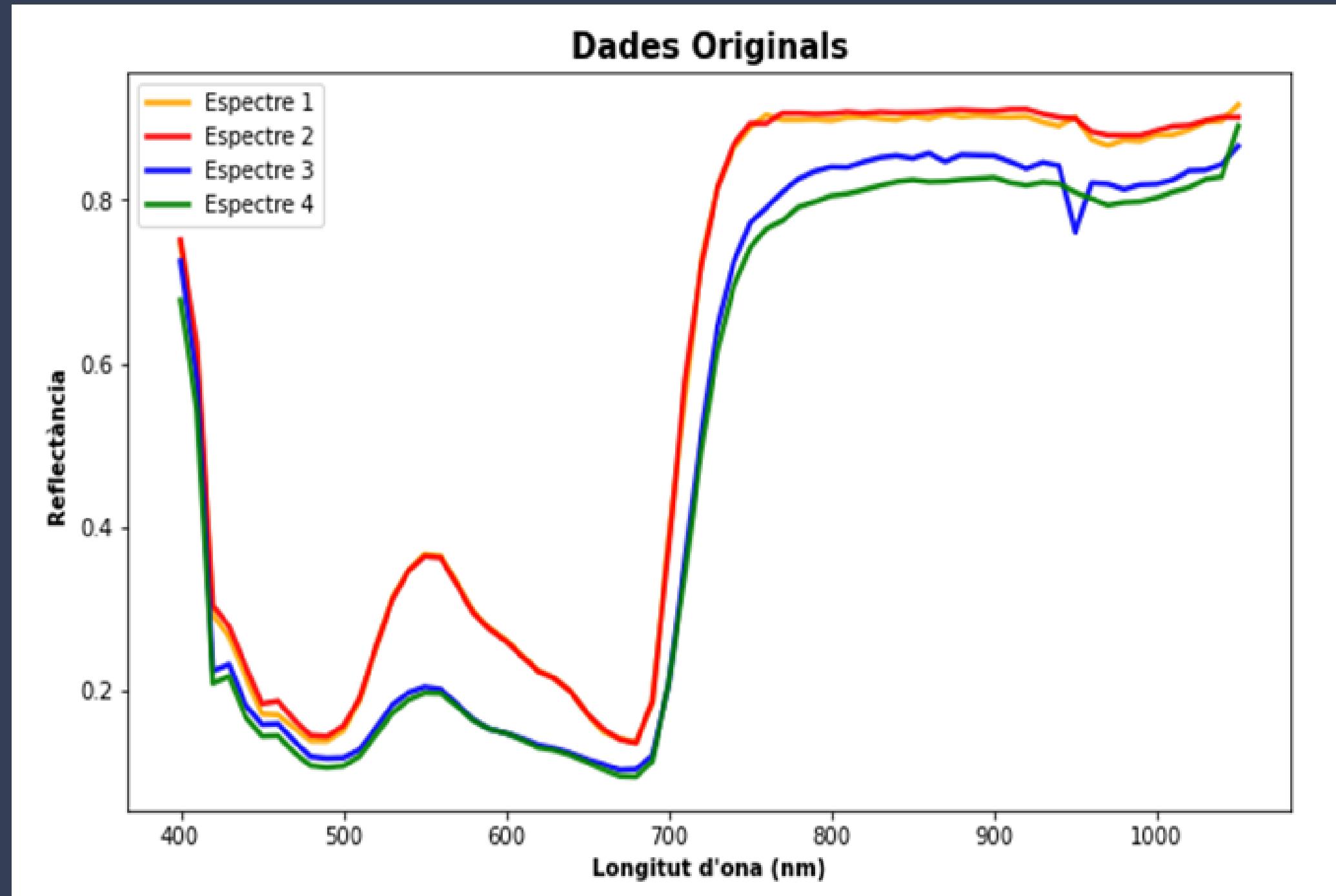


Representació gràfica de 4 mostres espectrals transformades amb la diferència respecte al espectre promig



Representació gràfica de 4 mostres espectrals originals

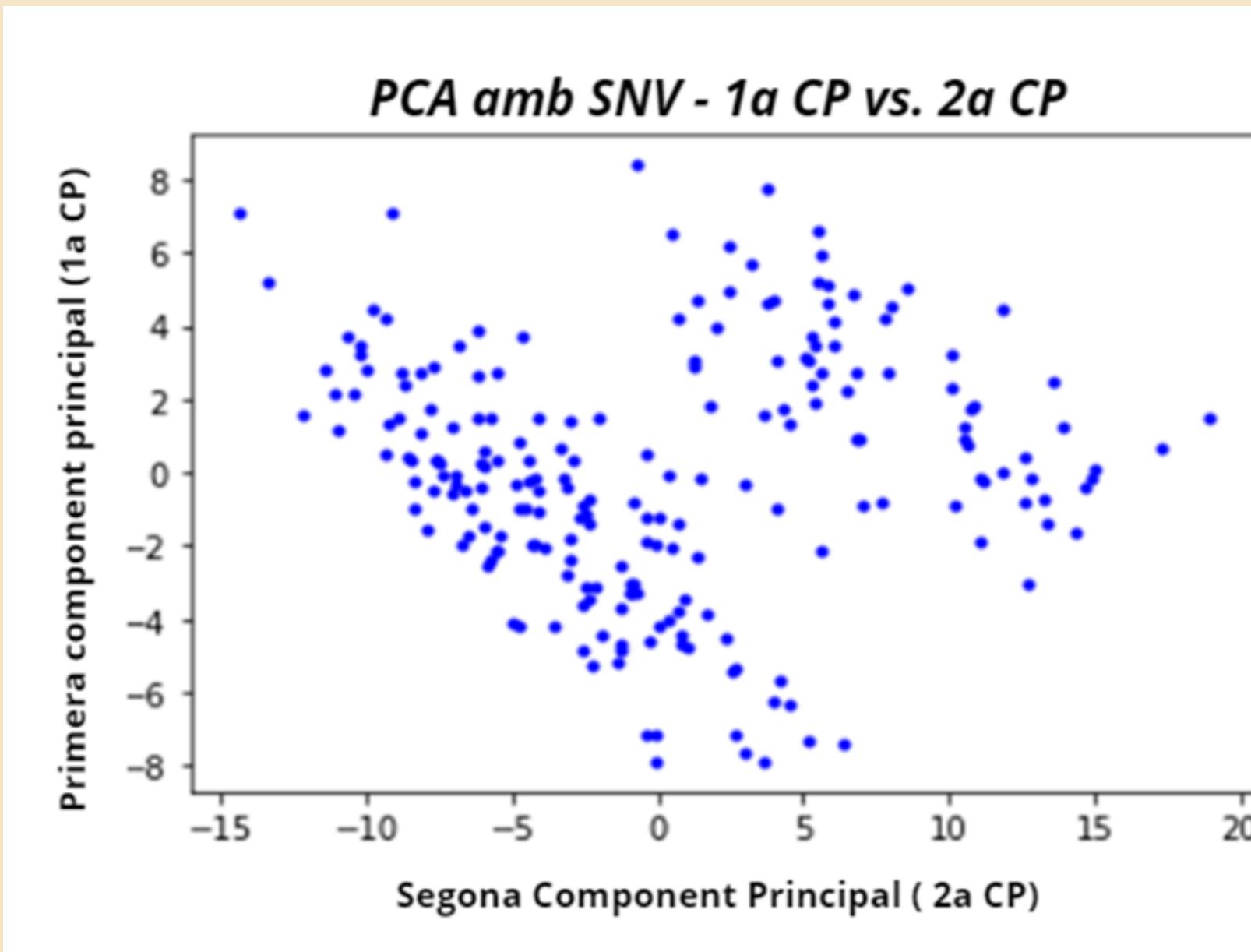
PREPROCESSAMENT ESPECTRAL: SENSE TRANSFORMACIÓ



Representació gràfica de 4 mostres espectrals originals

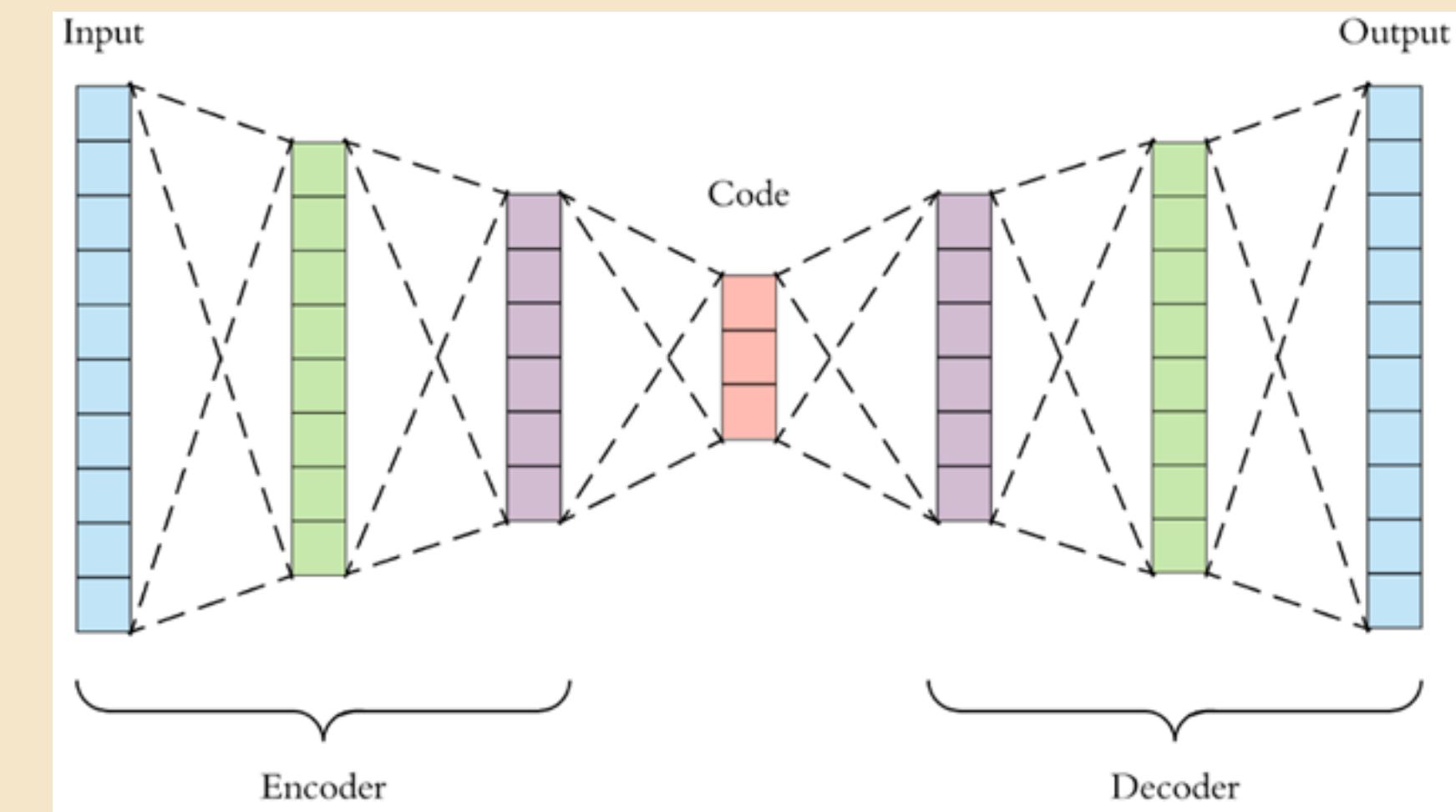
REDUCCIÓ DE LA DIMENSIONALITAT - TOTES LES BANDES

PCA: ANÀLISI DE COMPONENTS PRINCIPALS



Primera component vs. segona component de les dades estandarditzade

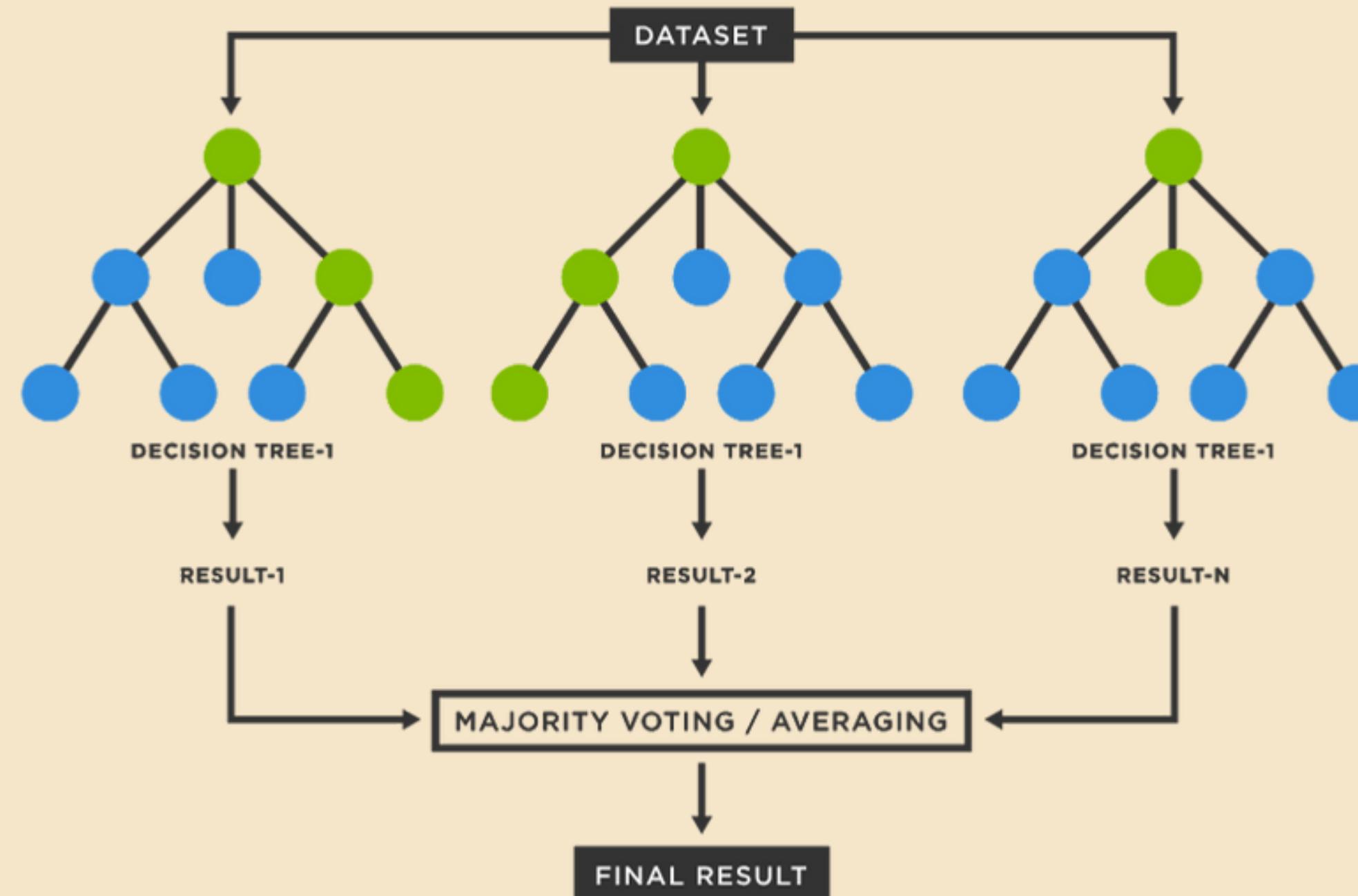
AUTOENCODER



Estructura General d'un Autoencoder. Font: Fernando Sancho Caparrini Blog.

REDUCCIÓ DE LA DIMENSIONALITAT - SELECCIÓ DE BANDES

RANDOM FOREST: 3, 6 / 12 BANDES



Estructura de un Random Forrest

MODELS

► MODELS LINEALS

- Regresió Lineal
- Ridge Regresion
- Bayesian Ridge Regression
- SGD Regressor

► CROSS- DESCOMPOSITION

- PLS Regressor

► NEAREST NEIGHBOURS

- K – Neighbors Regressor

► SUPPORT VECTOR MACHINE (SVM)

- Support Vector Regressor (SVR)
- Linear SVR (LSVR)
- NuSVR

► DECISION TREES

- Decision Tree Regressor
- Random Forest Regressor
- Extra Tree Regressor

► ENSEMBLE MODELS

- Ada Boost Regressor
- Gradient Boosting Regressor

+ Bagging

MÈTRIQUES

Coeficient de determinació (R^2)

$$R^2 = \frac{\sum_{i=1}^n (y_i - z_i)^2 - \sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (y_i - z_i)^2}$$

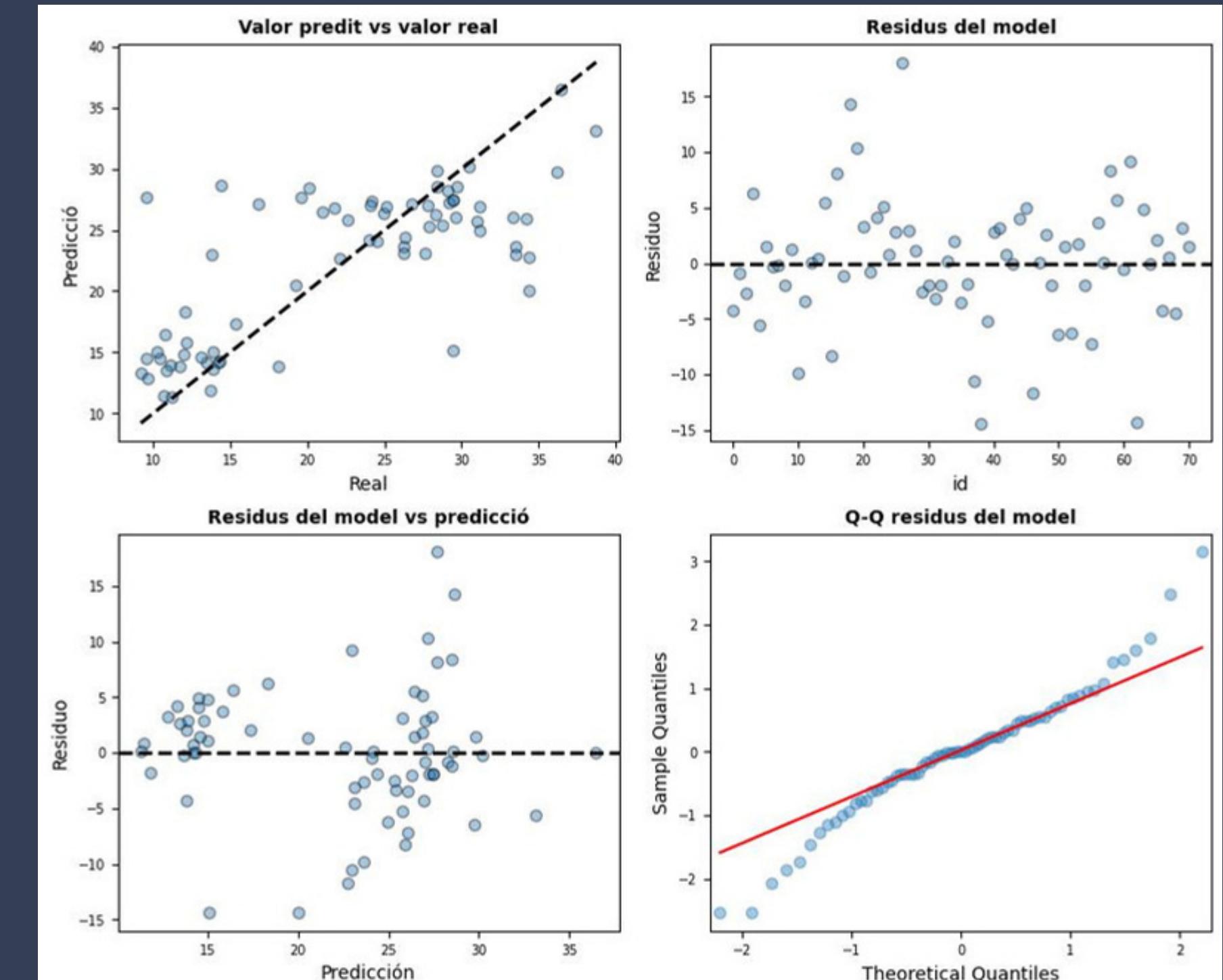
Mean Absolute Error (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

GRÀFIC DE RESIDUS



Residus ExtraTree en predicción del Mn

CONSTRUCCIÓ DE MODELS

Grid Search

La **cerca dels millors paràmetres d'un model d'aprenentatge automàtic** és un punt clau.

S'utilitza el mètode de Grid Search per definir una quadrícula de **valors** de paràmetres i trobar la millor combinació segons el rendiment mitjançant Validació Creuada

PROVES

Amb totes les bandes:

Un determinat model, prova les transformaciones i:

- PCA de 5, 7, 12, 15, 18, 20, 25 i 30 components
- Autoencoder de 5, 7, 12, 15, 18, 20, 25 i 30 components.
- Sense reduir la dimensionalitat.

Seleccionant les bandes:

- S'entrena un Random Forest per a cada nutrient.
- Es seleccionen les 3, 6 i 12 bandes més importants.
- S'utilitza el preprocesament Savinsky Golay + SNV.

MACRONUTRIENTS PRIMARIS

Totes les bandes

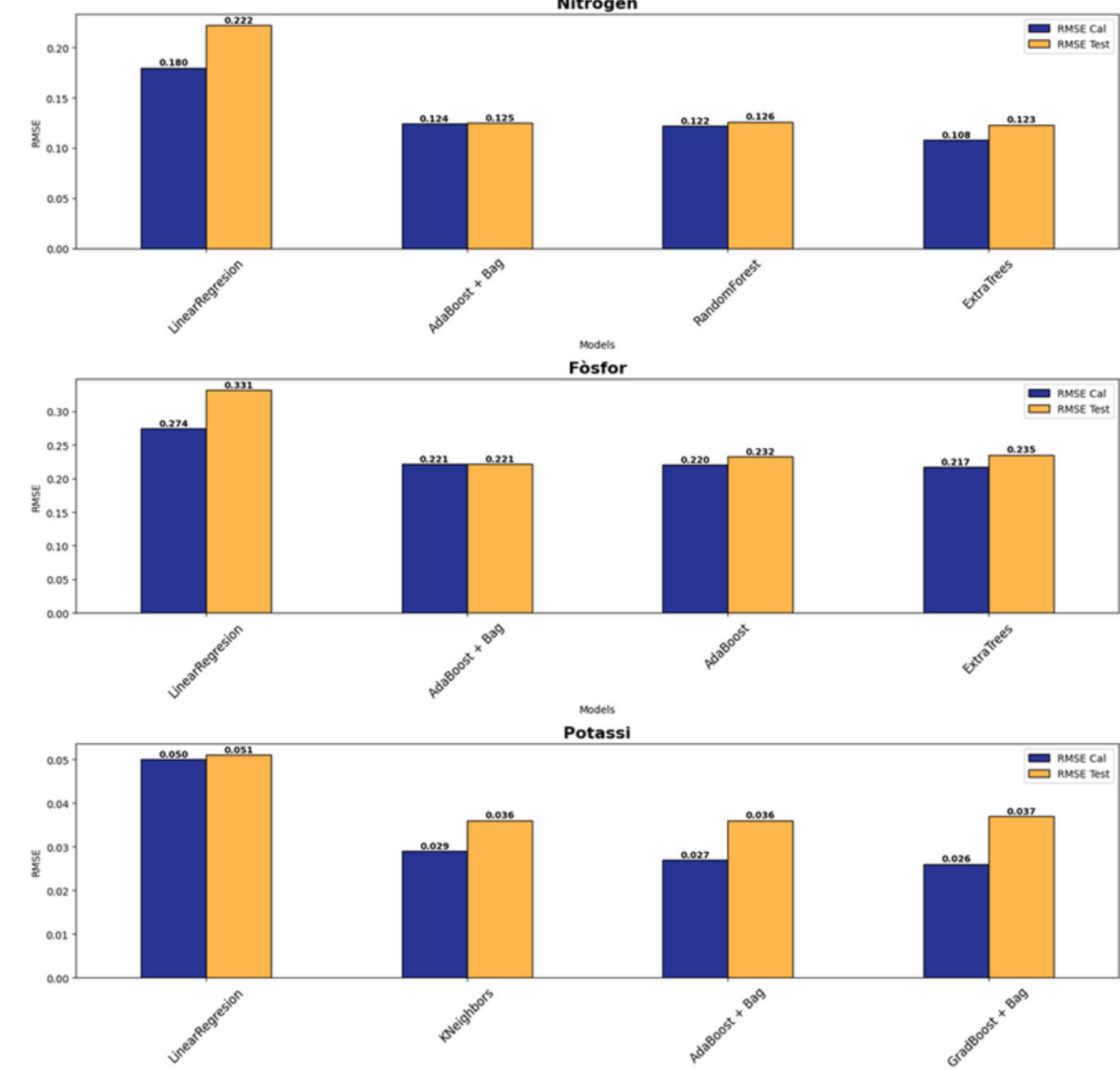
	MODEL	PREP	ICR	R ² Cal	MAE	RMSE	R ²	MAE	RMSE
					Cal	Cal	Test	Test	Test
N	Linear <u>Regresion</u>	SNV	PCA 5	0,584	0,132	0,1796	0,334	0,144	0,222
	Extra <u>Trees</u>	Original	PCA 20	0,853	0,078	0,108	0,794	0,098	0,123
	Ada <u>Boost + Bag</u>	<u>SavGol Diff</u>	PCA 20	0,807	0,098	0,124	0,788	0,100	0,125
	Random Forest	<u>SavGol Diff</u>	PCA 20	0,812	0,093	0,122	0,785	0,096	0,126
P	Linear <u>Regresion</u>	SNV	PCA_5	0,119	0,031	0,050	0,196	0,038	0,051
	Ada <u>Boost + Bag</u>	<u>SavGol Diff</u>	No	0,762	0,017	0,027	0,601	0,021	0,036
	K-Neighbors	<u>SavGol Diff</u>	PCA 25	0,718	0,020	0,029	0,593	0,022	0,036
	Grad <u>Boost + Bag</u>	<u>SavGol Diff</u>	No	0,777	0,018	0,026	0,583	0,023	0,037
K	Linear <u>Regresion</u>	SNV	PCA_5	0,482	0,195	0,274	0,267	0,256	0,331
	Ada <u>Boost + Bag</u>	<u>SavGol Diff</u>	No	0,651	0,140	0,221	0,672	0,144	0,221
	Ada <u>Boost</u>	<u>SavGol Diff</u>	No	0,659	0,142	0,220	0,639	0,162	0,232
	Extra <u>Trees</u>	SNV	PCA 9	0,674	0,136	0,217	0,629	0,155	0,235

MACRONUTRIENTS PRIMARIS

Utilitzant totes les bandes

RESULTATS

Comparació del RMSE obtingut calibració i prova dels diferents millors models que prediuen cada macronutrient primari.



MACRONUTRIENTS PRIMARIS

Selecció de bandes

	MODEL	BANDES	R ² Cal	MAE Cal	RMSE Cal	R ² Test	MAE Test	RMSE Test
N	Nu SVR	3	0,815	0,115	0,136	0,675	0,116	0,163
	K- <u>Neighbors</u>	6	0,811	0,102	0,133	0,753	0,102	0,131
	Extra <u>Trees</u>	12	0,828	0,098	0,131	0,750	0,103	0,132
P	Bayesian Ridge	3	0,680	0,024	0,033	0,488	0,029	0,043
	PLS <u>Regression</u>	6	0,756	0,019	0,029	0,529	0,026	0,041
	Grad <u>Boost</u> + Bag	12	0,759	0,017	0,028	0,530	0,025	0,038
K	Extra <u>Trees</u>	3	0,693	0,147	0,222	0,631	0,142	0,240
	K- <u>Neighbors</u>	6	0,693	0,138	0,227	0,628	0,140	0,241
	Random <u>Forest</u>	12	0,654	0,147	0,232	0,621	0,155	0,243

MACRONUTRIENTS SECUNDARIS

Totes les bandes

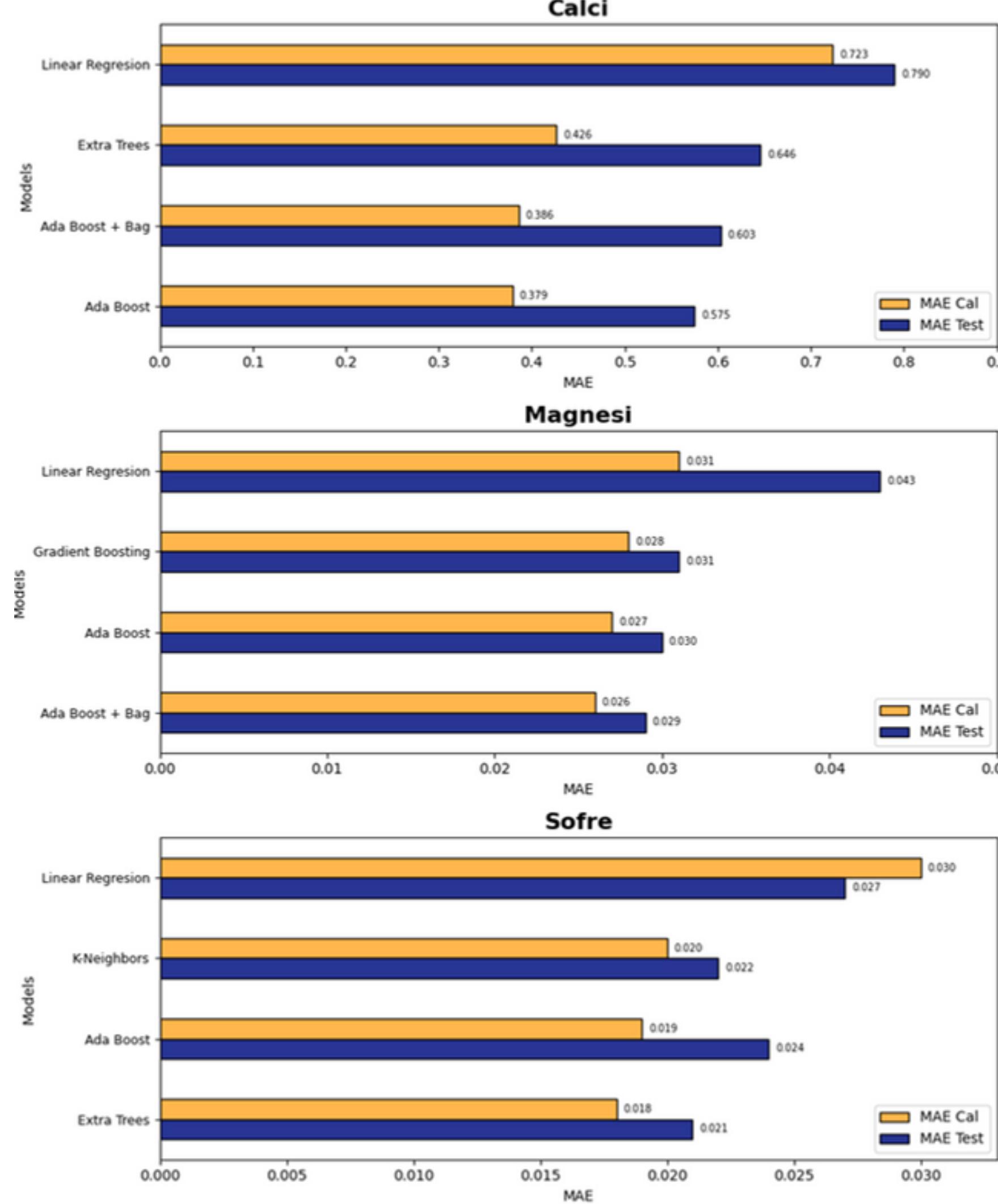
	MODEL	PREP	ICR	R ² Cal	MAE Cal	RMSE Cal	R ² Test	MAE Test	RMSE Test
Ca	Linear <u>Regresion</u>	SNV	PCA 5	0,339	0,723	1,214	0,528	0,790	1,123
	Ada <u>Boost</u>	Original	PCA 15	0,769	0,379	0,686	0,645	0,575	0,974
	Ada <u>Boost + Bag</u>	Original	PCA 30	0,762	0,386	0,702	0,613	0,603	1,017
	Extra <u>Trees</u>	Original	PCA 25	0,769	0,426	0,708	0,604	0,646	1,029
Mg	Linear <u>Regresion</u>	SNV	PCA_5	0,457	0,031	0,042	0,079	0,043	0,056
	Ada <u>Boost + Bag</u>	SavGol SNV	PCA 15	0,607	0,026	0,037	0,552	0,030	0,039
	Ada <u>Boost</u>	SavGol SNV	PCA 20	0,601	0,027	0,040	0,538	0,030	0,040
	Gradient <u>Boosting</u>	SavGol SNV	ACC 10	0,587	0,028	0,041	0,514	0,031	0,041
S	Linear <u>Regresion</u>	SNV	PCA_5	0,199	0,030	0,042	0,543	0,027	0,034
	Extra <u>Trees</u>	SavGol Diff	No	0,661	0,018	0,027	0,670	0,021	0,028
	K- <u>Neighbors</u>	SavGol	PCA 12	0,608	0,020	0,028	0,649	0,022	0,030
	Ada <u>Boost</u>	SavGol	PCA 12	0,595	0,019	0,029	0,606	0,024	0,031

MACRONUTRIENTS SECUNDARIS

Utilitzant totes les bandes

RESULTATS

Comparació del MAE obtingut calibració i prova dels diferents millors models que prediuen cada macronutrient secundari.



MACRONUTRIENTS SECUNDARIS

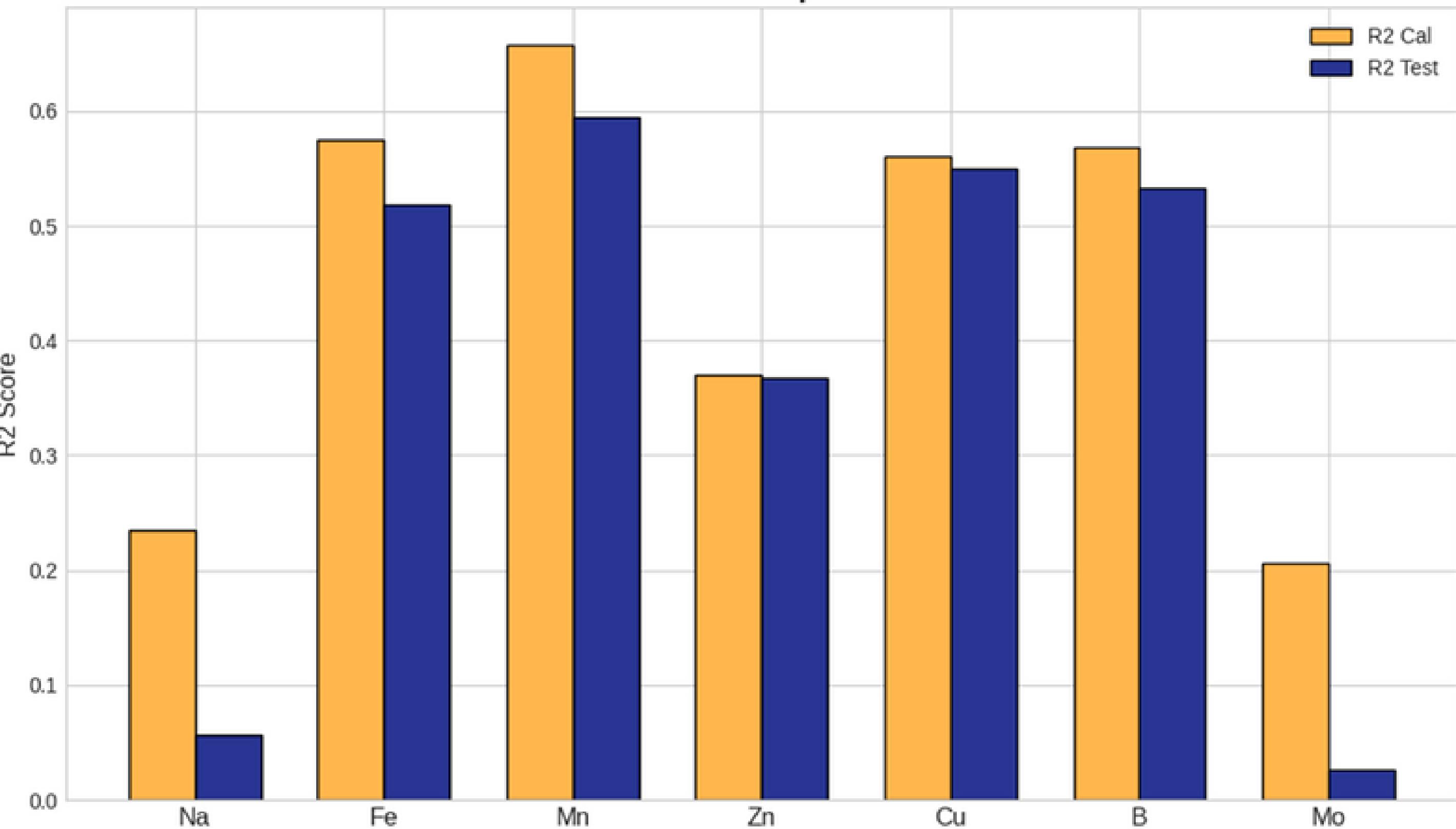
Selecció de bandes

	MODEL	BANDES	R ² Cal	MAE Cal	RMSE Cal	R ² Test	MAE Test	RMSE Test
	SVR	3	0,791	0,368	0,740	0,578	0,568	1,068
Ca	PLS Regression	6	0,778	0,457	0,742	0,587	0,665	1,057
	PLS Regression	12	0,809	0,397	0,712	0,595	0,606	1,059
	Random Forest	3	0,658	0,025	0,037	0,547	0,029	0,040
Mg	K-Neighbors	6	0,613	0,025	0,037	0,573	0,027	0,038
	K-Neighbors	12	0,566	0,027	0,039	0,572	0,028	0,039
	NuSVR	3	0,539	0,023	0,034	0,635	0,023	0,032
S	Ada Boost + Bag	6	0,590	0,022	0,033	0,648	0,020	0,029
	Linear SVR	12	0,607	0,024	0,034	0,653	0,022	0,030

MICRONUTRIENTS

Utilitzant totes les bandes

R2 Calibració vs R2 Prova per als micronutrients



MICRONUTRIENTS

Selecció de bandes

	MODEL	BANDES	R ² Cal	MAE Cal	RMSE Cal	R ² Test	MAE Test	RMSE Test
Ca	GradBoost + Bag	3	0,345	0,012	0,015	0,054	0,012	0,016
	K-Neighbors	6	0,381	0,010	0,014	0,055	0,013	0,018
	PLS Regression	12	0,377	0,011	0,014	0,025	0,013	0,017
Mg	Nu SVR	3	0,388	14,345	17,798	0,422	14,899	19,136
	Nu SVR	6	0,420	13,031	17,239	0,418	15,149	19,198
	Linear SVR	12	0,430	13,167	17,202	0,438	15,056	18,875
Mn	Ridge Regresion	3	0,522	4,076	5,664	0,460	4,325	5,820
	Ada Boost	6	0,590	3,614	5,041	0,578	4,003	5,409
	Ada Boost + Bag	12	0,522	4,059	5,388	0,564	4,148	5,493
Zn	SGD Regressor	3	0,201	7,430	9,392	0,138	7,262	8,916
	Extra Trees	6	0,216	6,930	8,638	0,225	7,590	9,794
	Nu SVR	12	0,291	6,858	8,832	0,167	7,034	8,762
Cu	PLS Regression	3	0,513	13,444	17,875	0,432	14,526	19,591
	Nu SVR	6	0,5620	11,714	16,957	0,441	13,965	19,429
	Nu SVR	12	0,551	10,984	17,164	0,475	13,215	18,838
B	SVR	3	0,498	4,981	6,874	0,533	4,912	6,559
	K-Neighbors	6	0,525	4,536	6,506	0,544	4,620	6,487
	PLS Regression	12	0,564	4,626	6,353	0,534	4,843	6,554
Mo	SGD Regressor	3	0,063	0,035	0,044	0,009	0,036	0,043
	SGD Regresion	6	0,073	0,035	0,043	0,010	0,036	0,043
	Ridge Regresion	12	0,073	0,035	0,043	0,010	0,036	0,043

CONCLUSIÓ FUTURA PROJECCIÓ



- S'obtenen resultats de R2 mitjà superiors a 0,6 per a macronutrients i R2 mitjà de 0,48 per a micronutrients.
- **Alternativa no destructiva, econòmica, ràpida i precisa** en comparació amb les tècniques tradicionals, que era l'**objectiu de eficiència i reducció de l'impacte ambiental**.
- Es **recomana ampliar les dades i considerar altres fonts** de dades per millorar la precisió dels models.
- Amb més dades es pot **explorar l'ús de tècniques d'aprenentatge profund**.
- Cal destacar la importància de fer que els **models siguin accessibles i fàcils d'utilitzar mitjançant ferramentes per als agricultors** per millorar la producció i reduir l'impacte ambiental.

GRÀCIES

★ GUILLEM MIRALLES GADEA

*Estimació Mitjançant Anàlisi No Destructiu Dels Nivells Nutricionals De
Fulles De Cítrics Aplicant Tècniques De Machine i Deep Learning*