

Estimación de los niveles nutricionales de hojas de cítricos mediante análisis no destructivo aplicando técnicas de aprendizaje automático

Guillem Miralles¹, Isabel Rodríguez-Carretero², Sergio Cubero³, Marcelino Martínez¹,
Fernando Mateo¹, Ana Quiñones², José Blasco³, Juan Gómez-Sanchis^{1*}

¹ IDAL, Departamento de Ingeniería Electrónica, Universidad de Valencia, Av. de la Universidad, S/N, 46100Burjassot, Valencia (Spain). Email: Juan.Gomez-Sanchis@uv.es

² Centro para el Desarrollo de la Agricultura Sostenible. Instituto Valenciano de Investigaciones Agrarias (IVIA), CV-315, km 10,7, 46113 Moncada (Valencia), Spain.

³ Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), CV-315, km 10,7, Moncada (Valencia) 46113, Spain.

Abstract: One of the characteristics of all living beings is that adequate nutrition has a positive impact on health. In the case of plants, and specifically in fruit trees, adequate nutrition is also essential for them to grow healthy and produce fruits in the highest quantity and quality possible. Therefore, optimal nutrition is key for any farmer. However, excessive use of fertilisers can harm the environment and be a waste of resources for farmers. One of the keys to achieving adequate fertilisation is an accurate diagnosis of the nutritional status of the tree. Traditionally, this diagnosis is made by destructive ionomics analysis, which represents a high economic cost and a delay in obtaining the results. This work proposes Vis-NIR hyperspectral imaging and machine learning regression models to estimate the concentrations of macronutrients (N, P, K, and Ca) and micronutrients (Mn and Fe) in citrus leaves. The methodology involved the application of several machine learning regression methods (linear regression, partial least squares, random forest, support vector regression, and Ada Boost). Data were normalised with standard normal variable (SNV), and principal component analysis (PCA) was used to reduce dimensionality. The results were promising in estimating nutrients with R^2 greater than 0,50 in all cases, especially nitrogen (R^2 of 0.77).

Palabras clave: aprendizaje automático, imagen hiperespectral, espectroscopia, agricultura de precisión, cítricos

1. Introducción

La integración de las tecnologías de la información (TIC) en la agricultura ha optimizado la toma inteligente de decisiones basadas en datos, impulsando la productividad y minimizando el impacto ecológico de los procesos agrícolas [1]. Las plantas requieren nutrientes esenciales para crecer de manera saludable y proporcionar productos de calidad [2]. Por lo tanto, una nutrición óptima es clave para cualquier agricultor. Sin embargo, el uso excesivo de fertilizantes puede dañar el medio ambiente y representar una pérdida de recursos para los agricultores [3]. Una de las claves para lograr una fertilización adecuada es un diagnóstico preciso del estado nutricional del árbol. Tradicionalmente, este diagnóstico se realiza mediante análisis destructivos de la ionómica, lo que representa un costo económico elevado y un retraso en la obtención de los resultados.

La imagen hiperespectral es una técnica que captura información en múltiples bandas del espectro electromagnético. Esto permite identificar y distinguir variaciones en la reflectancia o emisión de energía de los objetos en la escena relacionada con el estado de la planta, lo que es útil

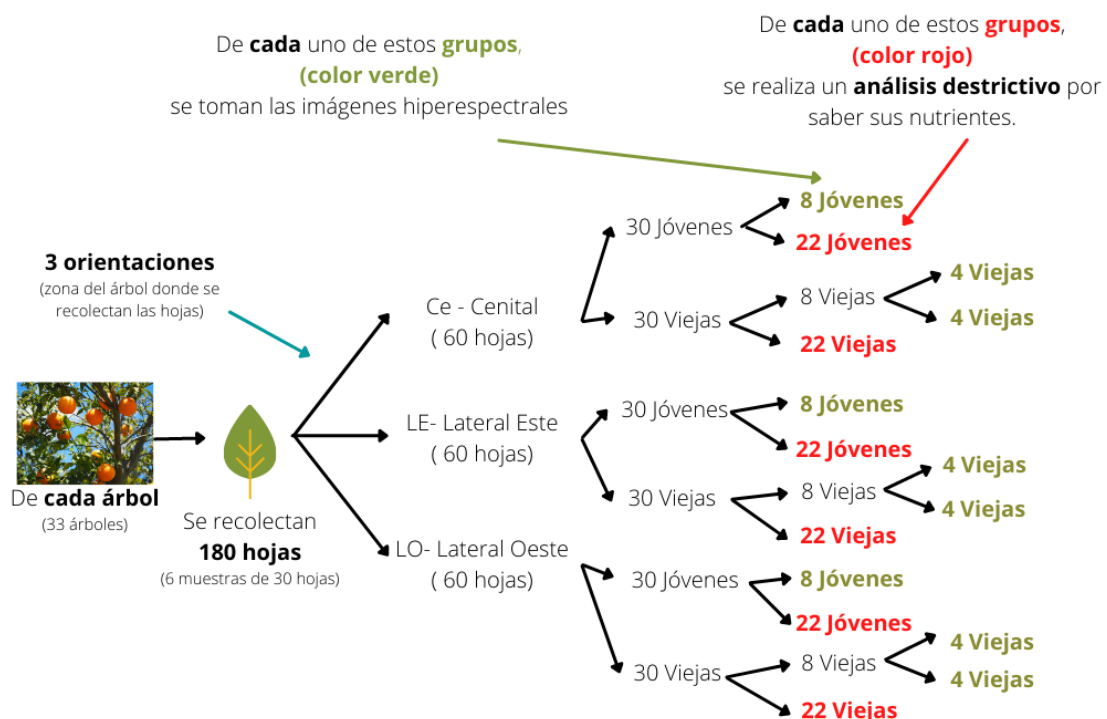
para el análisis de los cultivos [4]. Debido a su amplio rango espectral, las imágenes hiperespectrales pueden proporcionar información rápida sobre las propiedades bioquímicas de los cultivos, como la concentración de clorofila, contenido de humedad, estado vegetativo y contenido de nutrientes. Esto facilita la monitorización del estado de los cultivos, la detección temprana de enfermedades o estreses y la optimización de los programas de riego y fertilización. Este trabajo propone el uso de imágenes hiperespectrales Vis-NIR y modelos de regresión de aprendizaje automático para estimar las concentraciones de macronutrientes (N, P, K y Ca) y micronutrientes (Mn y Fe) en hojas de cítricos, a partir de los datos obtenidos de las imágenes y modelos de regresión de aprendizaje automático.

2. Materiales y métodos

2.1. Diseño Experimental

El material vegetal procede de una parcela de cítricos en Almenara (Castellón) de la variedad Clemenules. En mayo de 2021, se tomaron muestras de 33 árboles, 3 muestras de 60 hojas en diferentes orientaciones: este, oeste y cenital. De cada orientación, se seleccionaron 30 hojas jóvenes (brotación de primavera) y 30 de hojas viejas, con un total de 5940 hojas. De cada muestra, se destinaron 22 hojas para la obtención de los niveles de referencia de nutrientes mediante un análisis de la ionómica foliar y otras 8 se utilizaron para adquirir imágenes hiperespectrales. Debido a su diferente tamaño, en el caso de las hojas viejas, se tomaron 2 imágenes en grupos de 4 hojas; y en las jóvenes, una sola imagen de las 8. Así, se obtuvieron 3 imágenes hiperespectrales por árbol y orientación: hojas jóvenes, primera partición de hojas viejas y segunda partición de hojas viejas. El proceso se describe en la siguiente Figura 1.

Figura 1. Esquema del proceso de recolección de datos hiperespectrales y nutricionales.



2.2. Toma y procesamiento de imágenes

Se utilizó un sistema de adquisición de imágenes hiperespectrales en el rango Vis-NIR, consistente en una cámara industrial y dos filtros sintonizables de cristal líquido (LCTF). Se

capturaron imágenes de 1392 x 1040 píxeles con una resolución espacial de 0,14 mm/píxel en el rango espectral de 400 nm a 1050 nm, en pasos de 10 nm. Cada imagen se componía de 65 bandas.

Se utilizó un sistema basado en iluminación halógena difusa, compuesto por 12 lámparas halógenas de 35 W situadas dentro de una campana semiesférica. Para evitar problemas de imágenes borrosas debido a la dispersión cromática de la lente, se estableció el enfoque de la misma en la longitud de onda central (730 nm) del rango de trabajo. Se adquirió una imagen hiperspectral de un blanco de referencia certificada para optimizar el rango dinámico, evitar la saturación de la imagen y corregir la sensibilidad espectral del sistema.

Después de obtener las imágenes hiperspectrales, se realizaron tres etapas de procesamiento digital para extraer el espectro de cada hoja. Primero, se calibraron por referencia blanca y negra las imágenes registradas. Luego, se corrigió el desplazamiento espacial entre las imágenes. Debido a que las imágenes se obtuvieron con dos LCTF, existía una ligera variación entre ellas ocasionada por la geometría del sistema. Finalmente, se segmentó la hoja del resto de la escena mediante una máscara de umbralización y se obtuvo el espectro promedio de los píxeles de la hoja. Este espectro promedio constituyó la entrada a la etapa de preprocesado espectral y aprendizaje automático. La metodología de preprocesado se detalla en Fazari et al. [5].

2.3. Conjuntos de calibración y prueba, preprocesado espectral e ingeniería de características.

Para evaluar el rendimiento de los modelos de regresión, el conjunto inicial se dividió en un conjunto de calibración (Cal) (75 %) y un conjunto de prueba independiente (Test) (25 %). Se realizó un preprocesado espectral simple. En particular, se normalizaron las bandas espectrales estandarizándolas con el fin de uniformizar el rango de variación de las mismas. El método empleado fue standard normal variate (SNV) [6]. También se exploró una opción simple de ingeniería de características (ICR) usando el análisis de componentes principales (PCA) [7]. Se evaluaron modelos de regresión empleando como entrada diferente número de componentes principales (5, 7, 9, 12, 15, 18, 20, 25 y 30, estos números se determinaron de forma empírica hasta no conseguir mejora en el rendimiento de los métodos de regresión) y con todas las bandas sin aplicar PCA.

2.4. Modelos

En este estudio, se utilizaron cinco técnicas de regresión del campo del aprendizaje automático entre los que se incluyen la Regresión Lineal, partial least squares (PLS), support vector regression (SVR), Random Forest y Ada Boost. Todos ellos se implementaron en el lenguaje de programación Python con la ayuda de la librería Scikit-learn.

- **Regresión Lineal**, se usó como método de referencia para comparar con las otras técnicas empleadas en el trabajo. Basa su funcionamiento en suponer una relación lineal entre las variables independientes y la variable dependiente, su principal ventaja radica en la interpretabilidad del modelo [8].
- **PLS**, es un método de regresión especialmente indicado cuando las variables de entrada del problema están altamente correlacionadas. El método permite modelizar el problema reduciendo el conjunto de variables a un conjunto más pequeño de componentes no correlacionados [9].
- **SVR**, es una técnica basada en las máquinas de vectores de soporte. Esta técnica de regresión es capaz de gestionar relaciones no lineales entre las variables entrada y la variable independiente del problema, proporcionando una solución robusta frente a posibles *outliers* [10].

- **Random Forest**, es un método de regresión basado en la combinación (*ensemble methods*) de árboles de decisión. Se caracteriza por agregar múltiples soluciones para mejorar la predicción y la robustez de la solución [11].
- **Ada Boost** es un método de regresión basado en *boosting* que se basa en la idea de crear un modelo fuerte a partir de varios modelos débiles [12].

2.5. Construcción de los modelos.

Para garantizar una correcta construcción de los modelos, se empleó el conjunto de calibración. Con el fin de evitar el sobreajuste de los modelos y conseguir una combinación óptima de hiperparámetros en la fase de entrenamiento, se adoptó un enfoque de validación cruzada *K-Fold* [13] con 3 subconjuntos (*folds*). Se empleó técnica *Grid Search* [14] para conseguir una combinación adecuada de hiperparámetros. Los rangos de variación de los hiperparámetros (en aquellos modelos con hiperparámetros libres) escogidos de forma empírica para cada modelo fueron:

- **PLSRegression**: *n_components* [1-25], *scale* [False, True], *max_iter* [100-1000], *tol* [0,1-0,00001].
- **SVR**: *kernel* ['linear', 'rbf', 'sigmoid'], *gamma* ['scale', 'auto'], *C* [0,1-250], *epsilon* [0,1-0,00001].
- **RandomForest**: *n_estimators* [5-500], *max_depth* [10, 50, 100, None], *min_samples_split* [2-10], *min_samples_leaf* [1-4], *max_features* ['auto', 'sqrt', 'log2', None], *bootstrap* [True, False].
- **AdaBoost**: *n_estimators* [5-500], *learning_rate* [0,001-1], *loss* ['linear', 'square', 'exponential'].

Los análisis y experimentos fueron llevados a cabo en el entorno de *Google Colab* en la nube, que proporciona recursos computacionales para la ejecución de tareas de procesamiento de datos. Este entorno en la nube permitió realizar cálculos y ejecutar modelos de manera eficiente y escalable.

3. Resultados y discusión

3.1 Macronutrientes.

La Tabla 1 muestra los resultados relativos a la estimación de los macronutrientes nitrógeno, fósforo, potasio y calcio. Se muestran diferentes índices de rendimiento de los modelos, en particular el coeficiente de determinación R^2 , el error absoluto medio (MAE) y el error cuadrático medio (RMSE) [15], tanto en el conjunto de calibración como en el de prueba.

Se observa que, para el nitrógeno todos los modelos presentan una mejora respecto a la regresión lineal. Esta tendencia se va a mantener en todos los nutrientes incluidos en el estudio. La metodología basada en *Random Forest* empleando como entrada todas las bandas, presenta los mejores resultados en el conjunto de prueba para el nitrógeno (R^2 de 0,77). Respecto al fósforo, la metodología que proporciona mejores resultados es la combinación de *Ada Boost* empleando como entradas 18 componentes principales consiguiendo un coeficiente de determinación R^2 de 0,66. En relación con el potasio, la mejor metodología es SVR empleando 9 componentes principales con una R^2 de 0,61. Del mismo modo, *Ada Boost* con 20 componentes principales es la metodología que mejores resultados proporciona para el calcio con un valor de R^2 de 0,62. La capacidad de estimación de las mejores metodologías para todos los macronutrientes es buena, con valores del coeficiente de determinación por encima de 0,6 en todos los casos.

Las mejores metodologías de estimación extraídas del análisis del coeficiente de determinación para los macronutrientes se mantienen si se analizan los otros índices de rendimiento calculados. Así, si se examinan los valores de MAE y RMSE para cada macronutriente, se observa que los valores mínimos de los errores de estimación coinciden con los valores máximos de R^2 . Esto pone

XII CONGRESO IBÉRICO DE AGROINENIERÍA
XII CONGRESSO IBÉRICO DE AGROENGENHARIA
4 – 6 septiembre 2023, Sevilla - España

de manifiesto la robustez de los resultados en cuanto a la estimación del rendimiento de los modelos.

Tabla 1. Resultados de los modelos para los macronutrientes N, P, K y Ca. La columna FE muestra la mejor opción de ingeniería de características (ICR), indicando el número de componentes principales (CP) o empleando todas las bandas (No). En negrita se muestra la mejor metodología para cada macronutriente en base al valor de R^2 en el conjunto de prueba.

	MODELO	ICR	R^2 Cal	MAE Cal	RMSE Cal	R^2 Test	MAE Test	RMSE Test
N	LinearRegresion	5 CP	0,585	0,132	0,180	0,335	0,145	0,223
	PLSRegression	7 CP	0,724	0,120	0,149	0,726	0,112	0,143
	SVR	No	0,792	0,102	0,129	0,677	0,122	0,155
	RandomForest	No	0,813	0,087	0,123	0,772	0,102	0,130
	AdaBoost	No	0,793	0,101	0,129	0,761	0,105	0,133
P	LinearRegresion	5 CP	0,115	0,030	0,050	0,210	0,035	0,055
	PLSRegression	7 CP	0,667	0,021	0,032	0,604	0,025	0,038
	SVR	9 CP	0,742	0,017	0,028	0,586	0,027	0,039
	RandomForest	12 CP	0,686	0,020	0,031	0,562	0,026	0,034
	AdaBoost	18 CP	0,700	0,020	0,030	0,657	0,025	0,033
K	LinearRegresion	5 CP	0,482	0,195	0,274	0,268	0,257	0,331
	PLSRegression	5 CP	0,575	0,171	0,246	0,618	0,153	0,241
	SVR	9 CP	0,623	0,144	0,232	0,614	0,144	0,239
	RandomForest	No	0,658	0,132	0,222	0,510	0,179	0,271
	AdaBoost	12 CP	0,596	0,156	0,241	0,597	0,167	0,246
Ca	LinearRegresion	5 CP	0,340	0,724	1,214	0,528	0,791	1,123
	PLSRegression	7 CP	0,730	0,504	0,763	0,537	0,675	1,113
	SVR	12CP	0,794	0,353	0,662	0,595	0,623	1,041
	RandomForest	No	0,747	0,399	0,732	0,533	0,693	1,118
	AdaBoost	20 CP	0,739	0,405	0,741	0,620	0,612	1,008

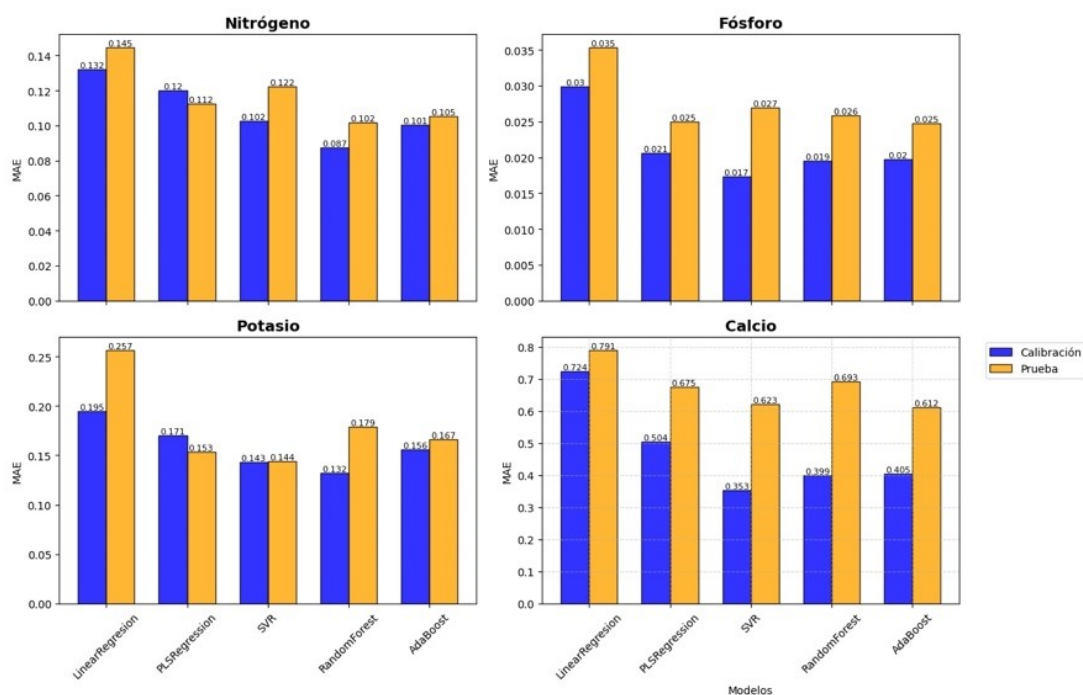


Figura 2. Comparación del R^2 obtenido calibración y prueba de los diferentes modelos que predicen los macronutrientes.

La figura 2 muestra, en un gráfico de barras, la comparación entre el MAE de los conjuntos de calibración y prueba para cada macronutriente con cada una de las metodologías empleadas. La pauta general es que el error producido en los conjuntos de calibración y prueba son muy similares (dentro de cada metodología). Este hecho pone de manifiesto que no ha habido sobreajuste en la construcción de los modelos y que la capacidad de generalización de los modelos es óptima.

3.2 Micronutrientes.

La Tabla 2 muestra los resultados relativos a la estimación de los micronutrientes manganeso, y hierro. Se muestran los mismos índices de rendimiento que en el caso de los macronutrientes. Se observa que, para el manganeso todos los modelos presentan una mejora respecto a la regresión lineal. Esta tendencia se mantiene también con el hierro. Las metodologías que mejores resultados proporcionan para el manganeso y el hierro son *Random Forest* y *Adaboost* respectivamente ($R^2=0,57$ y $R^2=0,52$). La capacidad de estimación de las mejores metodologías para los micronutrientes es menor que en el caso de los macronutrientes, no consiguiéndose valores del coeficiente de determinación mayores a 0,57 en el conjunto de prueba.

Al igual que en el caso de los macronutrientes, las mejores metodologías de estimación extraídas del análisis del coeficiente de determinación para los micronutrientes se mantienen si se analizan los otros índices de rendimiento calculados. Si se estudian los valores de MAE y RMSE para cada micronutriente, se observa que los valores mínimos de los errores de estimación coinciden con los valores máximos de R^2 , esto pone de manifiesto la robustez de los resultados en cuanto a la estimación del rendimiento de los modelos.

Tabla 2. Resultados de los modelos para los micronutrientes Mn y Fe. La columna ICR muestra la mejor opción de ingeniería de características, indicando el el número de componentes principales (CP) o empleando todas las bandas (No). En negrita se muestra la mejor metodología para cada micronutriente en base al valor de R^2 en el conjunto de prueba.

	MODELO	ICR	R^2 Cal	MAE Cal	RMSE Cal	R^2 Test	MAE Test	RMSE Test
Mn	LinearRegresion	5 CP	0,331	4,838	6,081	0,277	5,384	8,321
	PLSRegression	5 CP	0,525	3,866	5,206	0,569	3,988	4,810
	SVR	5 CP	0,517	3,713	5,236	0,566	3,861	4,862
	RandomForest	No	0,535	3,777	5,174	0,570	3,832	4,808
	AdaBoost	No	0,543	3,671	5,130	0,391	4,463	5,718
Fe	LinearRegresion	5 CP	0,318	14,337	18,542	0,367	16,070	20,537
	PLSRegression	25 CP	0,410	12,990	17,359	0,465	15,590	18,876
	SVR	5 CP	0,430	12,518	17,089	0,371	14,989	20,479
	RandomForest	25 CP	0,511	12,351	15,816	0,496	14,675	18,326
	AdaBoost	25 CP	0,495	12,973	16,091	0,518	14,298	17,917

La figura 3 muestra la comparación entre el MAE de los conjuntos de calibración y prueba para cada micronutriente con cada una de las metodologías empleadas. Al igual que ocurría en el caso de los macronutrientes, los modelos no muestran sobreajuste ya que el error de calibración y test son muy similares.

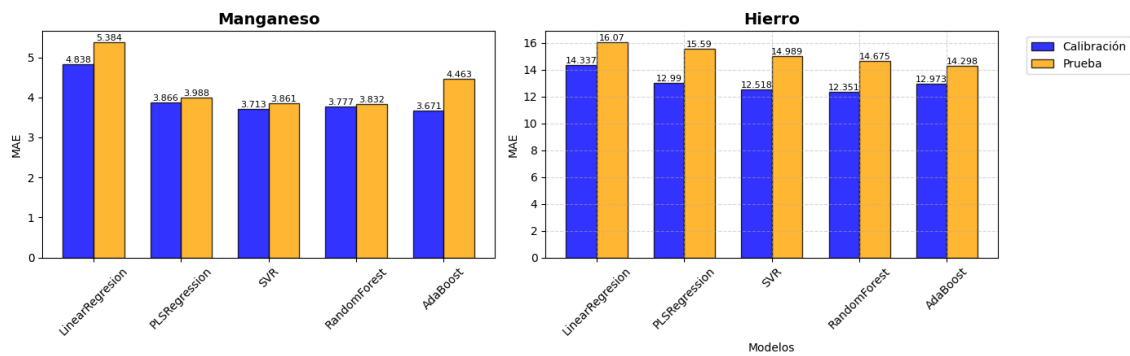


Figura 3. Comparación del R^2 calibración y prueba de los modelos que estiman el hierro y el manganeso.

La media de R^2 de los resultados en test en los micronutrientes es menor respecto a los macronutrientes (0,48 y 0,66 respectivamente). Esto se debe a la mayor complejidad y variabilidad en la determinación precisa de los niveles de micronutrientes. La estimación de micronutrientes presenta desafíos debido a su menor concentración y a la interacción con otros elementos. Aun así, los resultados obtenidos con las técnicas de análisis no destructivo muestran un progreso prometedor en la determinación de los niveles nutricionales de las hojas de cítricos, lo que respalda su potencial aplicación en futuras investigaciones agronómicas y nutricionales.

4. Conclusiones

La metodología propuesta, basada en el uso de un sistema de visión hiperspectral Vis-NIR y técnicas de regresión de aprendizaje automático, permite estimar el contenido del nitrógeno con un R^2 de 0,77. Para el fósforo, potasio y calcio, se obtuvieron valores de R^2 de 0,66, 0,61 y 0,62 respectivamente. En el caso de los dos micronutrientes probados, manganeso y hierro, se obtuvieron valores de 0,57 y 0,51 respectivamente, lo que indica una mayor dificultad en su estimación.

En el caso de los micronutrientes, la utilización de visión hiperspectral y los modelos de aprendizaje automático no demuestran ser una alternativa realmente útil para la estimación de los niveles de los nutrientes en las hojas de los cítricos, ya que en la mayoría de los casos los valores de R^2 se encuentran próximos a 0.5 e incluso por debajo. Esto indica que los modelos son capaces de explicar a lo sumo, el 50 % de la variabilidad en los datos de salida; es decir, que en torno al 50% de la variación en los valores predichos se puede atribuir a las variables empleadas en el modelo, mientras que el 50 % restante se debe a factores desconocidos.

La utilización del sistema de visión hiperspectral Vis-NIR y los modelos de regresión de aprendizaje automático se demuestran como una alternativa útil para estimar los niveles de nutrientes en las hojas de cítricos. Estos métodos ofrecen una solución no destructiva, más económica, rápida y precisa en comparación con las técnicas destructivas, costosas y lentas, cumpliendo así nuestro objetivo principal de eficiencia y reducción del impacto ambiental.

5. Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos TED2021-130117B-C31 y C33, financiados por MCIN/AEI/1 0,13039/501100011033 y por la UE "NextGenerationEU"/PRTR, y los proyectos GVA-IVIA 52203, 52204 y la UE a través del FEDER de la Generalitat Valenciana 2021-2027.

Referencias:

1. Geovanny Rambauth-Ibarra, "Agricultura de Precisión: La integración de las TIC en la producción Agrícola", J. Comput. Electron. Sci.: Theory Appl., vol. 3 no. 1 pp. 34-38. January - June, 2022, doi: <http://dx.doi.org/10.17981/cesta.03.01.2022.04>
2. Petra Marschner, "Marschner's Mineral Nutrition of Higher Plants", (3rd ed.), Elsevier, 2012, doi: <https://doi.org/10.1016/C2009-0-63043-9>
3. Jiao Chen, Shaoyu Lü, Zhe Zhang, Xuxia Zhao, Xinming Li, Piao Ning, Mingzhu Liu, "Environmentally friendly fertilisers: A review of materials used and their effects on the environment", Science of The Total Environment, Volumes 613–614, 2018, Pages 829-839, ISSN 0048-9697, doi: <https://doi.org/10.1016/j.scitotenv.2017.09.186>
4. Rafael Alejandro Casillas-Peñuelas, "Hyperspectral imaging analysis and applications for food quality", CRC Press, 2018, doi: <https://doi.org/10.1201/9781315209203>
5. Antonio Fazari, Oscar J. Pellicer-Valero, Juan Gómez-Sanchis, Bruno Bernardi, Sergio Cubero, Souraya Benalia, Giuseppe Zimbalatti, Jose Blasco, "Application of deep convolutional neural networks for the detection of anthracnose in olives using VIS/NIR hyperspectral images", Computers and Electronics in Agriculture, Volume 187, 2021, 106252, ISSN 0168-1699, doi: <https://doi.org/10.1016/j.compag.2021.106252>
6. Emily Grisanti, Maria Totska, Stefan Huber, Christina Krick Calderon, Monika Hohmann, Dominic Lingenfelser, Matthias Otto, "Dynamic Localized SNV, Peak SNV, and Partial Peak SNV: Novel Standardization Methods for Preprocessing of Spectroscopic Data Used in Predictive Modeling", Journal of Spectroscopy, vol. 2018, Article ID 5037572, 14 pages, 2018, doi: <https://doi.org/10.1155/2018/5037572>
7. Dong, Y., Shan, Y., Li, P., Jiang, L., & Zhang, X, "Nondestructive Characterisation of Citrus Fruit by near-Infrared Diffuse Reflectance Spectroscopy (NIRDRS) with Principal Component Analysis (PCA) and Fisher Linear Discriminant Analysis (FLDA)". Foods, 11(4), 1-16. 2022, doi: <https://doi.org/10.1080/00032719.2022.2063306>
8. Bishop, Christopher M., "Linear Models for Regression" in Pattern Recognition and Machine Learning, New York: Springer, 2006, doi: <https://doi.org/10.5555/1162264>
9. Vinzi, V. E., Chin, W. W., Henseler, J., & Wang, H, "Handbook of Partial Least Squares: Concepts, Methods and Applications", Springer, 2010, doi: <https://doi.org/10.1007/978-3-540-32827-8>
10. Bishop, Christopher M., "Support Vector Machines for Regression" in Pattern Recognition and Machine Learning, New York: Springer, 2006, doi: <https://doi.org/10.5555/1162264>
11. Trevor Hastie, Robert Tibshirani, Jerome Friedman, "Random Forests" in The elements of statistical learning: data mining, inference, and prediction, New York : Springer, Second Edition, 2009, doi: <https://doi.org/10.1007/978-0-387-84858-7>
12. Trevor Hastie, Robert Tibshirani, Jerome Friedman, "Boosting and Additive Trees" in The elements of statistical learning: data mining, inference, and prediction, New York : Springer, Second Edition, 2009, doi: <https://doi.org/10.1007/978-0-387-84858-7>

XII CONGRESO IBÉRICO DE AGROINENIERÍA
XII CONGRESSO IBÉRICO DE AGROENGENHARIA
4 – 6 septiembre 2023, Sevilla - España

13. Refaeilzadeh, P., Tang, L., Liu, H., "Cross-Validation " In: LIU, L., ÖZSU, M.T. (eds) Encyclopedia of Database Systems. Springer, Boston, MA, doi: https://doi.org/10.1007/978-0-387-39940-9_565
14. Feurer, M., Hutter, F., "Hyperparameter Optimization". In: Hutter, F., Kotthoff, L., Vanschoren, J. (eds) Automated Machine Learning. The Springer Series on Challenges in Machine Learning. Springer, Cham, 2019, doi: https://doi.org/10.1007/978-3-030-05318-5_1
15. Kuhn, M., Johnson, K. (2013). Measuring Performance in Regression Models. In: Applied Predictive Modeling. Springer, New York, NY, doi: https://doi.org/10.1007/978-1-4614-6849-3_5